UNIVERSITY OF CANTERBURY

DOCTORAL THESIS

Development and Evaluation of Mixed Reality-Enhanced Robotic Systems for Intuitive Tele-manipulation and Tele-manufacturing Tasks in Hazardous Conditions

Yunpeng Su

A thesis submitted in partial fulfillment of the requirements for the Degree
of Doctor of Philosophy of Mechanical Engineering
in the University of Canterbury
Christchurch, New Zealand

Jan 28, 2022

Dedicated to:

My dear parents and partner

For their love, encouragement and companionship

Acknowledgments

I would like to express my great gratitude to my senior supervisor, Professor Geoff Chase from the Mechanical Engineering Department at the University of Canterbury (UC) for his guidance and support throughout my PhD study and for being there supporting me whenever I have needed him. It is he who has given me hope when my life has had ups and downs. Without his support and inspiration, this achievement would not have happened.

Professor XiaoQi Chen from the Mechanical Engineering Department at the Swinburne University of Technology played a huge role in providing me with the opportunity to study in New Zealand and the right direction to walk into robotics and mixed reality with confidence. I would not have been able to turn my passion into doctoral research without his guidance. He has been my role model in working and communicating wisely across different cultures.

I also feel extremely grateful to my co-supervisor, Professor Chris Pretty from the Department of Mechanical Engineering at UC, for his invaluable advice and assistance over the years.

I would like to express my huge appreciation to the technical staff of the UC Mechanical Engineering workshop, Gerry Kirk, Julian Phillips, Rodney Elliott, and Garry Cotton, who contributed essential technical work and guidance throughout this study.

I wish to thank the visiting students I supervised during my study, Bertrand, Bastien, Sasha, Edgar, Yann, George and Benjamin for their great interest in this research project and for accompanying me along the journey.

But above all, I would not have been able to study here without the unyielding support of the rock of my family. I want to thank my parents for bringing me into this wonderful world and providing their endless love and care. I would love to thank my partner from the bottom of my heart. I am forever grateful for what she has sacrificed to support me. We have walked through many difficulties in our overseas life together. Her patience, kindness and tolerance have made me want to be a reliable and responsible person. I feel a deep sense of gratitude to my middle school English teacher, Cuiyu Zhang, for always being there for me and helping me out since I met her when I was 13. They are the ones who have encouraged me to go confidently in the direction of my dreams. My debt to them is beyond measure.

Abstract

In recent years, with the rapid development of space exploration, deep-sea discovery, nuclear rehabilitation and management, and robotic-assisted medical devices, there is an urgent need for humans to interactively control robotic systems to perform increasingly precise remote operations. The value of medical telerobotic applications during the recent coronavirus pandemic has also been demonstrated and will grow in the future. This thesis investigates novel approaches to the development and evaluation of a mixed reality-enhanced telerobotic platform for intuitive remote teleoperation applications in dangerous and difficult working conditions, such as contaminated sites and undersea or extreme welding scenarios. This research aims to remove human workers from the harmful working environments by equipping complex robotic systems with human intelligence and command/control via intuitive and natural human-robot-interaction, including the implementation of MR techniques to improve the user's situational awareness, depth perception, and spatial cognition, which are fundamental to effective and efficient teleoperation.

The proposed robotic mobile manipulation platform consists of a UR5 industrial manipulator, 3D-printed parallel gripper, and customized mobile base, which is envisaged to be controlled by non-skilled operators who are physically separated from the robot working space through an MR-based vision/motion mapping approach. The platform development process involved CAD/CAE/CAM and rapid prototyping techniques, such as 3D printing and laser cutting. Robot Operating System (ROS) and Unity 3D are employed in the developing process to enable the embedded system to intuitively control the robotic system and ensure the implementation of immersive and natural human-robot interactive teleoperation.

This research presents an integrated motion/vision retargeting scheme based on a mixed reality subspace approach for intuitive and immersive telemanipulation. An imitation-based velocity-centric motion mapping is implemented via the MR subspace to accurately track operator hand movements for robot motion control, and enables spatial velocity-based control of the robot tool center point (TCP). The proposed system allows precise manipulation of end-effector position and orientation to readily adjust the corresponding velocity of maneuvering.

A mixed reality-based multi-view merging framework for immersive and intuitive telemanipulation of a complex mobile manipulator with integrated 3D/2D vision is presented. The proposed 3D immersive telerobotic schemes provide the users with depth perception through the merging of multiple 3D/2D views of the remote environment via MR subspace. The mobile manipulator platform can be effectively controlled by non-skilled operators who are physically separated from the robot working space through a velocity-based imitative motion mapping approach.

Finally, this thesis presents an integrated mixed reality and haptic feedback scheme for intuitive and immersive teleoperation of robotic welding systems. By incorporating MR technology, the user is fully immersed in a virtual operating space augmented by real-time visual feedback from the robot working space. The proposed mixed reality virtual fixture integration approach implements hybrid haptic constraints to guide the operator's hand movements following the conical guidance to effectively align the welding torch for welding and constrain the welding operation within a collision-free area.

Overall, this thesis presents a complete tele-robotic application space technology using mixed reality and immersive elements to effectively translate the operator into the robot's space in an intuitive and natural manner. The results are thus a step forward in cost-effective and computationally effective human-robot interaction research and technologies. The system presented is readily extensible to a range of potential applications beyond the robotic tele-welding and tele-manipulation tasks used to demonstrate, optimise, and prove the concepts.

Publications

Journal Articles

Su, Yunpeng, Xiaoqi Chen, Tony Zhou, Christopher Pretty, and Geoffrey Chase. 2021. "Mixed Reality-Enhanced Intuitive Teleoperation with Hybrid Virtual Fixtures for Intelligent Robotic Welding" Applied Sciences 11, no. 23: 11280. https://doi.org/10.3390/app112311280

Su, Yunpeng, Xiaoqi Chen, Tony Zhou, Christopher Pretty, and Geoffrey Chase" Mixed Reality-Integrated 3D/2D Vision Merging and Mapping for Intuitive Teleoperation of Mobile Manipulator " Robotics and Computer-Integrated Manufacturing (In Revision)

Su, Yunpeng, Xiaoqi Chen, Tony Zhou, Christopher Pretty, and Geoffrey Chase" Mixed Reality-Enhanced Human-Robot Interaction with Imitation-Based Mapping Approach for Intuitive Teleoperation of Robotic Arm-Hand System" IEEE/ASME Transactions on Mechatronics (Submitted)

Su, Yunpeng, Xiaoqi Chen, Tony Zhou, Christopher Pretty, and Geoffrey Chase" Mixed Reality-Enhanced Intuitive Tele-welding with Robotic Haptic Assistance " IEEE/ASME Transactions on Mechatronics (Prepared)

Conference Proceedings

Su, Yunpeng, M. Ahmadi, Christoph Bartneck, Frank Steinicke and Xiaoqi Chen. “Development of an Optical Tracking Based Teleoperation System with Virtual Reality.” 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA) (2019): 1606-1611.

Book Chapters

Su, Yunpeng, and Xiaoqi Chen. “Chapter 8. Applications of Virtual Reality Based Intelligent Teleoperation.” Essay. In Mechatronic Systems: Design, Performance and Applications, edited by Mohamed Arezki Mellal, 277–310. Nova Science Publishers Incorporated, 2018.

Contents

List of Abbreviations

Abbreviation	Definition
3D	Three-Dimensional
3DP	Three-Dimensional Point Cloud
3DS	Three-Dimensional Stereoscopic Vision
AI	Artificial Intelligence
ANOVA	Analysis of Variance
API	Application Programming Interface
AR	Augmented Reality
AV	Augmented Virtuality
CAD	Computer-Aided Design
CAE	Computer-Aided Engineering
CAM	Computer-Aided Manufacturing
CTWD	Contact Tip-to-Work Distance
DC	Direct Current
DH	Denavit–Hartenberg
DOF	Degree of Freedom
GMAW	Gas Metal Arc Welding
GPU	Graphics Processing Unit
GTAW	Gas Tungsten Arc Welding
HIP	Haptic Interface Point
HITL	Human-in-the-Loop
HMD	Head-Mounted Display
HMM	Hidden Markov Model
HRI	Human-Robot Interaction
HRIT	Human-Robot Interactive Teleoperation
IPD	Interpupillary Distance
IR	Infrared
IRT	Immersive Robotic Teleoperation
JSON	JavaScript Object Notation
MHPR	Master-Controlled HIP and Proxy-Controlled Robot
MR	Mixed Reality
MRD	Mixed Reality 3D Mouse
MRS	Mixed Reality Subspace
MRVF	Mixed Reality Virtual Fixture
NASA	National Aeronautics and Space Administration
PLA	Polylactic Acid
RGB	Red Green Blue

RGB-D	Red Green Blue Depth
RMSE	Root Mean Square Error
ROS	Robot Operating System
RTWS	Robotic Tele-Welding Systems
TAM	Technology Acceptance Model
TCP	Tool Center Point
UR	Universal Robots
URDF	Unified Robot Description Format
VCMM	Velocity-Centric Motion Mapping
VF	Virtual Fixture
VR	Virtual Reality
XR	Extended Reality

Chapter 1

Introduction

1.1 Overview and Objectives

With the rapid development of space exploration, deep-sea discovery, nuclear rescue, radiation detection, and robot-assisted medical equipment in recent years, humans urgently need interactive control of 'slave' robots to complete remote operations [1], [2]. More recently, medical robotic applications during the coronavirus pandemic have proven valuable [3], [4]. Due to the highly contagious nature of the novel coronavirus, surgeons are at high risk of infection when examining and sampling patients face-to-face [5]. However, oropharyngeal swabbing is a commonly used technique for COVID-19 sampling and diagnosis in the pandemic worldwide [6], [7]. Thus, one application scenario of medical telerobotic systems is to teleoperate robots to conduct COVID-19 swab testing and provide other healthcare services, such as (1) robotics-assisted telesurgery, (2) tele-examination of patients before and after treatment, and (3) tele-training for surgical procedures [8], [9].

On the user side of biomedical telerobotic systems, surgeons can operate a Human-Robot Interaction (HRI) system with an MR Head-Mounted Display (HMD) and control robots from a distance to perform surgery [10]. Additionally, healthcare workers can telemanipulate robots for the care of infected patients or the collection of biological samples, which greatly reduces the risk of infection [11]. Telerehabilitation is another application scenario, the intuitive interaction system enables therapy providers to provide rehabilitation services/telerobotic

therapy to the patients remotely in their homes or other environments [12]. The providers can use their natural motion to guide the patients who are physically assisted by the robotic system placed by them [13]. Therefore, human-robot interactive teleoperation (HRIT) technology remains the main means to realize remote operations in a complex and dynamic environment.

Robotic teleoperation satisfies the demands of the scenarios in which human access is dangerous, but human intelligence is required [14]. Human-in-the-loop tele-control of robotic systems enables operators to remotely implement complex tasks to reduce risk without losing quality [15], [16]. An interactive teleoperation system consists of five main components, including human operator, master control loop, communication channel, slave control loop, and robotic agent. The operator passes commands to the slave loop via the communication channels [17], which also returns information on the robot's interaction with its environment [18], using visual and haptic feedback [19], [20]. An effective teleoperation system not only enables intuitive human-robot interaction, but ensures the robot can also be operated in a way allowing the operator to experience the "feel" of the robot working on the remote side [21], gaining a "sense of presence" [22], [23]. Mixed reality (MR) technology integrates real-world information with computer-generated graphics, and has the potential to enhance the effectiveness and performance of human-robot interaction (HRI) by providing depth perception and enabling judgment and decision-making while operating the robot in a dynamic environment [24].

This thesis presents the development and evaluation of an MR-enhanced telerobotic platform for intuitive remote manipulation and manufacturing applications in dangerous and unpleasant working conditions environments such as contaminated sites or welding scenarios. This research aims to remove human workers from the harmful working environments by equipping

complex robotic systems with human intelligence and command/control via human-robot-interaction, including the implementation of MR techniques to improve the user's situational awareness, depth perception, and spatial cognition, all of which are fundamental to effective and efficient teleoperation.

The emphasis is placed on intuitive and natural human-robot interaction and control in specific industry scenarios with an overall goal of a generalizable system applicable more broadly. The robotic platform is comprised of a UR5 robotic arm and a customized parallel gripper mounted on a differential-drive mobile platform for mobile manipulation tasks. The design and implementation process of such an intuitive telerobotic platform for immersive teleoperation requires hardware and software development. And intuitive mobile manipulation involves enabling unskilled users to interact with mobile manipulator platforms to perform manipulation and manufacturing tasks in an intuitive and natural manner.

The main challenges addressed in this thesis include:

- Design, development and intuitive control of an inexpensive and customized mobile manipulation robotic platform for remote operations that take full advantage of the simplicity, reliability and effectiveness of the robotic architectures.
- Developing effective methods for incorporating the multi-sensory feedback into the MR scenes for achieving immersive and interactive teleoperation schemes for multiple tasks.
- Implementation, improvement and experimental evaluation of MR-enhanced telerobotic platform for intuitive remote manipulation and manufacturing applications in specific industry scenarios.

1.2 Thesis Contribution

The main contributions of this thesis can be organized into three groups as follows:

Design and development of the inexpensive robotic platform for intuitive mobile manipulation

Design and control of a cost-effective robotic platform. As a fundamental component of the thesis, the mechatronic system design and development of a robotic mobile manipulation system as an intuitive human-robot interactive teleoperation platform is proposed. The robotic mobile manipulation platform consists of a UR5 industrial manipulator, 3D-printed parallel gripper, and customized mobile base. This base is envisaged to be controlled by non-skilled operators who are physically separated from the robot working space through an MR-based vision/motion mapping approach.

The platform development process involved CAD/CAE/CAM and rapid prototyping techniques, such as 3D printing and laser cutting. The differential-drive architecture is employed to build the cost-effective mobile module that provides mobility to the overall robotic platform and enables time-efficient prototyping. The robot gripper module was designed and manufactured using a cost-effective 3D printer using Prusament polylactic acid (PLA). The designed parallel gripper module utilizes the simplicity and reliability of the slider-crank mechanism to ensure the manipulation of objects in unstructured environments. Software development using Robot Operating System (ROS) and Unity 3D in the developing process enables the embedded system to intuitively control the robotic system and ensures the implementation of immersive and natural human-robot interactive teleoperation [25].

Modeling of kinematic models and implementation of velocity-based control is employed. The kinematic-based intuitive control approach is implemented based on the kinematic modeling of the mobile base and arm-hand modules of the robotic system. The translation and rotation of the mobile module concerning the world frame are derived by using forward kinematic, as mobile base odometry. The mobile base controller controls the desired speed of the two motors used for driving and steering the servo through the inverse kinematics. The use of different control strategies for the movement and manipulation modules leads to increased system complexity. However, and in contrast, velocity-based control provides smooth motion and is a suitable method of guiding the robotic systems intuitively by human operators [26]. In the proposed platform, both the mobile and manipulation modules are manipulated by velocity-based control in Cartesian coordinate space. The desired velocity commands given by the user are translated into differential joint velocity commands for the robot arm and driving and steering velocities for the mobile base.

Development of multiple human interaction schemes using various control input devices. Four human interaction schemes using various control input devices are incorporated into the proposed teleoperated robotic platform for the intuitive control of the arm-hand subsystem, which enhances the usability, compatibility and applicability of the system for a wide range of tele-manufacturing scenarios. The four types of intuitive teleoperation schemes include isometric-rate interaction, isotonic-position interaction, optical tracking-based interaction, and haptic effect-enhanced interaction.

Development and evaluation of mixed reality-enhanced natural and intuitive teleoperation in remote manipulation tasks

Mixed reality-based integrated vision and motion retargeting for intuitive spatial control for telemanipulation application is developed and performance evaluation is conducted for teleoperation efficiency and effectiveness. To achieve these goals, the 3D point cloud rendering architecture is deployed in the immersive interaction paradigm to form the incorporated 3D visualization of the remote robot working site, provide the desired depth perception of the workspace, and maintain the inspection of the workpieces, remote site and digital twin situated in the MR subspace as a whole. To analyze system performance, the isometric-rate tele-control paradigm through MR-based 3D motion and visualization retargeting scheme for telerobotics is compared to the isotonic-position 3D interaction condition and conventional 2D Baseline over two tabletop object manipulation experiments.

Mixed reality-integrated 3D/2D vision merging for vision mapping. Mixed reality-based multi-view merging for intuitive telemanipulation of mobile manipulator is investigated and metrics of operator performance are evaluated. An immersive spatially interactive MR subspace with multiple 3D/2D views was designed and evaluated to enable the effective remote manipulation of a robotic mobile manipulator. 3D point cloud display and stereo vision are employed as 3D visual inputs to augment the basic 2D views from the robot's wrist and external perspectives, respectively. Two immersive 3D/2D view merging modes are formed and compared to the Baseline situation in which only 2D views are displayed to the operator in MR. An experiment involving manipulation tasks of different difficulty levels was conducted to examine the potential benefits of MR-integrated 3D/2D vision merging and mapping approach for teleoperation.

Implementation and validation of mixed reality and virtual fixture integrated visual/haptic interaction in tele-welding scenarios

To enable the welders to conduct tele-welding in the same way it is performed onsite, the goal is to minimize the learning required by introducing a tele-welding robot in the loop, welding skill extraction while performing manual welding is carried out and the operational differences between unskilled and professional welders analyzed. As a basis for developing the intuitive tele-welding paradigm, the results of welding movement analysis are used to assist unskilled workers with integrated visual and haptic HRI modalities via MR to improve task performance and system usability in tele-welding and to achieve welding results from unskilled operators that are comparable to those of professional welders.

Mixed reality and virtual fixture enhanced tele-welding. An innovative integrated application of mixed reality and hybrid virtual fixtures to assist tele-welding tasks is developed. The mixed reality and virtual fixture integration approach implement hybrid haptic constraints to guide the operator's hand movements to effectively align the welding torch for tele-welding and constrain the welding operation within a collision-free area. Experimental study and statistical analysis on the potential of using virtual fixture-based haptic interaction for accuracy improvement and collision avoidance in tele-welding scenarios are carried out. The impact of an integrated visual/haptic perception in MR on a natural, 3D motion mapping, enhanced immersive, and the intuitive tele-welding process is investigated.

1.3 Thesis Outline

This thesis is presented in seven chapters. The first chapter presented here describes the motivation, research objectives, and contributions of this research. The following chapters are organized as below:

Chapter 2. *Literature Review* presents the overview of the recent research conducted in intuitive and immersive teleoperation and the challenges of developing the mixed reality-enhanced telerobotic platforms for operations in hazardous/unpleasant environments, particularly in tele-manipulation and tele-welding scenarios. The review involves research carried out in the design, implementation and assessment of teleoperation systems with a specific focus on natural human-robot interaction and mixed reality-enhanced teleoperation.

Chapter 3. *Development and Intuitive Control of Telerobotic Platform for Mobile Manipulation Tasks* presents the design and development of the proposed mobile robotic manipulation platform that is intuitively usable for users with different levels of operational proficiency and expertise, especially for unskilled users. The platform architecture, modeling of the modular components, control implementation and data communication are introduced. Mutilple command input devices and end-effectors are incorporated to ensure the system compatibility, applicability and generalizability.

Chapter 4. *Mixed Reality-Enhanced Human-Robot Interaction with Imitation-Based Mapping Approach for Intuitive Telemanipulation* presents the design and implementation of immersive an MR-based interaction interface and the integrated motion/vision visualization retargeting scheme based on mixed reality approach for intuitive telemanipulation of robotic systems. The system effectiveness and performance analysis of different control-feedback methods including isometric-rate and isotonic-position control schemes are validated and compared, aiming for finding out the optimal HRI schemes for intuitive telemanipulation application.

Chapter 5. *Mixed Reality-Integrated 3D/2D Vision Merging for Intuitive Teleoperation of Mobile Manipulator* presents Mixed reality-based multi-view merging for intuitive

telemanipulation of mobile manipulator. For telemanipulation, 2D views from robot end-effector and external camera perspectives are essential. And the 3D point cloud and stereoscopic are merged into the typical 2D views to augment the visualization in the immersive spatial interaction process.

Chapter 6. *Mixed Reality Enhanced Intuitive Teleoperation with Hybrid Virtual Fixtures for Intelligent Robotic Welding*

An on-site manual welding experiment is conducted to identify operational differences between unskilled and professional welders, and the hand movements of professional welders are tracked and compared to those of unskilled welders. The hybrid virtual fixtures are incorporated into a mixed reality enhanced telerobotic system to stabilize the user's hand, improve welding performance, and avoid collisions between the torch and workpieces for welding. Finally, a tele-welding experiment is performed using the developed robotic welding platform to evaluate user performance and system usability.

Chapter 7. *Conclusion and Future Work* presents a summary of the presented work and the key results. In addition, the future research directions are discussed and some recommendations are given in this chapter.

Chapter 2

Literature Review

2.1 Effective Robotic Teleoperation

Robotic teleoperation is a technological approach allowing a human operator to perform a task by controlling a device or machine at a remote location. The distance can vary from the micron level (micro-manipulation) to the kilometer level (space applications). Human operators manifestly have the ability to adapt to the irregular environment and employ a variety of strategies to place the telerobots at a favorable condition, despite the complexity in remote sites.

Teleoperation secures an enormous advantage over other robotic controlling methods. A teleoperation system benefits from taking advantage of both human and robotic capabilities. A typical teleoperation system consists of a master device and a slave device. A human operator manipulates the master device and issues commands to the slave device, which is manipulated to achieve a certain goal. A robot manipulator that follows the motion of the master arm in real-time rather than being pre-programmed is an example. The kinematics of the slave device can be identical to the kinematics of the master device, or different. It can also be specifically scaled or even completely customized according to the task and requirements.

Despite the emergence and growing application of artificial intelligence, autonomous robots cannot reach the same level of intuition and reasoning as humans. Practical control algorithms can barely be used to ensure complex interactions and kinematic aspects humans can perform

effortlessly in their daily lives [27]. One of the advantages of teleoperation is it empowers the robot to fulfill tasks in unstructured or hostile environments where situational perceptions, cognitive abilities, and professional experience have a predominant impact on the task execution [28]. Therefore, for tasks requiring a variety of human judgment and professional skills, it is practical and feasible to utilize teleoperation to accomplish the goals.

In general, application scenarios of teleoperation could include deep-sea robotics applications, space exploration and operations, and de-mining operations, search and rescue in disasters, inspection in restricted spaces, medical robot-assisted surgery, hazardous materials treatment, and micro-manipulation or minimally invasive surgery, among many possibilities.

More specifically, teleoperation is suitable for tasks meeting the following set of criteria:

- Tasks to be performed in unstructured and dynamic environments, such as deep-sea exploration and space applications.
- Tasks involving operations in hazardous situations where human health is severely harmed. For example, in the Fukushima nuclear disaster, there was a high demand for emergency treatment and rescue. Mining scenarios are also increasingly typical.
- Tasks requiring dexterity, especially the coordination between hands and eyes. Medical surgery performed with remote robots is a typical example.
- Tasks requiring object recognition, obstacle detection, or situational awareness, for example, inspection in confined spaces.
- Tasks requiring motion scaling or force amplification, such as those found in large space based operations, for one example.

2.1.1 Categories of Robotic Teleoperation

2.1.1.1 Collocated and Separated Teleoperation

Based on the relative locations of the human and robot, robot teleoperation can be categorized into collocated and separated teleoperation. In collocated teleoperation [29], [30], the operator can directly observe the robot and its working environment in the same space, and a visual feedback system is not necessary, but augmented reality technology can be used in collocated teleoperation to superimpose to provide users with informative virtual content superimposed on the user-robot shared space [31]. Collocated robot teleoperation presents users with a robotic platform sharing a physical environment with the users who have a natural and clear 3D view of the robot's workspace [32]. However, although humans and robots physically share the same space, users can often only observe from a third viewpoint alongside the robot, and cannot inspect the robot's environment from the robot's main viewpoint.

Spatially separated teleoperation is used in a wider range of scenarios where the user and the robotic system are far apart, or where the user must be separated from the robot for safety reasons and cannot directly view the robot's movements. This type of robot teleoperation system typically requires visual and haptic feedback systems to show the user how the robot is working at a distance, allowing the operator to make timely adjustments. However, in the separated telemanipulation, users can typically only perceive the robot space through 2D camera streams and have difficulty synthesizing the 2D information collected by the robot with contextual knowledge of robot movements in its working space [33].

2.1.1.2 Ego-centric and eco-centric teleoperation

Depending on the operator's perspective of observing the manipulation process, robot teleoperation can be divided into self-centered and eco-centered teleoperation. In ego-centric

robotic teleoperation, the user observes the remote manipulation process from the same perspective as the robot and perceives himself as one with the robot [34]. The operator controls the remote robot with human-level dynamics through immersive and multimodal control-feedback schemes. Ego-centric robotic teleoperation transfers human-level manipulation and proficiency in telemanipulation applications and equips robotic systems with human-level motor skills. The user does not need to consider the mapping relationship between the robot's actions and the user's input commands [35], [36].

Eco-centric teleoperation usually involves inspecting the robot and its interactive environment from a third-person perspective while manipulating the robotic system with a broader view of observation (**Figure 2.1**) [37]. With the help of MR technology, users can still experience coexistence with the robot in the same virtual space, but without the immersion and intuition of being one with the robot.

(a) (b)

Figure 2.1: The immersive teleoperation of the Baxter robotic system with two human-robot interaction modes [37]. (a) Ego-centric teleoperation mode. (b) Eco-centric teleoperation mode.

2.1.1.3 Robotic mechanism-based and motion sensor-based teleoperation

Robot teleoperation can be divided into robotic mechanism-based and motion sensor-based

teleoperation, according to the command input device used by the user. Teleoperation systems that use robots as master mechanisms usually apply robotic devices of the same or different structure as the remotely controlled robot as the user command input system [38]. Using the same robots to form the master-slave architecture for teleoperation, motion synchronization can be accomplished by simple joint space mapping, and the user can directly visualize the pose of the slave robot by observing the master robot being controlled [39]–[41]. When using robots of different configurations as master and slave components for teleoperation, motion synchronization can be achieved using joint space motion remapping or TCP motion in the Cartesian space, and a scaling factor can be set to match the working space ranges of the master and slave robots.

The motion sensor-based robotic teleoperation platforms eliminate the need to introduce expensive and complex robot platforms as human-robot interaction command input devices. Commercial optical and IMU-based sensors, depth cameras, motion controllers and gloves are often used to capture user motion information [42]. Compared to the robotic master devices, motion sensor-based robotic teleoperation platforms usually have a larger space for motion tracking, the user's movement is not constrained by the mechanical structure, and the relatively lower price reduces the overall cost of the system [43], [44].

2.1.2 Intuitive and Natural Teleoperation

In conventional telemanipulation systems, the user controls the remote robotic system with a joystick, gamepad, keyboard and mouse, or 3D mouse, and simultaneously receives visual feedback from 2D displays [45], [46]. Robot control is not intuitive and natural to the user [47], [48]. The mismatch between the range of user control space and the limits of input device workspace can increase the difficulty of telemanipulation and lead to poor operation [49], [50].

Another disadvantage of typical telerobotic systems is the lack of depth perception due to monocular, 2D visualization of the remote site [51]–[53], limiting operator performance [54], [55], and any feelings of immersion and telepresence in the remote workspace [56], [57].

Intuitive human-robot interaction schemes are needed so the operator can easily guide the robotic system using natural motion. A conventional intuitive teleoperation platform was developed in [58] (**Figure 2.2**). The user manipulates the 7 degree-of-freedom (DOF) master robot to intuitively drive the slave robot at a distance as if they were directly operating the slave robot. This platform requires a symmetrical relationship between the input (master) and output mechanisms (slave). However, for 2D visualization of the telemanipulation process, the user lacks a sense of presence. Monitoring the interaction process with a 2D monitor on the user side can also distract the user from effective robot control activities. In addition, using the isomorphic master robot as the motion input device at the user side in telemanipulation systems often leads to a significant hardware investment.

(a) (b)

Figure 2.2: A typical intuitive control approach for remote robotic manipulation. (a) The user manipulates the master robot of 7 DOF to intuitively drive the slave robot at a distance as if they were directly operating the slave robot. (b) The overview of the master-slave teleoperation architecture [58].

With the inevitable limitations in the field of view of remote on-site 2D cameras and the poor quality of the visual feedback sent back from the robot's workspace, users are often unable to achieve a level of situational awareness sufficient for effective and safe remote manipulation tasks [24]. The operator is unable to perceive the physical environment of the worksite in 3D due to the lack of stereoscopic depth information. Typical monocular visual feedback strategies greatly reduce the efficiency of remote-controlled robotic operations [59], [60].

More specifically, the lack of depth information from 2D vision causes ambiguity about the spatial positions of objects in static images. Operators cannot directly perceive depth in monoscopic vision (MV) streams and have difficulty in accurately determining the distance between the robot end-effector and workpieces. Thus, the user perception from 2D visual feedback makes it insufficient to accomplish remote tasks effectively. Hence, success and efficiency of teleoperated tasks mainly depends on operator skill and proficiency [61], [62].

Multi-camera arrangements can provide the relative position of objects and allow users to view and judge the distance from various perspectives, compensating for the limits of a single monocular 2D camera. However, as telerobotic systems become more complex, avoiding robot collisions with surrounding objects with only multiple 2D feedbacks inevitably leads to low robot control efficiency and increased operator workload.

Different levels of depth perception in telemanipulation can improve task performance for various manipulation and grasping tasks [63]–[65]. Using binocular cameras can provide additional depth information to assist operators. Stereoscopic vision and point cloud are two methods of providing stereo depth information and have been applied in various telerobotic systems for remote vision [66]–[68]. Binocular visual feedback for robotic teleoperation has

been widely studied on unmanned aerial vehicles and mobile robots [69], [70]. Leveraging depth perception in remote robot teleoperation tasks for general object manipulation is not well-studied on mobile manipulators. In addition, neither multi-camera nor binocular camera setups allow the user the flexibility to continuously change the viewing angle to observe the remote work without introducing additional mechanisms on the robot side [71].

This section on Effective Robotic Teleoperation highlighted two important requirements for an effective and intuitive teleoperation platform. First, natural motion retargeting, or the use of natural human behavior to teleoperate robotic systems, effectively reduces training costs and grants human-level dynamics to telerobotic systems and is an important component of natural and intuitive human-robot interaction systems. Second, adequate multimodal feedbacks, especially visual feedback that provides depth information and situational awareness to enable the operator to know the dynamic situation in the robot's workspace, are also critical in interactive telemanipulation systems. The human operator in the control loop team must be aware of the relative positions of the robot end-effector and the workpieces in the workspace to make decisions and avoid collisions or damage to the robotic system, working area, and/or the environment.

2.2 Mixed Reality for Human-Robot Interaction

2.2.1 Reality–Virtuality Continuum

Extended reality (XR) refers to all the mutual embedding and fusion of the physicals and virtual scenes and the human-computer interaction that happens in the generated environment [72], where X represents the degree of interpenetration and integration of real scenes and virtual content in spatial or immersive computing technologies, involving the user’s sense of presence

and acquisition of perception [73]. XR includes the complete spectrum from entirely realistic to fully virtual, including all representative forms and all possible variations of AR, VR, AV, MR, and other interpolations between the two reality and virtuality [74], [75]. The reality-virtuality continuum presented by Milgram et al. [76]outlines the XR ecosystem and the relationship between AR/VR/AV/MR and fully real and virtual environments, which are schematically illustrated in **Figure 2.3**.

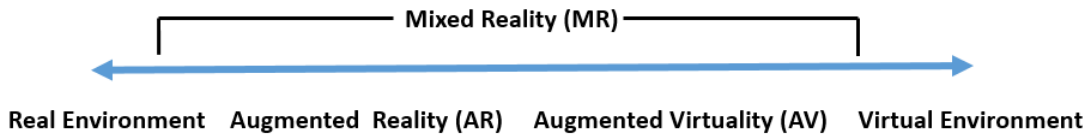


Figure 2.3: Milgram's reality–virtuality continuum [76].

Augmented reality (AR) overlays informative virtual 3D graphics on the physical world and allows real-time interaction with the 3D graphics, enabling the user to reach into the augmented world and manipulate 3D objects directly as if they were real-world objects [77]. The advantage of AR is it superimposes important cues on physical objects or hovers the display in real space to indicate the user's potential or actual future actions, or the robot's motion planning, while eliminating distractions caused by typical display methods [78]. AR-HRI interfaces provide users with an exo-centric view of the robot and its surroundings and allows operators to maintain situational awareness of the robot and ensure intuitive interaction and communication between the human and robot using multimodal interfaces [79].

Virtual reality (VR) uses fully immersive computer-generated graphics with scenes and objects that appear to be real to completely replace the physical surroundings in which the user is located [80]. In the virtual environment, the users perceive immersion in an identical or

completely different scene from the physical surroundings and perceive an imaginary environment with muti-modalities of visual, acoustic and haptic information [81]. VR allows users to move around in the scene and manipulate virtual objects by using wireless controllers or haptic robotic structures as input sources.

Mixed reality (MR) technology involves the merging of physical and imaginary spaces and does not occur exclusively in the physical or virtual world [82]. MR enables real-time visualization and interaction between physical and digital content [83] and takes full advantage of the visual information of real scenes and the three-dimensional immersion and interaction provided with virtual cues [84]. Head-mounted display (HMD) based MR can be achieved in two ways, either by displaying the camera captured video feedback of the real world in the HMD, or by allowing the user some direct visibility of the real world in the HMD [85], [86].

Augmented virtuality (AV) is a fully immersive computer-generated virtuality-dominated environment enhanced by sensory data from physical environments [87]. AV keeps the virtual environment central but superimposes the real-world elements on the virtual content [88]. AV is also classified as a subset of MR. The term AV is less commonly used in the literature for immersive telerobotic interfaces and is often generally referred to as MR.

The hierarchy of reality in MR ranges from partial sensory input to infinitely close to immersive reality [89]. The proposed MR-enhanced intuitive and immersive teleoperation scheme falls into the category of partially superimposing real-world visual inputs on the computer-generated virtual content.

2.2.2 Mixed Reality-Enhanced Teleoperation

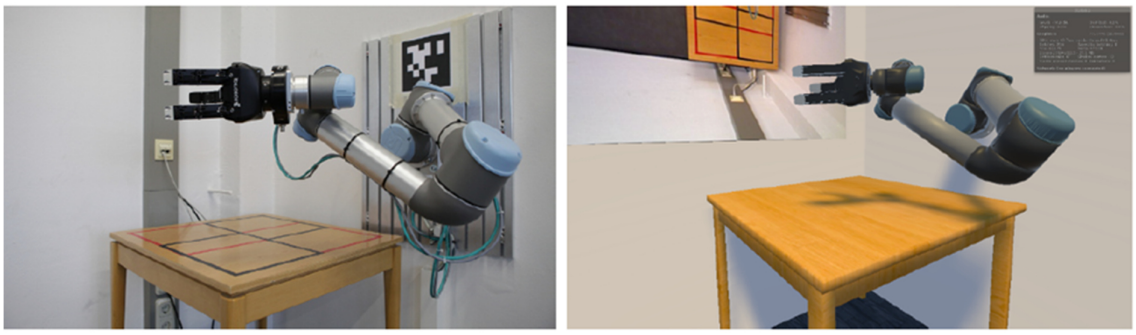
2.2.2.1 Mixed Reality-enhanced Intuitive Teleoperation

In recent years, MR technologies have been widely applied in the field of robotics, and these research directions can be divided into three categories by application aspects: (1) Human-robot interaction (HRI): immersive teleoperation, intuitive telemanipulation, collaborative interaction, wearable robots, haptic effects and virtual devices; (2) Medical robotics: Robot-assisted surgery (RAS), prosthetics, robot-assisted rehabilitation and training devices; (3) Robot motion planning and control: trajectory generation, robot programming, simulation, and manipulation.

MR has been applied to robotic teleoperation systems to enhance user perception of the remote side to enable immersive robotic teleoperation (IRT) [90]–[92]. A three-dimensional virtual world similar to the slave side can be simulated through MR and displayed to the user on the master side. By implementing MR as a human-robot interaction interface (HRI), the user experiences a physical presence in the remote environment, and co-existence with the robotic platform via MR subspace, while guiding and monitoring the robotic platform at the local user space [93], [94]. MR-enhanced teleoperation allows direct mapping of control commands and the actions between the user and the robot, and has the potential performance enhancements by serving as an intermediary for the integration of imitation-based motion mapping and 3D visualization [95].

Most recent research on MR-enhanced teleoperation systems focused on collecting demonstrations for robotic learning, solving long time delays, the development of immersive manipulation, and poor virtual transparency problems [96]–[98]. However, those telerobotic systems do not fully exploit the potential performance enhancements provided by using MR subspace as an intermediary for the integration of imitation-based motion mapping and 3D

visualization mapping. Dennis Krupke et al. [99] developed an MR-enhanced heterogeneous robot telemanipulation system (**Figure 2.4**) that presents the real robot working space and the corresponding virtual scene presented to the operator in the robot telemanipulation architecture for immersive and intuitive remote grasping. The pose of the robot replica in MR is synchronized with the current pose of the physical robot via messages. The communication between the robot and the MR scene is maintained via the ros-bridge [100]. A virtual screen on the left wall augments the virtual scene by displaying an image stream from the camera mounted above the physical robot. However, intuitive teleoperation is only applied to the robotic hand instead of the arm-hand system in its entirety, limiting the workspace and overall application significantly.



(a) (b)

Figure 2.4: The MR-based ROS-controlled telerobotic manipulation platform for immersive remote grasping [99]. (a) The physical laboratory environment with a robotic arm-hand system. (b) MR enhanced virtual corresponding scene displayed to the user in HMD.

In [96], researchers from the University of California, Berkeley built an MR-based teleoperation system on a PR2 robot through imitation learning. The system allows the operator to teleoperate robots to perform complicated tasks naturally and intuitively. In the proposed teleoperation system, a robot is controlled at a distance by the human operator through an MR-based telerobotic interface with overlaid information, which is an effective approach to collect high-quality demonstrations for training the robot. Imitation learning techniques allow the

robot to imitate human behavior and acquire skills through perceiving human demonstrations aiming at performing specific tasks.

An MR control room for dynamic vision and movement mapping between the operator and dual-arm robotic agent can be developed for tele-manufacturing (**Figure 2.5**) [23]. The multiple monocular sensor displays and motion mapping approach via MR outperforms telerobotic platforms with direct camera feeds. The control room has objects and controls floating in space, which allows the user to do movements relative to 3D markers to command the remote-controlled robot. However, the researchers did not determine if the immersive experience led to performance improvements compared to conventional 2D HRI platforms, so the impact remains unknown and unquantified.

