IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

Digital Twin-Driven Mixed Reality Framework for Immersive Teleoperation with Haptic Rendering

Wen Fan*, Xiaoqing Guo*, Enyang Feng*, Jialin Lin*, Yuanyi Wang, Jiaming Liang, Martin Garrad, Jonathan Rossiter, Zhengyou Zhang, Nathan Lepora, Lei Wei, Dandan Zhang*

***Abstract*—Teleoperation has widely contributed to many applications. Consequently, the design of intuitive and ergonomic control interfaces for teleoperation has become crucial. The rapid advancement of Mixed Reality (MR) has yielded tangible benefits in human-robot interaction. MR provides an immersive environment for interacting with robots, effectively reducing the mental and physical workload of operators during teleoperation. Additionally, the incorporation of haptic rendering, including kinaesthetic and tactile rendering, could further amplify the intuitiveness and efficiency of MR-based immersive teleoperation. In this study, we developed an immersive, bilateral teleoperation system, integrating Digital Twin-driven Mixed Reality (DTMR) manipulation with haptic rendering. This system comprises a commercial remote controller with a kinaesthetic rendering feature and a wearable cost-effective tactile rendering interface, called the Soft Pneumatic Tactile Array (SPTA). We carried out two user studies to assess the system's effectiveness, including a performance evaluation of key components within DTMR and a quantitative assessment of the newly developed SPTA. The results demonstrate an enhancement in both the human-robot interaction experience and teleoperation performance. For more project details, please view our website: https://sites.google.com/view/hbts-brl/home**

***Index Terms*—Teleoperation, Haptic Rendering, Mixed Reality.**

I. INTRODUCTION

TELEOPERATION has been widely used in many applications, such as robotic surgery, remote space or ocean exploration, and remote manipulation in hazardous or contaminated environments [1]–[3]. Most teleoperation systems adopt the leader-follower control paradigm, in which, the human operator serves as the leader, generating control commands through motion tracking devices, while the robot working in the remote area replicates the same motions [4]–[6]. Traditional teleoperation systems typically provide visual feedback to the operator but often lack adequate tactile and force information [7]. Therefore, a bilateral teleoperation system is desirable as it offers the operator the capacity to manipulate remote objects and receive force/tactile feedback through the control interface [8]. However, few bilateral teleoperation systems incorporate tactile rendering. Tactile feedback devices can replicate the sensation of shape, texture, and other contact information, giving the operator a better understanding of the manipulation task. Therefore, we aim to integrate low-cost tactile feedback devices into a commercial motion caption device to provide tactile feedback at the operator's fingertip, working in parallel with kinaesthetic rendering to support intuitive bilateral teleoperation.

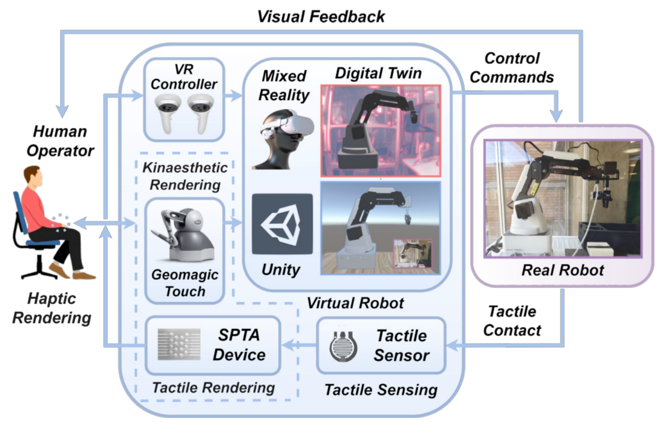

Fig. 1. The architecture of the Digital Twin-driven Mixed Reality (DTMR) framework for immersive manipulation with haptic rendering, combining kinaesthetic rendering through Geomagic Touch and tactile rendering through Soft Pneumatic Tactile Array (SPTA).