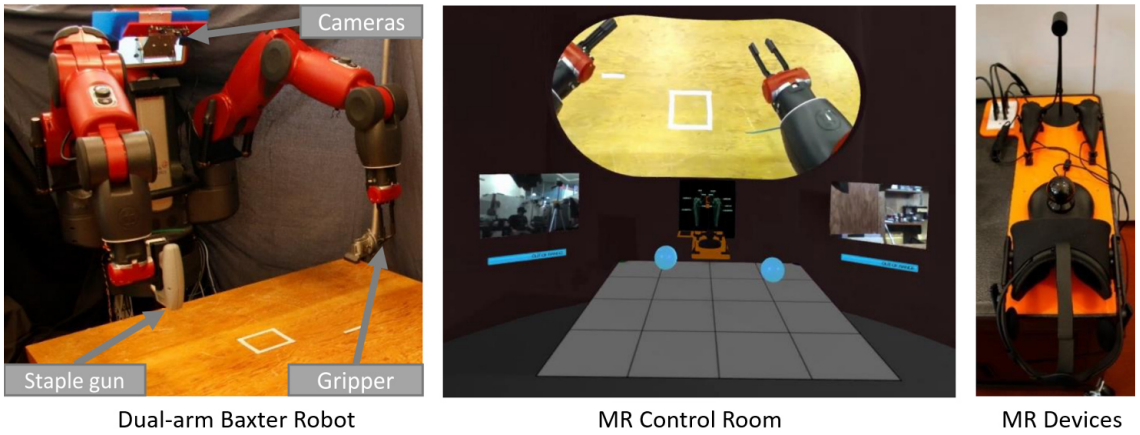
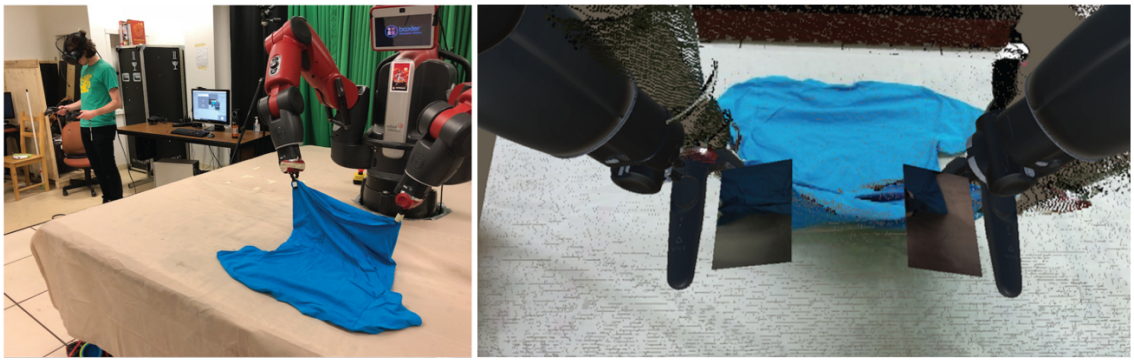


Figure 2.5: The MR-based Baxter's homunculus telerobotic system for a wide range of tele-manufacturing tasks [23].

ROS Reality [91], [92]is an open-source MR-based telerobotic manipulation framework that has been developed at Brown University (**Figure 2.6**). This work enables communication and interaction between ROS-based robots and Unity-compatible MR systems. A total of 24 dexterous telerobotic manipulation tasks using ROS Reality were conducted by expert users

compared to direct kinematic manipulation of the Baxter robot. the remote-controlled robotic platform targeted at expert teleoperators. However, user efficiency and system functionality were not verified with novice users for manufacturing-related tasks. This system can be used as data acquisition and validation platform for learning from demonstration (LfD) and other machine learning approaches to transfer human expertise and skills to robots.



(a) (b)

Figure 2.6: The ROS Reality-controlled telerobotic manipulation platform. (a) The intuitive and immersive control of the dual-arm Baxter robot; (b) the point cloud enhanced MR scene displayed to the users [91], [92].

In [101], the authors developed an MR-based telemanipulation system to control a robotic arm-hand system. The MR scene is augmented by real-time data from the robot task space, to enhance the operator's visual perception. The system incorporates a new interactive agent to control the robot and reduce the operator's workload. Two control algorithms are introduced to the MR-based teleoperation system to improve the long-distance and fine motions of the robot. Telemainpualtion experiments using a UR 10 robot arm and Robotiq-85 gripper demonstrate the feasibility of the proposed telerobotic paradigm.

2.2.2.2 Mixed Reality-based Vision Mapping and Merging

MR scenes feature immersive integration of multiple 3D/2D visual display modes to the users. At present, various teleoperation platforms have focused on MR-enhanced robot control [85]. It is not yet known whether using stereoscopic vision and point cloud within an MR environment can enhance users' stereoscopic perception and task performance in separated telerobotic operations. In [23], an MR control room for dynamic vision and movement mapping between the operator and dual-arm robot is introduced. The multiple monocular sensor displays and motion mapping approach via MR outperforms telerobotic platforms with direct camera feeds. The control room has objects and controls floating in space, which allows the user to do movements relative to 3D markers to command the remote-controlled robot. However, they did not determine if the immersive experience led to performance improvements.

In [51], the authors presented an MR teleoperation interface for mobile manipulation tasks with visual inputs from a monoscopic and stereoscopic camera setup for remote mobile manipulation tasks in hazardous production environments. This system is equipped with two monocular cameras and a stereoscopic camera at the robot's working site. Users acquire multi-view 2D images and stereo vision with depth cues in a Unity-generated MR control room. However, the intuitive control of the robotic platform as a whole was not assessed or presented. In addition, no comparative tests with typical telemanipulation systems were performed to verify the performance and efficiency of the proposed system. To reduce the training time for teleoperation, a multi-view merging method via MR was designed (**Figure 2.7**) in [102], and the platform provides users with intuitive control of the robot’s motion by using commercial VR controllers.

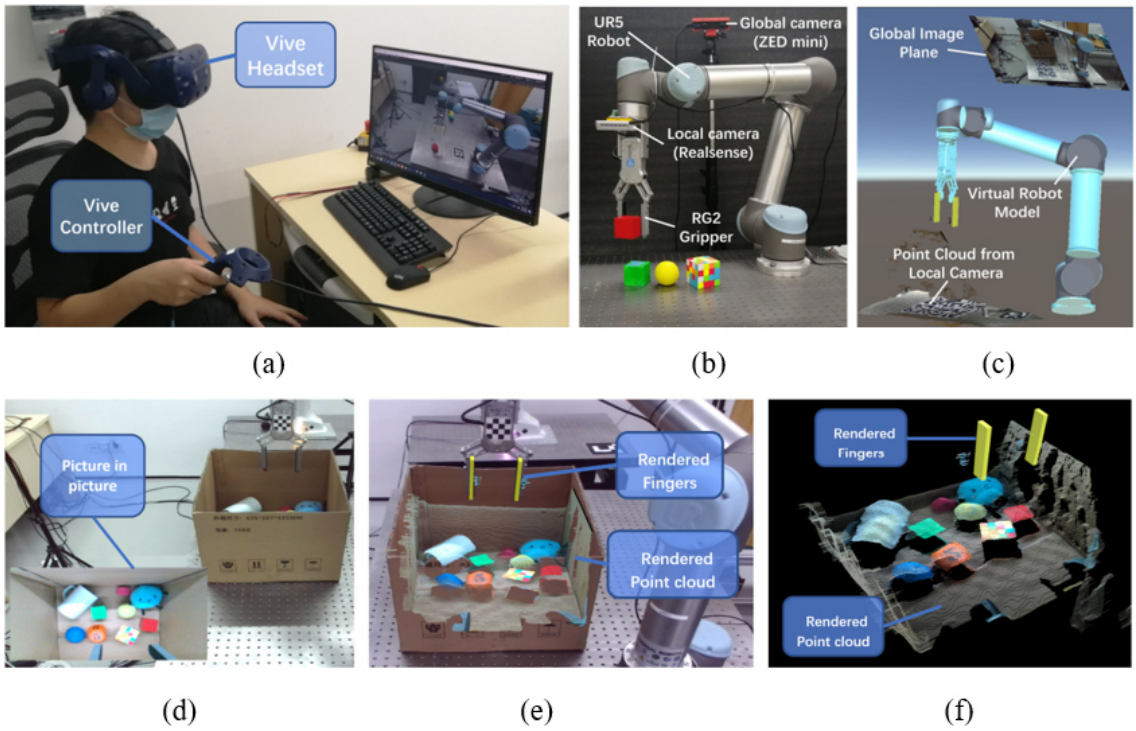


Figure 2.7: The MR-based multiple views merging approach for robotic telemanipulation [102]. (a) The user controls the robot's motion by using commercial VR devices. (b) The real robot working environment. (c) The MR-based view merging scene. (d-f) The three different vision merging setups.

Yiming Luo et al. [103] explored the use of stereoscopic view in an immersive manner for mobile robot teleoperation and navigation, and the results showed the stereoscopic view and immersive perception via virtual reality head-mounted displays (VR-HMDs) provided the user with depth cues and improved user performance and system usability. However, the effect of stereoscopic perception on telerobotic manipulation and the depth cues provided by other 3D visualization resources such as point cloud was not studied.

2.3 Mixed Reality-Enhanced Robotic Tele-Welding

2.3.1 Robotic Tele-Welding

Welding has been used extensively in the maintenance of nuclear plants, the construction of

underwater structures, and the repair of spacecraft in outer space [104]. In these hazardous situations in which human welders have no effective access, the judgment and intervention of the human operators is still required [105]. Customized production is also an application scenario for tele-welding, where welders often work in environments with dust, strong light, radiation, and explosion hazards [106]. Human-in-the-loop (HITL) robotic tele-welding strategies have become a feasible approach for bringing humans out of these dangerous, harmful, and unpleasant environments while performing welding operations [107]. Robotic tele-welding systems (RTWSs) combine the advantages of humans and robotics and coordinate the functions of all system components efficiently and safely [108]. For example, RTWSs can diminish geographical limitations for scarce welding professionals and bring a remote workforce into manufacturing.

Welding training is a time-consuming and costly process. Intensive instruction and training are usually required to bring unskilled welders to an intermediate skill level [109]. It is important to analyze the differences between the operating skills of professional and novice welders to facilitate the professional welding level of unskilled welders and to further improve the feasibility, efficiency, and welding quality of RTWSs used by novice welders during remote welding operations.

Thus, the expertise and skill extraction of professional welders as well as the application of robot assistance in on-site welding operations have become popular research topics [110]. The implementation of interactive robots can stabilize the hand movements of novice welders for improved welding quality, but robot-assisted welding has not been studied in teleoperated welding scenarios. Welding motion capture systems were used in [111] and [112] to differentiate between professional and unskilled welders in terms of operational behavior in

the gas tungsten arc welding (GTAW) process, providing an experimental basis for the development of robot-assisted tele-welding schemes. The experiments in [113] revealed the differences between professional and unskilled welders in the trajectory of the GTAW hand movements and indicated the main cause of the unsatisfactory welding results is novice welders make abrupt movements in the direction perpendicular to the weld surface. However, the operational difference in gas metal arc welding (GMAW) between professional and novice users was not well examined [114].

2.3.2 Mixed Reality-Enhanced Robot-Assisted Welding

Recent research on human-centered robotic welding has focused on the development of MR-based robot-assisted welding training platforms, intuitive programming for telerobotic welding, interactive telerobotic welding design, and MR-enhanced tele-welding paradigms. A VR-based haptic-guided welder training system was introduced in [115]. This system provides guidance force to welders, simulating a human welding trainer. Both novice and skilled welders can use this platform to improve their welding skills in a virtual environment. However, this system does not integrate real welding scenarios into the virtual environment to allow welders to adjust their movements in real-time according to the welding pool status, nor does it transfer human movements to the robot for actual tele-welding operations.

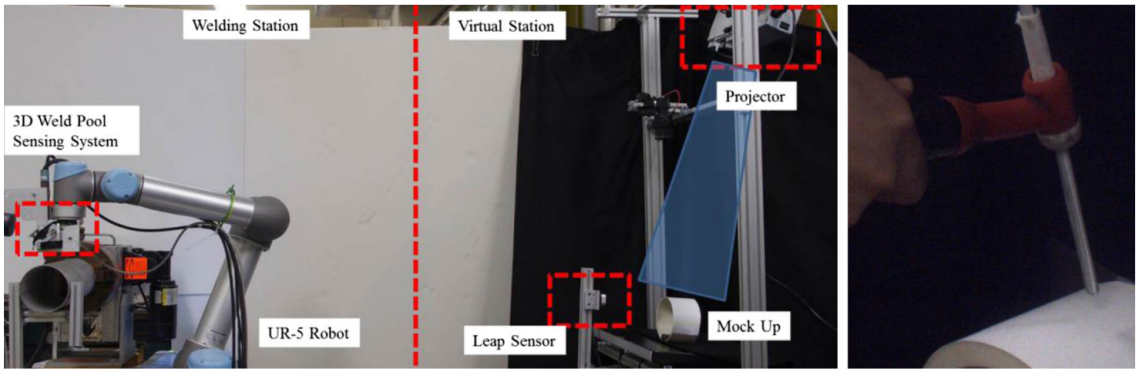
Olaf Ciszak et al. [116] proposed a vision-guided approach for programming automated welding robot paths in 2D, where the programmer draws the target weld pattern in the user presentation space, a low-cost camera in the system captures the image, and an algorithm detects and processes the geometry (contour lines) drawn by the human. This intuitive remote programming system for welding is limited by programming the contour lines in two-dimensional planes only and does not have the real-time capability of a telerobotic welding

system. In [117], the authors analyzed the integration of advanced technologies, such as MR, robot vision, intuitive and immersive teleoperation, and artificial intelligence (AI) to build an interactive telerobotic welding system. This paradigm enables efficient human-centered collaboration between remote welding platforms and operators through multi-channel communication.

A teleoperated wall-climbing robotic welding system was developed to demonstrate the application of various technologies in an innovative robotic interaction system to best achieve natural human-robot interaction. However, the mobile wall-climbing welding robot presented in this system has a simple structure and does not have a flexible robot manipulator to mimic welders' human-level manipulation and make dexterous welding adjustments. Natural human movement signals were not used to improve the system intuitiveness and control the robot for tele-welding tasks.

More recent research attention focused on MR-enhanced tele-welding paradigms [118]. It was verified in [119] that there were no statistically significant differences in the total welding scores between participants in the physical welding group and the MR-based welding groups. The MR welding user interface gives operators the ability to perform welding at a distance while maintaining a level of manipulation [120]. An optical tracking-based telerobotic welding system was introduced in [121]. The Leap Motion sensor captures the trajectory of a virtual welding gun held by a human welder in userspace to control the remote welding robot for the welding task. However, this welding system requires the use of a physical replica of the welding workpiece in the userspace to superimpose a real-time weld pool state and guide the welders to adjust their hand movements to the shape of the workpiece [122]. Thus, this MR welding system is not suitable for a wide range of workpieces.

Yukang Liu et al. [123] developed a projection-based MR telerobotic welding system for transferring welder skills and human-level dynamics to the welding robot (**Figure 2.8**). Human welders interact with the welding robot by moving a tracked virtual welding torch in 3D space. A UR 5 industrial manipulator equipped with vision sensors is operated remotely to undertake the welding task. The weld pool stream from the welding site is transmitted back to the user and projected on a mock-up of the workpiece. The human welder can monitor the work process and adjust the movements according to the projected weld pool status from the worksite. However, the operator does not have sufficient visual feedback to check the status of the robot during operation.



(a) (b)

Figure 2.8: The projection-based MR remotely-controlled robotic welding system for transferring welder skills and human-level dynamics to the welding robot [123]. (a) Overview of the welding station and virtual station. (b) The replica of a welding torch used in the MR welding station.

Wang et al. [124] developed an MR-based human-robot collaborative welding system (**Figure 2.9**). The collaborative tele-welding platform combines the strengths of humans and robots to perform weaving gas tungsten arc welding (GTAW) tasks. The welder can monitor the welding process through an MR display without the need to be physically present. Welding experiments indicated collaborative tele-welding has better welding results compared to welding performed

by humans or robots independently.

MR-based robot-assisted remote welding platforms were developed in [125] to provide the welders with more natural and immersive human-robot interaction (HRI). However, in these systems, the users rely on visual feedback for movement control and have no haptic effects to completely prevent accidental collisions between the robot and the workpiece when the operator controls the robot for welding from a distance. Hence, they are still limited.

A visual and haptic robot programming system based on MR and force feedback was developed in [126]. However, the system was not suitable for real-time remote welding operations and was inefficient in unstructured and dynamic welding situations. Haptic feedback provides the welders with additional scene modality and increases the sense of presence in the remote environment, thereby improving the ability to perform complex tasks [127]. In contrast, it can be difficult to implement effectively.

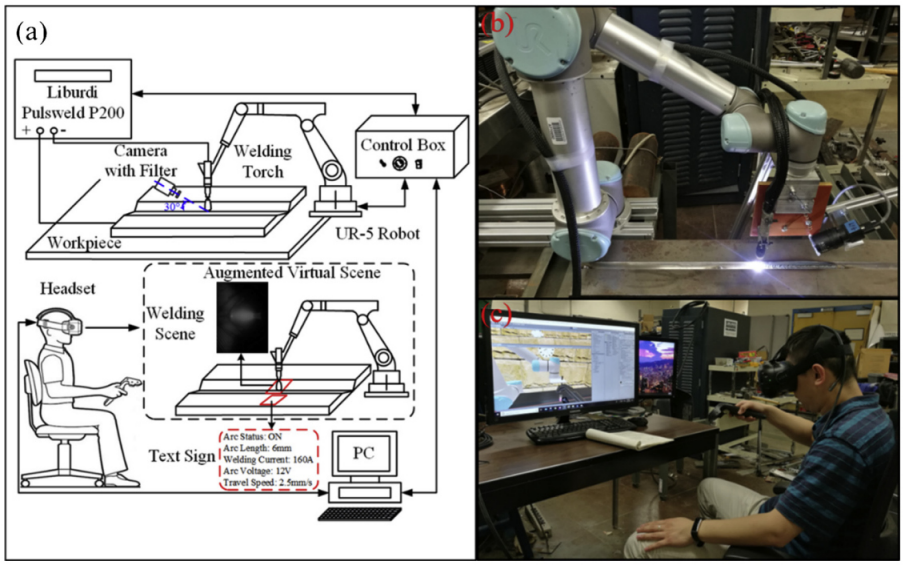


Figure 2.9: The overview of MR-based telerobotic system for remote welding tasks. (a) System configuration of MR-based human-robot collaborative welding platform. (b) The telerobotic welding site. (c) The MR scene displayed to the operator is augmented by a real-time on-site welding stream [124].

The primary benefit of incorporating haptic effects is to enhance the tele-welding task performance and operator's perception. These existing remote robotic welding systems do not take sufficient advantage of the potential performance improvements that various forms of haptic effects can bring to the user. The rapid development of MR-enhanced teleoperation has led to the integration of MR and virtual fixtures to improve task performance and user perception. The integration of MR and VF in RTWSs can effectively address the defects and problems that exist in the above telerobotic welding systems. The immersive and interactive MR environment allows for the effective generation of virtual workpieces in the user space and can be combined with VF technology to provide force feedback and guidance to users, thereby improving the accuracy of robot movements and effectively preventing accidental collisions [128], [129].

The main weakness of the published studies on tele-welding is existing remote-controlled robotic welding systems do not adequately incorporate MR technology and virtual fixtures to effectively eliminate potentially harmful collisions in the tele-welding process and grant welding robots human-level dynamics for dexterous GMAW welding tasks. No attempt has been made to reduce operational complexity to assist inexperienced welders to perform welding quickly and address the time-consuming training and shortage of a qualified workforce. This thesis addresses the limitations summarised in this literature review.

Chapter 3

Development and Intuitive Control of Teleoperation Platform for Mobile Manipulation Tasks

3.1 Introduction

This chapter describes the mechatronic development of a mobile manipulator robotic platform and the implementation of intuitive human-robot interaction for intelligent remote manufacturing. The robotic system is built from a UR5 robotic arm and a customized parallel gripper mounted on a differential-drive mobile platform for mobile manipulation tasks. The goal of intuitive telemanipulation is to enable the human-in-the-loop interaction in the execution of complex manipulation tasks in unstructured and dynamic environments, in which cooperation with humans is required. To achieve this goal, the open-source robot control middleware ROS has been employed as a framework to develop robotics software for each robotic module of the platform and to implement intuitive human-in-the-loop robot control. And an innovative integrated application of multiple human interaction schemes using various control input devices has been implemented. The system is designed to be compatible with a wide range of commercial intuitive control devices and to perform a variety of tele-manufacturing tasks. Currently, there is no complete set of ROS-based frameworks for the

intuitive interaction with such mobile manipulation robotic platforms. Therefore, one of the goals of this chapter is to fill this gap. The proposed robotic platform solution is characterized by cost-effectiveness, extensibility and generality.

3.2 Mechanical Design and Kinematic Modeling

3.2.1 Robotic System Overview

For the system design of the robot platform, the goal was to prototype an inexpensive and effective mobile manipulation robotic platform for intuitive remote human-robot interaction that takes full advantage of the simplicity, reliability and effectiveness of the robotic architectures. The developed platform consists of three modular components: a differential-drive mobile platform, a UR5 robotic arm and a 3D printed parallel gripper. **Figure 3.1** shows the overview of the mobile manipulation robotic platform.

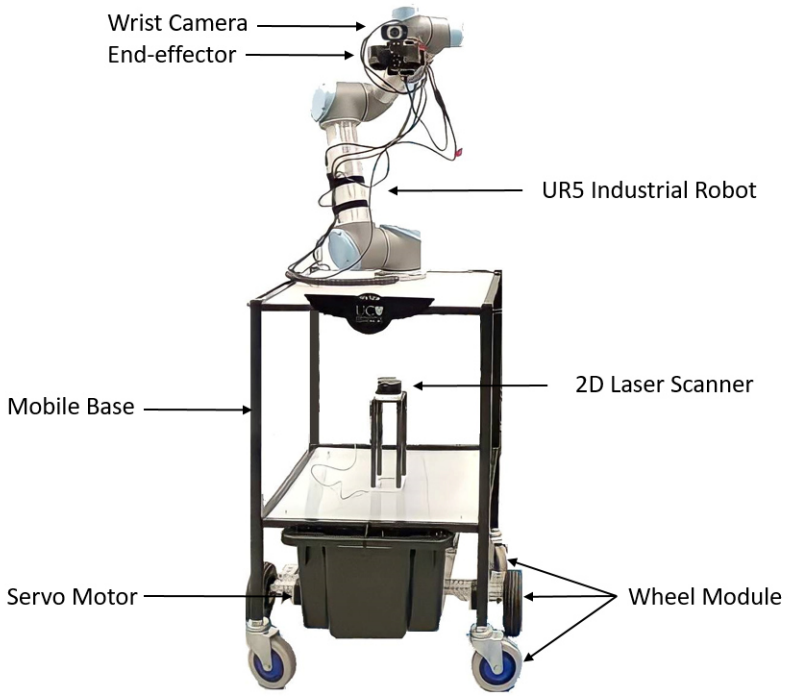


Figure 3.1: The robotic platform and the main components.

The development and integration of the three modular components, especially the mobile base and low-cost parallel gripper, were the most important part of the design, as they were not commercially available. The system design process includes mathematical modeling of the modular components, design and manufacturing of customized parts, selection of off-the-shelf parts, and final assembly of the robot platform. 3D printing and laser cutting are the main means of rapid prototyping for customized part manufacturing.

For the application of intuitive control of the robotic platform, different interaction strategies with various control input devices have been implemented. The UR5 manipulator and robot hand were integrated as an arm-hand system and considered as a whole during the control design and implementation process.

3.2.2 Robot Drive Development

3.2.2.1 Mobile Module Architeture

The differential-drive system is one of the simple and cost-effective solutions for the mobile robotic mechanism and is often employed as a mobile architecture to compose complex robotic systems. The differential-drive robot developed in this work consists of two motorized wheels and four non-driven caster wheels. The two driving wheels are mounted on a common plane and rotate around the same axis and are controlled by two independent actuators. The four symmetrically mounted non-driven casters ensure the stability and weight distribution of the robot. **Figure 3.2** shows the proposed differential mobile base providing mobility to the overall robotic manipulation system.

Figure 3.2: The two-wheeled drive base with independent actuators for each wheel.

3.2.2.2 Structural Analysis of Mobile Module

The differential-drive mobile robot architecture is selected as a mobile base for the proposed intuitive telerobotic platform because of its structural simplicity, high maneuverability and reliability of operation. The motion of the two-wheeled drive base is provided by 2 active joints and 4 passive joints, with the active joints driven by 2 servo motors for both steering and driving. **Table 3.1** shows the requirements for the joint torque validation of the whole robotic manipulation platform. The required total drive torque was calculated as 4 $N \cdot m$ through the equations in [130]:

Table 3.1: Robot parameter for joint torque validation.

Parameter	Value
Weight	15 kg
Payload	30 kg
Maximum incline	5°
Maximum Speed	0.5 m/s
Maximum acceleration	0.25 m/s^2
Driving Wheel Diameter	0.165 m
Servo motor speed	77 RPM

The total torque applied to drive the whole robotic platform can be written as:

$$T_{total} = F_d \frac{d}{2} \tag{3.1}$$

$$F_{total} = F_d - F_p = Ma \tag{3.2}$$

$$T_{total} = \frac{M(a + g\sin\theta)d}{2} \tag{3.3}$$

The torque applied by each motor is written:

$$T_{per\ motor} = \frac{M(a + \mathrm{g}\sin\theta)d}{2n} \tag{3.4}$$

where T_{total} is the torque applied to drive the whole robotic platform, F_d is the driving force provided by the motors, d is the diameter of the driving wheel. F_{total} denotes the resultant force that moves the robot, F_p is the force pulling the robot and θ is the maximum incline to climb. M is the mass of the whole robotic system, a is the acceleration of the robot system and g is the acceleration of gravity. n denotes the number of motors and in this case $n = 2$. The required driving torque for each motor $T_{per\ motor}$ is 2 $N \cdot m$.

The maximum speed of the two-wheeled drive system is 0.67 m/s as found by calculating:

$$V_{max} = \frac{\pi d \cdot RPM}{60} \tag{3.5}$$

3.2.2.3 Mobile Module Kinematic Model

The developed drive robot consists of two independently driven wheels and four non-driven wheels. The mobile base can be steered in place, driven in a straight line, or moved along a circular path around the instantaneous center of rotation (ICR). The kinematic formulation of the two-wheeled mobile module governs how the human operator's control inputs are mapped to the robot's wheel speeds and is essential to the follow-up development of the intuitive control system.

As shown in **Figure 3.3**, the two motorized wheels and four low-friction casters rotate around the same axis. The reference point (x, y) represents the midpoint of the common axis of the driving wheels, and ϕ denotes the robot heading direction representing the orientation of the robot about the axis O, and ϕ is defined by the angle of the longitudinal axis y_{robot} with

respect to the ground axis $\hat{x}$. θ_{L} and θ_{R} are the angular speeds of the two independent motorized wheels. A positive angular speed of each wheel corresponds to forward motion at that wheel. The width between the two motorized wheels is denoted as l, r denotes the radius of rotation for independently motorized wheels, and the rolling speed of the mobile module motors is $U_m = [u_1\ u_2\ ]^T$, with u_1 and u_2 the respective rotating speed of the left and right wheels.

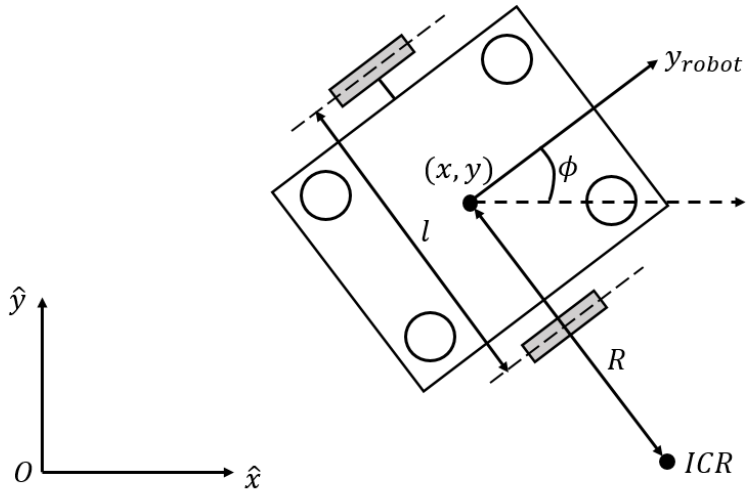


Figure 3.3: Kinematic representation of the non-holonomic mobile module consisting of two motorized wheels (shaded gray) and four caster wheels on a Cartesian coordinate system.

The control vector fields matrix $C(q)$ is employed to define the mobile module kinematic model, and $c_1(q)$ and $c_2(q)$ are the velocity vector functions with respect to the robot posture $q = (x, y, \phi, \theta_{\mathrm{L}}, \theta_{\mathrm{R}})$. The robot motion representation $\dot{q}$ is expressed [131]:

$$\dot{q} = C(q)V_m = c_1(q)u_1 + c_2(q)u_2 \tag{3.6}$$

The kinematic configurations of the mobile base can be written in the matrix form:

$$\dot{q} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \\ \dot{\theta}_{\mathrm{L}} \\ \dot{\theta}_{\mathrm{R}} \end{bmatrix} = \begin{bmatrix} \frac{r\cos\phi}{2} & \frac{r\cos\phi}{2} \\ \frac{r\sin\phi}{2} & \frac{r\sin\phi}{2} \\ -\frac{r}{l} & \frac{r}{l} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_m \\ u_m \end{bmatrix} \tag{3.7}$$

For simplicity and intuitiveness, the user directly inputs the linear and rotational speed of the mobile base to manipulate the robot movement and does not need to be concerned about the rotation of the wheels. The last two rows of Equation (3.7) can be dropped and the simplified kinematic equation for intuitive control is defined [132], [133]:

$$\dot{q} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} \cos\phi & 0 \\ \sin\phi & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_m \\ \omega_m \end{bmatrix} \tag{3.8}$$

where vector $V_m = [v_m\ \omega_m\ ]^T$ is the control inputs transmitted from the human operator to the mobile robot, with v_m the robot's driving speed and ω_m the steering speed. The translational and rotational input commands to the robot are subject to the interval $[-v_{m,max}, v_{m,max}]$ andl $[-\omega_{m,max}, \omega_{m,max}]$.

3.2.3 Arm-Hand Module Design

3.2.3.1 Gripper Mechanism Design

The parallel gripper assembly that includes all the inexpensive actuating mechanism parts is designed for additive manufacturing (AM) with engineering plastic material. **Figure 3.4** shows the 3D printed parts and main electronics and sensory components of the low-cost parallel

robotic gripper. The Prusament polylactide (PLA) was used to manufacture the parts and **Table 3.2** shows the main mechanical properties of PLA. The two fingers are connected to the servo motor through a slider-crank mechanism and driven to interact with objects. The UR5 manipulator is employed to hold the base of the gripper providing arm-level dynamics.

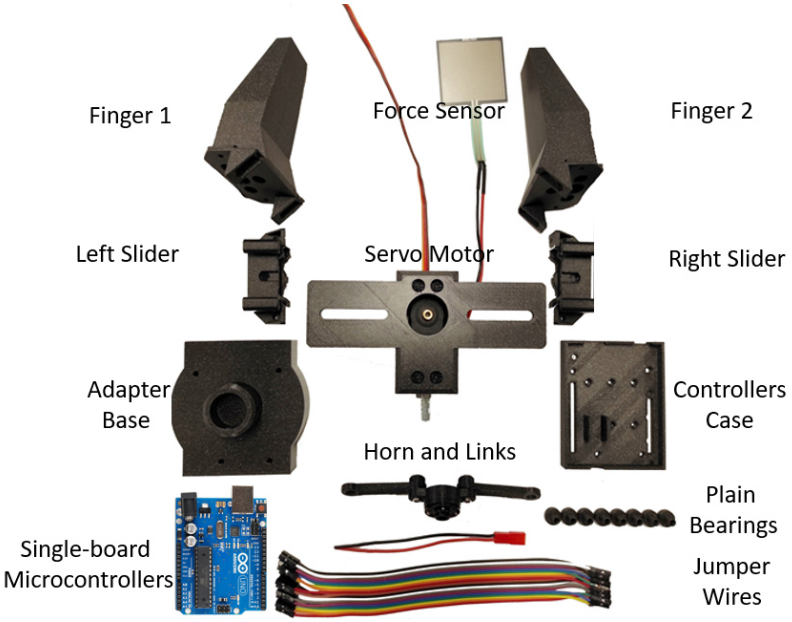


Figure 3.4: The 3D printed parts and main electronics and sensory components of the low-cost parallel robotic gripper.

Table 3.2: Mechanical properties of Prusament PLA [134].

Property / Print direction	Horizontal	Vertical X,Y	Vertical Z
Tensile Modulus $[GPa]$	2.2 ± 0.1	2.4 ± 0.1	2.3 ± 0.1
Tensile Yield Strength $[MPa]$	50.8 ± 2.4	59.3 ± 1.9	37.6 ± 4.0
Elongation at Yield Point [%]	2.9 ± 0.3	3.2 ± 0.1	1.9 ± 0.3
Impact Strength Charpy $[kj/m^2]$	12.7 ± 0.7	13.7 ± 0.7	5.0 ± 1.4

3.2.3.2 Gripper Kinematic Model

The position of a single finger with respect to the motor angle can be expressed:

$$x = L_1 \cos\alpha + \sqrt{{L_2}^2 - {L_1}^2 \sin^2\alpha} \tag{3.9}$$

where x is the distance between the connecting point of the slider and the motor shaft, α is the rotating angle of the servo motor measured from the gripper's maximum opening position. L_1 is the length of link 1 and L_2 is the length of link 2, as shown in **Figure 3.5 (c)**.

The stroke length of the parallel gripper d is calculated:

$$d = 2L_1 \cos\alpha + 2\sqrt{{L_2}^2 - {L_1}^2 \sin^2\alpha} \tag{3.10}$$

The finger velocity x' with respect to motor angle α is expressed:

$$x' = \left(-L_1 \sin\alpha - \frac{{L_1}^2 \sin\alpha \cos\alpha}{\sqrt{{L_2}^2 - {L_1}^2 \sin^2\alpha}} \right) \frac{d\alpha}{dt} \tag{3.11}$$

The force F exerted by the holding fingers is expressed [135]:

$$F = 2F_j = \frac{T_s}{L_1 \sin(\alpha + \beta)} \tag{3.12}$$

where T_s is the torque on the servo motor shaft and F is the gripping force exerted by each gripper jaw.

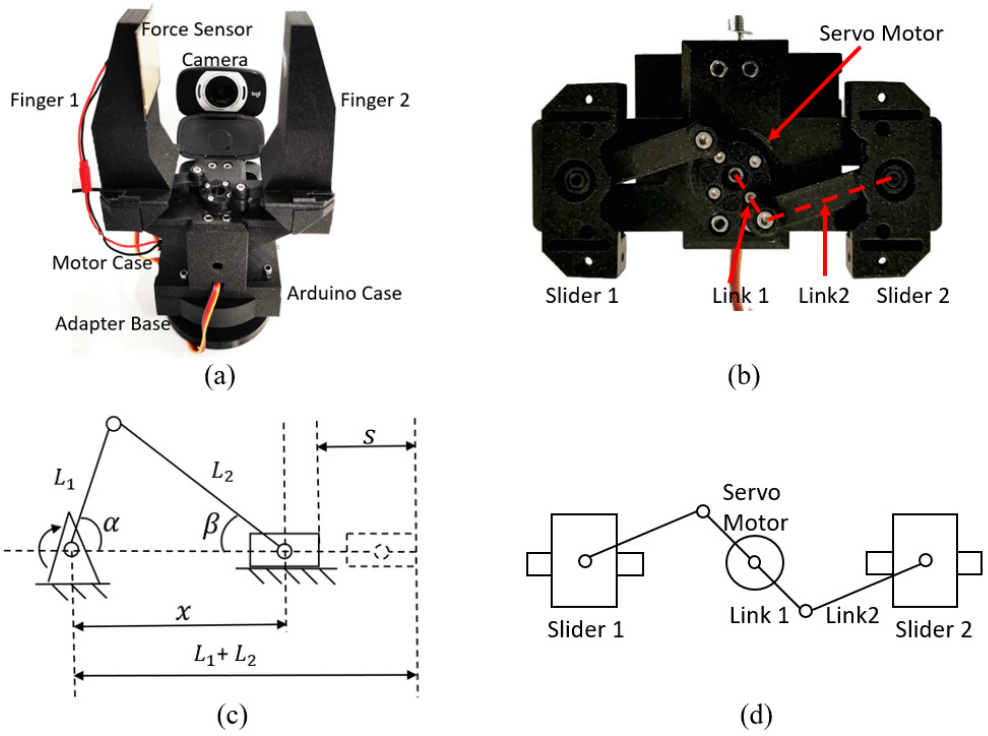


Figure 3.5: The overall gripper structure and parallel mechanism design. (a) The two fingers of the gripper are mounted on linear slider 1 and slider 2, which are driven by a servo motor. (b) The corresponding view of the slider-crank mechanism with fingers removed. (c) Diagram showing the geometric layout of the slider pin, link pin and motor shaft. The length of the crank is L_1, the length of the link connecting the finger and crank is L_2, the motor angle is α and the slider angle is β. (d) Sketch of the in-line slider-crank linkage of the parallel gripper.

3.2.3.3 Arm-Hand Module Overview

The arm-hand module consists of the ur5 manipulator and parallel gripper which are integrated as a system in terms of modeling and control design. **Figure 3.6** shows the UR5 manipulator at its zero position $\boldsymbol{q} = [0,0,0,0,0,0]^T$. UR5 robot is a 6-DOF robotic manipulator developed by the Universal Robots Company based in Odense, Denmark. The UR5 manipulator is widely used in teaching, research and industrial production for its reliability, repeatability, safety, ease of use and performance. As shown in **Table 3.3**, the UR5 robot weighs only 18.4 kg, but has

a payload of 5 kg. Each joint of the arm can rotate 360° and the last three joints of it move with non-coincident axes at the wrist. All the six joints contribute to the transformational and rotational movements of its tool-center point (TCP).

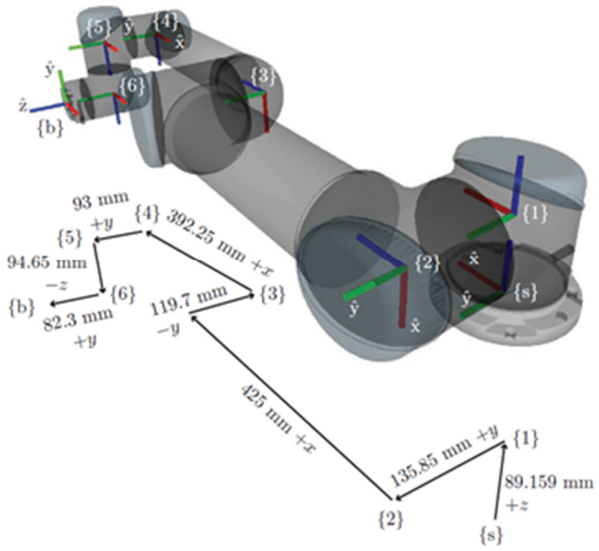


Figure 3.6: The UR5 manipulator at its zero position $\boldsymbol{q} = [0,0,0,0,0,0]^T$ with the join axes indicated [131].

Table 3.3: Mechanical properties of UR5 robot [136].

System Parameter	UR5
DOF	6
Payload	5 kg
Repeatability	± 0.1 mm
Weight with cable	18.4 kg
Joint ranges	± 360°
Reach	0.85 m

The UR5 robot has the following advantages: easy programming, the operator can program the robot through a graphical interface, no professional programmer is needed; the compact and lightweight structure of the UR5 robot make it flexible in tight workspaces or production sites; the modular design allows for quick repairs in case of malfunction, reducing production line downtime and increasing productivity; The overall dimensions of the UR5 are close to those of a human arm, which facilitates natural human-robot interaction/cooperation and intuitive control using human movement. Its small footprint makes it ideal to be mounted directly on mobile platforms for mobile manipulation applications [137].

3.2.3.4 Arm-Hand Module Kinematic Model

The Denavit-Hartenberg (D-H) representation approach is used for kinematic modeling of the robotic arm-hand system. Frames are affixed to links of the UR5 manipulator to describe the kinematic parameters between two neighboring coordinate systems, as shown in **Figure 3.7**. The Definition of D-H parameters in terms of the link frames is shown in **Table 3.4**. Only the four kinematic parameters are needed to determine the homogeneous transformation matrix.

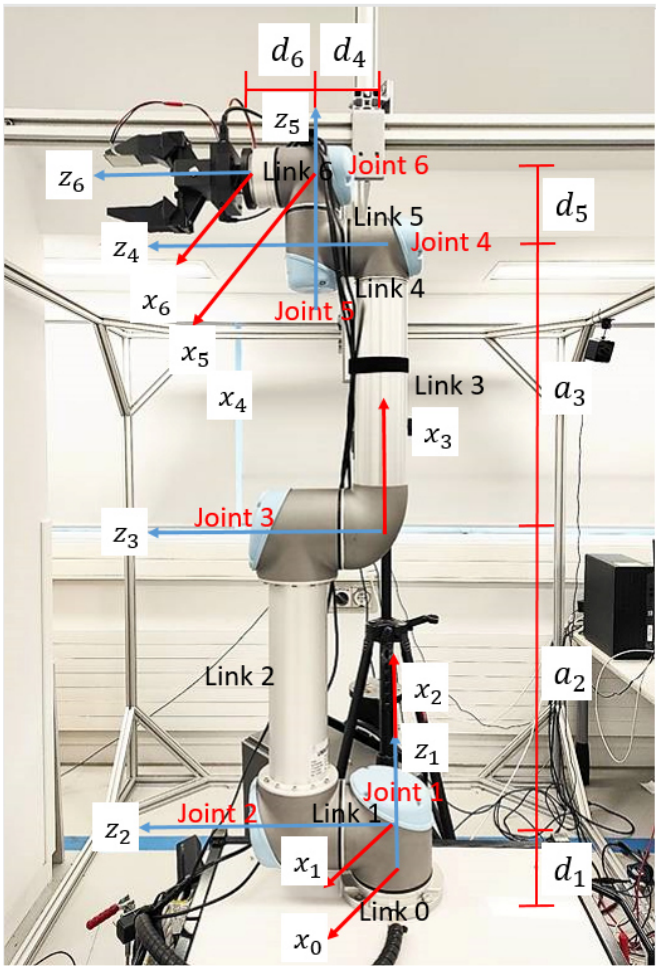


Figure 3.7: The schematic of the UR5 manipulator at position $\boldsymbol{q} = \left[0, -\frac{\pi}{2}, 0, -\frac{\pi}{2}, 0, 0\right]^T$ and the frame assignment of each joint indicated by the right-hand rule.

Table 3.4: Definition of transformation parameters in terms of the link frames.

DH-parameters	Description
a_i	distance between Z_i to Z_{i+1} measured along axis X_i
α_i	angle between Z_i to Z_{i+1} measured around axis X_i
d_i	distance between X_{i-1} to X_i measured along axis Zi
θ_i	angle between X_{i-1} to X_i measured around axis Zi

The general the homogeneous transformation matrix T_i^{i-1} of UR5-based arm-hand system according to the determined D-H parameters can be expressed [138]:

$$T_i^{i-1} = \begin{bmatrix} cos\,\theta_i & -sin\,\theta_i & 0 & a_{i-1} \\ sin\,\theta_i\,cos(\alpha_{i-1}) & cos\,\theta_i\,cos(\alpha_{i-1}) & -sin(\alpha_{i-1}) & -sin(\alpha_{i-1})d_i \\ sin\,\theta_i\,sin(\alpha_{i-1}) & cos\,\theta_i\,sin(\alpha_{i-1}) & cos(\alpha_{i-1}) & cos(\alpha_{i-1})d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.13)$$

The D-H parameters of the specified joints and frames of the arm-hand system are listed in **Table 3.5**.

Table 3.5: The D-H parameters of the UR5-based arm-hand architecture.

i	α_{i-1}(rad)	a_{i-1} (mm)	d_i (mm)	θ_i
1	0	0	d_1	θ_1
2	$\alpha_1 = \frac{\pi}{2}$	0	0	θ_2
3	0	a_2	0	θ_3
4	0	a_3	d_4	θ_4
5	$\alpha_4 = \frac{\pi}{2}$	0	d_5	θ_5
6	$\alpha_5 = \frac{\pi}{2}$	0	d_6	θ_6

The resulting homogeneous transformation matrix T_6^0 depicting the configuration of the robot gripper concerning the ur5 base frame is defined [139]:

$$T_6^0 = T_1^0 T_2^1 T_3^2 T_4^3 T_5^4 T_6^5 = \begin{bmatrix} R_6^0 & P_6^0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.14)$$

3.2.3.5 Inverse Velocity Kinematics

Velocity-based motion control in Cartesian space provides smooth motion and is a suitable method of guiding the robotic arm-hand system intuitively by human operators [26]. To convert a user-given velocity command into a differential joint command for the robot controller, it is essential to derive the instantaneous relationship between the arm-hand system joint rate vector $\dot{\theta}$ and the end-effector's spatial twist V_r (also referred to as Cartesian velocity), as shown in **Figure 3.8**. The Cartesian velocities of the tool center point (TCP) of the arm-hand system related to joint rates are expressed:

$$V_r = J_{r1}(\theta)\dot{\theta}_1 + \cdots + J_{rn-1}(\theta)\dot{\theta}_{n-1} + J_{rn}\dot{\theta}_n \tag{3.15}$$

where each Jacobian $J_{ri}(\theta) = (v_{ri}(\theta), \omega_{ri}(\theta))$ depends explicitly on the joint values θ for $i = 1, \ldots, n$. Since the UR5 manipulator is a six-jointed robot, the Jacobian matrix $J_r(\theta) \in \mathbb{R}^{6\times6}$ governs the joint velocity vector $\dot{\theta} \in \mathbb{R}^{6\times1}$ to the spatial twist V_r is computed:

$$V_r = J_r(\theta)\dot{\theta} = [J_{r1}(\theta) \quad J_{r2}(\theta) \quad \cdots \quad J_{r6}(\theta)] \begin{bmatrix} \dot{\theta}_1 \\ \vdots \\ \dot{\theta}_6 \end{bmatrix} \tag{3.16}$$

The 6×1 spatial twist vector $V_r = [v_r\ \omega_r\ ]^T$ consists of the 3×1 linear velocity v_r and 3×1 rotational velocity v_r. The translational and rotational input commands to the robot are limited to the robot speed constraints $[-V_{r,max}, V_{r,max}]$ with $V_{r,max} > 0$.

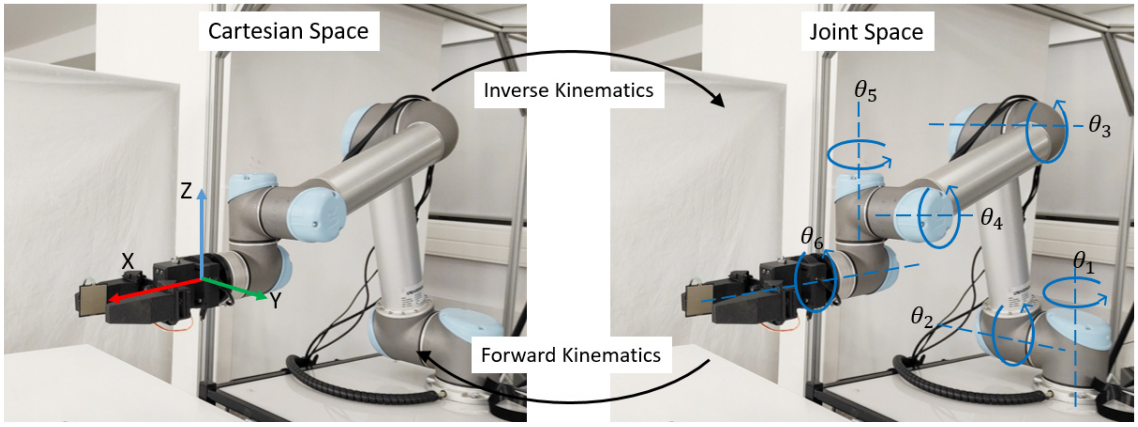


Figure 3.8: Motion mapping between Cartesian space and joint space for the UR5-based arm-hand system.

The approach for intuitively controlling the robot so its end-effector smoothly follows a user's hand trajectory, $T_u(t)$, is to calculate and control the required n-vector of joint velocities $\dot{\theta}$ directly from the relationship $V = J(\theta)\dot{\theta}$, where the desired robot end-effector twist $V_r = [v_r\ \omega_r]^T$ and $J(\theta) \in \mathbb{R}^{6\times n}$ are expressed in the same frame, yielding:

$$\dot{\theta} = J^{\dagger}(\theta)\, V_r \tag{3.17}$$

The use of the pseudo-inverse in Equation (3.17) implicitly weights the cost of each of the 6 joint velocities identically, returns the minimal two-norm of joint velocities, and reduces the energy consumption of the robot [140].

At time k, the motion tracking system measures and determines the configuration of the user's hand, and converts this calculated spatial velocity $V_u = [v_u\ \omega_u]^T$ to a spatial twist command $V_r = [v_r\ \omega_r]^T$ to the robot in the Cartesian coordinate system, employs the Inverse-Jacobian solver to determine the appropriate joint rate vector, $\dot{\theta}$, according to the desired twist

representation of end-effector motion, V_r, and derives the joint configuration sent to the UR5 robot controller:

$$V_u = \begin{bmatrix} v_u \\ \omega_u \end{bmatrix} = K_S V_r = \begin{bmatrix} v_r \\ \omega_r \end{bmatrix} = J_r(\theta)\dot{\theta} = \begin{bmatrix} J_v(\theta) \\ J_\omega(\theta) \end{bmatrix} \dot{\theta} \tag{3.18}$$

The mapping between the system incoming commands and the robotic end-effector velocities is defined:

$$v_u = \begin{bmatrix} X_u' \\ Y_u' \\ Z_u' \end{bmatrix} = K_S v_r = \begin{bmatrix} K_S & 0 & 0 \\ 0 & K_S & 0 \\ 0 & 0 & K_S \end{bmatrix} \begin{bmatrix} X_r' \\ Y_r' \\ Z_r' \end{bmatrix} \tag{3.19}$$

where K_S is a diagonal matrix and serves as a positive scaling factor directly affecting the operator's interaction resolution. $v_u = [X_u' Y_u' Z_u']^T$ and $v_r = [X_r' Y_r' Z_r']^T$ are linear velocities of the user input commands and robot TCP velocities to the frames of the respective workspaces.

Similar techniques apply to scaling the rotational motion ω_r of the arm-hand system. When $K_S = M$ is selected as the scaling factor, the scaling elements along the diagonal are all the same for the 6-vector of robot TCP velocities, where $M \in \mathbb{R}^{6\times6}$ is a scalar matrix. $K_S = I$ is chosen in the applications for precisely mimicking the user's movement, where $I \in \mathbb{R}^{6\times6}$ is an identity matrix. When the entries of K_S are relatively small, the robot performs a large motion when the operator input motion is fine-scale. Conversely, when entries of K_S are large, the operator needs to input a large movement to produce a delicate motion of the robot. A higher interaction resolution is desired when conducting micro-telemanipulation [141].

By applying inverse-Jacobian-based kinematic techniques, real-time control of the spatial velocity of the TCP is achieved, instead of creating discrete path plans. Thus, a smooth trajectory and quality manipulation behavior can be achieved. The proposed teleoperated robotic system provides the operator with a more natural and intuitive scheme to interact with the remote robot, in comparison to conventional robotic teleoperation systems, where the robot is manipulated in joint space or the end-effector is driven at a certain specified velocity.

3.3 Intuitive Control Implementation

3.3.1 User Command Input Schemes

In telemanipulation, telerobotic systems must perform a variety of tasks and be equipped with various end-effectors and sensors for application-specific teleoperation scenarios. In addition, certain command input approaches are particularly advantageous for application-specific tasks. In this section, multiple human interaction schemes using various control input devices are incorporated into the proposed teleoperated robotic platform for the intuitive control of the arm-hand subsystem, which enhances the usability, compatibility and applicability of the system for a wide range of tele-manufacturing scenarios. Subsequent sections evaluate the effectiveness of different input and output combinations in detail.

Typical human-machine intuitive interaction approaches can be divided into isometric-rate and isotonic-position schemes [142]. The isometric-rate interaction constraints the user's input motion through its structure and does not require the significant movement of the operator to control the robots. The human interaction devices under this domain have poor human-level proprioceptive capabilities. The benefits are they are not constrained by the workspace of input

devices on the user side and can reduce user fatigue due to minimal movement requirements [143].

The isotonic-position scheme enables operators to interact with the robot using natural movements/gestures and fully take advantage of human-level dynamic performance to command the robotic systems. The isotonic-position interaction has the benefit of utilizing the operator's proprioception while manipulating the robotic platform. The operator can perceive the motion of the robot through their body movement. However, these interaction schemes typically involve the incorporation of mechanically complicated and expensive command input mechanisms into the user-side environment. In isotonic position-based teleoperation, a scaling factor is usually applied between command input on the master side and the motion output on the slave side, and the robot can reproduce the human action accurately and mimic the user realistically when the scale factor is an identity matrix. The operator’s motion input can be scaled down to perform the high precision operation on the robot side.

The master space requires a relatively large workspace compared to the isometric human-robot interaction mode, but this is not feasible in the case of space-constrained situations. Consequently, when the operational constraints of the master device are smaller than the workspace limits of the robot being controlled, the operator must decouple the master device from the slave robot using a clutching mechanism, reposition the master device to a suitable position/orientation and reconnect the master-slave communication to continue manipulating the robot system [144].

In addition, this section also presents optical tracking-based interaction and haptic interaction approaches. By using non-contact optical tracking-based tele-control approaches, the

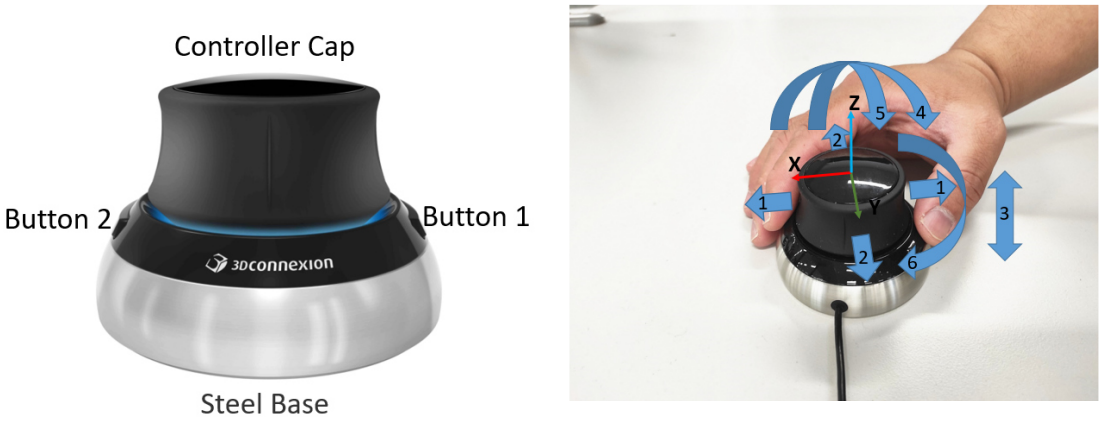
operator's motion is not constrained by the mechanical structure of the contact control devices used in typical interaction cases such as robotic systems based on mechanical master-slave devices. Unlike typical master-slave teleoperation structures, the optical tracking-based haptic interaction approaches do not require the introduction of expensive and complex mechanical or robotic mechanisms on the user side. And the robot moves directly based on optically tracked human motion information [145]. The advantage of the haptic interaction solution is that it provides the operator with kinesthetic haptic feedback during the human-robot remote interaction.

3.3.2 The Isometric-Rate Intuitive Interaction

The control input device used in this setup is the SpaceMouse Compact sensor manufactured by the German company 3Dconnexion, as shown in **Figure 3.9 (a)**. The SpaceMouse Compact is a 6-DOF control input component widely used for intuitively guiding mobile robots, precise 3D navigation in Computer-aided design (CAD), or 3D analysis and review of digital models. The SpaceMouse has a dimension of $77\ mm \times\ 77\ mm \times\ 54\ mm$ and consists of a pressure-sensitive controller cap and two programmable buttons. The user can manipulate the spring-mounted controller cap to intuitively interact with robots in an isometric-rate control manner, and customize the two buttons for additional functionality. The SpaceMouse allows intuitive manipulation of the TCP motion of the robotic arm-hand system. The natural control inputs through the SpaceMouse are shown in **Figure 3.9 (b)**. The mapping between the user command motion and the movement of the robot TCP is shown in **Table 3.6**.

The operator can use the 3D mouse to control the robot at the remote site accordingly. The mouse transmits the control messages to the ubuntu computer in the user space via USB serial

port or 2.4 GHz wireless communication protocol, and the system generates control commands, which are then sent to the robot side via the Internet communication system to control the arm-hand system to complete the telemanipulation tasks.



(**a**) (**b**)

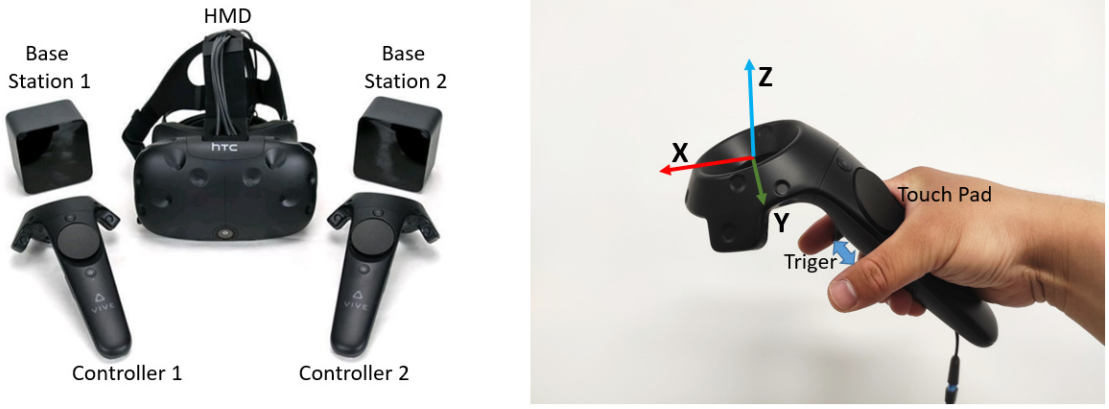
Figure 3.9: The SpaceMouse isometric-rate control input device used in the robotic system to command the robotic arm-hand subsystem. (a) The components of the SpaceMouse sensor; (b) description of the natural control inputs using the SpaceMouse device.

Table 3.6: Command and Function mapping from SpaceMouse to the robotic arm-hand.

No.	Control Input	Robot Function
1	Push or pull along X-axis	Translational motion of TCP along X-axis
2	Push or pull along Y-axis	Translational motion of TCP along Y-axis
3	Pull or press along Z-axis	Translational motion of TCP along Z-axis
4	Tilt around X-axis	Rotational movement of TCP about X-axis
5	Tilt around Y-axis	Rotational movement of TCP about Y-axis
6	Twist around Z-axis	Rotational movement of TCP about Z-axis
7	Button 1 pressed or released	Gripper half-closed or open
8	Buttons 1 and 2 pressed or released	Gripper fully closed or open

3.3.3 The Isotonic-Position Intuitive Interaction

The HTC Vive platform (HTC Corporation, Taiwan) is composed of a binocular head-mounted display (HMD), two laser-based positional tracking base stations, and two wireless motion-tracking controllers, as shown in **Figure 3.10 (a)**. The HMD is tracked by the two HTC Vive base stations and used to display the MR manipulation scene generated in Unity (Unity Technologies, USA) to the user.



(a) (b)

Figure 3.10: The HTC Vive isotonic-position motion tracking platform is used in the robotic system to command the robotic arm-hand subsystem. (a) The components of the HTC Vive tracking platform; (b) description of the natural control inputs using the Vive motion controller.

The operator moves the handheld Vive controller as an isotonic-position motion input device to grant human-level dynamic performance to the robotic arm-hand, and accomplish the remote telerobotic spatial manipulation tasks. The touchpad embedded in the Vive controller is used as an activation key, and only when the user touches the touchpad with the thumb, the captured information of the tracked controller is transmitted to the Ubuntu PC driving the robot. If no touch is made, the output command is $\boldsymbol{v}_u = [0,0,0,0,0,0]^T$ and the robot remains stationary. This clutch-based setting can effectively prevent the occurrence of misoperation caused by full-time tracking and commanding.

For workspace isomerism between the user and robot sides, the clutching mechanism interrupts the instantaneous velocity signals from the input device, effectively repositioning and remapping. If the robot's workspace has a significantly different scale than that of the human arm, the tracked user's hand velocity signal can be scaled appropriately by using a scaling factor. **Figure 3.11** shows the flowchart for implementing the clutching mechanism and repositioning the master device in telemanipulation tasks.

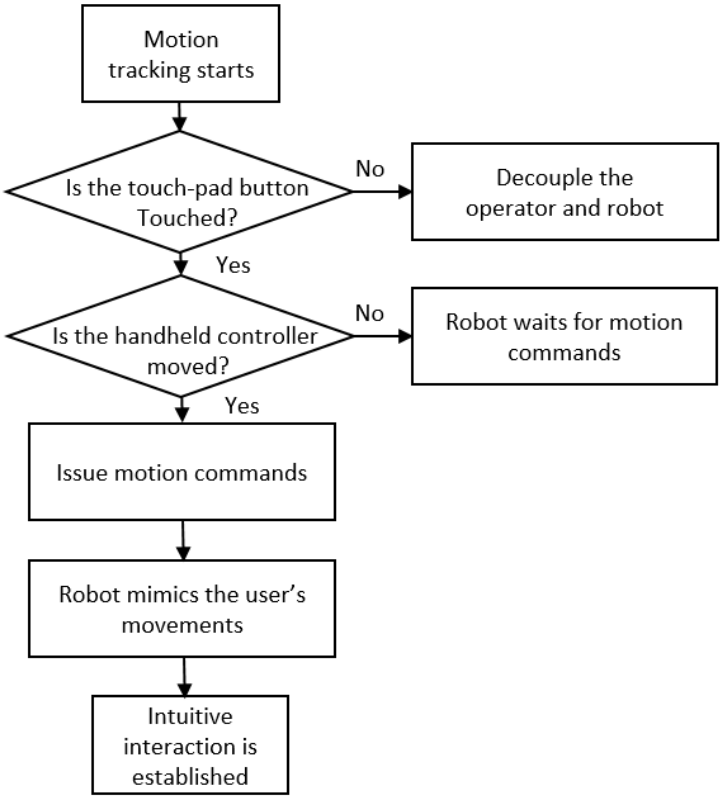


Figure 3.11: The flowchart of the algorithm for implementing the clutching mechanism and repositioning the master device during telemanipulation tasks.

The operator can also use the proposed clutching mechanism to release control of the robot and reengage control again after repositioning the motion input device to a user-desired position in the userspace. This human-machine interaction strategy allows the user to intuitively

manipulate the robot and navigate its workspace in situations where the robot's working range is significantly greater than the reaching limits of the operator's arm. The proposed clutch-based motion retargeting scheme is achieved by simply transmitting a zero matrix to the robot when the connection between the user and the robot needs to be temporarily discontinued. Since the user's input updates are based on their instantaneous motion state, when the operator resumes commands to the robot, the position and orientation of the end-effector continue smoothly from the point where the operator disconnected the operation.

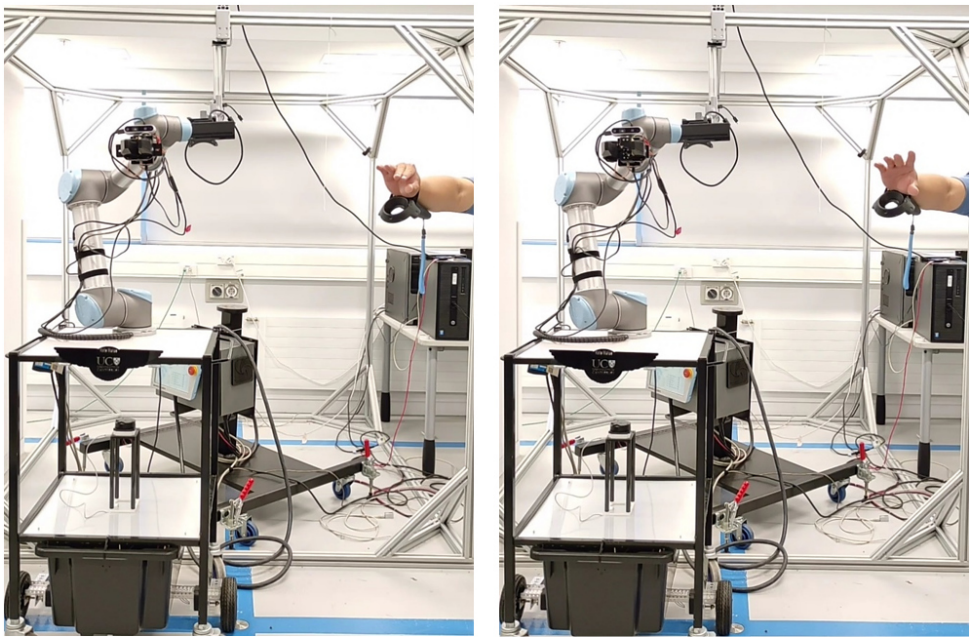
The HTC Vive base stations are linked to the Unity computer and track the position and orientation of the user’s head and hands to a sub-millimeter level accuracy. The base stations use alternating horizontal and vertical lasers to scan across the HTC Vive headset and handheld controllers, which are equipped with sensors that detect the lasers as they pass by. The user's hand movements are tracked by two base stations that trace the handheld HTC Vive controllers at a refresh rate of 90 Hz. The acquired position and orientation data is calculated and transmitted to the ROS-based controller of the UR5 robot, where the input velocity values are converted into robot joint velocities.

3.3.4 Optical Tracking-Based Interaction

The remotely operated arm-hand system consists of a 6-DOF UR5 industrial robot manipulator and a parallel robot gripper. The optical tracking-based command input system at the user side is composed of an HTC Vive motion tracker and a Leap Motion hand tracking sensor (Ultraleap, USA) which track the user’s arm and hand movements respectively. The Leap Motion controller can accurately capture the user's finger movements for robot hand control [146], and the Vive motion tracking module has relatively large capturing coverage of the user's

arm movements which is suitable for robot arm control, the combination of the two can achieve intuitive and accurate teleoperation of the robotic arm-hand system.

The Leap Motion tracking data is utilized to classify the closing and opening movement of the user's fingers, which are then automatically executed by the ROS-controlled two-finger robotic hand. The Vive motion tracker sensor measures the operator's hand pose in real-time, and the user's hand velocity is used to directly control the velocity of the tool center point (TCP) of the robotic arm. **Figure 3.12** shows the proposed optical tracking-based interaction approach which integrates the Vive motion tracker and Leap Motion sensor.



(a) (b)

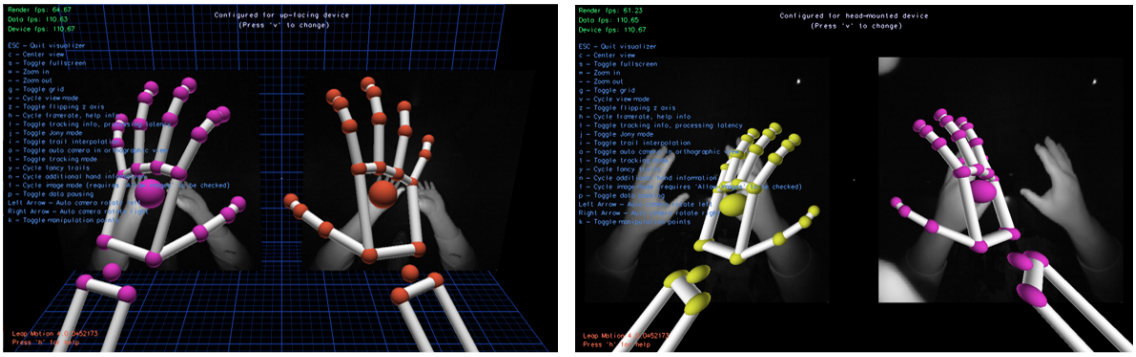
Figure 3.12: The robotic system is controlled with the user's arm and hand movements by using the Vive motion tracker and Leap Motion sensor integrated motion tracking approach. The Leap Motion sensor is mounted on the user's forehead. (a) The robot's hand closes as the user's hand pinch distance decreases. (b) The robot's hand opens as the user's hand pinch distance increases.

Leap Motion controller is a non-contact motion tracking device that primarily focuses on hand gesture recognition and finger movement tracking. The module has a dimension of $80\ mm \times 30\ mm \times 11.3mm$ $(L \times W \times H)$. The Leap Motion Controller has a tracking depth from $100\ mm$ to $800\ mm$ and a field of view of $140\ \times\ 120°$ [147], as shown in **Figure 3.13**.

(a) (b)

Figure 3.13: The Leap Motion tracking device and its interaction zones [147]. (a) The Leap Motion hand tracking hardware. (b) The tracking zone and field of view where the user's hand can be detected.

The Leap Motion device features hand/finger motion tracking and hand posture recognition functionalities, which can be used to produce the control signals to provide a more natural and intuitive method for human-robot interaction. **Figure 3.14** shows the two tracking modes used in robotics to command the robotic arm-hand subsystem. However, the user's hand movements used to control the UR5 manipulator are constrained by the limited field of view of the Leap Motion controller sensor. The control of the robotic arm-hand movements in the optical tracking-based teleoperation scheme is performed by utilizing a Leap Motion controller and Vive tracker for finger tracking and hand tracking respectively.



(a) (b)

Figure 3.14: The Leap Motion tracking modes used in the robotic system to command the robotic arm-hand subsystem. (a) Desktop tracking mode. (b) Head-mounted tracking mode.

On the user side, the teleoperation command input system includes an HTC Vive tracker and two base stations which captures the exact position and orientation of the user's hand, a Leap Motion hand tracking sensor is mounted on the user's forehead with an adjustable strap to track the motion of the user's finger movements. The Leap Motion controller is separately connected to Unity on the PC to sense the user's hand gestures, which are converted into motion commands afterward. The user can issue commands by presenting simple and intuitive hand gestures in the Leap Motion field of view. Once the user's hand is detected in the Leap Motion sensors field of view and the data generated are within the given range.

The Cartesian position and orientation of the user's fingers are retrieved from the Leap Motion sensor. As shown in **Figure 3.15**, the Leap Motion sensor measures the pinch distance D_p between the thumb and index finger of the user's hand pose in millimeters. The tracking data are returned by the Leap Motion sensor in the form of frames and certain motor motion commands θ_p are issued to the servo motor driver which is embedded in the base of the robot hand and moves the robot fingers for object manipulation. The user's pinch range is mapped to the servo range $[c, d]$ that controls the full opening and closing of the robot hand. The input commands to the robot are limited to the operator's hand pinch motion constraints $[a, b]$ with

$a \neq b$. The motion retargeting function $f(x)$ is used for shifting the user's pinch distance into the servo range. The motion drives the robot hand to interact with a remote environment in a natural manner using the function defined:

$$f(x) = c + \left(\frac{d-c}{b-a}\right)(x-a) \tag{3.20}$$

where x is the pinch distance value D_p within the interval $[a, b]$, with $a \leq x \leq b$ and $a \neq b$.

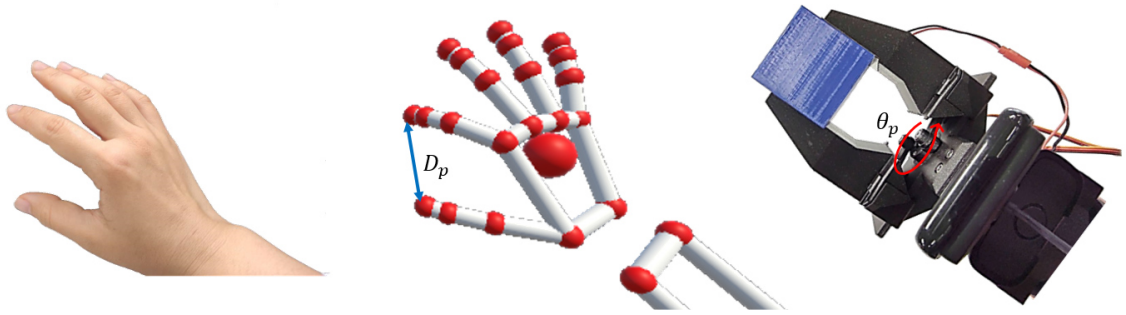


Figure 3.15: The non-contact intuitive hand control through user hand pinch distance and robotic hand motion retargeting

By using contact control devices, such as 3D mouse buttons and controller triggers to control the robotic hand, the user can accurately determine the beginning and stop of the robot's movements. However, for non-contact components, such as leap motion, once the user's hand appears in the sensor's field of view, the pinch measurement begins and the corresponding control command is transmitted to the robotic hand actuator to control its motion. To eliminate motion commands generated by the user's hand inadvertently entering the sensor's capture range and filter out unwanted motion commands and prevent misoperation, a hand control algorithm is developed to translate sensor messages into servo commands and enable intuitive and safe human-robot interaction, as shown in **Figure 3.16**.

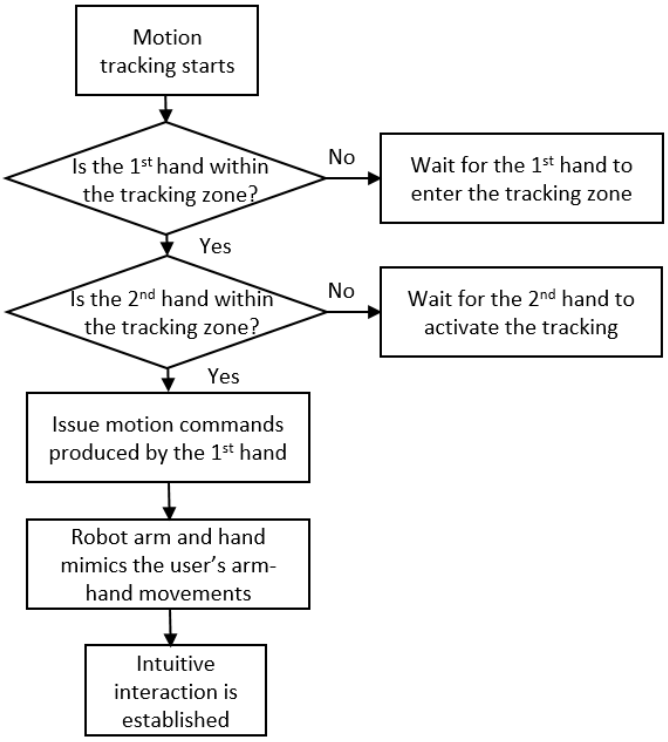


Figure 3.16: The flowchart of the hand control algorithm for implementing the second hand as a switch for sending and stopping commands during telemanipulation tasks.

The user controls the robotic arm-hand system with two hands, with only one hand sending commands. If only one hand appears in the Leap Motion tracking area, the system does not send commands to the robot hand. When the second hand enters the interaction area, the motion of the first hand is mapped to the robot hand. When the second hand leaves the interaction zone, the first hand discontinues issuing commands to the robot hand. The second hand is utilized only to connect and decouple the first hand from the robot hand for the control of command transmission, which increases the security of the system while maintaining intuitiveness.

3.3.5 Haptic Intuitive Interaction

The Phantom Omni haptic device (SensAble Technologies Inc., USA) is a commercial and portable haptic device with 6 DOF. The Phantom Omni has a workspace of $160 \times 120 \times 70$ mm (W × H × D) and can provide force feedback of up to 3.3 N. The Phantom Omni system is a robotic device based on a serial architecture with six revolute joints, three of which are actuated by direct current (DC) motors. The tip of the stylus, which is the end effector of the haptic device, is driven by motors equipped with encoders and potentiometers to span and measure over the full workspace. The haptic stylus is linked to the housing through a single serial chain, as shown in **Figure 3.17**.

Figure 3.17: The typical intuitive interaction approach using the Phantom Omni haptic device.

The encoders embedded in the Phantom Omni haptic device track the user's hand motion or position along the X, Y and Z axes. Measurement of rotations about these axes (roll, pitch and yaw) is achieved using potentiometers. The motors also provide the user with haptic feedback by exerting force on the stylus along the X, Y and Z axes. The updated coordinates of the hand

motions are transmitted to a Windows PC for mapping from the user space to the robot space. This mapping involves linear velocities and angular velocities. The operator moves the haptic handheld stylus as an isotonic-position motion input device to grant human-level dynamic performance to the robotic arm-hand, and to perceive the haptic effects to enhance performance while conducting the tele-manufacturing tasks.

The Denavit-Hartenberg (D-H) representation approach is used for kinematic modeling of the Phantom Omni robotic system. Frames are affixed to links of the haptic robot to describe the kinematic parameters between two neighboring coordinate systems, as it is shown in **Figure 3.18**. The Definition of D-H parameters in terms of the link frames is shown in **Table 3.7**. Only the four kinematic parameters are needed to determine the homogeneous transformation matrix.

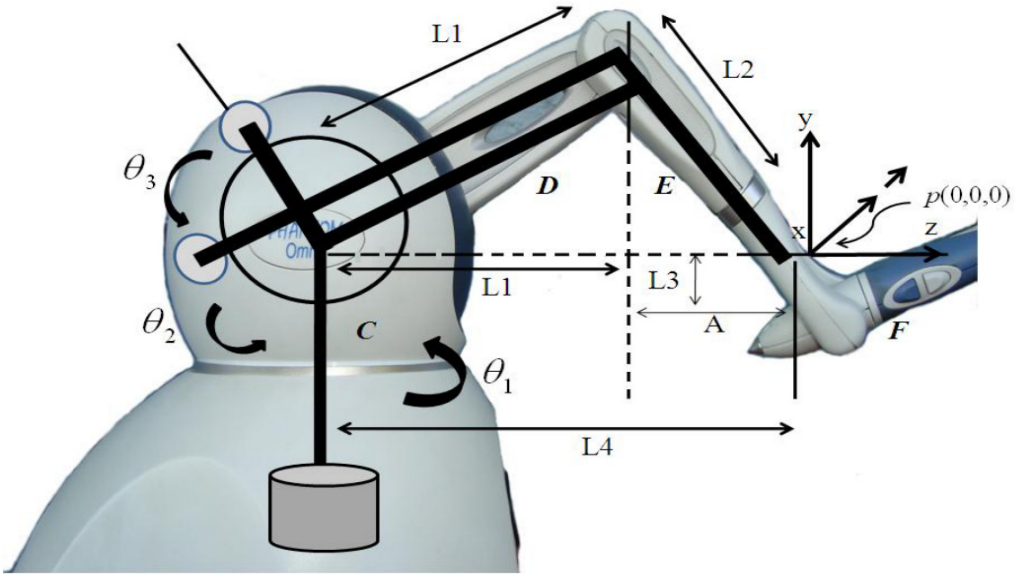


Figure 3.18: Diagram for mathematical modeling of the Phantom Omni haptic device [148].

Table 3.7: D-H parameters of Phantom Omni haptic robot.

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	90°	0	0	θ_2
3	0	L_1 = 135.0 mm	0	θ_3
4	90°	0	L_2 = 135.0 mm	θ_4
5	-90°	0	0	θ_5
6	90°	0	0	θ_6

The forward kinematics calculates the transformation relationship from the joint coordinates θ of the haptic device to the operational position X of the end-effector (haptic stylus), given link parameters including the link lengths and offsets and joint variables.

The 3×1 vector $X \in \mathbb{R}^{3\times1}$ represents the haptic stylus coordinates and $\theta \in \mathbb{R}^{3\times1}$ denotes the joint coordinates vector. And $c_i = cos\theta_i$ and $s_i = sin\theta_i$ for $i = 1, \dots, n$. Given the joint values θ in the robot joint space, the haptic stylus coordinates for position $X = [x\ y\ z]^T$ in Cartesian space can be expressed [148]:

$$X = f(\theta) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -s_1(L_1c_2 + L_2s_3) \\ L_1s_2 - L_2c_3 + L_3 \\ c_1(L_1c_2 + L_2s_3) - L_4 \end{bmatrix} \tag{3.21}$$

where the length of the link $L_1 = L_2 = 135.0\,mm\ \ L_3 = 25.0\,mm\ L_4 = 170.0\,mm$.

The Phantom Omni haptic robot is used as a motion input device at the user side, also referred to as the master robot, for velocity-based motion mapping. To convert the haptic stylus velocity,

which represents the operator's hand motion into a differential joint command for the robot controller, it is essential to derive the instantaneous relationship between the haptic robot joint rate vector $\dot{\theta}$ and the end-effector's Cartesian velocity V_p. The Cartesian velocities of the stylus tip related to the robot joint rates are expressed:

$$V_p = J_{p1}(\theta)\dot{\theta}_1 + \cdots + J_{pn-1}(\theta)\dot{\theta}_{n-1} + J_{pn}\dot{\theta}_n \tag{3.22}$$

where each jacobian $J_{pi}(\theta)$ depends explicitly on the joint values θ for $i = 1, \dots, n$. Since the Phantom Omni is a serial 6-DOF input, 3-DOF output device, the Jacobian matrix $J_p(\theta) \in \mathbb{R}^{3\times3}$ governs the joint velocity vector $\dot{\theta} \in \mathbb{R}^{3\times1}$ to the desired spatial twist $V_p = [\dot{x}\ \dot{y}\ \dot{z}]^T \in \mathbb{R}^{3\times1}$ is computed:

$$V_p = J_p(\theta)\dot{\theta} = [\,J_{p1}(\theta) \quad J_{p2}(\theta) \quad J_{p3}(\theta)] \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} \tag{3.23}$$

The translational and rotational input commands to the haptic device are limited to the operator hand motion constraints $[-V_{p,max}, V_{p,max}]$ with $V_{p,max} > 0$. The matrix form of the Jacobian is derived:

$$V_p = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -c_1(L_1c_2 + L_2s_3) & L_1s_1s_2 & -L_2s_1c_3 \\ 0 & L_1c_2 & L_2s_3 \\ -s_1(L_1c_2 + L_2s_3) & -L_1c_1s_2 & L_2c_1c_3 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} \tag{3.24}$$

The solution to controlling the haptic device so that it follows the desired stylus tip motion is to calculate the inverse differential kinematics $\dot{\theta}$ from the relationship $V_p = J_p(\theta)\dot{\theta}$, and defined:

$$\dot{\theta} = {J_p}^{-1}(\theta)\, V_p \tag{3.25}$$

where $\dot{\theta}$ is the 3×1 vector of joint velocities, and ${J_p}^{-1}(\theta) \in \mathbb{R}^{3\times3}$ is the corresponding inverse matrix to be extracted from the geometric Jacobian $J_p(\theta)$. V_p is the 3×1 vector of the haptic stylus velocities of concern for the specific manipulation task.

3.4 Software Implementation

3.4.1 Software System Overview

ROS is an open-source framework of libraries and tools for robotic prototyping and controlling. The ROS ecosystem is composed of several highly interconnected and dependent nodes, each of which represents the module of executable code and communicates with the other nodes by sending and receiving messages through specific topics.

ROS messages are data formats with typed fields and organized into particular topic names for identification. The publish/subscribe messaging protocol enables a node to register with the ROS master node, receive subscribed data through a topic, execute the program contained in the node and send the processed output data through another topic. ROS nodes are primarily based on C++ or Python programming languages and can talk to each other across different languages. ROS communication protocol is shown in **Figure 3.19**.

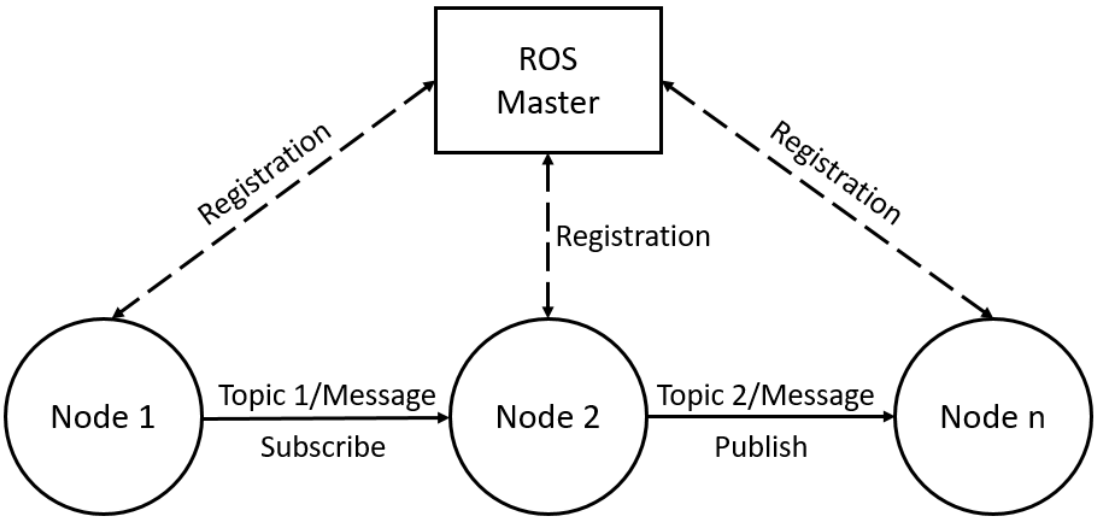


Figure 3.19: The ROS communication protocol between the nodes registered at the ROS Master.

ROS encompasses a full range of device drivers, sensor libraries, package management, efficient algorithms, user interfaces and visualization software to prototype, control and monitor the robotics platform. Robot development using ROS is highly extensible, allowing the application of intensive software packages to enhance the robot's functionality. The software development for interacting with the mobile manipulation system has been implemented in ROS, which is a flexible framework for writing robot software, using the C++/Python programming language. For this application, different human-robot interaction strategies with multiple input devices are implemented for the control of the arm-hand subsystem as well as the mobile module.

3.4.2 ROS Implementation

The 3D-printed gripper is implemented as a node of the ROS ecosystem, which subscribes to servo motor commands and publishes sensor data in the message stream. The Arduino board embedded in the robot gripper contains a microcontroller handling one ROS node at a time.

The real-time data exchange between the Ubuntu PC and the parallel gripper is achieved using the rosserial_arduino package and applying ROS communication protocol over the Arduino serial ports.

For the UR5 manipulator, the commands generated by the various control input devices are all converted in the geometry_msgs/Twist message type. A twist message is composed of geometry_msgs/Vector3 linear and geometry_msgs/Vector3 angular and it is used to control the corresponding linear and angular velocities of the UR5 TCP. The teleoperation control of the UR5 manipulator is provided through the jog_arm package which is a Jacobian-based driver of ROS-compatible robot arms. The sensor_to_twist node subscribes to the /geometry_msgs/Twist messages and publishes /jog_arm_server/joint_delta_jog_cmds message as the velocity control input to the jog_arm package of ROS. And the jog_arm_server node subscribes to the /jog_arm_server/joint_delta_jog_cmds messages and publishes /ur_drvier/joint_speed message as the input signal to the ur_driver node which controls the UR5 arm, as shown in **Figure 3.20**.

For the mobile module, the geometry_msgs/Twist message type is used for publishing motion commands to the base controller node which governs the motor driver and robot controller. The node converts the received messages from the keyboards/controllers/motion sensors into geometry_msgs/Twist messages, which contain six-dimensional values representing the linear velocity along each axis and the angular velocity around each axis of the robots. The differential drive type robots move only forward or backward along its horizontal axis and rotate only around its vertical axis. Only the linear X-component and the angular Z-component are required to portray the robots' motion. Hence, the geometry_msgs/Twist messages only convey the linear velocity of the mobile module along its forward X-axis and the angular velocity

around its vertical Z-axis (the other values remain zero). The base controller node subscribes to the /cmd_vel topic and translates geometry_msgs/Twist messages into motor signals that drive the wheels.

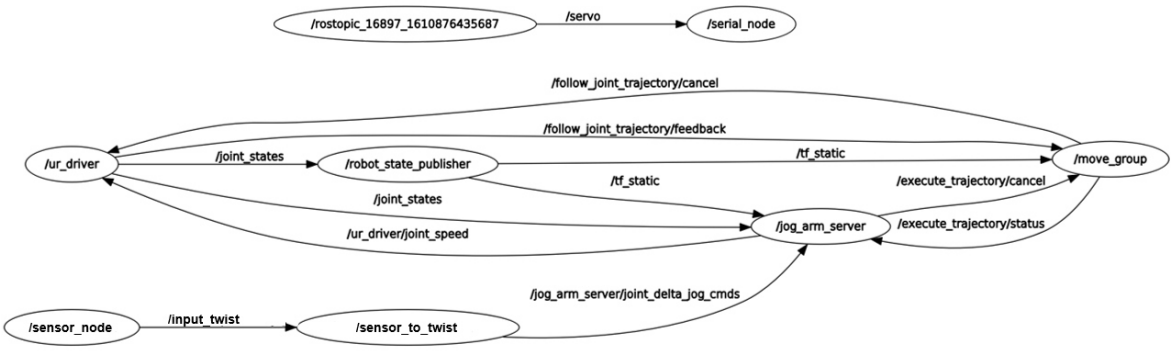


Figure 3.20: The ROS communication between the nodes for the implementation of intuitive robot control.

In ROS, the robot modeling is represented in Unified Robot Description Format (URDF), which is an XML-based robot description and specification to represent all kinematic and dynamic elements of a robot in the ROS ecosystem. The robot_sate_publisher node is used for updating the spawned URDF robot model in both ROS simulation tool Gazebo and visualization tool Rviz. This node takes robot URDF parameters and robot joint status as input and publishes the pose of the robot links. In this application, the robot_sate_publisher node is used to update the URDF robot model in the MR environment.

3.4.3 Unity Implementation

Unity 3D is a fully integrated development engine for the creation of interactive 3D content. Unity 3D is utilized in the intuitive teleoperation platform to interface with the command input devices and generate the MR environment for human-robot interaction. The implementation of

ros-bridge as an intermediate integrates the communication seamlessly between ROS running on Linux PC and Unity 3D on Windows system. **Figure 3.21** shows the implementation schema of the bi-directional communication between Unity and ROS for interfacing the user command input devices and the robot entities.

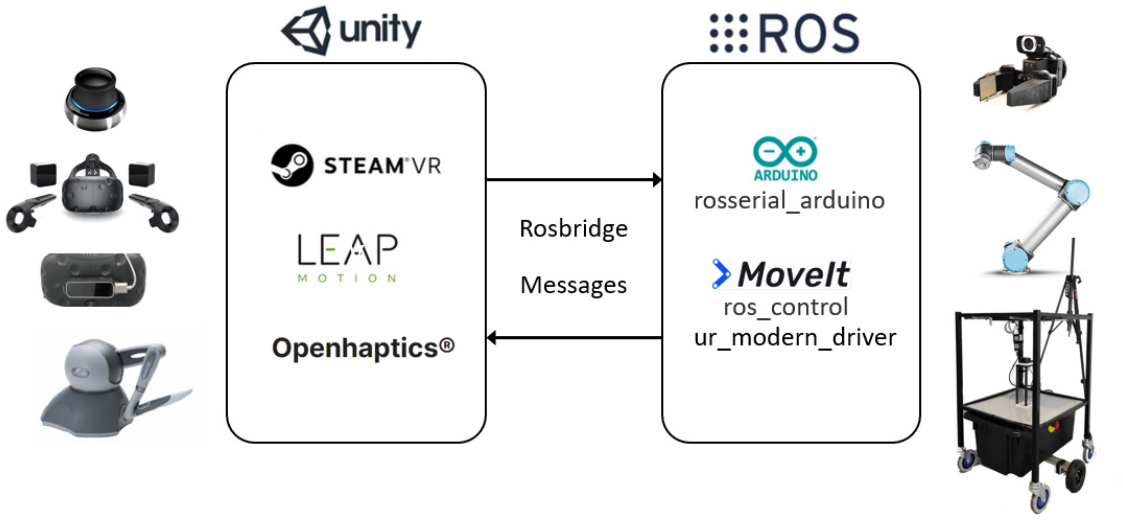


Figure 3.21: The implementation schema of the bi-directional communication between the user command input devices and the robot entities.

When the communication between the Unity 3D engine and ROS system is built up, an immersive user interface can be developed with the library and tools provided by the 3D engine and the different levels of robot control can be achieved by taking advantage of the ROS system. The bi-directional communication facilitates the implementation of MR with ROS to integrate its modular robotic functionalities. In addition, the URDF model of the robotic platform is imported into Unity by publishing data to the parameter server in ROS. Hence, the poses and movements of the digital twin of the robot in Unity adapt to the ones of its real counterpart in the remote environment simultaneously.

3.5 Summary

In this chapter, a comprehensive mechatronic design and development procedure of a mobile manipulator robot platform for intuitive remote manipulation is presented. To achieve this goal, an integrated application of CAD/CAE/CAM and rapid prototyping manufacturing has been implemented. This robot platform enables the development of a mobile manipulation platform meeting design requirements including modularity, manufacturability and manipulability. With the integrated application, the robotic platform has been successfully developed at a relatively low cost in terms of expenses and time, and provides a novel technology platform directly designed for this research.

The intuitive control schemes of the robotic platform for mobile manipulation based on kinematic modeling are implemented. The goal is to allow unskilled users with minimal pre-knowledge of robotic systems to intuitively interact with the proposed platform for different remote manufacturing scenarios. For this purpose, multiple interaction schemes with four different input devices are developed. ROS integrates all the modular robotic functionalities and ensures the communication between the nodes driving the robotic components. Unity 3D is interfaced with the control input devices for the user interaction, which allows the development and implementation of the 3D immersive human-robot interaction environments.

This overall framework presented allows the development and incorporation of MR-based interactive interfaces and enables the implementation and experimental validation for diverse remote manufacturing tasks on the developed robotic platform. The incorporation of MR technologies and experiments for validation of different remote manufacturing tasks will be presented in the following chapters.

Chapter 4

Mixed Reality-Enhanced Human-Robot Interaction with Imitation-Based Mapping Approach for Intuitive Teleoperation

4.1 Introduction

This chapter presents an integrated mapping of motion and visualization scheme based on the MR subspace approach for intuitive and immersive telemanipulation of robotic arm-hand systems. The effectiveness of different control-feedback methods for the teleoperation system is validated and compared. The robotic arm-hand system consists of a 6-DOF industrial manipulator and a low-cost 2-finger gripper, which can be manipulated in a natural manner by novice users physically distant from the working site, as presented in Chapter 3.

By incorporating MR technology, the user is fully immersed in a virtual operating space augmented by real-time 3D visual feedback from the robot working site. An imitation-based velocity-centric motion mapping is implemented via the MR subspace to accurately track operator hand movements for robot motion control, and enables spatial velocity-based control of the robot tool center point (TCP). The proposed system allows precise manipulation of end-effector position and orientation to readily adjust the corresponding velocity of maneuvering.

The proposed MR subspace integrated mapping approach implements point cloud-based 3D visual feedback to provide visual inspection with depth information and perception of the working site, while the robot agent works in the remote environment where manipulation tasks are performed. The user control space and robot working space are overlaid through the MR subspace, and the local user and a digital twin of the remote robot share the same environment in the MR subspace. The MR-based motion and visualization mapping scheme for telerobotics is compared to the conventional 2D Baseline and MR tele-control paradigms over two tabletop object manipulation experiments. A user survey on user performance and experience demonstrates the effectiveness and performance enhancements enabled by the proposed system.

Overall, the MR subspace integrated 3D mapping of motion and visualization scheme presented in this chapter reduced the aggregate task completion time by 48% compared to the 2D Baseline module and 29% compared to the MR SpaceMouse module, while operation and perceived workload decreased.

4.2 Methodology

4.2.1 Teleoperation System Overview

The proposed robotic manipulation system and the related mapping are shown in **Figure 4.1**. On the MR side (local space), a binocular head-mounted display (HMD) is tracked by the HTC Vive MR platform (HTC Corporation, Taiwan) and used to display the virtual manipulation scene generated in Unity (Unity Technologies, USA) to the user. Motion tracking devices (HTC lighthouses) are linked to the Unity computer and track the pose of the user's head, while they inspect the workpiece from different perspective angles and perform teleoperation tasks.

4.2.1.1 Robot Space and User Space Communication

A Kinect V2 depth sensor (Microsoft Corporation, USA) publishes compressed RGB-D images from the perspective of the robot's head to produce the point cloud-based 3D visual feedback in the MR subspace. Compressed images enter the MR subspace through a custom image subscriber and depth image subscriber in Unity 3D. In Unity, RGB images and depth images are decompressed using OpenCV for C# (Microsoft Corporation, USA). A custom "material" shader in Unity combines these images into a 3D point cloud from the RGB image and the depth image, and the computation is done in parallel on the GPU.

Onsite sensory information from the remote workspace is transferred back to the human operators. The operators act as part of the control loop to interact with the remote robotic arm and hand system via a computer-generated virtual environment and overlaid multisensory data. The 3D virtual scene is rendered according to the actual robot working environment and displayed to the user via the HMD. A force sensor attached to the robot finger is primarily dedicated to monitoring the grasping process. The user can directly observe the workpiece and

end-effector status with the real-time overlaid sensory information in MR during operation. The depth sensor information is mapped and reconstructed in the point cloud form on the master side to be completely consistent with the environment of the slave side. The proposed teleoperation system allows information mapping between the user side and the robot side using an MR subspace [80].

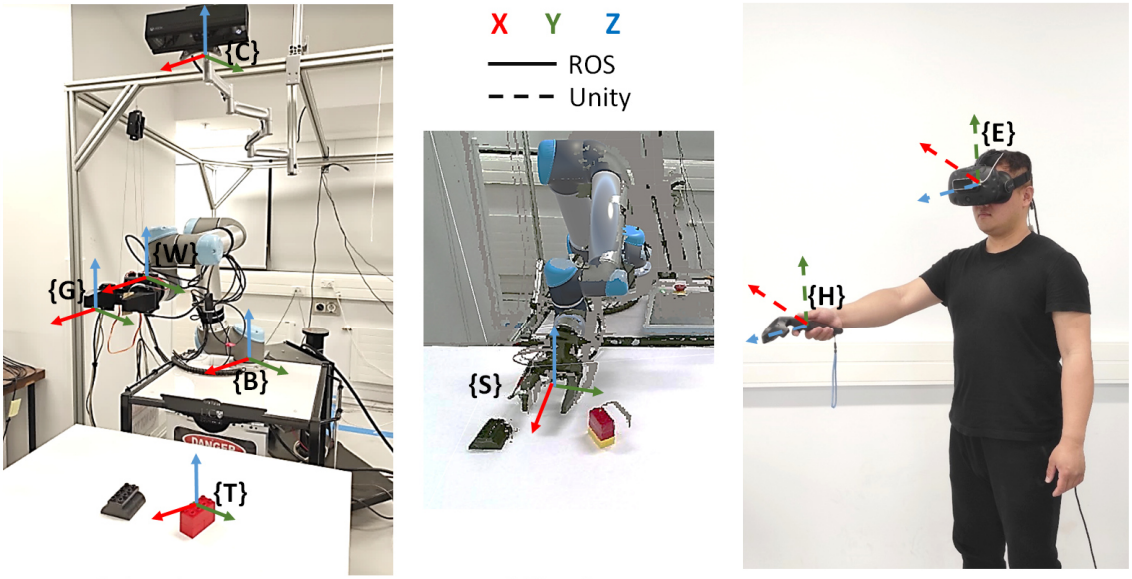


Figure 4.1: Schematic of the motion mapping approach for the MR subspace-enhanced imitation-based human-robot interaction system of teleoperated robotic arm-hand. The MR subspace decouples the human from the robot, and the mappings for sensors and grippers are not necessarily direct and identical. The user's eyes {E} and hand {H} and the state of the camera {C} the robot arm's grippers {H} and are linked to each other through MR subspace {S}, with {T}, {W} and {B} denoting reference frames of the target objects, robot wrist and robot arm base.

4.2.1.2 MR Subspace

The MR subspace serves as an intermediary between the user command loop and robot control loop. A digital twin system is generated in the virtual environment and features a synchronous representation of the physical UR5 robot (Universal Robotics, Denmark) employed. The

remote human operator can inspect the sensory information, robot control command input, and robot pose in its work environment through a 3D MR subspace interface. The virtual replica allows the user to inspect how the physical robot is situated in the remote environment without deploying an array of static cameras. The digital twin subscribes to the joint state data of the physical UR5 robot and updates its configuration accordingly in MR using ros-bridge software provided by the ROS. In the MR scene, the user and the remote robotic platform share the same space virtually. The digital twin enhanced MR subspace interface is shown in **Figure 4.2**.

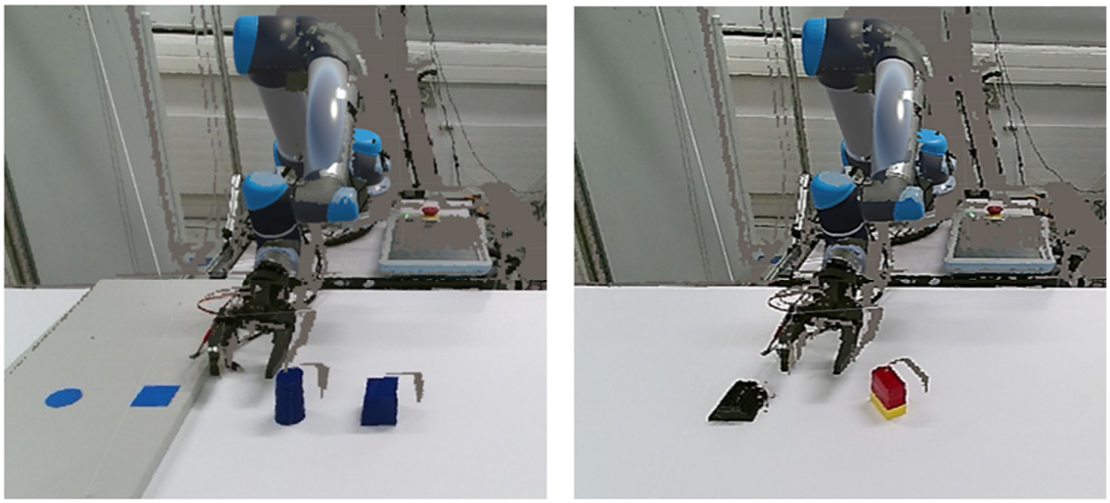


Figure 4.2: Overview of the MR subspace interface with the digital twin of the UR5 robot superimposed on the 3D Point Cloud of the physical UR5 robot in the MR environment. (a) The pick-and-place task of two cubical and cylindrical objects given a goal position with different heights, (b) the assembly task involved grabbing one LEGO subassembly from a predefined spot and stacking it onto a fixed LEGO base on the table.

In typical telerobotic systems with a 2D camera feed, the user is provided with a monocular RGB stream of the remote scene. In the virtual subspace, a human operator monitors the spatial manipulation process using the real-time 3D point clouds for remote environmental visualization. The digital twin can either be situated in front of the user or superimposed on the user depending on their requirements.

The operator uses the handheld HTC Vive controller as a motion input device to grant human-level dynamic performance to the robotic arm-hand for remote telerobotic spatial manipulation tasks. The Vive tracking system tracks the controller position to a sub-millimeter level accuracy. User hand movements are tracked by two base stations tracking the handheld HTC Vive controllers at a refresh rate of 90 Hz. The HTC Vive base stations use alternating horizontal and vertical lasers to scan across the HTC Vive headset and handheld controllers, which are equipped with sensors to detect the lasers as they pass. The HTC Vive system integrates all this data to determine the position and rotation of the device being tracked components in the 3D MR subspace. The acquired position and orientation data is calculated and transmitted to the ROS-based controller of the UR5 robot, where the input velocity values are converted into robot joint velocities.

4.2.2 Velocity-Centric Motion Mapping

For telemanipulation applications, the remote robot should maintain smooth, accurate tracking of user hand movement. However, raw hand movement data provided by the Vive tracking system contains intended hand motion, as well as tremors and noise. Thus, direct velocity mapping causes aggressive control maneuvers and jittering robot motion. Single Exponential Smoothing (SES) is applied to remove unintended short-term fluctuations and reduce hand tremors and noise. The SES for filtering out the noise from the hand motion series can be calculated recursively:

$$V_{t+1|t} = \alpha V_t + \alpha(1-\alpha)V_{t-1} + \alpha(1-\alpha)^2 V_{t-2} \ldots + \alpha(1-\alpha)^k V_{t-k} \qquad (4.1)$$

where V_t are velocity measurements and $0 \leq \alpha \leq 1$ is the smoothing parameter. The decrease rate of the weighting terms is controlled by the smoothing parameter. If α is large and close to 1, more weight is given to the more recent hand motion observations. If α approaches 0, the output velocity signal tends to be the average of the historical input velocity data. A value of $\alpha = 0.9$ is given in previous human motion tracking-based teleoperation research and verified in our experiments, which gives the optimal performance in the motion smoothing application according to experiments [149]. The k value indicates the number of time steps the algorithm looks back while smoothing the date. The k is set as 1 in our system, which means one-step-back motion data V_{t-1} is used for the one-step-ahead smoothing $V_{t+1|t}$. And k = 1 is given in previous human motion tracking-based teleoperation research and verified in our experiments [149].

The approach for intuitively controlling the robot so its end-effector smoothly follows a user's hand trajectory, $T_u(t)$, is to calculate and control the required n-vector of joint velocities $\dot{\theta}$ directly from the relationship $J(\theta)\dot{\theta} = V$, where the desired robot end-effector twist $V_r = [v_r\ \omega_r\ ]^T$ and $J(\theta) \in R^{6\times n}$ are expressed in the same frame. At time k, the motion tracking system measures and determines the configuration of the user's hand, converts this calculated spatial velocity $V_u = [v_u\ \omega_u\ ]^T$ to a Cartesian twist command $V_r = [v_r\ \omega_r\ ]^T$ to the robot in the ROS coordinate system, employs the Inverse-Jacobian solver to determine the appropriate joint rate vector, $\dot{\theta}$, according to the desired twist representation of end-effector motion, V_r, and derives the joint configuration sent to the UR5 robot controller, as defined:

$$\dot{\theta} = J^{\dagger}(\theta)\, SV_r \tag{4.2}$$

$$V_u = [v_u\ \omega_u]^T = kV_r = [v_r\ \omega_r]^T = J_r(\theta)\dot{\theta} = [J_v(\theta)\ J_\omega(\theta)]\dot{\theta} \quad (4.3)$$

The use of the pseudo-inverse in Equation (4.2) implicitly weights the cost of each of the 6 joint velocities identically, returns the minimized two-norm of joint velocities, and reduces the energy consumption of the robot [140]. S is a positive scaling factor, and $S = I$ is chosen in the application for precisely mimicking the user's movement, where $I \in \mathbb{R}^{6\times6}$ is an identity matrix.

Fingertip movements of the index finger collected from the hall-effect sensor embedded in the Vive controller are projected onto the X-Z plane perpendicular to the palm and contain index finger movements without abduction in the Y direction. The projected tip positions are transformed and scaled to the MR subspace through a standard frame transformation. When the user extends their index finger away from or towards the palm, the press depth detected by the sensor is updated accordingly. The press depth is mapped to the workspace limits of the robotic hand while it approaches the open and close pose for the manipulation of objects. A grasping depth index (GDI) is proposed as the criterion for grasping modulation of the robotic gripper, rather than reading binary values of 0 or 1 directly.

4.3 Experiments

To validate whether a teleoperation intermediary using MR subspace enhanced spatial mapping of the human motion facilitates control of robotic arm-hand systems by unskilled operators, a set of spatial manipulation tasks, shown in **Figure 4.3**, including pick-and-place and assembly, were developed. A user study was designed and conducted to verify task performance and user experience with the proposed HRI system compared to two typical teleoperation modes.

4.3.1 Hypothesis

The null hypothesis (H0) for a repeated-measures ANOVA is the MR subspace enhanced imitation-based motion mapping approach for telemanipulation and the two alternatives, commonly used teleoperation methods have identical effects in task efficiency, user performance and system usability for unskilled users, in terms of effectiveness, intuitiveness, and learnability.

4.3.2 Experimental Setup

A within-subjects experiment was developed to verify the null hypothesis by presenting novice participants with three HRI conditions (Baseline/2D SpaceMouse; MR direct control/MR SpaceMouse; and MR subspace) in a random order to guide the robotic arm-hand system and complete two spatial manipulation tasks: **1**) pick-and-place; and **2**) assembly.

A total of 24 participants were run through the user study, 13 males and 11 females. Subjects were all from the University of Canterbury and ranged in age from 20 to 31. In general, participants were unfamiliar with robotic systems, HRI, or MR. None had experience with controlling a robot, and they were regarded as novice users of the teleoperated robotic system. The tasks for the user study were to guide the robotic arm-hand system to conduct pick-and-place and assembly telemanipulation processes through three different teleoperation intermediaries**.** Each participant completes the user study within 60 minutes.

The proposed MR-HRI mapping approach was compared to a SpaceMouse 6-DOF control input device sensor (3Dconnexion, USA) with 2D camera feeds, and the SpaceMouse with 3D point cloud interface of telemanipulation. The 3D connexion SpaceMouse Compact is a 6-DOF

control input component widely used for guiding robots, precise 3D navigation in Computer-aided design (CAD), or 3D analysis and review. The 6-DOF SpaceMouse was chosen as a Baseline because prior work indicated that despite the complex operations, the SpaceMouse interface was highly effective for the Cartesian teleoperation of a robot end-effector. The experiment goal is to test whether the isotonic-position control scheme with MR subspace outperforms isometric-rate control using the SpaceMouse input with either 2D direct camera feeds or 3D visual feedback.

The Baseline HRI is a typical teleoperation mode providing participants with 2D visual feedback displaying the working view through the RGB camera in the Kinect V2 depth sensor. The user is presented with a 2D ego-centric view of the robot's working space in the monitor and unable to change perspective. The piloting metaphor was used for human-robot interaction with a stationary 6-DOF SpaceMouse. The user interaction included manipulating input from a SpaceMouse for TCP translation and rotation of the UR5 robot arm and the grasp and release of the robot hand.

The MR direct control/MR SpaceMouse module can be regarded as MR subspace without velocity-centric motion mapping (VCMM) and is a condensed version of the proposed MR-HRI system that enables the user to inspect the robotic arm-hand situation in the remote environment via a 3D point cloud in MR subspace. The user can interact with the robotic system through an isometric-rate control scheme by using the SpaceMouse input with 3D visual feedback, but without perceived correspondence between the input device motion and the robot end-effector movement.

The MR subspace experiment is a full MR-HRI system with VCMM using the tracked Vive motion controller in a grasping metaphor and allows the user to use hand motion to directly guide the robotic arm-hand system through the isotonic-position control scheme, and view the robot's space through the point cloud generated within MR subspace, conducting a set of spatial manipulation tasks. The full MR-HRI system features the perceived movement correspondence between the input device (Vive controller) and the robot end-effector. In MR subspace, the movement of the user's arm and hand commands that of the controlled robotic arm and hand system so that the user can perceive the movement of the robot through his/her own movement.

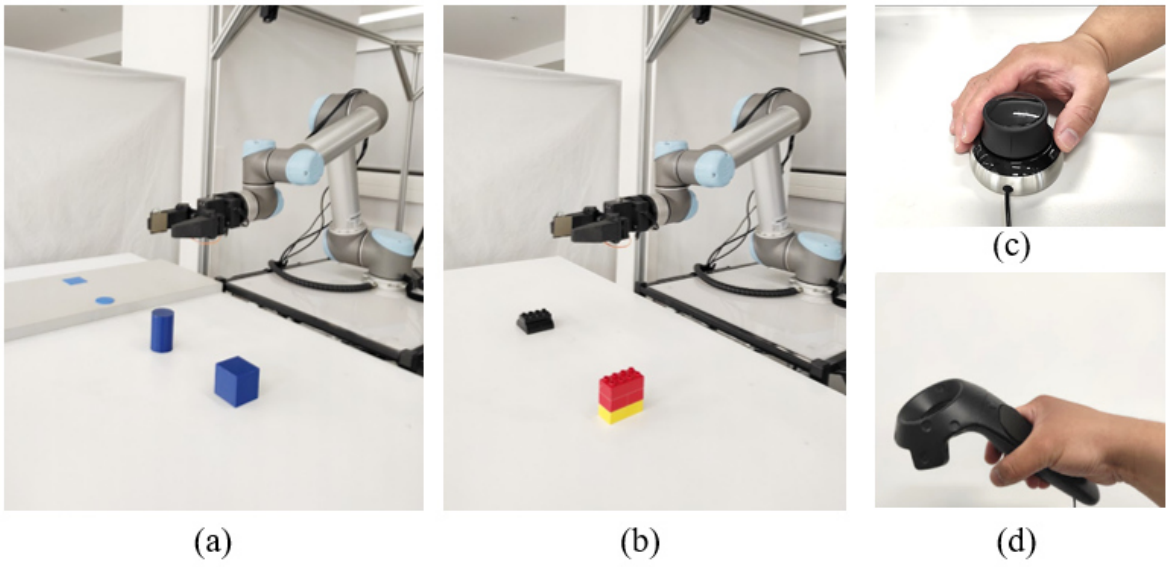


Figure 4.3: An overview of spatial manipulation tasks and command input methods. (a) Pick-and-place task of two cubical and cylindrical objects given a goal position with different heights; (b) Assembly task involved grabbing one LEGO subassembly from a predefined spot and stacking it onto a fixed LEGO base on the table; (c) Isometric-rate HRI scheme using the 6-DOF input device; (d) Isotonic-position imitative HRI scheme using the Vive motion controller.

4.3.3 Experimental Procedure

Each subject picked up two cubic and cylindrical items and placed them on target positions on the tabletop in the first pick-and-place manipulation task. By requiring both translation and

rotation to place the objects, the process assesses the translational and rotational functionality of the three HRI schemes. The second task grabbed one LEGO subassembly from a predefined spot and stacked it onto a fixed LEGO base on the table. As an assembly task with increased complexity, this task required the subject to rearrange the in-hand object's orientation by rotating the red LEGO cuboid subassembly and aligning the two components before mating, which restricted the orientation of the end-effector in comparison to placing the cubical and cylindrical items.

The trial began with each subject filling out a pre-experiment questionnaire to record age and gender, and assess prior knowledge and experience with MR, robotics, and HRI. Following the pre-experiment questionnaire, participants were given health and safety guidelines and provided with details on the experiment tasks and objectives. Prior to using each HRI method, each subject was presented with a tutorial video demonstrating the user interaction module. The video-based introduction was utilized to equalize the training subjects were offered on each HRI module.

After indicating preparedness, participants run through a training task of grasping a cylindrical object with the robotic arm-hand module to get used to each control method of the teleoperation system. The object grasping task was chosen for training because of its simplicity and the preparation it provides for the experimental tasks [150]. Each participant completed a training activity once to standardize the training process. When the user expressed readiness to conduct the experiment tasks, the cumulative training time for each HRI module was recorded to assess the learnability and efficiency of each HRI module. The locations of target objects and the robotic pose changed, respectively. To reduce order effects during the execution of the experimental tasks, the sequence of the HRI modules and the target position of the objects were

randomly allocated, and pick-and-place and assembly manipulation task orders were randomized, which minimizes the learning effects between tasks. The teleoperation task was suspended when the subject placed the object to the goal position and mated the components together in an assembly.

To standardise the initial position, the experimenter returned the robot to the identical home configuration after each trial. Participants completed a post-trial questionnaire on each HRI module after completing the training and experimental tasks. Participants remained outside the robotic platform operating zone at all times to guarantee physical safety and one experimenter closely observed all operations with an emergency stop. If a collision occurred on the objects or tabletop, the experimenter immediately activated the emergency stop button to halt all robot motion. The robot was then sent back to its original pose, demanding the subjects redo the task. In the MR subspace, virtual safety grids were also established to surround and remind the user. When participants reached predetermined boundaries, warning grids showing the MR subspace edge were displayed. The experiment requested all users avoid these edges if they could.

Task performance and work efficiency were evaluated by measuring the time for each task and the total time for both tasks. Participant perception of the different HRI modalities was measured using a questionnaire based on a previous study assessing user preferences and system usability, including the NASA Task Load Index (NASA-TLX) and Technology Acceptance Model (TAM) for evaluating the user's acceptance [151]–[153].

4.3.4 Analyses

The goal of the research is to investigate the effect of the MR subspace-enhanced imitation-based HRI module on the operator’s ability to interact with objects in a remote environment.

To evaluate the effectiveness of the robotic manipulation system presented, human operator performance is assessed under three HRI modes denoted: B, MRD and MRS. B denotes the Baseline using only a 6-DOF SpaceMouse and monocular RGB display. MRD (MRnoVCMM/MR SpaceMouse) represents the MR direct control module using SpaceMouse and MR-enhanced 3D point cloud visual method without deploying VCMM. MRD is a limited version of the proposed MR-HRI system. MRS (MRwithVCMM) provides the user with the proposed MR subspace module using the VCMM approach using a Vive controller and MR-assisted 3D point cloud display.

User performance is measured by time for each manipulation task and aggregate total time for both tasks. User effort and workload during teleoperation experiments are evaluated by the NASA task load index (NASA-TLX) score, where from 0 to 100 (most difficulty), participants rated qualitative experiences of mental demands, physical demands, time demands, performance, effort, and frustration at the end of each experimental case. User acceptance and perception including usefulness and ease-of-use is assessed by post-task questionnaires based on the technology acceptance model (TAM). This survey uses a 0 to 7 (best) scale to measure the acceptance and ease of use of different HRI modules. A one-way within-subjects ANOVA with repeated measures analyses data from all measures. The Greenhouse-Geisser correction is applied to assess the difference in the survey responses with B, MRD and MRS HRI modules as within-subjects variables.

4.4 Results and Discussion

The MR subspace-based motion-mapping HRI technique was compared to Baseline and MR direct control schemes across two tasks. Results in **Tables 4.1** and **4.2** show the data and statistical findings.

Table 4.1: Statistical results for objective and subjective measures for each method compared.

	Baseline		MRD		MRS	
Measure	**Mean**	**Std.Dev**	**Mean**	**Std.Dev**	**Mean**	**Std.Dev**
Pick and Place (s)	136.06	20.72	98.27	17.00	58.52	9.75
Assembly (s)	162.60	23.61	122.88	16.84	97.51	12.18
Aggregate Time (s)	298.66	26.42	221.15	23.39	156.03	17.19
Physical Demand	76.04	11.32	70.88	11.51	55.63	12.54
Mental Demand	81.38	13.29	68.46	11.12	54.54	13.15
NASA TLX	75.81	6.23	66.89	5.39	51.92	8.15
Usefulness	2.83	1.17	3.08	1.10	4.75	1.33
Ease of Use	2.67	0.87	3.25	1.39	5.50	0.89

Table 4.2: One-way ANOVA Statistics for all measures. B, MRD, and MRS denote Baseline, MR SpaceMouse, and MR subspace.

				Post-hoc Tests		
Measure	**Partial Eta Squared**	**F**	**P**	**MRS-MRD**	**MRS-B**	**MRD-B**
Pick and Place (s)	0.87	F (1.382, 31.778) =148.198	<.001	<.001	<.001	<.001
Assembly (s)	0.76	F (1.553, 35.725) =74.080	<.001	<.001	<.001	<.001
Aggregate Time(s)	0.93	F (1.875, 43.128) =303.197	<.001	<.001	<.001	<.001
Physical Demand	0.45	F (1.971, 45.339) =18.478	<.001	<.001	<.001	.387
Mental Demand	0.59	F (1.995, 45.874) =32.638	<.001	<.001	<.001	.002
NASA TLX	0.76	F (1.663, 38.247) =74.408	<.001	<.001	<.001	<.001
Usefulness	0.41	F (1.846, 42.449) =15.794	<.001	<.001	<.001	1.000
Ease of Use	0.69	F (1.832, 42.133) =50.205	<.001	<.001	<.001	.299

4.4.1 Objective Measures

The analysis rejected the null hypothesis (H0) that the MR subspace enhanced spatial motion mapping approach for telemanipulation and the other typical teleoperation modules have identical effects on task performance (**Table 4.2**). However, the results indicated the MR subspace motion-centric HRI approach significantly outperformed both the 2D Baseline and MR SpaceMouse HRI schemes on both tasks for all pairwise comparisons (**Table 4.1**). Guiding a robotic arm-hand system using the natural arm motion mapping through MR subspace improved task performance for both tasks.

As shown in **Figure 4.4**, a one-way within-subjects ANOVA with repeated measures with a Greenhouse-Geisser correction indicated the time taken to complete the pick and place tasks were statistically significantly different ($F(1.382, 31.778) = 148.20$, $p < 0.001$, Partial = 0.87).

The post-hoc test revealed the completion time for the pick and place tasks significantly reduced from the Baseline (M = 136.06) compared to the MR SpaceMouse module (M = 98.27) and the MR subspace module (M = 58.52). Statistical significance was also seen for the assembly task completion time, between the three HRI modules (F (1.553, 35.73) = 74.08, p < 0.001, Partial = 0.76). Pairwise comparisons indicated the time to complete the assembly tasks significantly reduced from the Baseline (M = 162.60) compared to the MR SpaceMouse module (M = 122.88) and the MR subspace module (M = 97.51).

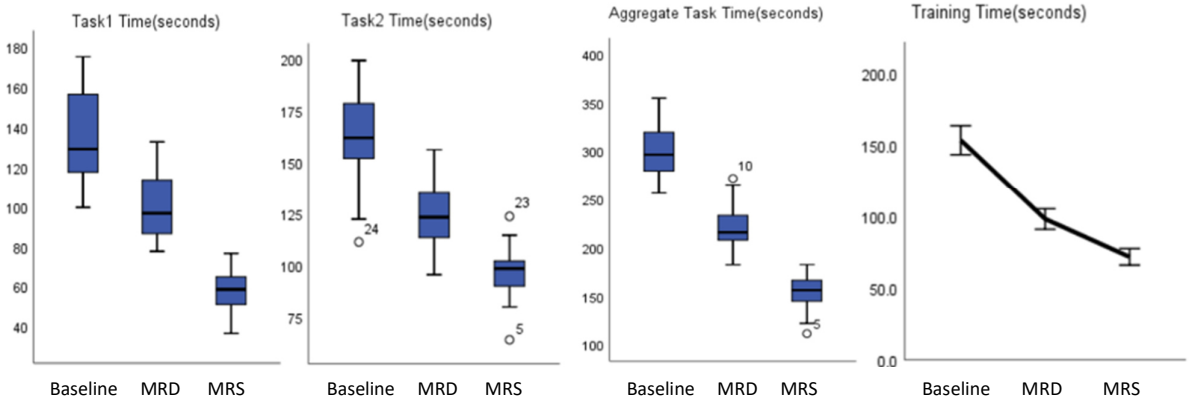


Figure 4.4: Boxplots of quantitative measures on the user performance for each HRI scheme across two manipulation tasks.

The F value for the HRI factor of overall task completion time and its related significance level and the magnitude of the effect (Partial Eta Squared) in the Greenhouse-Geisser correction reported the mean aggregate task completion time for each HRI module was statistically significantly different, (F (1.875, 43.128) = 303.197, p < 0.001, Partial = 0.93). The pairwise comparisons indicated aggregate time significantly decreased from the Baseline (M = 298.66) compared to the MR SpaceMouse module (M = 221.15) and the MR subspace module (M = 156.03).

Overall, the aggregate task completion time reduced by 48% compared to the 2D Baseline module and 29% compared to the MR SpaceMouse module. With the MR subspace enhanced motion and vision mapping, a comparable rate of improvement in completion time was attained for operators with minimal technical knowledge.

Mean training completion time was 153.7s, 98.3s and 71.8s for the Baseline, and MR with and without the imitation-based motion mapping modes, respectively. The learning time of the training tasks with the MR subspace approach decreased by 53% and 36%, compared to the Baseline and MR SpaceMouse approaches, indicating extra learning was required for the two typical HRI modules to reach the same competency as the proposed MR subspace-enhanced imitation-based HRI module. It is also observed that, even at the end of the training, subjects did not reach the same proficiency and dexterity as they immediately did while using the MR subspace-enhanced imitation-based module. The MR subspace module has an important effect in saving training time even for the non-skilled operators.

4.4.2 Subjective Measures

The NASA task load index (NASA-TLX) score showed participants rated their qualitative experiences of mental demands, physical demands, temporal demands, performance, effort, and frustration. The Technology Acceptance Model (TAM) was used for evaluating the acceptance and measuring participant perception of different HRI modules and scales to measure usefulness and ease of use were presented to the subjects in each post-task questionnaire. Three items on a seven-point rating scale (1 = strongly disagree; 7 = strongly agree) were included in each of the two aspects. As shown in **Figure 4.5**, all average NASA-TLX scores were lower for the MR subspace enhanced motion-mapping tasks compared with

the 2D Baseline and MR SpaceMouse cases. In particular, the MR subspace motion-mapping module significantly reduced the physical and mental demand and frustration of participants.

In general, the overall workload (OW) decreased from Baseline (M = 75.81) compared to the MR SpaceMouse module (M = 66.89) and the MR subspace module (M = 51.92), as shown in **Figure 4.5 (f)**. As a result, the average score of NASA-TLX decreased significantly by 32% (F (1.663, 38.247) = 74.408， p < 0.001, Partial = 0.87) when the MR subspace mapping was used.

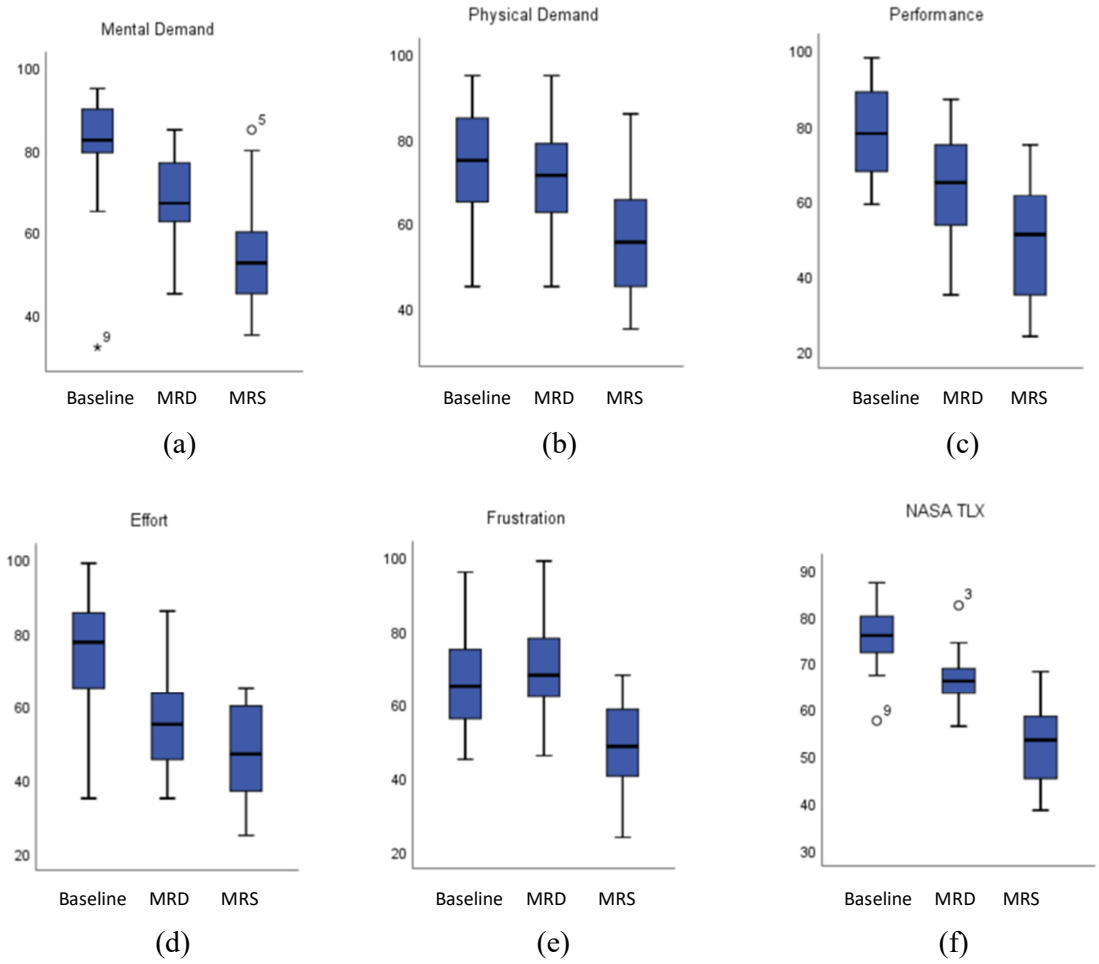
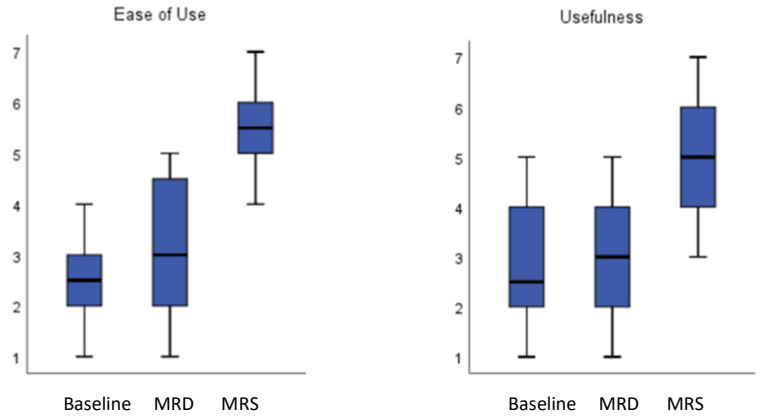


Figure 4.5: Boxplots of workload measures for each HRI scheme across two tasks.

The TAM results indicated there was a substantial disparity between the means of the users' appeal at the three different HRI modules. Participants found the MR subspace imitation-based HRI module (M = 5.50) had better usability than the 2D Baseline (M = 2.67), and is marginally easier to use than the MR SpaceMouse module (M = 3.25), as shown in **Figure 4.6 (a)**. The MR subspace method (M = 4.75) was reported to be slightly more acceptable than the MR SpaceMouse module (M = 3.08) and 2D Baseline interface (M = 2.83), in terms of perceived usefulness, as shown in **Figure 4.6 (b)**. The subjective measures analysis proved the MR subspace enhanced spatial motion mapping approach for telemanipulation outperformed the other typical teleoperation modules in task workload and user perception.



(a) (b)

Figure 4.6: Boxplots of subjective measures about ease of use and usefulness for each HRI scheme across tasks.

4.5 Summary

This work presents the design of an MRS-based intuitive telemanipulation paradigm and a user-study evaluation of three control and visual feedback HRI modes on a practical robotic arm-hand platform. The particular interest is the potential benefits of deploying an MRS-enhanced 3D vision/motion mapping approach to improve the work efficiency and situation

awareness of unskilled operators in teleoperated pick-and-place and assembly tasks. The IRT system introduced in this chapter enabled novice users to intuitively and naturally perform high-quality manipulation tasks at a distance. The proposed MRS HRI interface for robotic tele-control is designed and implemented by leveraging the 3D mapping of motion and vision through MR subspace.

The intuitive and natural interaction scheme is achieved by mapping the user's hand motions to the robot movements and applying spatial velocity-centric control techniques. A VCMM approach is implemented to accurately track the operator's hand movements and minimize aggressive velocity commands, while generating smooth movement in the two typical manipulation tasks. The 3D point cloud rendering architecture is deployed in the MRS paradigm to form the incorporated 3D visualization of the remote site, provide the desired depth perception of the workspace, and maintain the inspection of the workpieces, remote site and digital twin situated in the MR subspace as a whole. Telemanipulation experiments of novice operators are carried out to test the proposed intuitive teleoperation of the robotic platform.

Telemanipulation experiments found the MRS integrated scheme reduced aggregate task completion time reduced 48% compared to the 2D Baseline module and 29% compared to the MR SpaceMouse module. The MRS enhanced 3D mapping of motion and vision paradigm improved completion time for operators with minimal technical knowledge. Further, the learning time of the training tasks with the MRS scheme decreased by 53% and 36%, compared to the 2D Baseline and MR SpaceMouse approaches, indicating extra learning was required for the two typical HRI modules to reach the same competency as the proposed MRS imitation-based HRI module; and finally, accomplished desired telemanipulation results for novice users

and significantly reduced the physical and mental demand and frustration of participants while offering higher user acceptance.

Overall, the MRS teleoperation scheme for robotic arm-hand teleoperation presented improved remote pick-and-place and assembly performance of operators with minimal technical knowledge. The proposed teleoperation scheme using integrated 3D mapping of vision and motion through MR subspace involving intuitive movement control provides improved tele-control performance of manipulation tasks and reduces the workload on operators.

Chapter 5

Mixed Reality-Integrated 3D/2D Vision Mapping for Intuitive Teleoperation of Mobile Manipulator

5.1 Introduction

Depth cues are crucial to increase user perception and spatial awareness of the remote environment when remotely guiding complex robotic systems. A MR integrated 3D/2D vision merging framework for immersive and intuitive telemanipulation of a complex mobile manipulator is presented. The proposed 3D immersive telerobotic schemes provide the users with depth perception through the merging of multiple 3D/2D views of the remote environment via MR subspace. The mobile manipulator platform consists of a 6-DOF industrial manipulator, 3D-printed parallel gripper, and mobile base, which can be controlled by non-skilled operators who are physically separated from the robot working space through a velocity-based imitative motion mapping approach.

This work evaluates the impact of depth perception and immersion provided by integrated 3D/2D vision and motion mapping schemes on teleoperation efficiency and user experience in an MR environment. In particular, the MR enhanced systems maintain spatial awareness and

perceptual salience of the remote scene in 3D, facilitating intuitive mixed reality human-robot interaction (MR-HRI). This study compared two MR-integrated 3D/2D vision and motion mapping schemes against a typical 2D Baseline visual display method through pick-and-place, assembly, and dexterous manufacturing tasks. The MR-integrated 3D/2D vision and motion mapping schemes of teleoperation reduced overall task completion times by 34% and 17%, compared to the MR-2D Baseline, while minimizing training effort and cognitive workload.

5.2 Methodology

5.2.1 Teleoperation System Design and Implementation

This study explores the effects of integrated 2D/3D visualization in MR with natural 3D motion mapping methods for immersive and intuitive telemanipulation. There are three main elements comprising the robot manipulator and visualization system, the MR subspace implementation, and the robot and userspace communication implementation.

5.2.1.1 Robot Manipulator and Visualization System

As shown in **Figure 5.1 (a)**, the robotic platform consists of a 6-DOF UR5 industrial manipulator (Universal Robotics, Denmark), a 3D-printed parallel gripper for interacting with the remote environment, a mobile robot base for providing mobility and an onboard Ubuntu computer for robot control. **Figure 5.1 (b)** shows the low-cost 2-finger gripper and the Logitech C615 webcam (Logitech International S.A, Switzerland) mounted on the gripper base. The wrist camera is useful for pre-approaching and positioning the robot gripper in preparation for the object grasping process. The 50 mm stroke allows the gripper to handle a variety of object sizes, as shown in **Figure 5.1 (c)**. Multi-camera systems including a binocular camera/Kinect V2 depth camera and two monocular 2D webcams are integrated into the robotic platform for

providing hybrid 3D/2D visual feedback to the operator via a head-mounted display (HMD) during the telemanipulation process.

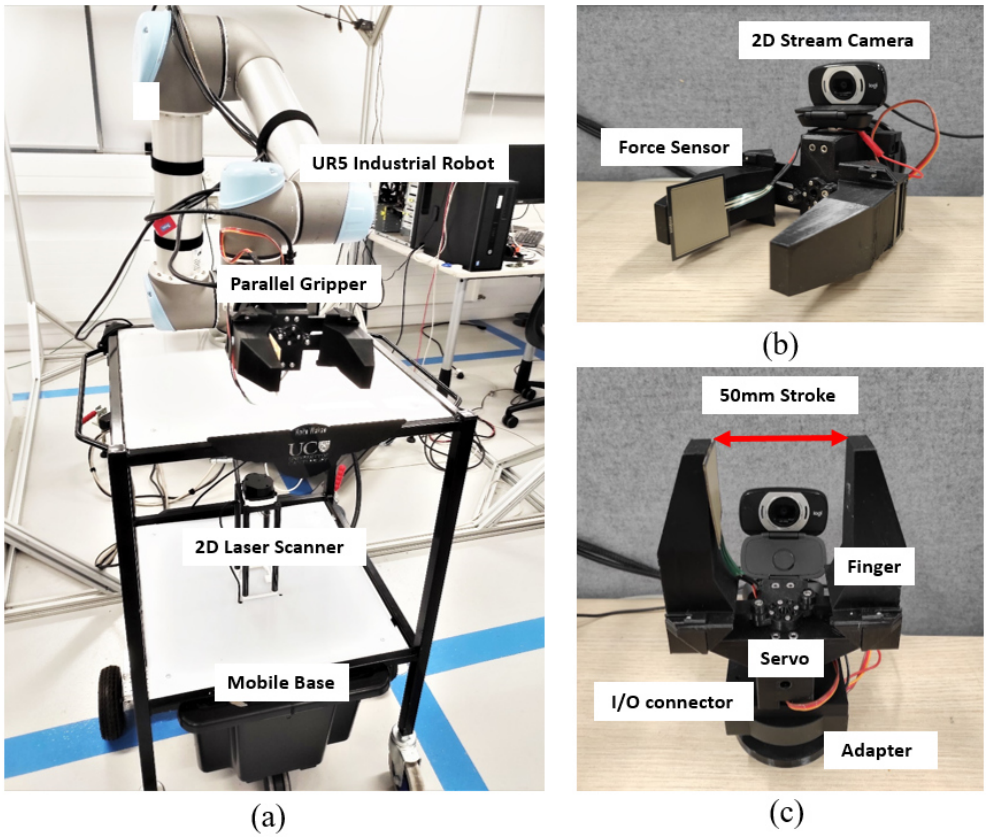


Figure 5.1: The hardware components of the proposed telerobotic platform. (a) Mobile manipulator platform including UR5 arm, low-cost gripper and mobile base, (b) the isometric view of the parallel 2-finger gripper with the attached Logitech C615 webcam, (c) the motorized structure and the stroke length of the 3D printed gripper.

To validate the proposed MR-based hybrid 3D/2D visualization of intuitive and immersive teleoperation paradigms, three visualization modules from robot-assisted telemanipulation paradigms are designed for direct comparison. In particular, the following teleoperation multi-camera setups and the corresponding visual feedback modules via MR are developed:

- **Baseline (MR-2D)**: multi-perspective 2D monocular RGB camera views. The Baseline interface displays two 2D monocular RGB camera views from robot wrist

and workspace perspectives. The user does not feel immersed or present in the remote workspace, and no depth cues are provided for object manipulation.

- **MR-3DS**: 3D stereoscopic vision (SV) and monocular RGB (Baseline) integrated vision. The main advantage of MR-3DS over the monocular vision used in Baseline is the depth cues supplied to the user by the camera pair embedded in the stereoscopic camera. In the MR-3DS mode, the user can guide the virtual camera representing the user's viewpoint to the desired location and perceive the stereoscopic view, but the user's viewpoint is constrained for stereoscopic vision.
- **MR-3DP**: 3D point cloud (PC) and monocular RGB (Baseline) integrated vision. The MR-3DP module publishes compressed RGB-D images from the perspective of the user's head to produce PC-based 3D visual feedback in the MR subspace. The MR-3DP paradigm provides the user with a more natural vision system for depth perception.

Comparing MR-3DS and MR-3DP to the Baseline setup, levels of immersion are increased and depth cues are supplied by introducing stereo vision and point cloud. In addition, point cloud display allows the user the change their viewpoint as needed without introducing a moving camera system that imitates the position and orientation of the user's eyes.

With a wrist camera alone, users have difficulty in judging the distance between objects and avoiding unwanted collisions during operation [153], [154], necessitating a second workspace camera on the side of the robot to provide distance information. The Baseline condition is a limited version of the MR-3DS and MR-3DP paradigms allowing the user to inspect the robot in its work environment from two monocular displays in MR and interact with it using the VCMM approach, but without 3D visualization.

The MR-3DS module (**Figure 5.2 (a)**) adds on an extra stereoscopic camera to the Baseline camera arrangement. The customized stereoscopic camera consists of two Logitech C615 monocular cameras with up to 1920 x 1080 pixels, featuring a 74-degree field-of-view. Each Logitech camera views the remote scene and obtains image information from a slightly different position because human vision can produce stereoscopic vision mainly based on the binocular disparity [155].

The two monocular cameras are positioned inside a 3D-printed camera case and the interpupillary distance (IPD) is set to 67 mm to produce a stereoscopic first-person perspective [156]. The stereoscopic camera feed is transmitted from the robot side to the HMD in userspace via the Internet. The usb_cam node of ROS is used to interface with the two webcams embedded in the stereoscopic camera, and the Ubuntu computer running ROS publishes the captured images in sensor_msgs/Image format while using the image transport library for image compression and transfer. The compressed image topics are subscribed by Unity 3D using ros-bridge protocol, which provides a ROS functionality for sending JSON-based commands to Unity. Finally, the captured image messages from the left and right webcams in the stereoscopic camera are projected onto two corresponding planes in the MR space for users' stereoscopic perception.

The MR-3DP vision module incorporates a Kinect V2 depth camera (Microsoft Corporation, USA) with the two webcams from the Baseline module (**Figure 5.2 (b)**). The depth camera is mounted on the base of the mobile manipulator. One of the two webcams is attached to the wrist of the UR5 robotic arm-hand system providing an ego-centric perspective, and the other is positioned in the workspace for an exo-centric view during the manipulation tasks. The MR-

3DP module publishes compressed RGB-D images from the perspective of the user's head to produce PC-based 3D visual feedback in the MR subspace. Compressed images enter the MR subspace through a custom image subscriber and depth image subscriber in Unity 3D. In Unity, RGB images and depth images are decompressed using OpenCV for C# (Microsoft Corporation, USA). A custom material shader combines the RGB image and depth image into a PC for 3D visualization. The MR-3DP module simultaneously publishes 2D streams from wrist and workspace webcams from the Baseline module to the MR subspace to couple with the PC visualization.

To validate the applicability and generalizability of the proposed MR-integrated 3D/2D vision schemes, three different manipulation tasks in a remote manufacturing scenario were devised. The manipulation tasks used to validate and compare the three MR 3D/2D vision schemes are defined:

- **Task A**: pick-and-place task stacking a cylindrical item on a cubic object (**Figure 5.2 (c)**). Pick-and-place is a typical task for evaluating telerobotic systems with metrics including task performance, user workload and system usability. Although grasping a cylindrical object with the robot was a simple task, the height difference between the object's initial position and the target position (the height of the cube) made depth information very important.
- **Task B**: horizontal assembly task grabbing one LEGO subassembly from a predefined spot and joining it onto a horizontal fixture on the table (**Figure 5.2 (d)**). In the assembly experiment, the subassembly consisted of three layers of 2×4 BRICK to facilitate the grasping of the parallel gripper. The fixture was composed of one 2×4 BRICK with a base. The base was mounted on a horizontal workbench for assembly Task B. Compared with

the pick-and-place task, the assembly task further constrained the orientation of the robot end-effector. Because the block could fit in the fingers of the parallel gripper only when the axis of the workpiece was orthogonal to the orientation vector of the end effector, and the workpiece and the fixture must be well aligned before mounting.

- **Task C**: tilted assembly fitting a LEGO component to a fixture at 30° (**Figure 5.2 (e)**). This task tests the rotational usability of the proposed intuitive motion mapping approach by requiring the rotation of the LEGO component to assemble it correctly on the tilted fixture at 30o. Compared with Task A and B, Task C required both translational and rotational movement in three dimensions to manipulate and assemble the objects. Assembly on the slope not only involved high-precision operation but also required the user to observe the assembly process from multiple perspectives, which increased the complexity of the task.

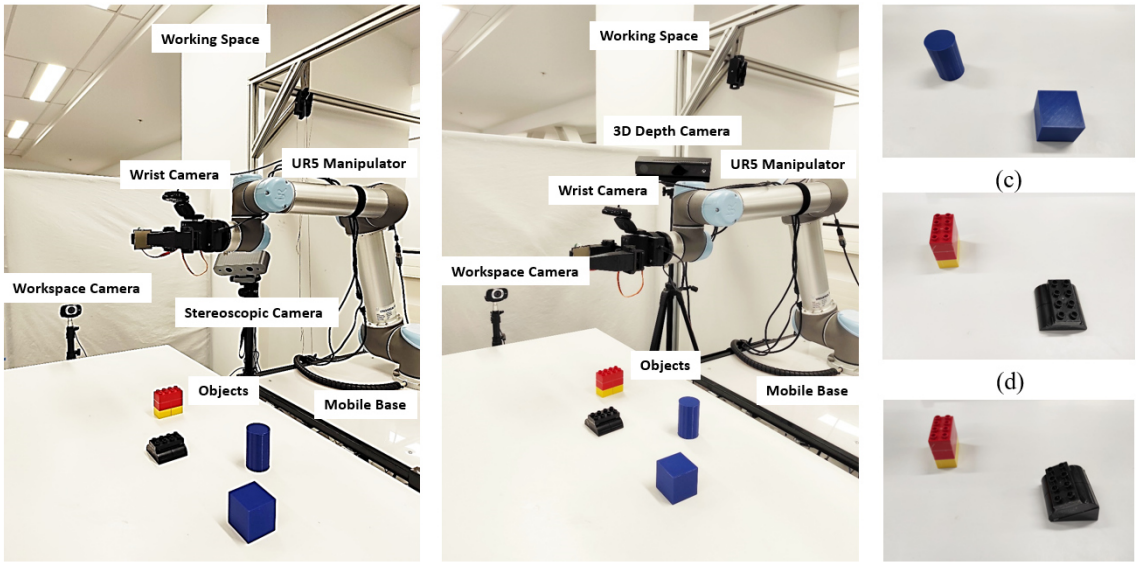


Figure 5.2: The camera setup of the proposed MR-3DS and MR-3DP telerobotic platforms and the remote manipulation tasks used in the experiments to test the feasibility of the MR-based modules for teleoperation. (a) A static binocular camera module is enhanced by a workspace camera and robot wrist camera for stereo vision feedback and (b) an RGB-D camera integrated with a workspace camera and robot wrist camera for the 3D PC visualization in the

MR scene. (c) The pick-and-place task of stacking a cylindrical item on a cubic object, (d) the horizontal assembly task involved grabbing one LEGO subassembly from a predefined spot and joining it onto a horizontal fixture on the table, (e) the tilted assembly task of fitting the LEGO component on a fixture on a 30-degree slope.

5.2.1.2 MR Subspace

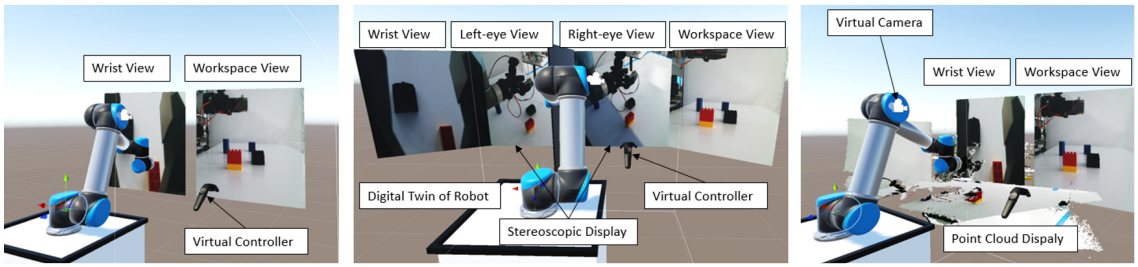
The proposed teleoperation paradigm allows vision and motion mapping between the user side and the robot side by introducing the MR subspace based on a computer-generated virtual environment in Unity 3D. Multiple on-site data and information from remote sites are transmitted back to users who serve as part of the control loop to manipulate the remote robotic system. The user can directly observe the workpiece and robot status with real-time overlaid sensory information in the MR subspace during operation. The three-dimensional virtual scene and overlaid 2D/3D visualization according to the physical robot working environment is rendered and displayed to the user via the tracked HMD.

The MR-2D Baseline scene consists of a virtual surrogate of the physical UR5 robot and two virtual screens to display the ego-centric and exo-centric views for direct vision mapping from the remote robot site to the human operator at the local site, as shown in **Figure 5.3 (a)**. Within the MR subspace, the operator is positioned at an equal distance to the two virtual ($1m \times 1m$) screens. The left screen displays the streaming feed from a single camera on the gripper wrist, and the right displays the video feed from the workspace camera. The wrist camera view shows the gripper fingers and objects approaching an object for a grasping task, and the workspace camera image displays the distance between the gripper, objects and tabletop while grasping an object to reduce collisions.

The MR-3DS scene uses two monocular RGB cameras delivering ego-centric and exo-centric perspectives to supplement the binocular camera and enhance the stereoscopic view with two

typical 2D displays. As shown in **Figure 5.3 (b)**, the displays in the MR-3DS scene consists of two symmetric virtual screens. The stereoscopic video streams incorporate the left and right camera feed captured by the two Logitech webcams inside the stereoscopic camera and are projected on the displays to each of the user's respective eyes for stereo visual inspection of the teleoperation process. In addition, two monocular RGB streams are positioned to the left and right side of the stereoscopic visualization respectively so users can avoid moving their eyes off the stereoscopic image to view the 2D images from the robot wrist and workspace perspectives.

In the MR-3DP scene, the centralized PC visual feedback is enhanced by two monocular RGB cameras providing two typical 2D perspectives from the robot wrist and workspace. The MR-3DP augments the PC via the Baseline metaphor, attaching two image panels to the live stream feed of the 3D PC, as shown in **Figure 5.3 (c)**. MR-3DP has the advantage of overlaying the virtual surrogate of the physical robot on the PC display. The virtual replica allows the user to examine the pose of the physical robot in the remote environment without deploying an array of static cameras for the same purpose. The virtual surrogate subscribes to the joint state data of the physical UR5 robot and updates its configuration in MR accordingly using the ros-bridge software provided by ROS. Thus, the user and the virtual robot share the same space and work together in a collaborative manner via the MR subspace.



(a) (b) (c)

Figure 5.3: The three MR-based telemanipulation modules the participants tested in the experiments while interacting with virtual surrogates of physical robotic platforms. **(a)**

Baseline: 2D displays in the MR subspace from the robot wrist and its workspace perspective; **(b) MR-3DS**: the integration of stereoscopic visualization and 2D monocular robot wrist and workspace views through the MR subspace; **(c) MR-3DP**: PC visualization complemented by the 2D robot wrist camera and its workspace views.

5.2.1.3 Robot Space and User Space Communication

Figure 5.4 shows the MR subspace-assisted robotic teleoperation system consists of four parts: 1) the human-robot interaction unit; 2) the MR subspace for vision and motion mapping; 3) the remotely controlled robot working unit; and 4) the bi-directional communication links.

On the user side, the human-robot interaction unit includes an HTC Vive HMD and monitor connected to a desktop with an i7-8700k CPU, 32 GB RAM, and a GeForce GTX 1080 Graphics processor. It also includes the MR environment generated in Unity 3D, Vive controllers command input devices, and the Vive motion tracking system (Vive base stations). One Vive controller is attached to the user's waist as a body tracker for remotely driving the mobile base of the robotic platform. Users manipulate an HTV Vive wireless controller as the input device to control the UR5 arm and robotic hand. Finally, the multi-camera enhanced scene inside the MR subspace is generated in Unity3D to display integrated 3D visualization and auxiliary monoscopic image streams of the teleoperation processing the MR scene.

To increase system safety and reliability, the motion tracking of the user's hand and waist are only activated while the controller touchpads are touched by the user's fingers to avoid accidental and unwanted motion input. An advantage of this technique is once the user's hand reaches a physical limitation or the waist approaches the margin of the user operating space, the user can release the controller touchpads, discontinuing transmitting commands to the robotic platform, move back to the initial pose, and then continue guiding the robotic arm and mobile base to the target poses. This approach eliminates limitations due to dissimilarity

between the spatial range of robot movements and the spatial range of user operations.

On the robot side, an Ubuntu 16.04 computer with an i7-10700 CPU, 64 GB RAM, and GeForce RTX 2060 Graphics was used to control the robotic platform, and capture and process on-site sensor data. ROS (Robot Operating System) is a widely used framework for robotic systems, is installed on the Ubuntu computer, and serves as the core of the robot control and image processing for bi-directional communication. The ros-bridge WebSockets server with a JSON API transmits control command and image transmission between the robot working unit and user operating space.

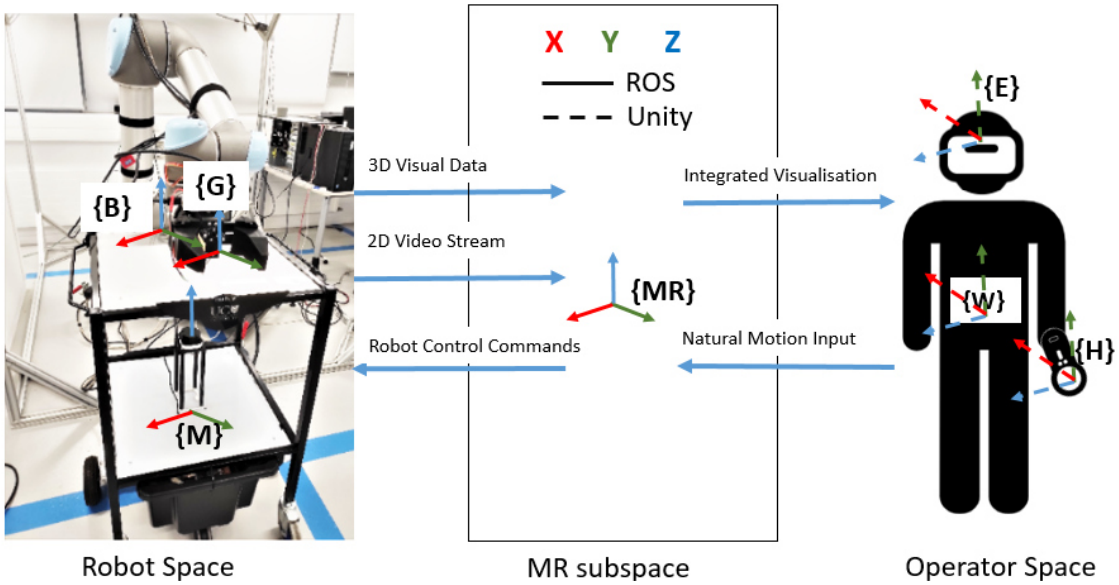


Figure 5.4: The overview diagram and communication structure of the proposed intuitive and immersive teleoperation system. (a) The robot working space platform including the mobile manipulator and its environment for interaction, (b) the MR subspace as a medium for 2D and 3D vision and motion mapping, and (c) the motion tracking system on the user side updates the pose change of user's head, hand and waist and motion command is transmitted to the robotic system through the MR subspace.

2.2 Velocity-centric motion mapping

A whole-body velocity-centric motion mapping (VCMM) framework is used to intuitively

control the robot so its end-effector smoothly follows a user's hand trajectory, $T_u(t)$. Specifically, the goal is to calculate and control the required n-vector of joint velocities, $\dot{\theta}$, directly from the relationship $J(\theta)\dot{\theta} = V$, where the desired robot end-effector twist $V_e = [\omega_e \; v_e]^T$ and $J(\theta) \in R^{6\times n}$ are expressed in the same frame. At time k, the motion tracking system obtains a measurement of the user's hand configuration, converts this calculated spatial velocity $V_h = [\omega_h \; v_h]^T$ to a Cartesian twist command $V_e = [\omega_e \; v_e]^T$ to the robot in the ROS coordinate system, employs the Inverse-Jacobian solver to determine the appropriate joint rate vector, $\dot{\theta}$, according to the desired twist representation of end-effector motion, V_e, and derives the joint configuration sent to the UR5 robot controller. Similarly, for a given mobile base 2D velocity $V_b = [\omega_b \; v_b]^T$ with the current pose (ϕ, x, y), the motion tracking system measures the user's waist motion and converts this calculated waist velocity $V_w = [\omega_w \; v_w]^T$ to a Cartesian command $V_b \in R^2$, including translational and rotational components of the base velocities:

$$\dot{\theta} = J^{\dagger}(\theta)\, SV_e \tag{5.1}$$

$$V_h = [\omega_h \; v_h]^T = kV_e = V_e = [\omega_e \; v_e]^T = J_e(\theta)\dot{\theta} = [J_v(\theta)\, J_\omega(\theta)]\dot{\theta} \tag{5.2}$$

$$V_w = [\omega_w \; v_w]^T = kV_b = [\omega_b \; v_b]^T = \begin{bmatrix} \dot{\phi} \\ \dot{x} \\ \dot{y} \end{bmatrix} \tag{5.3}$$

Single Exponential Smoothing (SES) is applied to filter unintended short-term fluctuations from noise or body tremors, and is recursively calculated:

$$u_{t+1|t} = \alpha u_t + \alpha(1-\alpha)u_{t-1} + \alpha(1-\alpha)^2 u_{t-2} \ldots + \alpha(1-\alpha)^k u_{t-k} \tag{5.4}$$

where u_t are human body velocity measurements and $0 \leq \alpha \leq 1$ is the smoothing parameter. If α is large, more recent hand motion observations are weighted more highly, and as α approaches 0, the output velocity signal tends to be the average of the historical input velocity data. A value of $\alpha = 0.9$ gives optimal performance based on experimental testing for this application.

5.3 Experiments

5.3.1 Experiment Design

A 3 ×1 within-participants experiment was designed to validate whether the MR-3DS and MR-3DP designs facilitate novice users' control of a complex mobile manipulator for manipulation tasks at a distance and affect the teleoperation performance and user experience. These results are all assessed relative to the Baseline case for the spatial manipulation tasks described in **Section 2.1**. Participants are not required to navigate the robot base to the location.

5.3.1.1 Hypothesis

The null hypothesis (H0) of the repeated-measures ANOVA is the Baseline, MR-3DS, and MR-3DP teleoperation paradigms are equally effective in task performance and user experience for novices, in terms of effectiveness, intuitiveness, and learnability, using the MR subspace-enhanced imitation-based motion mapping approach as the basis for teleoperation.

5.3.1.2 Experimental Procedure

Fifteen participants were recruited at the University of Canterbury (10 male and 5 female) from undergraduates and postgraduates, with ages from 21 to 31 years (mean: 23.9 years). Ten participants were from engineering departments, and the rest had non-scientific backgrounds. All of the fifteen participants rated themselves as not familiar with robotic systems, teleoperation, and MR. Participants run through a total of three MR-based hybrid 3D/2D visualization setups distinguished by increasing levels of immersion with depth cues and flexibility in the MR-based visualization.

The first step of the experiment was to have each participant fill out a pre-task questionnaire to

record their age, gender and their familiarity with robotics and MR experience. The objective of each trial was then explained. To test system learnability and efficiency, each subject watched a tutorial video demonstrating the proposed intuitive teleoperation platform with visual feedback modules they were going to use before testing, which ensured standardized, consistent training for all subjects. After the video, participants were given 2 minutes to experience the MR-enhanced telerobotic system to familiarize themselves with the motion controller, MR imagery, and the robotic platform.

After indicating readiness, each participant ran through a training exercise of grasping a cylindrical object with the robotic system to adapt to the remote manipulation system. Participants were given a training task of grasping the object due to its simplicity and the preparation provided for the following experimental tasks. Each participant performed the training task once to equalize the training process. The training time was recorded for post-experiment analysis of the learnability and efficiency of the telerobotic platform.

Each participant completed three tasks including the pick-and-place task, horizontal assembly task and tilted assembly task on a tilted fixture of 30 degrees under each experimental condition including MR-2D Baseline, MR-3DS and MR-3DP. Condition order was randomized across subjects to mitigate the potential order effects. After completing all the experimental tasks using one control-feedback condition, participants filled out a questionnaire about the HRI module to directly compare the three conditions.

The MR-HMD and motion controllers were fitted on each participant at the user site. When the subject sent verbal confirmation the MR subspace scene appeared as intended, they started completing the manipulation tasks as required. Participants were explained prior to the

experiment completion time for each task would be measured to evaluate the task performance. Participants were given unbounded time to complete each task, but were instructed to complete the full tasks as effectively as they could. Task completion time for each task and the aggregate time it took participants to complete full tasks were measured to evaluate improvement in their work efficiency. Participant workload and acceptance of each MR-HRI module were measured using the NASA Task Load Index (NASA-TLX) and Technology Acceptance Model (TAM).

Participants stayed at a distance from the robotic system workspace at all times, and one researcher closely observed the experimental procedure and remained alert to stop the robot by pressing the emergency button. In case of a collision with an object or table, the researcher immediately activated the emergency stop button, halted all robot motion and recorded the number of collisions for post-experiment analyses. The robot then returned to its original configuration, and the participant was required to restart the task. Additionally, four virtual safety planes around the user were defined in the MR subspace. When the participant reached a predefined boundary during the interaction, a warning grid showing the edges of the MR subspace was prompted. The researcher ensured participants stayed away from the safety planes during all trials.

All the participants were able to complete the telemanipulation tasks and did not fail during any of our experiments. In the pilot trial conducted to help design and assess the feasibility of the proposed telemanipulation system, the risk of test failure due to damage to the end-effector caused by the end-effector crashing into the workbench was identified. This test failure occurred when the operator accidentally hit the workbench with the end-effector and continued moving the robot in the wrong direction. However, we did not observe this failure in the

experiments using the proposed methods and all participants completed the test task without damage to the end-effector occurring throughout the experiments.

5.3.2 Analyses

To evaluate the effectiveness of the described robotic manipulation system, human operator performances across three manipulation tasks are assessed under three MR-HRI modes: Baseline, MR-3DS, and MR-3DP. User performance is measured by time for each manipulation task and aggregate time representing the total time it took participants to complete the three tasks. The number of collisions was also recorded. At the end of each operation, operator effort and workload during telemanipulation tasks were assessed by the NASA task load index (NASA-TLX) score [92], [157], [158], which evaluates qualitative mental demand, physical demand, time demand, performance, effort, and frustration (score range: 1-100 most demanding). User acceptance and system usability including usefulness and ease-of-use were evaluated by a questionnaire based on the technology acceptance model (TAM) measuring acceptance and ease-of-use (score range: 0-7 best) [150].

A one-way within-subjects ANOVA with repeated measures analyzed data from all measures. Bonferroni correction indicated which mean values were significantly different, and was used in this analysis when the ANOVA test showed a significant main effect of the experimental condition. The Greenhouse-Geisser correction was applied to assess the difference in the operator survey responses with Baseline, MR-3DS and MR-3DP modules as within-subject variables.

5.4 Results and Discussion

In the brief interviews following the telemanipulation experiment, most participants preferred the point cloud paradigm and stated that the point cloud-based 3D vision allowed them to change the head pose as needed to observe the remote manipulation process from different angles, and seamlessly transition from the point cloud display to the 2D auxiliary screens when the wrist view or workspace view was needed for reference. Some participants commented the depth information provided by the stereoscopic and point cloud displays was very helpful for completing manipulation tasks from a distance, especially for tasks that require viewing the workpieces from multiple angles (tilted assembly task). Some users stated the stereo-vision in the MR-3DS mode provides depth perception and improves the sense of presence, but the perspective is restricted. Five of the participants stated using the Baseline setup to perform telemanipulation was not intuitive and it was difficult to control the remote robot only through the view the wrist and workspace views. One user indicated that he was not well adapted to the stereo-vision in the MR-3DS mode and had difficulty perceiving the stereoscopic vision.

The experimenter noticed differences in the continuity of user movements when controlling the robot using the three different modes. In the case of baseline and stereo vision (MR-3DS), the users exhibited waiting and motion discontinuity between action inputs. In contrast, the operators' control movements were generally more continuous when using the point cloud-based feedback paradigm (MR-3DP). This was the result of non-intuitive and unnatural transitions between different various displays supplied to the user in the baseline and MR-3DS modes.

The analyses of the experimental data were outlined, comparing the two MR subspace enhanced vision/motion-mapping based HRI paradigms to the Baseline over three tasks. The

statistical results are given in **Tables 5.1**, **5.2** and **5.3**. Inter-task and within-mode comparisons were performed over the three tasks. In the same mode, the inter-task comparison of the workload reflects the various difficulty of the three tasks.

As **Table 5.1** shows, In the Baseline condition, tilted assembly was the most time-consuming (206 s), followed by the horizontal assembly (174 s), while pick-and-place took the least time (135 s). In the stereo vision (MR-3DS) condition, tilted assembly consumed the most time (176 s), horizontal assembly took the second most time (133 s), and pick-and-place required the least time (101 s). The point cloud paradigm (MR-3DP) indicates the same trend as the other two modes, with tilted assembly being the most time-consuming and pick-and-place being the most efficient. The within-mode comparison between different tasks indicates the increasing difficulty levels over the three tasks, which proves the purpose of the experimental design.

Table 5.1: Descriptive statistics for objective measures (time required for each task and training) for each method compared.

	Baseline	MR-3DS	MR-3DP
Measure	Mean Std.Dev	Mean Std.Dev	Mean Std.Dev
Pick and Place (s)	134.78 ± 16.72	101.29 ± 13.23	**83.04 ± 13.33**
Assembly (s)	173.79 ± 24.38	132.79 ± 16.18	**108.35 ± 13.35**
Tilted Assembly (s)	205.96 ± 25.46	176.39 ± 12.45	**147.70 ± 24.74**
Aggregate Time (s)	514.53 ± 33.84	410.46 ± 30.06	**339.09 ± 30.48**
Training Time (s)	167.07 ± 36.58	128.60 ± 18.04	**123.93 ± 14.50**

Table 5.2: Descriptive statistics for subjective measures (workload and usability) for each method compared.

	Baseline		MR-3DS		MR-3DP	
Measure	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Mental Demand	80.87 ± 10.76		62.20 ± 14.41		**47.80 ± 18.76**	
Physical Demand	79.47 ± 11.94		43.73 ± 13.61		**41.33 ± 11.75**	
Temporal Demand	79.87 ± 10.58		62.53 ± 12.40		**47.73 ± 10.15**	
Performance	85.00 ± 8.29		72.40 ± 17.36		**50.33 ± 14.41**	
Effort	79.67 ± 11.12		52.60 ± 13.27		**49.67 ± 18.86**	
Frustration	74.20 ± 9.86		63.13 ± 14.42		**41.73 ± 15.63**	
Overall Workload	79.84 ± 4.72		59.43 ± 5.41		**46.43 ± 6.55**	
Usefulness	3.07 ± 1.22		3.33 ± 1.11		**4.67 ± 1.23**	
Ease of Use	2.60 ± 0.83		3.93 ± 1.10		**5.53 ± 0.92**	
TAM	2.83 ± 0.75		3.63 ± 0.70		**5.17 ± 0.88**	

Table 5.3: Inferential statistics for all measures, where B = Basline, 3DS = MR-3DS, and 3DP = MR-3DP.

				Post-hoc Tests		
Measure	**Partial Eta Squared**	**F**	**P**	**3DP-3DS**	**3DP-B**	**3DS-B**
Pick-and-Place (s)	0.80	F (1.277, 17.874) =56.053	<.001	<.001	<.001	<.001
Assembly (s)	0.80	F (1.920, 26.883) =54.156	<.001	0.03	<.001	<.001
Tilted Assembly (s)	0.69	F (1.864, 26.095) =31.413	<.001	.001	<.001	.006
Overall Time(s)	0.90	F (1.939, 27.150) =128.407	< .001	<.001	<.001	<.001
Training Time(s)	0.53	F (1.644, 23.017) = 15.596	< .001	1.000	.002	.001
Physical Demand	0.74	F (1.889, 26.444) =39.608	<.001	1.000	<.001	<.001
Mental Demand	0.58	F (1.923, 26.922) =19.209	<.001	<.001	.063	.005
Temporal Demand	0.73	F (1.775, 24.843) =38.302	<.001	.010	<.001	.001
Frustration	0.65	F (1.706, 23.884) =26.485	<.001	.004	<.001	.031
Usefulness	0.33	F (1.739, 24.344) =6.916	.006	.007	.021	1.000
Ease of Use	0.72	F (1.747, 24.456) =36.794	<.001	<.001	<.001	.012

5.4.1 Objective Measures

The analysis rejected the null hypothesis (H0) that the MR subspace-integrated 3D/2D vision mapping approach for intuitive telemanipulation and the typical MR-assisted 2D teleoperation module have identical effects on task performance. The results indicate the MR-3DP motion-centric HRI approach significantly outperformed both the MR-assisted 2D Baseline and MR-3DS HRI methods on all tasks for all pairwise comparisons. Controlling a mobile manipulator using natural body motion mapping through hybrid 3D/2D visualization in the MR subspace improved task performance for all three tasks.

As shown in **Figure 5.5 (a)**, a one-way within-subjects ANOVA with repeated measures with a Greenhouse-Geisser correction indicated the time taken to complete the pick-and-place tasks were statistically significantly different, $F(1.277, 17.874) = 56.053$, $p < 0.001$, Partial = 0.80. The post-hoc test revealed the time to complete the pick-and-place tasks decreased significantly from the Baseline (M=134.8) compared to the MR-3DS module (M=101.3) and the MR-3DP module (M=83.0). Statistical significance was also seen for the horizontal assembly task completion time, between the three HRI modules, $F(1.920, 26.883) = 54.156$, $p < 0.001$, Partial = 0.80. The pairwise comparisons indicated the time to complete the horizontal assembly tasks significantly reduced from the Baseline (M=173.8) compared to the MR-3DS module (M=132.8) and the MR-3DP module (M=108.3), as shown in **Figure 5.5 (b)**. There was a significant main effect between the three HRI conditions on the completion times for the tilted assembly task, with an ANOVA test finding ($F(1.864, 26.095) = 31.413$, $p < 0.001$, Partial = 0.76). As shown in **Figure 5.5 (c)**, the pairwise comparisons indicated the time to complete the horizontal assembly tasks significantly reduced from the Baseline (M=206.0) compared to the MR-3DS module (M=176.4) and the MR-3DP module (M=147.7).

The F value for the HRI factor of overall task completion time and its associated significance level and effect size (Partial Eta Squared) in the Greenhouse-Geisser correction reported the mean aggregate task completion time for each HRI module was statistically significantly different, (F (1.939, 27.150) =128.407, $p < 0.001$, Partial = 0.93). As shown in **Figure 5.5 (d)**, the pairwise comparisons indicated aggregate time significantly decreased from the Baseline (M=514.5) compared to the MR-3DS module (M=410.5) and the MR-3DP module (M=339.1). The aggregate task completion time for MR-3DP and MR-3DS modules reduced by 34% and 17% respectively compared to the MR-2D Baseline module.

The mean training completion times for the Baseline, MR-3DS, and MR-3DP modules were 167.1 s, 128.6 s, and 123.9 s, respectively, with an imitation-based VCMM mapping approach, as shown in **Figure 5.6 (a)**. The learning times for the training task using the MR-3DS/3DP subspace approaches were reduced by 23% and 26% compared to the MR-2D Baseline case, indicating the typical 2D HRI module requires additional learning to achieve the same capabilities as the proposed MR-integrated 3D/2D visual and motion mapping HRI modules.

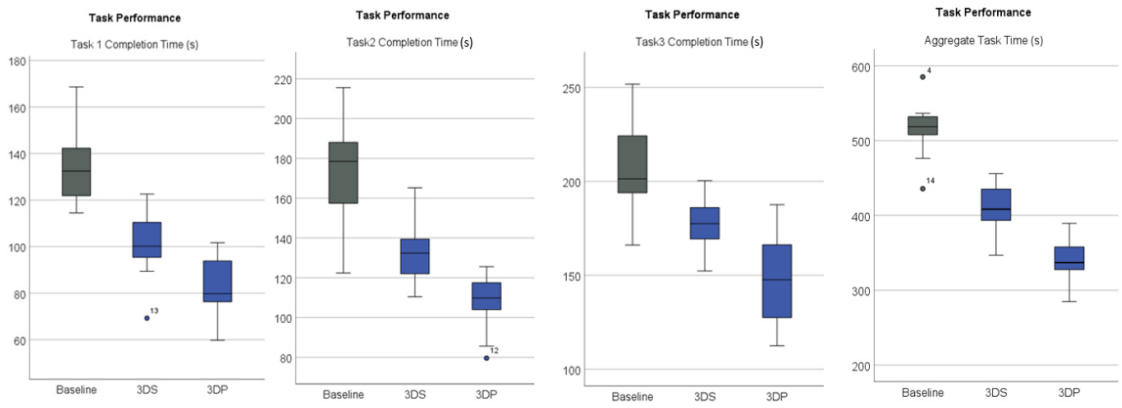


Figure 5.5: Boxplots of quantitative measures on the users' task performance for each MR-HRI module across three manipulation tasks.

Finally, it can also be observed that even at the end of the training, participants did not achieve the same level of proficiency and dexterity they immediately achieved with the MR subspace-integrated 3D/2D vision module. The similar training times for the MR-3DS and MR-3DP modules have an important role in saving training time even for unskilled operators in comparison with the MR-2D Baseline. A significant improvement was observed in collision reduction by mapping the PC in the MR-3D scene, which provides the operators with salient visual depth cues and flexible viewpoints. The total count of undesired collisions reduced from 10 for the MR-2D Baseline to 4 for MR-3DS and 3 for MR-3DP, respectively.

5.4.2 Subjective Measures

The Technology Acceptance Model (TAM) measures system acceptance including usefulness and ease of use (scores range from 0-7, from worst to best). Lower values (close to 0) indicate poor system usability, while relatively high values (close to 7) indicate high ease of use and usability.

As shown in **Table 5.2**, for usefulness, the point cloud-based paradigm (MR-3DP) was rated higher (4.67 ± 1.23) than stereo-vision mode (3.33 ± 1.11) and Baseline (3.07 ± 1.22), indicating that users perceived that using a system with point cloud-based paradigm increased their work performance; for ease of use, point cloud-based paradigm (MR-3DP) was rated higher (5.53 ± 0.92) than stereo-vision mode (3.93 ± 1.10) and Baseline (2.60 ± 0.83), indicating that the point cloud-based paradigm was easier to use and required the minimum amount of effort; for the overall TAM assessment, the point cloud-based paradigm (MR-3DP) had higher values (5.17 ± 0.88) than stereo-vision mode (3.63 ± 0.70) and Baseline (2.83 ± 0.75), showing that the users preferred using MR-3DP for the tasks and had a more positive attitude towards the MR-3DP system. As shown in **Figure 5.6 (b)**, the MR-3DP vision/motion

mapping method (M=4.7) was reported to be slightly more acceptable than the MR-3DS module (M=3.3) and 2D Baseline interface (M=3.1), in terms of perceived usefulness. The TAM results indicated there was an overall significant difference between the means of the users' appeal at the three different HRI modules. Participants found the MR subspace-integrated PC/2D vision and motion-mapping module (MR-3DP) (M=5.1) to be significantly easier to use than the MR-2D Baseline (M=2.8), and marginally easier to use than the MR-3DS module (M=3.6), as shown in **Figure 5.6 (d)**. The subjective measures analysis proved the MR subspace-integrated PC/2D mapping approach for telemanipulation outperformed the MR-3DS and MR-2D Baseline modules in task workload and user perception.

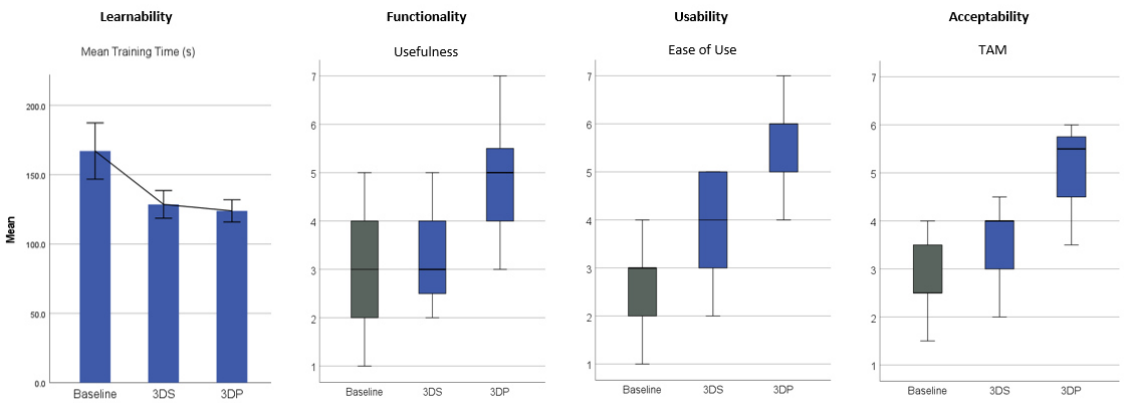


Figure 5.6: System learnability and usability assessment by the telemanipulation conditions across three manipulation tasks. The MR-3DS and MR-3DP designs demonstrated improvements in learnability and several additional subjective measurements on usability.

The NASA task load index (NASA-TLX) assessed cognitive workload. At the end of each task, subjects scored their qualitative experience of mental demand, physical demand, temporal demand, performance, effort, and frustration, on a scale from 0 to 100, with 100 being the most demanding. **Figure 5.7** shows all average NASA-TLX scores were lower for the MR subspace-integrated PC/2D vision and motion-mapping module (MR-3DP) compared with the MR-2D

Baseline and MR-3DS cases. The MR-3DP vision/motion-mapping module significantly reduced the physical and mental demands and frustration of participants. In particular, the physical workload was reduced from the Baseline (M=79.5) compared to the MR-3DS module (M=43.7) and the MR-3DP module (M=41.33). The mental workload was reduced from the Baseline (M=80.9) compared to the MR-3DS module (M=62.2) and the MR-3DP module (M=47.8). In addition, the average frustration score in the NASA-TLX decreased significantly by 44% and 15% in comparison with the Baseline and MR-3DP ($F(1.706, 23.884) = 26.485$, $p < 0.001$, Partial = 0.76) when the MR-3DP mapping was used.

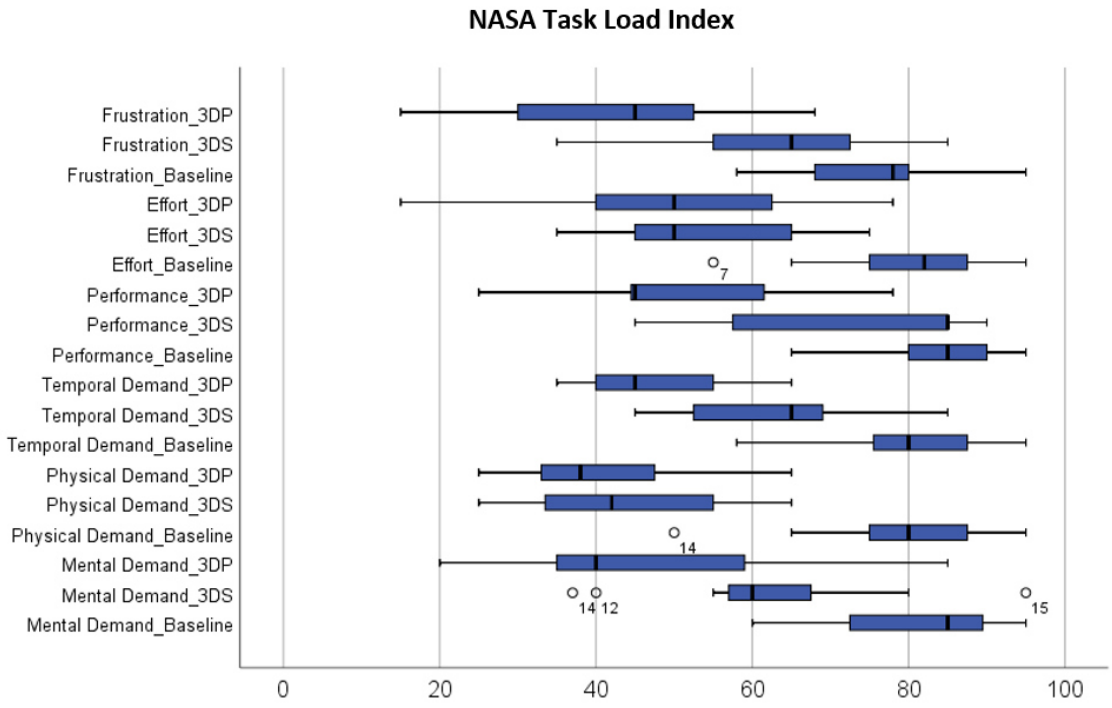


Figure 5.7: Boxplots of quantitative measures on cognitive workload scores across three MR-HRI conditions.

5.5 Summary

This chapter investigated using an MR subspace as a teleoperation framework to integrate 3D/2D multiple visual information and natural motion mapping. Using the MR-integrated vision mapping for immersive remote viewing of robot-assisted manipulation improves task performance and user skill acquisition efficiency.

An experiment involving manipulation tasks of different difficulty levels was conducted to examine the potential benefits of MR integrated 3D/2D vision mapping for teleoperation. In comparison with the MR-2D Baseline, results showed both MR-3DS and MR-3DP approaches can assist telemanipulation by reducing task completion times, shortening the time needed for training, reducing unwanted collisions, and decreasing cognitive workload. The stereopsis and immersion provided by the MR-3D systems allow even non-skilled users to effectively interact with a mobile manipulator in a natural manner. Additionally, it is observed the MR-3DP mapping scheme for HRI outperforms the MR-3DS module and MR-2D Baseline, in terms of task performance, user workload, and system usability.

The proposed MR-3D designs allowed users to complete the manipulation tasks at a distance more effectively than the Baseline MR-2D teleoperation scheme modeled after existing systems in common use. The experiments also indicated the MR-3DP teleoperation mapping scheme allows novice users to complete tasks more quickly than using either MR-3DS or MR-2D Baseline conditions. In future work, we will explore between-task and within-mode comparisons to further analyze the differences between tasks under the same experimental condition.

Chapter 6

Mixed Reality-Enhanced Intuitive Teleoperation with Hybrid Virtual Fixtures for Welding

6.1 Introduction

This chapter presents an integrated scheme based on an MR and haptic feedback approach for the intuitive and immersive teleoperation of robotic welding systems. By incorporating MR technology, the user is fully immersed in a virtual operating space augmented by real-time visual feedback from the robot working space. The proposed robotic tele-welding system features imitative motion mapping from the user's hand movements to the welding robot motions, and it enables the spatial velocity-based control of the robot tool center point (TCP).

The proposed mixed reality virtual fixture (MRVF) integration approach implements hybrid haptic constraints to guide the operator's hand movements following the conical guidance to effectively align the welding torch for welding and constrain the welding operation within a collision-free area. Onsite welding and tele-welding experiments identify the operational

differences between professional and unskilled welders and demonstrate the effectiveness of the proposed MRVF tele-welding framework for novice welders.

Overall, the MRVF-integrated visual/haptic tele-welding scheme reduced the torch alignment times by 56% and 60% compared to the MRnoVF and baseline cases, with minimized cognitive workload and optimal usability. The MRVF scheme effectively stabilized welders' hand movements and eliminated undesirable collisions while generating smooth welds.

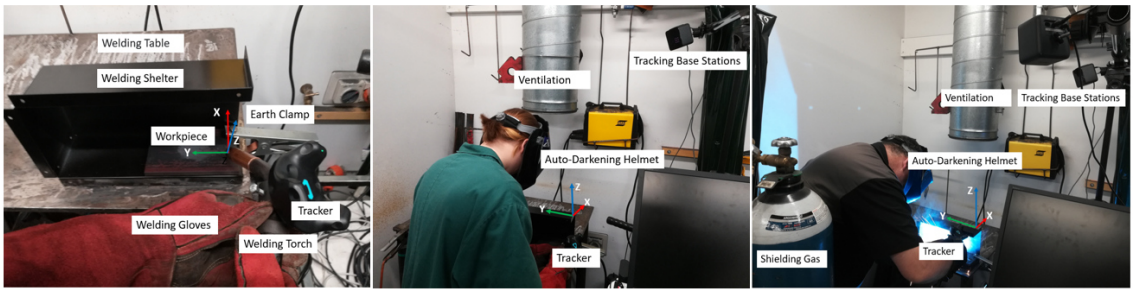
6.2 Methodology

6.2.1 Welding Skill Extraction System Design

The objectives of this research were to: (**1**) remove human welders from hazardous and unpleasant working environments without increasing operational complexity or sacrificing the welding quality; (**2**) enable the welders to conduct tele-welding in the same way it is performed onsite, minimizing the learning required by introducing a tele-welding robot in the loop; (**3**) analyze the operational techniques and welding expertise distinguishing professional welders from unskilled welders; and (**4**) further assist unskilled workers with integrated visual and haptic HRI modalities via MR to improve task performance and system usability in remote-controlled tele-welding and to achieve welding results comparable to those of professional welders. These objectives address key issues in remote tele-welding.

To identify operational differences between unskilled and professional welders, hand movements of professional welders performing manual welding tasks were tracked and compared to those of unskilled welders. **Figure 6.1** shows the hardware components of the gas metal arc welding (GMAW) motion tracking platform, including an HTC Vive tracking system,

welding shelter, welding torch, and extra welding gas/wire/electricity supplies. A 6-DOF Vive tracker was mounted on the welding torch and exposed to the two surrounding Vive tracking base stations for tracking the translational and rotational motion of the welder's torch hand by generating a wireless connection between the tracked welding torch and the base stations. A metal welding shelter enables more precise motion tracking and covers the torch tip and workpieces to prevent infrared (IR) light exposure, which may interfere with the IR-sensitive tracking sensors. Auto-darkening welding helmets and welding gloves were used by all participants.



(a) (b) (c)

Figure 6.1: The hardware components of the GMAW motion tracking platform. (a) Welding motion-tracking platform, including welding shelter, welding torch, and the attached motion tracker; (b) an unskilled welder performing manual welding for motion data collection and analysis; (c) a professional welder performing on-site GMAW operation for expertise and skill extraction.

6.2.2 MRVF Tele-Welding System Overview

In this study, we investigated the impact of an integrated visual/haptic perception in MR on a natural, 3D motion mapping, enhanced immersive, and intuitive tele-welding process. **Figure 6.2** shows the MR-incorporated virtual fixture (MRVF) telerobotic system consisting of four main elements: (1) the welding robot and visualization system; (2) the haptic master robot; (3) the MR workspace implementation; (4) the robot and operator space communication implementation.

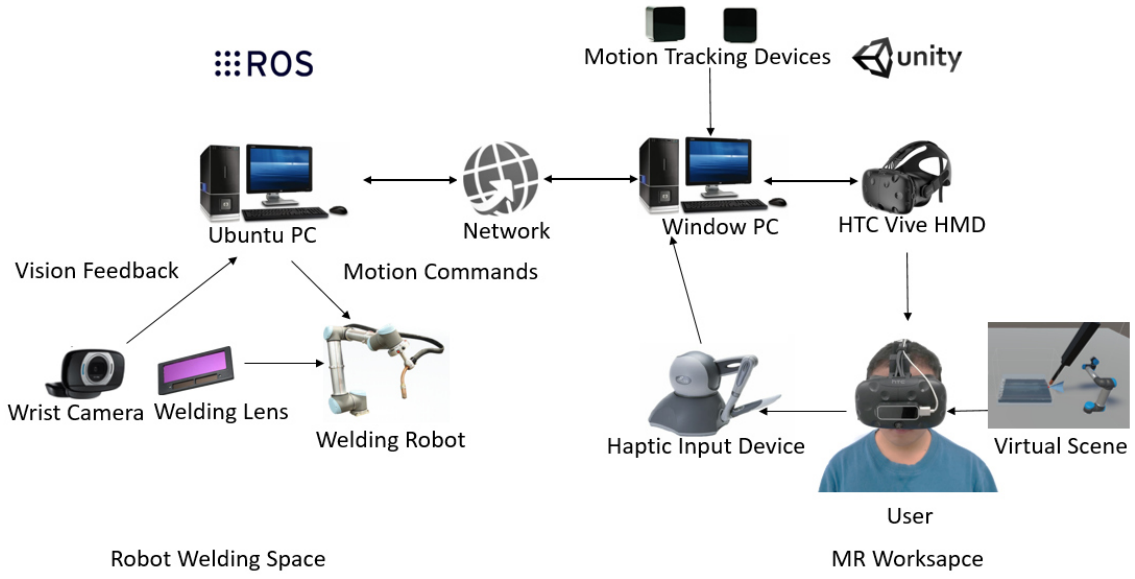
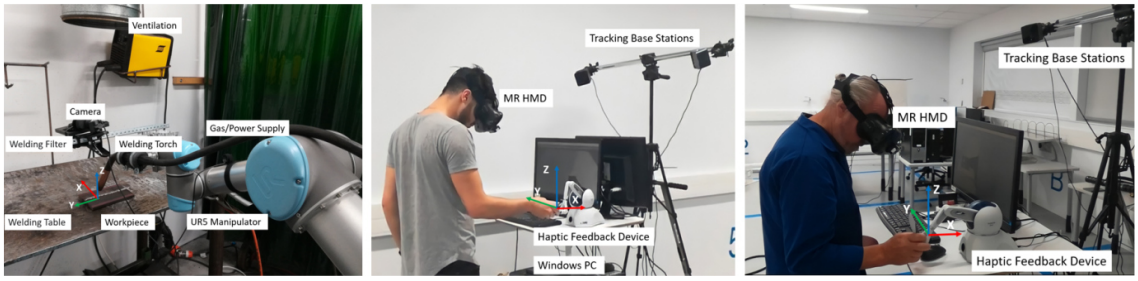


Figure 6.2: The communication scheme of the MRVF tele-welding hardware apparatus.

The remote robotic welding platform consisted of a UR5 industrial manipulator with 6 DOF, gas metal arc welding (GMAW) equipment, a welding camera, and an auto-darkening filter. The UR5 industrial manipulator was equipped with an arc welding torch to perform the remote welding process, as shown in **Figure 6.3**. A monocular Logitech C615 webcam (Logitech International S.A, Lausanne, Switzerland) with an auto-darkening lens was mounted on the robot to observe the welding process and provide the operator with a direct view of the workpieces. A ROS middleware-supported driver for the UR5 robot ran on a computer with an Ubuntu 16.04 operating system. The Ubuntu computer was equipped with an i7-10700 CPU, 64 GB RAM, and GeForce RTX 2060 graphics to command the UR5 robot controller through TCP/IP and process the on-site welding streams. The TCP/IP-based ROS communication protocol was capable of fast control rates at 125 Hz, which is sufficient for teleoperated robotic welding tasks, where real-time control is required.



(a) (b) (c)

Figure 6.3: The imitation-based and robot-assisted teleoperation of the GMAW process. (a) Telerobotic welding platform including the UR5 manipulator, welding torch, and the attached vision system; (b) a novice welder conducting immersive and intuitive robot-assisted welding with haptic guidance; (c) a professional welder performing remote-controlled robotic GMAW operation for comparison.

A Phantom Omni haptic robot (SensAble Technologies Inc., USA) was utilized as the motion input device to remotely operate the welding robot in a manual welding manner. The MRVF system features VCMM from the user's hand movements to the robot motions and enables spatial velocity-based control of the robot tool center point (TCP). The welder uses the stylus of the haptic robot with the same motion and manner as when performing manual welding. This approach enables intuitive and precise user control of the position and orientation of the UR5 end effector to achieve the desired travel speed, and travel/work angles as if the user was directly present.

The operator space for the MRVF tele-welding consisted of an HTC Vive HMD and 27-inch monitor connected to a desktop with an i7-8700k CPU, 32 GB RAM, and a GeForce GTX 1080 graphics processor. The immersive MRVF environment was generated in Unity 3D to display an integrated 3D visualization with overlaid monoscopic image streams and corresponding haptic feedback during the welding process. The ros-bridge provided a network intermediate, enabling the exchange of messages between nodes, and it was used to establish communication between the master and slave robot sides.

6.2.3 MRVF Visual/Haptic Workspace

Digital twin technology was used to capture the physical UR5 robot pose during operation and allowed the welders to view the rotation status of each joint [159]. The combination of the virtual twin and onsite video streams in MR provided comprehensive real-time monitoring of the robot's operating status. It also assisted in accurately and efficiently amending the welding motion based on data from the robot model. The scale ratio for the virtual UR5 robot was 1:5 so the digital twin data and motions fit the user's view in the MR welding workspace.

Virtual fixtures (VFs) can be divided into guidance fixtures and prevention fixtures [160]. The proposed MRVF presented uses a combination of both to guide the users to efficiently navigate to the initial welding point and effectively prevent the torch tip from colliding with the workpiece.

During the welding process, the electrode needs to contact the molten weld pool to ensure the filler metal can be transferred from the electrode to the work. However, contact between the torch tip and the workpiece must be prevented to avoid damage. In the MRVF tele-welding workspace (**Figure 6.4(d)**), a transparent prevention VF panel remains overlaid on the virtual workpiece with a 2D display of the actual welding process to minimize collisions of the torch tip manipulated by the user and the workpiece.

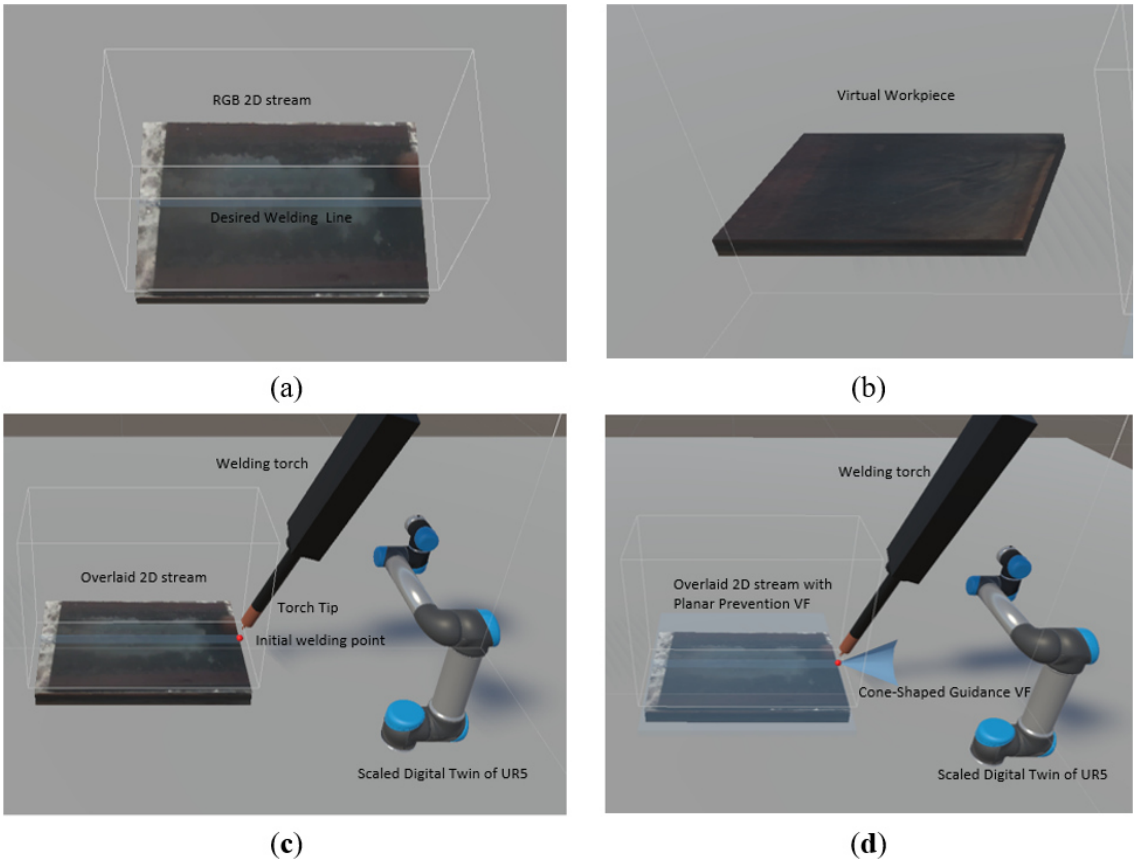


Figure 6.4: The visualization of tele-welding interaction modules used in the tele-welding experiments. (a) Typical 2D visual feedback for remote-controlled robotic welding where the user uses a monitor to observe the welding process, without the immersive HMD usage; (b) the virtual replica of the physical welding workpiece; (c) tele-welding MR platform without haptic effects including welding virtual workpiece, overlaid RGB stream, virtual welding torch, and the scaled digital twin of UR5; (d) the MRVF module involving hybrid guidance and prevention VFs.

The welding experiments revealed it is relatively difficult to move the torch to the exact weld starting point for novice users, and this torch alignment process is often time-consuming and increases overall task completion time. In the MR workspace, a conical guidance fixture is installed with the tip aligned to the welding start point, as shown in **Figure 6.4(d)**. The user simply moves the torch head to the wide end and then quickly moves the virtual torch tip to the cone tip position by following the resistance of the inner wall of the conical shape, and the actual torch is simultaneously driven to the intended welding start point.

Interaction between the haptic robot and the MR environment occurs at the haptic interface point (HIP), representing the corresponding position of the physical haptic probe of the master haptic. The force exerted on the haptic stylus is calculated by simulating a spring between the proxy and the HIP. The resistance force exerted by the haptic stylus to the user's hand is proportional to the distance between the proxy point and the HIP. **Figure 6.5** illustrates how haptic rendering and robot control are implemented using a master-controlled HIP and proxy-controlled robot (MHPR) architecture. Considering the proxy point never violates the constraints imposed by the virtual fixtures, the welding robot will not collide with the workpiece, even though the operator overcomes the resistance force. This architecture forms a hard prevention fixture, allowing the user to maintain the desired contact tip-to-work distance (CTWD), preventing unwanted collisions and increasing the precision and stability of tele-welding operations.

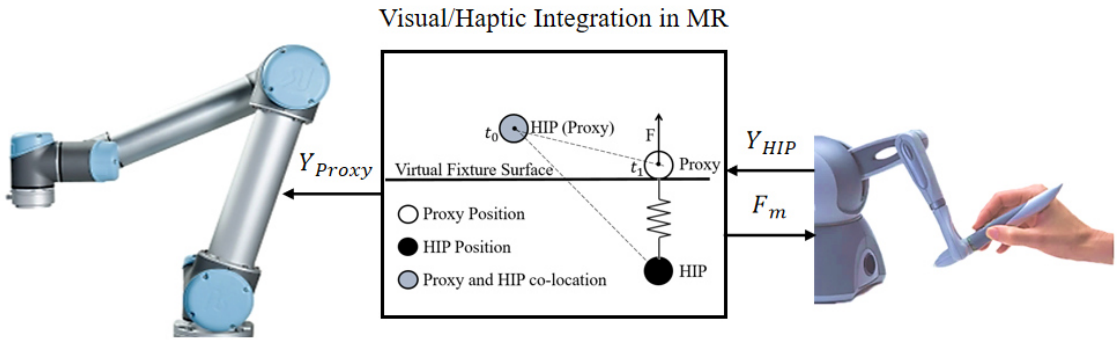


Figure 6.5: Block diagram of the master-controlled HIP and proxy-controlled robot architecture.

6.2.4. Welding Experiments

6.2.4.1. Experimental Design

Two experiments were conducted with novice and professional welders. Experiment 1 (the onsite welding experiment) investigated the motion difference between the expert and unskilled

welders and extracted the expertise and skills of professional welders to optimize the robot-assisted welding platform. The experimental results further served as the "ground truth" for the development of MRVF robot-assisted welding platforms when facilitating novice welders with better weld control by incorporating MR and VF. Experiment 2 (the tele-welding experiment) was carried out to verify the effect of MR-integrated visual and haptic cues on the tele-welding performance of unskilled welders. The study was focused on novice participants to assess improvements and quality relative to professional on-site welding. Experiments were also conducted with professional welders to produce the criteria for the desired welding results.

6.2.4.2 On-Site Welding Experiment

Sixteen (16) student volunteers and four (4) technical staff members were recruited at the University of Canterbury (Christchurch, New Zealand). All participants were right-handed. The 16 students were unskilled welders who self-rated as having no prior experience, and the 4 technical staff members were very experienced welders who perform welding regularly and train undergraduates with no welding experience. Due to the relatively small number of professional welders, each professional welder was asked to weld multiple times to produce a comparable sample size.

Prior to the welding experiments, the workshop technician provided the novice subjects with the same standardized welding face-to-face instructions on the manual GMAW process, including the use of the welding torch, melting conditions, and the desired molten weld pool status for quality welding results. To observe the workshop safety precautions, the professional welders and experimenters remained close onsite and supervised the novice welders throughout

the experiment. The novice welding results are, thus, safe and the best-case results for this cohort.

The welding experiments were conducted using a single-phase welding GMAW machine. A steel workpiece plate was placed in a horizontal position on the welding table for typical flat welding. The dimensions of the plates were $150\,mm \times 100\,mm \times 10\,mm$. The centerline line of the workpiece was set as the intended welding trajectory. Each professional welder was required to perform onsite flat welding four times for GMAW operation expertise and skill extraction. Each novice welder performed once for motion data collection and analysis. The corresponding hand movements and welding results were used to assess the absolute and relative welding performance, distinguishing the gap between experts and novices.

6.2.4.3. Tele-Welding Experiment

A 3 × 1 within-participants experiment was designed to validate whether the designed MRVF scheme facilitated novice welder control of a robotic tele-welding platform to achieve quality welding results and to assess the user experience. The null hypothesis (H0) of the repeated-measures ANOVA was that the baseline, MRnoVF, and MRVF tele-welding paradigms are equally effective in welding quality and welder experience for novices, in terms of effectiveness, intuitiveness, and learnability, using the VCMM imitation-based motion mapping approach as the basis for teleoperation.

In this work, three visualization modules in the tele-welding HRI platform, shown in **Figure 6.4**, were tested to validate the effectiveness of the proposed MRVF tele-welding paradigm. In

particular, to show the differences between the 2D baseline, MR, and MRVF settings for remote tele-welding. Specifically, the three modules are defined:

- **Baseline**: Perform the tele-welding operation with a non-immersive display using monoscopic streams (**Figure 6.4(a)**). The display screen was a standard 27-inch PC monitor. The 2D visualization was transmitted from the monoscopic camera mounted on the welding robot. The welder manipulated the master haptic robot for the welding robot control without haptic effects. The non-immersive 2D display was used as the baseline condition, as it is commonly used for visual feedback in typical remote-controlled welding systems.
- **MRnoVF**: Conduct the tele-welding task with immersive MR-HMD with overlaid monocular images on the top of the virtual workpiece (**Figure 6.4(c)**). The MRnoVF scheme is a limited version of the proposed MRVF module because it does not provide the participants with haptic cues to support hand maneuvering. The haptic device was deployed to command the UR5 arm for welding but provided no force feedback to the operator.
- **MRVF**: MRVF incorporates combined planar prevention and conical guidance haptic cues in the immersive MR workspace (**Figure 6.4 (d)**). The user maneuvered the haptic device within the constraints provided by guidance and prevention VFs while welding with the remotely placed robot. The user inspected the real-time pose of the physical welding robot via the scaled virtual replica in the scene.

The participants ran through all three experimental setups distinguished by increasing levels of visual and haptic HRI modalities. First, each participant read the instructions and completed a pre-task questionnaire recording age, gender, and familiarity with welding, robotics, and MR

experience. The objective of each trial was then explained. Each subject was given the same introduction that demonstrated the proposed intuitive tele-welding platform with the visual/haptic feedback modules they were going to use before testing, ensuring standardized, consistent training for all subjects. After a demonstration, the participants were given 2 min to experience the MR-enhanced telerobotic welding system to familiarize themselves with the haptic robot, MR imagery, and robotic welding platform.

The MR-HMD and haptic stylus were fitted on each participant at the user site. When the subject sent verbal confirmation, the MR welding workspace appeared as intended, and they started completing the teleoperated robotic welding tasks as required. Each participant completed the typical horizontal flat welding task under each experimental condition (2D baseline, MRnoVF, MRVF). The condition sequence was randomized to mitigate learning and fatigue effects. After completing each experimental task using one control-feedback condition, the participants filled out a questionnaire about the HRI module to directly compare the three conditions.

The participants were given unbounded time to complete the welding tasks but were instructed to navigate the torch from a given pose to the desired welding starting pose as effectively as they could. The alignment time participants spent to position the torch tip significantly influences the overall tele-welding completion time compared to the welding itself. Thus, alignment times were measured to evaluate improvement in participant work efficiency with each condition as the VFs aid this process in particular. The number of accidental collisions between the torch tip and the metal was recorded as a performance metric. User effort and workload during teleoperation experiments were evaluated by the NASA task load index (NASA-TLX) score at the end of each task, assessing the qualitative mental demand, physical

demand, time demand, performance, effort, and frustration (score range of 1–100, from the least to the most demanding) [158]. User acceptance and system usability, including usefulness and ease-of-use, was assessed by a questionnaire based on the technology acceptance model (TAM) measuring acceptance and ease-of-use (score range of 0–7, from worst to best) [150], [161].

A one-way within-participants ANOVA with repeated measures analyzed the statistical differences among the means of all measurements [162]. Bonferroni correction indicated which mean values were significantly different and was used in this analysis when the ANOVA test showed a significant main effect of the experiment condition [163]. The Greenhouse–Geisser correction was applied to assess the difference in the welder reports of the baseline, MRnoVF, and MRVF modules as within-subject variables [52].

6.3. Results

6.3.1 Onsite Welding Results

The experiment identified the difference between the welding motion trajectories of the skilled and unskilled welders to assist unskilled welders in achieving better control of the weld in telerobotic welding operations. **Figure 6.6** shows the welding results of the skilled and unskilled welders; the expert welds exhibited consistent uniformity, with a smooth weld surface and even thickness across the weld axis. The results of the unskilled welders were heterogeneous, abrupt, variable, and uneven in thickness and direction. Analysis of the tracked torch motion data was performed to determine the causes of these discrepancies.

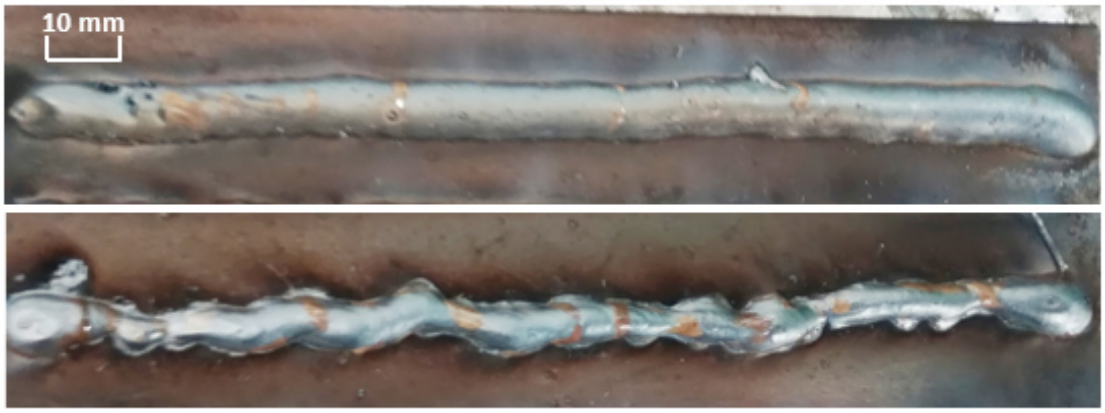


Figure 6.6: Manual welding results of a professional (above) and unskilled welder (below).

Figure 6.7 compares the motions and velocities between the professional and unskilled welders, showing that the unskilled welders had difficulty stabilizing the torch hand movement in the X and Z directions. **Figure 6.7 (b)** shows that both the professional and novice welders could manipulate the torch smoothly in the target direction, Y. Significantly aggressive hand velocities were observed in the X and Z directions, which indicates that the unskilled welders could adjust the motion velocity in the welding direction according to the real-time weld pool status just as the professional welders did, but they had more velocity and motion due to instability. The motion analysis for the hand motion differences was summarized by variance and root mean square error (RMSE) and are shown in **Table 6.1** and **Figure 6.8**, in which the overall results match those in **Figure 6.7**.

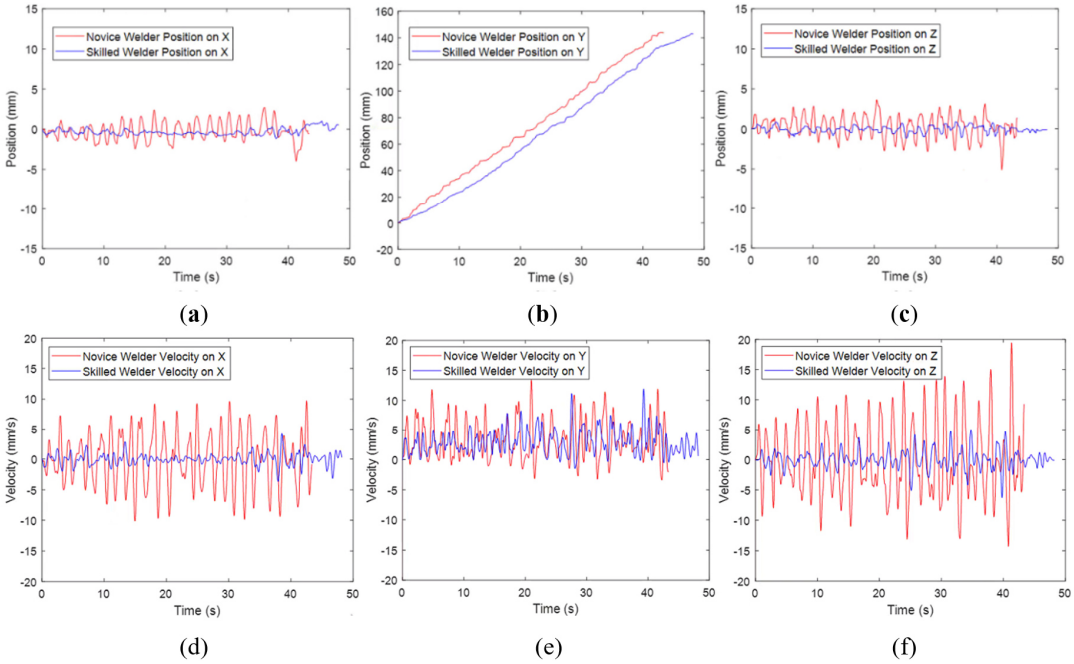


Figure 6.7: Sample welding motion of skilled and novice welders. Direct comparison of (a) back and forth movement perpendicular to the direction of welding movement (X); (b) movement in the direction of welding movement (Y); (c) up and down movement perpendicular to the direction of welding movement (Z). (d-f) show the associated (X, Y, Z) velocities.

Table 6.1: Mean statistic results for objective movement measures for each direction.

	X-Direction		**Z-Direction**		**Y-Direction**
Measured Groups	**Variance**	**RMSE**	**Variance**	**RMSE**	**Mean Velocity in Y**
Skilled Welder	0.26	0.58	0.57	0.90	3.07
Novice Welder	0.96	1.17	1.40	1.27	3.17

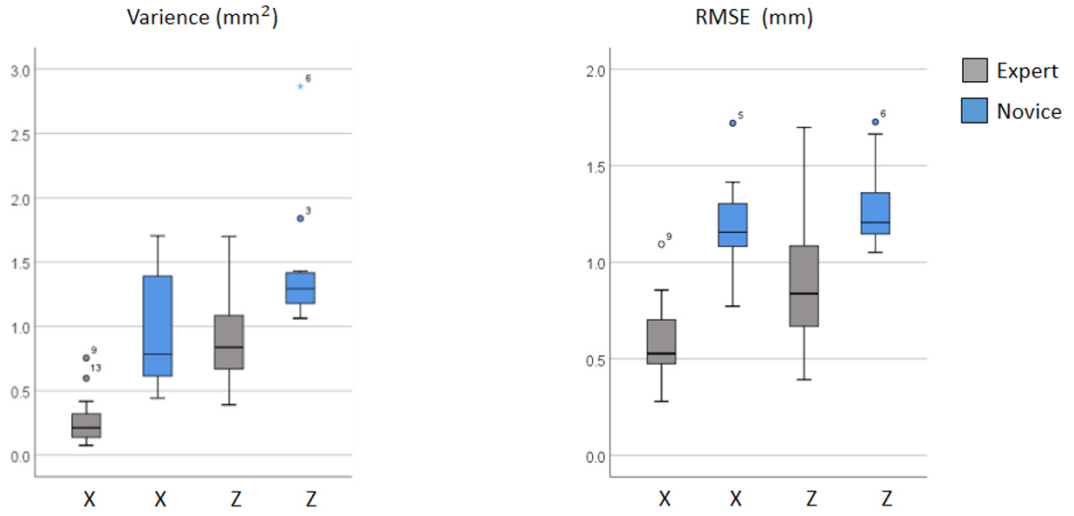


Figure 6.8: Boxplots of quantitative measures in the X and Z directions for variance (a) and RMSE (b) of the users' torch hand motions for both professional and unskilled welders in the onsite welding experiment.

6.3.2 Tele-Welding Results

Overall, all the subjects completed the tele-welding experiments under the three conditions. In the post-experiment questionnaire, the baseline case was rated as the most difficult welding task condition by the majority of participants. Most subjects commented that the MRVF VFs supported their suspended torch hands and reduced fatigue during the robotic welding process. **Figure 6.9** presents a comparison between the sample welding results of the expert and novice welders for the MRVF-integrated visual/haptic scheme, which reduced undesirable deviations of the unskilled welder. The welding results show the gap between the unskilled and professional welders was significantly reduced, and the MRVF condition was intuitive enough to enable experienced welders to quickly transfer their skills from onsite welding to remote tasks.

Figure 6.9: Results of the MR-integrated visual/haptic tele-welding system from a professional welder (above) and unskilled welder (below).

Statistical analysis results that compared the MR-enhanced visual/haptic tele-welding frameworks for HRI paradigms to the baseline and MRnoVF cases are given in **Tables 6.2** and **6.3**. **Table 6.2** lists the mean scores and standard deviations of all measurements and ratings for all participants under each condition. **Table 6.3** lists the p-values and statistical significance of the three modules using one-way ANOVA. The results indicate significant differences between the three visual/haptic integration levels in tele-welding tasks.

Table 6.2: Mean values and standard deviations of all objective and subjective measurements.

	Baseline		MRnoVF		MRVF	
Measure	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Time	46.72	8.75	42.32	11.40	18.60	5.37
Collisions	0.50	0.73	0.25	0.45	0.00	0.00
Mental Demand	80.31	9.41	75.31	10.08	45.00	15.71
Physical Demand	77.94	16.02	74.44	10.58	32.25	13.87
Temporal Demand	78.00	15.56	73.44	13.15	66.81	11.36
Performance	64.75	24.33	78.50	11.80	43.88	18.06
Effort	78.94	14.36	63.56	19.43	34.38	10.78
Frustration	92.81	5.47	69.94	12.92	38.31	11.46
Average Workload	78.79	5.80	72.53	5.13	43.44	4.86
Usefulness	2.25	0.78	2.69	1.08	4.19	1.42
Ease of Use	2.63	1.03	3.13	0.96	5.50	0.82
TAM	2.44	0.48	2.91	0.78	4.85	0.77

Table 6.3: Statistical *p*-values for all quantitative metrics, where B = baseline and MR = MRnoVF.

				Post-Hoc Tests		
Measure	Partial Eta Squared	F	*p*	MRVF-MR	MRVF-B	MR-B
Time	0.76	F (1.866, 27.995) = 47.279	<0.001	<0.001	<0.001	0.478
Collisions	0.21	F (1.424, 21.353) = 4.091	0.043	0.123	0.046	0.783
Mental Demand	0.75	F (1.905, 28.580) = 45.449	<0.001	<0.001	<0.001	0.584
Physical Demand	0.79	F (1.972, 29.580) = 57.679	<0.001	<0.001	<0.001	1.000
Temporal Demand	0.15	F (1.486, 22.292) = 2.594	0.109	0.222	0.118	1.000
Performance	0.49	F (1.505, 22.581) = 14.660	<0.001	<0.001	0.061	0.046
Effort	0.65	F (1.590, 23.852) = 27.782	<0.001	0.001	<0.001	0.156
Frustration	0.88	F (1.703, 25.552) = 112.067	<0.001	<0.001	<0.001	<0.001
Overall Workload	0.94	F (1.703, 25.552) = 228.777	<0.001	<0.001	<0.001	0.023
Usefulness	0.44	F (1.683, 25.241) = 11.719	<0.001	0.017	0.002	0.559
Ease of Use	0.72	F (1.641, 24.617) = 38.829	<0.001	<0.001	<0.001	0.684
TAM	0.78	F (1.916, 28.742) = 54.141	<0.001	<0.001	<0.001	0.180

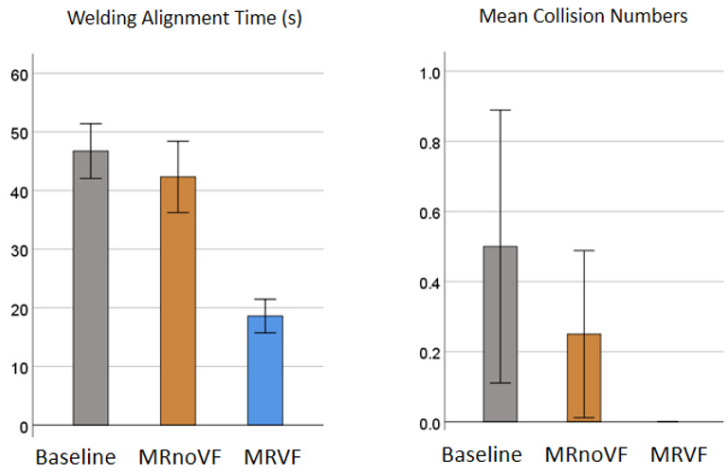
6.3.2.1 Objective Measures

The analysis rejected the null hypothesis (H0) that the MRVF visual/haptic HRI approach for intuitive tele-welding, the MRnoVF, and 2D baseline modules have identical effects on welder performance. In particular, the results show that the MRVF visual/haptic HRI approach significantly outperformed both the 2D baseline and MRnoVF HRI methods on the welding tasks for all pairwise comparisons. Guiding a welding robot using natural welding motion through MR with hybrid guidance/prevention VFs in the MR workspace improved remote welding performance and reduced novice, unskilled welder effort.

As shown in **Figure 6.10 (a)**, a one-way within-subjects ANOVA with repeated measures and a Greenhouse–Geisser correction indicated the time taken to position the torch to the desired welding pose was statistically significantly different—$F(1.866, 27.995) = 47.279$, $p < 0.001$, Partial = 0.76. The post-hoc test revealed the time to position the torch decreased significantly with the MRVF (M = 18.60) compared to the MRnoVF module (M = 42.32) and the baseline module (M = 46.72). Torch alignment times for the welding tasks using the MRVF-integrated visual and haptic tele-welding framework were reduced by 56% and 60% compared to the MRnoVF and baseline cases, respectively, indicating that the typical 2D tele-welding module and the MRnoVF case require additional time to achieve the same capabilities as the proposed MR-integrated visual/haptic HRI module.

Statistical significance was also seen for the average number of collisions between the three HRI modules with $F(1.424, 21.353) = 4.091$, $p < 0.05$, Partial = 0.21. The pairwise comparisons indicated the mean collision numbers to complete the welding task were significantly reduced in the baseline (M = 0.50) compared to the MRnoVF module (M = 0.25) and the MRVF module (M = 0), as shown in **Figure 6.10 (b)**. The statistical results demonstrate

that following through the cone-shaped guidance fixture provided by the MRVF can reduce the welding completion time by minimizing the time used for navigating the torch tip to the initial welding pose. In addition, the prevention VF greatly reduced the likelihood of a collision occurring.



(a) (b)

Figure 6.10: (a) The amount of time participants spent aligning the torch tip across all conditions. The alignment time serves as a major component of task completion time. (b) The average number of collisions for each of the three conditions (Baseline, MRnoVF and MRVF) in the flat position tele-welding tasks. The prevention effect in MRVF eliminates all unintentional contact.

6.3.2.2 Subjective Measures

The NASA task load index (NASA-TLX) assessed the cognitive workload. On a scale of 0 to 100, with 100 being the most difficult, the participants rated their qualitative experience of mental demand, physical demand, temporal demand, performance, effort, and frustration after completing each task. **Figure 6.11** shows all average NASA-TLX scores were lower for the MR-integrated visual and haptic HRI module (MRVF) compared to the baseline and MRnoVF cases. The MRVF visual-haptic mapping module significantly reduced the mental and physical

demands and efforts of participants. In particular, the mental workload was reduced from baseline (M = 80.31) compared to the MRnoVF module (M = 75.31) and the MRVF module (M = 45.00). The physical workload was reduced from baseline (M = 77.94) compared to the MRnoVF module (M = 74.44) and the MRVF module (M = 32.25). In addition, the average effort score in the NASA-TLX decreased significantly by 56% and 19% in comparison to the baseline and MRnoVF, respectively (F (1.590, 23.852) = 27.782, p < 0.001, Partial = 0.65), when the visual and haptic feedback were incorporated in the MRVF.

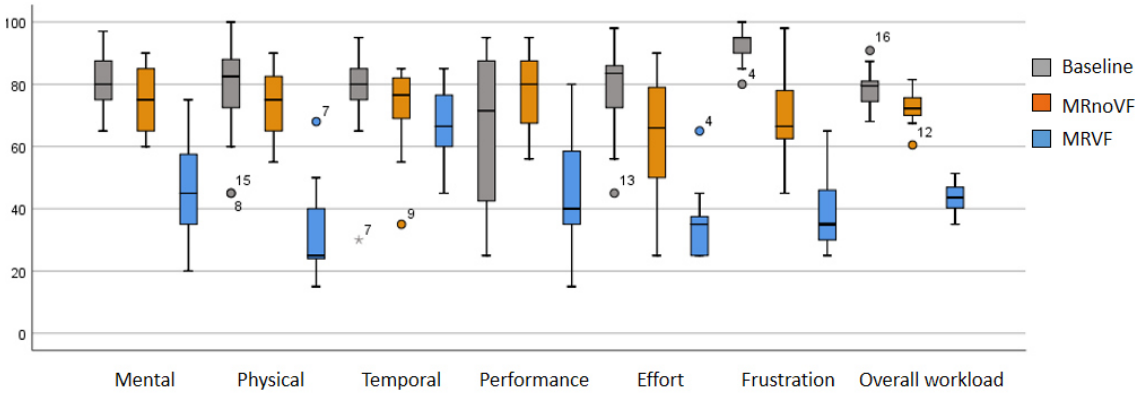


Figure 6.11: Subjective NASA-TLX ratings of task workload across all conditions in the tele-welding tasks.

The technology acceptance model (TAM) evaluated the system functionality, usability, and user's acceptance and perception of the three tele-welding modules. Each scale consisted of three items measured on a seven-point scale (1 = strongly disagree; 7 = strongly agree). The MRVF visual/haptic HRI method (M = 4.19) was reported to be more acceptable than the MRnoVF module (M = 2.69) and baseline case (M = 2.25) in terms of perceived usefulness, as shown in **Figure 6.12**. The TAM results indicate there was an overall significant difference between the means of the users' appeal with the three different HRI modules. The participants

found the MR-integrated visual/haptic tele-welding framework (MRVF) (M = 5.50) to be significantly easier to use compared to the 2D baseline (M = 2.63), and marginally easier to use than the MRnoVF module (M = 3.13). The subjective measures analysis proved the MRVF vision/force mapping approach for tele-welding outperformed the MRnoVF and 2D baseline modules in task workload and user perception.

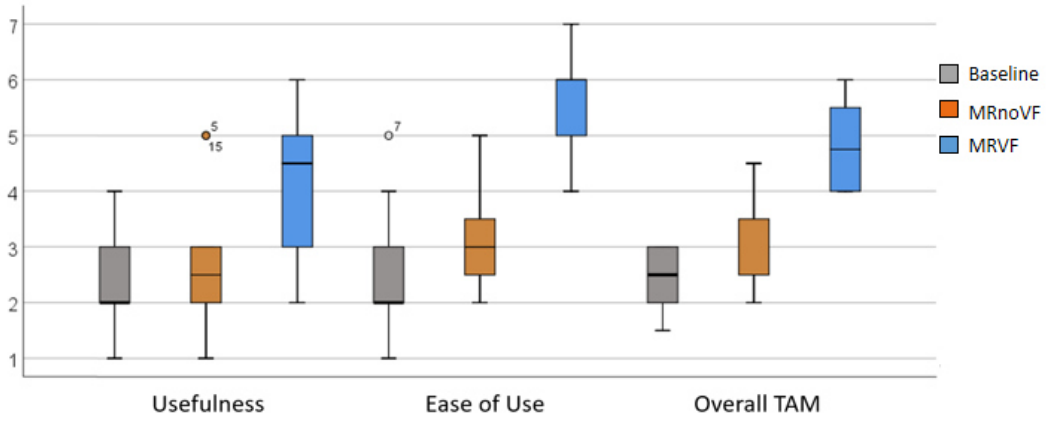


Figure 6.12: Subjective scores on system functionality, usability and user's acceptance and perception of the three tele-welding modules. Higher scores represent more preferences in all cases. The MRVF design demonstrated improvements in perceived usefulness and perceived ease of use.

6.3.3 Limitations

The overall system experiment was conducted at short range. Thus, time lags were relatively very small. Research conducted in other studies indicated that lag between user motion and robot motion causes increasing errors [164], [165]. Ongoing work using Markov models and other forecasting methods can address this issue in future work, given the results in our proof-of-concept system.

The MRVF system presented relatively low-cost and readily available components, where faster or more precise systems could provide greater accuracy, potentially reducing the

improvements seen here. One purpose of this study was to use commercial off-the-shelf products to demonstrate the potential of a relatively low-cost system to achieve tele-welding. Hence, while performance can be improved with better components, it also raises the cost, for which economic feasibility is application-dependent.

Subject numbers were limited in this study, and future work should replicate this effort with a larger study if feasible. However, the relatively low number of unskilled welders had consistent results. Thus, while greater numbers would more accurately quantify the gains to be obtained by an MRVF approach, the consistently large differences seen in both objective and subjective assessments indicate that the results should be replicable. This study aimed to enable inexperienced welders to perform quality remote welding tasks. Unskilled welders do not have frequent contact with the physical welding torch and are not reliant on its weight. It is feasible in future work to replace the handheld stylus on the haptic device with an actual welding torch or a 3D-printed torch model of the same weight to improve the professional welder's experience.

6.4 Summary

This research was focused on immersive and intuitive human–robot interaction with visual and haptic cues, specifically focusing on the MRVF framework for tele-welding scenarios. The MRVF visual/haptic mapping framework provided the welders with an intuitive approach to control the movement of the complex robotic welding system in a manner similar to conventional handheld manual welding via using a single-point grounded haptic robot. The users felt they could access the physical welding scenario from the MR-based operator space, as indicated in the subjective assessments. The MRVF allowed the unskilled, novice welders to rest their suspended torch hands against the VF surface during the robotic welding process,

stabilizing the torch hand movements in the X and Z directions. With the integrated visual and haptic perception, the MRVF tele-welding scheme enabled the non-professional welders to achieve welding results in remote control tele-welding that were comparable to those of professional welders both remotely and on-site, reducing the dependence of remote welding on welder experience and specialized skills. The prevention haptic structure enabled in the MRVF module using VFs successfully eliminated collisions that can damage the robot and/or workpiece. The proposed MRVF visual/haptic framework for remote-controlled welding also enabled professional welders to retain a professional level of operation in the tele-welding process, indicating its intuitive ease of use. Overall, this approach improved the task performance of unskilled, novice welders, increased work efficiency, was intuitive and easy to use, and prevented unwanted collisions.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

The work presented in this thesis contributes to the mechatronic design and development of the inexpensive robotic platform for intuitive mobile manipulation. The development and evaluation of MR-enhanced natural and intuitive teleoperation in telemanipulation tasks were conducted and the implementation and validation of MRand virtual fixture integrated visual/haptic interaction in tele-welding scenarios were carried out. From the research presented, the following conclusions can be drawn.

For the remote manipulation application, the integration of spatial motion and vision retargeting through MR subspace is applied in robotic teleoperation as an interface to enable novice users to remotely control the robotic arm-hand system for performing manipulation tasks in harsh, unstructured situations. Telemanipulation experiments found the MRS integrated scheme reduced aggregate task completion time reduced 48% compared to the 2D Baseline module and 29% compared to the MR SpaceMouse module. The MRS enhanced 3D mapping of motion and vision paradigm improved completion time for operators with minimal technical knowledge. Further, the learning time of the training tasks with the MRS scheme decreased by 53% and 36%, compared to the 2D Baseline and MR SpaceMouse approaches, indicating extra learning was required for the two typical HRI modules to reach the same competency as the proposed MRS imitation-based HRI module. Finally, and overall, the

overall system achieved the desired telemanipulation results for novice users and significantly reduced the physical and mental demand and frustration of participants, while offering higher user acceptance.

The MR-integrated 3D/2D multi-view merging and mapping framework for immersive and intuitive telemanipulation of a complex mobile manipulator for non-skilled operators is presented. The use of MR subspace for the teleoperation of a mobile manipulator is a novel solution for 3D/2D vision and velocity-centric motion mapping. The work directly compares SV and PC enhanced 3D vision via MR for robot-assisted telemanipulation, specifically comparing MR-integrated 3D/2D vision and motion mapping teleoperation schemes against a conventional MR-2D method via three dexterous manipulation tasks. This study compared two MR-integrated 3D/2D vision and motion mapping schemes against a typical 2D Baseline visual display method through pick-and-place, assembly, and dexterous manufacturing tasks. The MR-integrated 3D/2D vision and motion mapping schemes of teleoperation reduced overall task completion times by 34% and 17%, compared to the MR-2D Baseline, while minimizing training effort and cognitive workload.

For the welding application, an on-site welding experiment was designed to investigate the motion difference between the expert and unskilled welders, extracting the expertise and skills of professional welders to optimize the robotic tele-welding platform. Onsite welding and tele-welding experiments identified the operational differences between professional and unskilled welders and demonstrated the effectiveness of the proposed MRVF tele-welding framework for novice welders. With the integrated visual and haptic perception, the MRVF tele-welding scheme enabled the non-professional welders to achieve welding results in remote control tele-

welding comparable to those of professional welders both remotely and on-site, reducing the dependence of remote welding on welder experience and specialized skills. The MRVF-integrated visual/haptic tele-welding scheme reduced the torch alignment times by 56% and 60% compared to the MRnoVF and baseline cases, with minimized cognitive workload and optimal usability. The MRVF scheme effectively stabilized welders' hand movements and eliminated undesirable collisions while generating smooth welds.

Overall, the proposed intuitive and immersive human-robotic interaction framework improved the task performance of unskilled users, increased work efficiency, was intuitive and easy to use, and improved the user's situational awareness, depth perception, and spatial cognition. The system is the first to be fully and quantifiably evaluated from both technical and human subjective interaction perspectives. Finally, the system, supporting methods, and models are fully generalisable to other tele-robotic applications, where the results presented on accuracy and usability could be expected to translate almost directly.

7.2 Future Work

The following recommendations are presented for the extension of the current work. They include two major aspects of improvements. The first part is regarding using the haptic effect to further improve the developed immersive teleoperation platform, while the second part is about using machine learning strategies to learn the operator's behavior and predict their intention for the latency mitigation for long-distance scenarios and employing cloud computing to achieve better management of the developed telerobotic infrastructure.

- *Mixed Reality-Based Intuitive Teleoperation with Robotic haptic/*tactile *effect for telemanipulation*. In the proposed telemanipulation schemes, the haptic effect on user

performance, system usability and collision avoidance was not investigated. Both haptic and tactile feedback has the potential to further improve the task performance and user's situational awareness

- *Mixed Reality-Based Intuitive Teleoperation with Robotic Force Assistance for Intelligent Robotic Welding*. For intuitive tele-welding, the robotic assistance force given by the haptic master robot has the potential to stabilize the user's hand. In this case, the user experiences resistance force given by the haptic master robot all the time when he/she manipulate the robot to drive the remote mobile manipulator for welding.

- *Latency Mitigation Using Applied HMMs for Virtual Teleoperated Welding System*. Using machine learning to learn and predict human motion. Teleoperation enables the most productive utilization of scarce expertise. However, latency is a major issue in long-distance teleoperation. HMMs as a general-case latency mitigation protocol has the potential to deal with error inducing time-delays inherent in MR-based teleoperated welding systems

- *Cloud-based Intuitive Teleoperation platform*. Building cloud architecture and moving the developed platform to the cloud. The integration of Cloud with the developed intuitive teleoperation platform ensures that better management can be conducted on the telerobotic infrastructure and application.

References

[1] H. Su, W. Qi, C. Yang, J. Sandoval, G. Ferrigno, and E. de Momi, "Deep Neural Network Approach in Robot Tool Dynamics Identification for Bilateral Teleoperation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2943–2949, 2020, doi: 10.1109/LRA.2020.2974445.

[2] S. Li, R. Rameshwar, A. M. Votta, and C. D. Onal, "Intuitive Control of a Robotic Arm and Hand System with Pneumatic Haptic Feedback," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4424–4430, 2019, doi: 10.1109/LRA.2019.2937483.

[3] D. Conte, S. Leamy, and T. Furukawa, "Design and Map-based Teleoperation of a Robot for Disinfection of COVID-19 in Complex Indoor Environments," in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics, SSRR 2020*, 2020, pp. 276–282. doi: 10.1109/SSRR50563.2020.9292625.

[4] G. Yang *et al.*, "Keep Healthcare Workers Safe: Application of Teleoperated Robot in Isolation Ward for COVID-19 Prevention and Control," *Chinese Journal of Mechanical Engineering*, vol. 33, no. 1, p. 47, 2020, doi: 10.1186/s10033-020-00464-0.

[5] C. Li, X. Gu, X. Xiao, C. M. Lim, X. Duan, and H. Ren, "A Flexible Transoral Robot Towards COVID-19 Swab Sampling," *Frontiers in Robotics and AI*, vol. 8, no. April, pp. 1–8, 2021, doi: 10.3389/frobt.2021.612167.

[6] Y. Chen *et al.*, "A collaborative robot for COVID-19 oropharyngeal swabbing," *Robotics and Autonomous Systems*, vol. 148, p. 103917, 2022, doi: https://doi.org/10.1016/j.robot.2021.103917.

[7] J. Zhou, W. Chen, S. S. Cheng, L. Xue, M. C. F. Tong, and Y. Liu, "Bio-inspired Soft (BIS) Hand for Tele-operated COVID-19 Oropharyngeal (OP) Swab Sampling," in *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2021, pp. 80–86. doi: 10.1109/ROBIO54168.2021.9739351.

[8] L. Yiming *et al.*, "Electronic skin as wireless human-machine interfaces for robotic VR," *Science Advances*, vol. 8, no. 2, p. eabl6700, Apr. 2022, doi: 10.1126/sciadv.abl6700.

[9] N. Feizi, M. Tavakoli, R. v. Patel, and S. F. Atashzar, "Robotics and AI for Teleoperation, Tele-Assessment, and Tele-Training for Surgery in the Era of COVID-19: Existing Challenges, and Future Vision," *Frontiers in Robotics and AI*, vol. 8, no. April, pp. 1–9, 2021, doi: 10.3389/frobt.2021.610677.

[10] S. Elmoghazy, E. Yaacoub, N. v Navkar, A. Mohamed, and A. Erbad, "Survey of Immersive Techniques for Surgical Care Telemedicine Applications," in *2021 10th Mediterranean Conference on Embedded Computing (MECO)*, 2021, pp. 1–6. doi: 10.1109/MECO52532.2021.9460135.

[11] R. Ebad, "Telemedicine: Current and Future Perspectives," *International Journal of Computer Science Issues (IJCSI)*, vol. 10, no. 6, pp. 242–249, 2013.

[12] J. Varela-Aldás, J. Buele, P. Ramos Lorente, I. García-Magariño, and G. Palacios-Navarro, "A Virtual Reality-Based Cognitive Telerehabilitation System for Use in the COVID-19 Pandemic," *Sustainability*, vol. 13, no. 4, p. 2138, 2021, doi: 10.3390/su13042183.

[13] G. Bauer and Y.-J. Pan, “Review of Control Methods for Upper Limb Telerehabilitation With Robotic Exoskeletons,” *IEEE Access*, vol. 8, pp. 203382–203397, 2020, doi: 10.1109/ACCESS.2020.3036596.

[14] C. Li *et al.*, “A novel master-slave teleoperation robot system for diaphyseal fracture reduction: A preliminary study,” *Computer Assisted Surgery*, vol. 21, pp. 163–168, 2016, doi: 10.1080/24699322.2016.1240304.

[15] Z. Wang, Z. Chen, and B. Liang, “Fixed-time velocity reconstruction scheme for space teleoperation systems: Exp Barrier Lyapunov Function approach,” *Acta Astronautica*, vol. 157, pp. 92–101, 2019, doi: 10.1016/j.actaastro.2018.12.018.

[16] Y. Shen *et al.*, “Robots under COVID-19 Pandemic: A Comprehensive Survey,” *IEEE Access*, vol. 9, pp. 1590–1615, 2021, doi: 10.1109/ACCESS.2020.3045792.

[17] K. H. Lee, V. Pruks, and J. H. Ryu, “Development of shared autonomy and virtual guidance generation system for human interactive teleoperation,” *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2017*, pp. 457–461, 2017, doi: 10.1109/URAI.2017.7992775.

[18] K. T. Song, S. Y. Jiang, and M. H. Lin, “Interactive Teleoperation of a Mobile Manipulator Using a Shared-Control Approach,” *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 6, pp. 834–845, 2016, doi: 10.1109/THMS.2016.2586760.

[19] L. Gao, Z. Xu, W. Huang, and A. Song, “Design and application of experimental platform for interactive teleoperation robot,” *Journal of Southeast University (Natural Science Edition)*, vol. 34, no. 6, pp. 775–779, 2004.

[20] H. Saeidi, J. R. Wagner, and Y. Wang, “A mixed-initiative haptic teleoperation strategy for mobile robotic systems based on bidirectional computational trust analysis,” *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1500–1507, 2017, doi: 10.1109/TRO.2017.2718549.

[21] J. E. Solanes, A. Muñoz, L. Gracia, A. Martí, V. Girbés-Juan, and J. Tornero, “Teleoperation of industrial robot manipulators based on augmented reality,” *International Journal of Advanced Manufacturing Technology*, vol. 111, no. 3–4, pp. 1077–1097, 2020, doi: 10.1007/s00170-020-05997-1.

[22] F. Navarro, J. Fdez, M. Garzón, J. J. Roldán, and A. Barrientos, “Integrating 3D Reconstruction and Virtual Reality: A New Approach for Immersive Teleoperation,” *Advances in Intelligent Systems and Computing*, vol. 694, pp. 606–616, 2018, doi: 10.1007/978-3-319-70836-2_50.

[23] J. I. Lipton, A. J. Fay, and D. Rus, “Baxter’s Homunculus: Virtual Reality Spaces for Teleoperation in Manufacturing,” *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 179–186, 2018, doi: 10.1109/LRA.2017.2737046.

[24] X. Wang and P. S. Dunston, “Mixed reality - Enhanced operator interface for teleoperation systems in unstructured environment,” *Earth and Space 2006 - Proceedings of the 10th Biennial International Conference on Engineering, Construction, and Operations in Challenging Environments*, vol. 2006, p. 93, 2006, doi: 10.1061/40830(188)93.

[25] A. Koubaa, *Robot Operating System (ROS)*, vol. 1. Cham: Springer, 2017. doi: 10.1007/978-3-319-54927-9.

[26] A. Zelenak, C. Peterson, J. Thompson, and M. Pryor, “The advantages of velocity control for reactive robot motion,” *ASME 2015 Dynamic Systems and Control Conference, DSCC 2015*, vol. 57267, p. V003T43A003), 2015, doi: 10.1115/DSCC2015-9713.

[27] A. Kheddar, E. S. Neo, R. Tadakuma, and K. Yokoi, “Enhanced teleoperation through virtual reality techniques,” *Springer Tracts in Advanced Robotics*, vol. 31, pp. 139–159, 2007, doi: 10.1007/978-3-540-71364-7_10.

[28] M. Ismet Can Dede, Ö. Selvi, T. Bilgincan, and Y. Kant, “Design of a haptic device for teleoperation and virtual reality systems,” *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, pp. 3623–3628, 2009, doi: 10.1109/ICSMC.2009.5346857.

[29] H. Hedayati, M. Walker, and D. Szafir, “Improving Collocated Robot Teleoperation with Augmented Reality,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 78–86, 2018, doi: 10.1145/3171221.3171251.

[30] Y. C. Liu, M. H. Khong, and T. W. Ou, “Nonlinear bilateral teleoperators with non-collocated remote controller over delayed network,” *Mechatronics*, vol. 45, pp. 25–36, 2017, doi: 10.1016/j.mechatronics.2017.05.005.

[31] Z. H. Lai, W. Tao, M. C. Leu, and Z. Yin, “Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing,” *Journal of Manufacturing Systems*, vol. 55, no. February, pp. 69–81, 2020, doi: 10.1016/j.jmsy.2020.02.010.

[32] M. E. Walker, H. Hedayati, and D. Szafir, “Robot Teleoperation with Augmented Reality Virtual Surrogates,” in *ACM/IEEE International Conference on Human-Robot Interaction*, 2019, pp. 202–210. doi: 10.1109/HRI.2019.8673306.

[33] H. Liu and L. Wang, “Remote human – robot collaboration : A cyber – physical system application for hazard manufacturing environment,” *Journal of Manufacturing Systems*, vol. 54, no. November 2019, pp. 24–34, 2020, doi: 10.1016/j.jmsy.2019.11.001.

[34] P. Stotko *et al.*, “A VR System for Immersive Teleoperation and Live Exploration with a Mobile Robot,” in *IEEE International Conference on Intelligent Robots and Systems*, 2019, pp. 3630–3637. doi: 10.1109/IROS40897.2019.8968598.

[35] L. Almeida, P. Menezes, and J. Dias, “Improving robot teleoperation experience via immersive interfaces,” *Proceedings of 2017 4th Experiment at International Conference: Online Experimentation, exp.at 2017*, pp. 87–92, Jul. 2017, doi: 10.1109/EXPAT.2017.7984414.

[36] F. de Pace, G. Gorjup, H. Bai, A. Sanna, M. Liarokapis, and M. Billinghurst, “Leveraging Enhanced Virtual Reality Methods and Environments for Efficient, Intuitive, and Immersive Teleoperation of Robots,” pp. 12967–12973, Oct. 2021, doi: 10.1109/ICRA48506.2021.9560757.

[37] T. Zhou, Q. Zhu, and J. Du, “Intuitive robot teleoperation for civil engineering operations with virtual reality and deep learning scene reconstruction,” *Advanced Engineering Informatics*, vol. 46, no. May, pp. 10170–10190, Oct. 2020, doi: 10.1016/J.AEI.2020.101170.

[38] M. D. Duong, C. Teraoka, T. Imamura, T. Miyoshi, and K. Terashima, “Master-slave system with teleoperation for rehabilitation,” *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 16, pp. 48–53, 2005, doi: 10.3182/20050703-6-CZ-1902.01410.

[39] N. V. Q. Hung, T. Narikiyo, H. D. Tuan, and K. Fuwa, “A Dynamics-Based Adaptive Control for Master-Slave System in Teleoperation,” *IFAC Proceedings Volumes*, vol. 34, no. 6, pp. 237–242, Jul. 2001, doi: 10.1016/S1474-6670(17)35180-7.

[40] Y. Ji, D. Liu, and Y. Guo, “Adaptive neural network based position tracking control for Dual-master/Single-slave teleoperation system under communication constant time delays,” *ISA Transactions*, vol. 93, pp. 80–92, Oct. 2019, doi: 10.1016/J.ISATRA.2019.03.019.

[41] G. LIU, X. GENG, L. LIU, and Y. WANG, “Haptic based teleoperation with master-slave motion mapping and haptic rendering for space exploration,” *Chinese Journal of Aeronautics*, vol. 32, no. 3, pp. 723–736, Mar. 2019, doi: 10.1016/J.CJA.2018.07.009.

[42] H. Jin, L. Zhang, S. Rockel, J. Zhang, Y. Hu, and J. Zhang, “A novel optical tracking based tele-control system for tabletop object manipulation tasks,” *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015-December, pp. 636–642, Dec. 2015, doi: 10.1109/IROS.2015.7353439.

[43] I. Cerulo, F. Ficuciello, V. Lippiello, and B. Siciliano, “Teleoperation of the SCHUNK S5FH under-actuated anthropomorphic hand using human hand motion tracking,” *Robotics and Autonomous Systems*, vol. 89, pp. 75–84, Mar. 2017, doi: 10.1016/J.ROBOT.2016.12.004.

[44] F. Suligoj, B. Jerbic, M. Svaco, B. Sekoranja, D. Mihalinec, and J. Vidakovic, “Medical applicability of a low-cost industrial robot arm guided with an optical tracking system,” *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015-December, pp. 3785–3790, Dec. 2015, doi: 10.1109/IROS.2015.7353908.

[45] T. Q. Dinh, J. il Yoon, J. Marco, P. Jennings, K. K. Ahn, and C. Ha, “Sensorless force feedback joystick control for teleoperation of construction equipment,” *International Journal of Precision Engineering and Manufacturing*, vol. 18, no. 7, pp. 955–969, 2017, doi: 10.1007/s12541-017-0113-5.

[46] D. Q. Truong, B. N. M. Truong, N. T. Trung, S. A. Nahian, and K. K. Ahn, “Force reflecting joystick control for applications to bilateral teleoperation in construction machinery,” *International Journal of Precision Engineering and Manufacturing*, vol. 18, no. 3, pp. 301–315, 2017, doi: 10.1007/s12541-017-0038-z.

[47] J. Nakanishi, S. Itadera, T. Aoyama, and Y. Hasegawa, “Towards the development of an intuitive teleoperation system for human support robot using a VR device,” *Advanced Robotics*, vol. 34, no. 19, pp. 1239–1253, 2020, doi: 10.1080/01691864.2020.1813623.

[48] C. Meeker, T. Rasmussen, and M. Ciocarlie, “Intuitive Hand Teleoperation by Novice Operators Using a Continuous Teleoperation Subspace,” in *Proceedings - IEEE International Conference on Robotics and Automation*, 2018, pp. 5821–5827. doi: 10.1109/ICRA.2018.8460506.

[49] S. R. Ellis, B. D. Adelstein, and R. B. Welch, “Kinesthetic Compensation for Misalignment of Teleoperator Controls through Cross-Modal Transfer of Movement Coordinates,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2002, vol. 46, no. 17, pp. 1551–1555. doi: 10.1177/154193120204601705.

[50] G. Li, F. Caponetto, E. del Bianco, V. Katsageorgiou, I. Sarakoglou, and N. G. Tsagarakis, “Incomplete Orientation Mapping for Teleoperation with One DoF Master-Slave Asymmetry,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5167–5174, 2020, doi: 10.1109/LRA.2020.3006796.

[51] B. Bejczy *et al.*, “Mixed reality interface for improving mobile manipulator teleoperation in contamination critical applications,” *Procedia Manufacturing*, vol. 51, pp. 620–626, 2020, doi: 10.1016/j.promfg.2020.10.087.

[52] E. Triantafyllidis, C. McGreavy, J. Gu, and Z. Li, “Study of multimodal interfaces and the improvements on teleoperation,” *IEEE Access*, vol. 8, pp. 78213–78227, 2020, doi: 10.1109/ACCESS.2020.2990080.

[53] A. W. W. Yew, S. K. Ong, and A. Y. C. Nee, “Immersive Augmented Reality Environment for the Teleoperation of Maintenance Robots,” *Procedia CIRP*, vol. 61, pp. 305–310, 2017, doi: 10.1016/j.procir.2016.11.183.

[54] R. Komatsu, H. Fujii, Y. Tamura, A. Yamashita, and H. Asama, “Free viewpoint image generation system using fisheye cameras and a laser rangefinder for indoor robot teleoperation,” *ROBOMECH Journal*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.1186/s40648-020-00163-4.

[55] L. G. Ribeiro, O. J. Suominen, A. Durmush, S. Peltonen, E. R. Morales, and A. Gotchev, “Retro-reflective-marker-aided target pose estimation in a safety-critical environment,” *Applied Sciences (Switzerland)*, 2021, vol. 11, no. 1, pp. 1–26. doi: 10.3390/app11010003.

[56] B. Illing, M. Westhoven, B. Gaspers, N. Smets, B. Bruggemann, and T. Mathew, “Evaluation of Immersive Teleoperation Systems using Standardized Tasks and Measurements,” *29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020*, pp. 278–285, 2020, doi: 10.1109/RO-MAN47096.2020.9223497.

[57] J. Chen, M. Glover, C. Yang, C. Li, Z. Li, and A. Cangelosi, “Development of an immersive interface for robot teleoperation,” in *Annual Conference Towards Autonomous Robotic Systems*, 2017, pp. 1–15. doi: 10.1007/978-3-319-64107-2_1.

[58] G. Cody, R. Brian, W. Allison, M. Miller, and A. Stoytchev, “An Effective and Intuitive Control Interface for Remote Robot Teleoperation with Complete Haptic Feedback,” *proceedings of the Emerging Technologies Conference-ETC. sn*, 2009.

[59] C. Yang, X. Wang, Z. Li, Y. Li, and C. Y. Su, “Teleoperation Control Based on Combination of Wave Variable and Neural Networks,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 8, pp. 2125–2136, 2017, doi: 10.1109/TSMC.2016.2615061.

[60] V. Girbes-Juan, V. Schettino, Y. Demiris, and J. Tornero, “Haptic and Visual Feedback Assistance for Dual-Arm Robot Teleoperation in Surface Conditioning Tasks,” *IEEE Transactions on Haptics*, vol. 14, no. 1, pp. 44–56, 2021, doi: 10.1109/TOH.2020.3004388.

[61] M. Zolotas, M. Wonsick, P. Long, and T. Padır, “Motion Polytopes in Virtual Reality for Shared Control in Remote Manipulation Applications,” *Frontiers in Robotics and AI*, vol. 8, no. September, pp. 1–14, 2021, doi: 10.3389/frobt.2021.730433.

[62] X. Gao, J. Silverio, E. Pignat, S. Calinon, M. Li, and X. Xiao, “Motion mappings for continuous bilateral teleoperation,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5048–5055, 2021, doi: 10.1109/LRA.2021.3068924.

[63] M. Ferre, S. Cobos, R. Aracil, and M. A. S. Urán, “3D-image visualization and its performance in teleoperation,” in *Second International Conference, ICVR 2007, Held as part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings (2007)*, 2007, vol. 4563 LNCS, pp. 21–27. doi: 10.1007/978-3-540-73335-5_3.

[64] S. Arévalo Arboleda, T. Dierks, F. Rücker, and J. Gerken, “Exploring the Visual Space to Improve Depth Perception in Robot Teleoperation Using Augmented Reality: The Role of Distance and Target’s Pose in Time, Success, and Certainty,” in *IFIP Conference on Human-Computer Interaction. Springer, Cham.*, 2021, vol. 12932 LNCS, pp. 522–543. doi: 10.1007/978-3-030-85623-6_31.

[65] T. Guzsvinecz, C. Kovacs, … D. R.-2018 9th I., and U. 2018, “Developing a virtual reality application for the improvement of depth perception,” in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2018, pp. 17–22. doi: 10.1109/CogInfoCom.2018.8639935.

[66] F. Chenf *et al.*, “A framework of teleoperated and stereo vision guided mobile manipulation for industrial automation,” in *2016 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2016*, 2016, pp. 1641–1648. doi: 10.1109/ICMA.2016.7558810.

[67] N. McHenry *et al.*, “Predictive XR Telepresence for Robotic Operations in Space,” *IEEE Aerospace Conference Proceedings*, vol. 2021-March, pp. 1–10, 2021, doi: 10.1109/AERO50100.2021.9438161.

[68] N. Smolyanskiy and M. Gonzalez-Franco, “Stereoscopic first person view system for drone navigation,” *Frontiers Robotics AI*, vol. 4, no. MAR, pp. 1–10, 2017, doi: 10.3389/frobt.2017.00011.

[69] S. Livatino, G. Muscato, and F. Privitera, “Stereo viewing and virtual reality technologies in mobile robot teleguide,” *IEEE Transactions on Robotics*, vol. 25, no. 6, pp. 1343–1355, 2009, doi: 10.1109/TRO.2009.2028765.

[70] L. Niu, L. Aha, J. Mattila, A. Gotchev, and E. Ruiz, “A stereoscopic eye-in-hand vision system for remote handling in ITER,” *Fusion Engineering and Design*, vol. 146, pp. 1790–1795, 2019, doi: 10.1016/j.fusengdes.2019.03.036.

[71] Z. Guo *et al.*, “Applications of virtual reality in maintenance during the industrial product lifecycle: A systematic review,” *Journal of Manufacturing Systems*, vol. 56, no. May 2019, pp. 525–538, 2020, doi: 10.1016/j.jmsy.2020.07.007.

[72] J. Ratcliffe, F. Soave, N. Bryan-Kinns, L. Tokarchuk, and I. Farkhatdinov, “Extended Reality (XR) Remote Research: A Survey of Drawbacks and Opportunities,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–13. doi: 10.1145/3411764.3445170.

[73] S. Doolani *et al.*, “A Review of Extended Reality (XR) Technologies for Manufacturing Training,” *Technologies 2020, Vol. 8, Page 77*, vol. 8, no. 4, p. 77, Dec. 2020, doi: 10.3390/TECHNOLOGIES8040077.

[74] T. Morimoto *et al.*, "XR (Extended Reality: Virtual Reality, Augmented Reality, Mixed Reality) Technology in Spine Medicine: Status Quo and Quo Vadis," *Journal of Clinical Medicine*, vol. 11, no. 2, p. 470, Jan. 2022, doi: 10.3390/JCM11020470.

[75] G. Zwolinski *et al.*, "Extended Reality in Education and Training: Case Studies in Management Education," *Electronics (Basel)*, vol. 11, no. 3, p. 336, 2022, doi: 10.3390/ELECTRONICS11030336.

[76] P. Milgram and F. Kishino, "Taxonomy of mixed reality visual displays," *IEICE Transactions on Information and Systems*, vol. E77-D, no. 12, pp. 1321–1329, 1994, doi: 10.1.1.102.4646.

[77] X. Han, Y. Chen, Q. Feng, and H. Luo, "Augmented Reality in Professional Training: A Review of the Literature from 2001 to 2020," *Applied Sciences*, vol. 12, no. 3, p. 1024, Jan. 2022, doi: 10.3390/APP12031024.

[78] Z. Makhataeva and H. A. Varol, "Augmented reality for robotics: A review," *Robotics*, vol. 9, no. 2, p. 21, Jun. 2020, doi: 10.3390/ROBOTICS9020021.

[79] C. Li, A. Fahmy, and J. Sienz, "An augmented reality based human-robot interaction interface using Kalman filter sensor fusion," *Sensors (Switzerland)*, vol. 19, no. 20, p. 4586, Oct. 2019, doi: 10.3390/S19204586.

[80] M. Wonsick and T. Padir, "A systematic review of virtual reality interfaces for controlling and interacting with robots," *Applied Sciences (Switzerland)*, vol. 10, no. 24, pp. 1–17, Dec. 2020, doi: 10.3390/APP10249051.

[81] E. Prati, V. Villani, M. Peruzzini, and L. Sabattini, "An Approach Based on VR to Design Industrial Human-Robot Collaborative Workstations," *Applied Sciences*, vol. 11, no. 24, p. 11773, Dec. 2021, doi: 10.3390/APP112411773.

[82] T. S. Sievers, B. Schmitt, P. Rückert, M. Petersen, and K. Tracht, "Concept of a mixed-reality learning environment for collaborative robotics," *Procedia Manufacturing*, vol. 45, pp. 19–24, 2020, doi: 10.1016/J.PROMFG.2020.04.034.

[83] J. A. Frank, M. Moorhead, and V. Kapila, "Mobile mixed-reality interfaces that enhance human-robot interaction in shared spaces," *Frontiers Robotics AI*, vol. 4, no. JUN, p. 20, Jun. 2017, doi: 10.3389/FROBT.2017.00020/FULL.

[84] R. Zhang, X. Liu, J. Shuai, and L. Zheng, "Collaborative robot and mixed reality assisted microgravity assembly for large space mechanism," *Procedia Manufacturing*, vol. 51, pp. 38–45, 2020, doi: 10.1016/J.PROMFG.2020.10.007.

[85] S. H. Choi *et al.*, "An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102258, Feb. 2022, doi: 10.1016/J.RCIM.2021.102258.

[86] M. Khatib, K. al Khudir, and A. de Luca, "Human-robot contactless collaboration with mixed reality interface," *Robotics and Computer-Integrated Manufacturing*, vol. 67, p. 102030, Feb. 2021, doi: 10.1016/J.RCIM.2020.102030.

[87] G. Palma, S. Perry, and P. Cignoni, "Augmented virtuality using touch-sensitive 3D-printed objects," *Remote Sensing*, vol. 13, no. 11, MDPI AG, p. 2186, Jun. 01, 2021. doi: 10.3390/RS13112186.

[88] R. Gralak, “A method of navigational information display using augmented virtuality,” *Journal of Marine Science and Engineering*, vol. 8, no. 4, p. 237, Apr. 2020, doi: 10.3390/JMSE8040237.

[89] M. Ostanin, R. Yagfarov, and A. Klimchik, “Interactive robots control using mixed reality,” *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 695–700, Sep. 2019, doi: 10.1016/J.IFACOL.2019.11.307.

[90] K. Nakamura, K. Tohashi, Y. Funayama, H. Harasawa, and J. Ogawa, “Dual-arm robot teleoperation support with the virtual world,” *Artificial Life and Robotics*, vol. 25, no. 2, pp. 286–293, 2020, doi: 10.1007/s10015-020-00587-6.

[91] D. Whitney, E. Rosen, D. Ullman, E. Phillips, and S. Tellex, “ROS Reality: A Virtual Reality Framework Using Consumer-Grade Hardware for ROS-Enabled Robots,” in *IEEE International Conference on Intelligent Robots and Systems*, 2018, pp. 5018–5025. doi: 10.1109/IROS.2018.8593513.

[92] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, “Comparing Robot Grasping Teleoperation Across Desktop and Virtual Reality with ROS Reality,” *Springer Proceedings in Advanced Robotics*, vol. 10, pp. 335–350, 2020, doi: 10.1007/978-3-030-28619-4_28.

[93] J. Delpreto *et al.*, “Helping Robots Learn: A Human-Robot Master-Apprentice Model Using Demonstrations via Virtual Reality Teleoperation,” *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 10226–10233, 2020, doi: 10.1109/ICRA40945.2020.9196754.

[94] N. Britton, K. Yoshida, J. Walker, K. Nagatani, G. Taylor, and L. Dauphin, “Lunar micro rover design for exploration through virtual reality tele-operation,” *Springer Tracts in Advanced Robotics*, vol. 105, pp. 259–272, 2015, doi: 10.1007/978-3-319-07488-7_18.

[95] A. Naceri *et al.*, “Towards a virtual reality interface for remote robotic teleoperation,” in *2019 19th International Conference on Advanced Robotics, ICAR 2019*, 2019, pp. 284–289. doi: 10.1109/ICAR46387.2019.8981649.

[96] T. Zhang *et al.*, “Deep Imitation Learning for Complex Manipulation Tasks from Virtual Reality Teleoperation,” *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 5628–5635, 2018, doi: 10.1109/ICRA.2018.8461249.

[97] D. Concannon, R. Flynn, and N. Murray, “A quality of experience evaluation system and research challenges for networked virtual reality-based teleoperation applications,” in *Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems, MMVE 2019*, 2019, pp. 10–12. doi: 10.1145/3304113.3326119.

[98] C. Stein and C. Stein, “Virtual Reality Design: How Head-Mounted Displays Change Design Paradigms of Virtual Reality Worlds,” *MediaTropes*, vol. 6, no. 1, pp. 52–85, 2016, doi: 10.1227/01.NEU.0000255485.65450.49.

[99] D. Krupke, L. Einig, E. Langbehn, J. Zhang, and F. Steinicke, “Immersive remote grasping: Realtime gripper control by a heterogenous robot control system,” in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 2016, vol. 02-04-Nove, pp. 337–338. doi: 10.1145/2993369.2996345.

[100] C. Crick, G. Jay, S. Osentoski, and O. C. Jenkins, “ROS and Rosbridge: Roboticists out of the Loop,” in *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2012, pp. 493–494. doi: 10.1145/2157689.2157846.

[101] D. Sun, A. Kiselev, Q. Liao, T. Stoyanov, and A. Loutfi, “A New Mixed-Reality-Based Teleoperation System for Telepresence and Maneuverability Enhancement,” *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 1, pp. 55–67, 2020, doi: 10.1109/THMS.2019.2960676.

[102] D. Wei, B. Huang, and Q. Li, “Multi-view merging for robot teleoperation with virtual reality,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8537–8544, 2021, doi: 10.1109/LRA.2021.3109348.

[103] Y. Luo, J. Wang, H.-N. Liang, S. Luo, and E. G. Lim, “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)*, 2021, pp. 418–425. doi: 10.1109/ro-man50785.2021.9515455.

[104] B. Fu *et al.*, “Towards virtualized welding: Visualization and monitoring of remote welding,” in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, 2014, pp. 1–6. doi: 10.1109/ICME.2014.6890297.

[105] S. Baklouti, G. Gallot, J. Viaud, and K. Subrin, “On the improvement of ros-based control for teleoperated yaskawa robots,” *Applied Sciences* , vol. 11, no. 16, pp. 7190–7203, Aug. 2021, doi: 10.3390/app11167190.

[106] B. Wang, S. J. Hu, L. Sun, and T. Freiheit, “Intelligent welding system technologies: State-of-the-art review and perspectives,” *Journal of Manufacturing Systems*, vol. 56, pp. 373–391, 2020, doi: https://doi.org/10.1016/j.jmsy.2020.06.020.

[107] Y. K. Liu, “Toward intelligent welding robots: virtualized welding based learning of human welder behaviors,” *Welding in the World*, vol. 60, no. 4, pp. 719–729, 2016, doi: 10.1007/s40194-016-0340-x.

[108] H. Ming, Y. S. Huat, W. Lin, and Z. Hui Bin, “On teleoperation of an arc welding robotic system,” *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2, no. May 1996, pp. 1275–1280, 1996, doi: 10.1109/robot.1996.506882.

[109] D. Ding *et al.*, “Towards an automated robotic arc-welding-based additive manufacturing system from CAD to finished part,” *CAD Computer Aided Design*, vol. 73, pp. 66–75, 2016, doi: 10.1016/j.cad.2015.12.003.

[110] J. van Essen *et al.*, “Identifying welding skills for robot assistance,” in *2008 IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications, MESA 2008*, 2008, pp. 437–442. doi: 10.1109/MESA.2008.4735675.

[111] M. S. Erden and A. Billard, “End-point impedance measurements across dominant and nondominant hands and robotic assistance with directional damping,” *IEEE Transactions on Cybernetics*, vol. 45, no. 6, pp. 1146–1157, 2015, doi: 10.1109/TCYB.2014.2346021.

[112] Y. K. Liu, Z. Shao, and Y. M. Zhang, “Learning human welder movement in pipe GTAW: A virtualized welding approach,” *Welding Journal*, vol. 93, no. 10, pp. 388–398, 2014.

[113] M. S. Erden and T. Tomiyama, “Identifying welding skills for training and assistance with robot,” *Science and Technology of Welding and Joining*, vol. 14, no. 6, pp. 523–532, 2009, doi: 10.1179/136217109X437150.

[114] Y. K. Liu and Y. M. Zhang, “Control of human arm movement in machine-human cooperative welding process,” *Control Engineering Practice*, vol. 32, pp. 161–171, 2014, doi: 10.1016/j.conengprac.2014.08.003.

[115] Y. Wang, Y. Chen, Z. Nan, and Y. Hu, “Study on Welder Training by Means of Haptic Guidance and Virtual Reality for Arc Welding,” in *2006 IEEE International Conference on Robotics and Biomimetics*, 2006, pp. 954–958. doi: 10.1109/ROBIO.2006.340349.

[116] O. Ciszak, J. Juszkiewicz, and M. Suszyński, “Programming of Industrial Robots Using the Recognition of Geometric Signs in Flexible Welding Process,” *Symmetry 2020, Vol. 12, Page 1429*, vol. 12, no. 9, p. 1429, Aug. 2020, doi: 10.3390/SYM12091429.

[117] H. Yu, J. Qin, and K. Zhao, “Innovation in Interactive Design of Tele-robotic Welding in the Trend of Interaction Change,” *Design Engineering*, pp. 322–330, Nov. 2020, doi: 10.17762/DE.VI.910.

[118] Q. Wang, W. Jiao, R. Yu, M. T. Johnson, and Y. M. Zhang, “Virtual Reality Robot-Assisted Welding Based on Human Intention Recognition,” *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 2, pp. 799–808, 2020, doi: 10.1109/TASE.2019.2945607.

[119] T. Wells and G. Miller, “The Effect of Virtual Reality Technology on Welding Skill Performance,” *Journal of Agricultural Education*, vol. 61, no. 1, pp. 152–171, 2020, doi: 10.5032/jae.2020.01152.

[120] A. P. Byrd, R. T. Stone, R. G. Anderson, and K. Woltjer, “The use of virtual welding simulators to evaluate experienced welders,” *Welding Journal*, vol. 94, no. 12, pp. 389–395, 2015.

[121] Y. Liu and Y. Zhang, “Human Welder 3-D Hand Movement Learning in Virtualized GTAW: Theory and Experiments,” *Transactions on Intelligent Welding Manufacturing*, vol. 1, no. 2, pp. 3–25, 2019.

[122] Y. K. Liu and Y. M. Zhang, “Supervised Learning of Human Welder Behaviors for Intelligent Robotic Welding,” *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 3, pp. 1532–1541, 2017, doi: 10.1109/TASE.2015.2453351.

[123] Y. K. Liu and Y. M. Zhang, “Fusing machine algorithm with welder intelligence for adaptive welding robots,” *Journal of Manufacturing Processes*, vol. 27, pp. 18–25, 2017, doi: 10.1016/j.jmapro.2017.03.015.

[124] Q. Wang, Y. Cheng, W. Jiao, M. T. Johnson, and Y. M. Zhang, “Virtual reality human-robot collaborative welding: A case study of weaving gas tungsten arc welding,” *Journal of Manufacturing Processes*, vol. 48, no. September, pp. 210–217, 2019, doi: 10.1016/j.jmapro.2019.10.016.

[125] Q. Wang, W. Jiao, R. Yu, M. T. Johnson, and Y. Zhang, “Modeling of Human Welders’ Operations in Virtual Reality Human–Robot Interaction,” *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2958–2964, 2019, doi: 10.1109/LRA.2019.2921928.

[126] D. Ni, A. W. W. Yew, S. K. Ong, and A. Y. C. Nee, “Haptic and visual augmented reality interface for programming welding robots,” *Advances in Manufacturing*, vol. 5, no. 3, pp. 191–198, 2017, doi: 10.1007/s40436-017-0184-7.

[127] M. Selvaggio, G. Notomista, F. Chen, B. Gao, F. Trapani, and D. Caldwell, “Enhancing bilateral teleoperation using camera-based online virtual fixtures generation,” *IEEE International Conference on Intelligent Robots and Systems*, vol. 2016-Novem, pp. 1483–1488, 2016, doi: 10.1109/IROS.2016.7759241.

[128] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S. M. Choi, “A review on mixed reality: Current trends, challenges and prospects,” *Applied Sciences (Switzerland)*, vol. 10, no. 2, pp. 636–662, 2020, doi: 10.3390/app10020636.

[129] M. M. Aygün, Y. Ç. Ögüt, H. Baysal, and Y. Taşcioglu, “Visuo-haptic mixed reality simulation using unbound handheld tools,” *Applied Sciences (Switzerland)*, vol. 10, no. 15, pp. 5344–5359, 2020, doi: 10.3390/APP10155344.

[130] M. Sharifi, “Enhanced Vision-Based Localization and Control for Navigation of Non-holonomic Omnidirectional Mobile Robots in GPS-Denied Environments,” 2017. [Online]. Available: https://ir.canterbury.ac.nz/bitstream/handle/10092/14660/Sharifi%2C Mostafa final PhD thesis.pdf?sequence=1

[131] Kevin Lynch and Frank Park, *Modern Robotics: Mechanics, planning, and Control, 1st ed.*, 1st ed. New York, NY, USA: Cambridge University Press, 2017.

[132] H. M. Becerra and C. Sagüés, *Visual Control of Wheeled Mobile Robots: Unifying Vision and Control in Generic Approaches*. Springer Verlag, 2014. doi: 10.1007/978-3-319-05783-5.

[133] Gerald Cook, *Mobile Robots: Navigation, control and Remote Sensing*. 2011. doi: 10.1002/9781118026403.

[134] Prusa Research, “TECHNICAL DATA SHEET Prusament PLA by Prusa Polymers,” 2018.

[135] K. Nie, W. Wan, and K. Harada, “A Hand Combining Two Simple Grippers to Pick Up and Arrange Objects for Assembly,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 958–965, Apr. 2019, doi: 10.1109/LRA.2019.2893153.

[136] “UR5 collaborative robot arm | Flexible and lightweight cobot.” https://www.universal-robots.com/products/ur5-robot/

[137] M. R. de Gier, “Control of a robotic arm: Application to on-surface 3D-printing,” 2015. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3Aa674a3fa-2534-44c4-b251-1e49a5194079

[138] Q. Liu, D. Yang, W. Hao, and Y. Wei, “Research on kinematic modeling and analysis methods of UR robot,” *Proceedings of 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference, ITOEC 2018*, pp. 159–164, Dec. 2018, doi: 10.1109/ITOEC.2018.8740681.

[139] J. J. Craig, “Introduction to Robotics: Mechanics and Control, 3rd Edition,” *Pearson*, 2004, [Online]. Available: https://www.mendeley.com/catalogue/43bfbab9-57d1-3dfe-93b8-1b0d7f43a6cc/

[140] A. Zelenak, M. W. Pryor, R. G. Reid, A. Pettinger, and M. Pryor, “Reactive Motion Control for Real-Time Teleoperation and Semi-Autonomous Contact Tasks,” *International Conference on Intelligent Robots and Systems. IROS*, 2019, doi: 10.13140/RG.2.2.18571.03362.

[141] A. Ghanbari, B. Horan, S. Nahavandi, X. Chen, and W. Wang, “Haptic microrobotic cell injection system,” *IEEE Systems Journal*, vol. 8, no. 2, pp. 371–383, 2014, doi: 10.1109/JSYST.2012.2206440.

[142] A. Franzluebbers and K. Johnsen, “Remote robotic arm teleoperation through virtual reality,” in *Proceedings - SUI 2019: ACM Conference on Spatial User Interaction*, Oct. 2019, pp. 1–2. doi: 10.1145/3357251.3359444.

[143] I. Kim, T. Kanno, K. Tadano, and K. Kawashima, “Research on a Master Manipulator Using an Isometric Interface for Translation in Robotic Surgery,” *International Journal of Advanced Robotic Systems*, vol. 12, no. 9, p. 128, Sep. 2015, doi: 10.5772/60074.

[144] D. Rakita, B. Mutlu, and M. Gleicher, “An Autonomous Dynamic Camera Method for Effective Remote Teleoperation,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 325–333, Feb. 2018, doi: 10.1145/3171221.3171279.

[145] Haiyang Jin, Liwei Zhang, S. Rockel, Jun Zhang, Ying Hu, and Jianwei Zhang, “A novel optical tracking based tele-control system for tabletop object manipulation tasks,” *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 636–642, 2015, doi: 10.1109/IROS.2015.7353439.

[146] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, “Analysis of the accuracy and robustness of the Leap Motion Controller,” *Sensors (Switzerland)*, vol. 13, no. 5, pp. 6380–6393, 2013, doi: 10.3390/s130506380.

[147] “Tracking | Leap Motion Controller | Ultraleap.” https://www.ultraleap.com/product/leap-motion-controller/

[148] A. Jarillo-Silva, O. A. Domínguez-Ramírez, V. Parra-Vega, and J. P. Ordaz-Oliver, “PHANToM OMNI haptic device: Kinematic and manipulability,” *CERMA 2009 - Electronics Robotics and Automotive Mechanics Conference*, pp. 193–198, 2009, doi: 10.1109/CERMA.2009.55.

[149] Y. K. Liu and Y. M. Zhang, “Toward Welding Robot With Human Knowledge: A Remotely-Controlled Approach,” *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 2, pp. 769–774, 2015, doi: 10.1109/TASE.2014.2359006.

[150] D. Rakita, B. Mutlu, and M. Gleicher, “A Motion Retargeting Method for Effective Mimicry-based Teleoperation of Robot Arms,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 361–370. doi: 10.1145/2909824.3020254.

[151] K. Huang, D. Chitrakar, F. Rydén, and H. J. Chizeck, “Evaluation of haptic guidance virtual fixtures and 3D visualization methods in telemanipulation—a user study,” *Intelligent Service Robotics*, vol. 12, no. 4, pp. 289–301, Oct. 2019, doi: 10.1007/s11370-019-00283-w.

[152] S. Livatino *et al.*, “Intuitive Robot Teleoperation through Multi-Sensor Informed Mixed Reality Visual Aids,” *IEEE Access*, vol. 9, pp. 25795–25808, 2021, doi: 10.1109/ACCESS.2021.3057808.

[153] B. P. Vagvolgyi *et al.*, “Scene Modeling and Augmented Virtuality Interface for Telerobotic Satellite Servicing,” *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4241–4248, Oct. 2018, doi: 10.1109/LRA.2018.2864358.

[154] B. Vagvolgyi, W. Niu, Z. Chen, P. Wilkening, and P. Kazanzides, “Augmented virtuality for model-based teleoperation,” in *IEEE International Conference on Intelligent Robots and Systems*, 2017, vol. 2017-Septe, pp. 3826–3833. doi: 10.1109/IROS.2017.8206233.

[155] K. Mikko, N. Mikko, and O. Pirkko, “Method for measuring stereo camera depth accuracy based on stereoscopic vision,” *Three-Dimensional Imaging, Interaction, and Measurement*, vol. 7864, pp. 168–176, 2011, doi: 10.1117/12.872015.

[156] K. Cho, K. Ko, H. Shim, and I. Jang, “Development of VR visualization system including deep learning architecture for improving teleoperability,” in *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2017*, 2017, pp. 462–464. doi: 10.1109/URAI.2017.7992776.

[157] J. Han, K. Cho, I. Jang, C. Ju, H. il Son, and G.-H. Yang, “Development of a shared controller for obstacle avoidance in a teleoperation system,” *International Journal of Control, Automation and Systems*, vol. 18, no. 11, pp. 2974–2982, Nov. 2020, doi: 10.1007/s12555-019-0410-0.

[158] F. de Pace, G. Gorjup, H. Bai, A. Sanna, M. Liarokapis, and M. Billinghurst, “Assessing the Suitability and Effectiveness of Mixed Reality Interfaces for Accurate Robot Teleoperation,” *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, pp. 7–9, 2020, doi: 10.1145/3385956.3422092.

[159] X. Tu, J. Autiosalo, A. Jadid, K. Tammi, and G. Klinker, “A Mixed Reality Interface for a Digital Twin Based Crane,” *Applied Sciences*, vol. 11, no. 20, p. 9480, Oct. 2021, doi: 10.3390/APP11209480.

[160] M. M. Marinho, H. Ishida, K. Harada, K. Deie, and M. Mitsuishi, “Virtual Fixture Assistance for Suturing in Robot-Aided Pediatric Endoscopic Surgery,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 524–531, 2020, doi: 10.1109/LRA.2019.2963642.

[161] Z. Wang and A. M. Fey, “Human-centric predictive model of task difficulty for human-in-the-loop control tasks,” *PLoS ONE*, vol. 13, no. 4, pp. 1–22, 2018, doi: 10.1371/journal.pone.0195053.

[162] A. Tavakkoli, B. Wilson, and M. Bounds, “An Immersive Virtual Environment for Teleoperation of Remote Robotic Agents for Everyday Applications in Prohibitive Environments,” *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VRW 2020*, pp. 371–375, 2020, doi: 10.1109/VRW50115.2020.00080.

[163] H. Dybvik, M. Løland, A. Gerstenberg, K. B. Slåttsveen, and M. Steinert, “A low-cost predictive display for teleoperation: Investigating effects on human performance and workload,” *International Journal of Human Computer Studies*, vol. 145, no. July 2019, pp. 102536–102554, 2021, doi: 10.1016/j.ijhcs.2020.102536.

[164] Z. Chen, F. Huang, W. Sun, and W. Song, “An Improved Wave-Variable Based Four-Channel Control Design in Bilateral Teleoperation System for Time-Delay Compensation,” *IEEE Access*, vol. 6, pp. 12848–12857, 2018, doi: 10.1109/ACCESS.2018.2805782.

[165] J. Guo, C. Liu, and P. Poignet, “A Scaled Bilateral Teleoperation System for Robotic-Assisted Surgery with Time Delay,” *Journal of Intelligent & Robotic Systems*, vol. 95, no. 1, pp. 165–192, 2019, doi: 10.1007/s10846-018-0918-1.