The design of intuitive control interfaces is essential for teleoperation, especially for complex manipulation tasks in unstructured, dynamic environments [9]. Recently, Mixed Reality (MR) has brought tangible benefits to human-robot interaction. It combines the real and virtual worlds, allowing operators to interact with both physical and virtual objects in real-time. MR enables human-robot interaction to be achieved in an immersive manner [10]. However, traditional MR systems for teleoperation suffer from the latency of visual feedback, and the lack of both kinaesthetic and tactile rendering. Digital twins (DT), which can provide real-time visualization of both physical and virtual robots, could be introduced into MR, reducing the latency of visual feedback [11]. Also, the virtual robot of DT can be customized freely to an appropriate view angle based on user preference, enhancing system ergonomics. Furthermore, it is crucial to stimulate human senses, such as touch, which enrich the virtual experience through comprehensive sensory feedback. Therefore, integrating haptic rendering devices with MR is also significant, as it can create more cues to enhance both the intuitiveness and efficiency of teleoperation.

Manuscript received: May, 10, 2023; Revised July, 17, 2023; Accepted October, 08, 2023. This paper was recommended for publication by Editor Angelika Peer upon evaluation of the Associate Editor and Reviewers' comments.

* Equal Contribution. W. Fan, X. Guo, E. Feng, J. Lin, M. Garrad, J. Rossiter, N. Lepora are with Bristol Robotics Laboratory, University of Bristol. J. Liang, Z. Zhang are with Tencent RoboticsX, Tencent, Shenzhen, China. Y. Wang is with Department of Mechanical and Aerospace Engineering, Hong Kong University of Science and Technology. L. Wei is with Institute for Intelligent Systems Research and Innovation, Deakin University. D. Zhang is with Department of Bioengineering, I-X Initiative, Imperial College London.

Digital Object Identifier (DOI): see top of this page.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

To this end, we construct a framework by combining MR for immersive manipulation and haptic rendering for bilateral teleoperation. The framework overview is shown in Fig. 1. The **contributions** of this paper include:

1) Integrating the DT technique into an MR-based immersive teleoperation system, leading to the development of the Digital Twin-driven Mixed Reality (DTMR) framework.

2) Building a novel haptic rendering system based on i) pneumatically-driven actuators for tactile feedback on human fingertips, and ii) a commercial haptic controller that enables kinaesthetic rendering and real-time motion tracking.

3) Constructing a seamless bilateral teleoperation framework through the integration of immersive manipulation via DTMR and the novel haptic rendering device, which includes both kinaesthetic and tactile rendering.

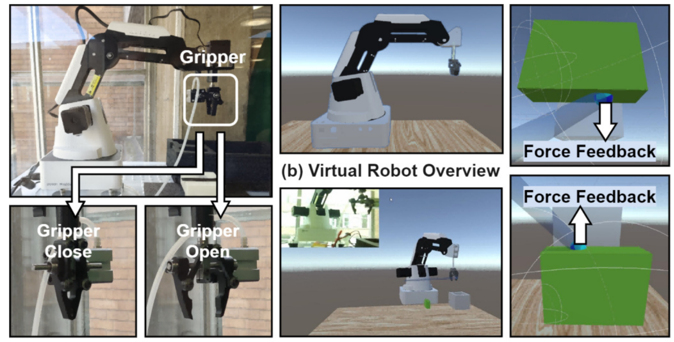

Fig. 2. Illustration of the digital twin and the kinaesthetic rendering mechanism. (a) Overview of the low-cost desktop robot (Dobot Magician). (b) The initial 3D rendering of the robot in the Unity environment. (c) A unified view showing both the real-world robot and its corresponding digital twin in Unity. (d) Illustration of the kinaesthetic rendering mechanism.

II. Related Work

Haptic Rendering Devices for Teleoperation: Traditional haptic controllers normally rely on mechanical linkage-based designs, such as exoskeleton controllers or desktop devices with rigid rods to realize motion capture for teleoperation [12]. Most of these devices use motors to provide force and torque feedback. CyberGrasp (CyberGlove Systems) [13], Dexmo (Dexta Robotics), and VRgluv Enterprise (VRgluv), are wearable kinaesthetic rendering exoskeletons enabled by electromagnetic actuators. Other examples can be found in the review paper [14]. Most of the devices mentioned above only provide kinaesthetic rendering but not tactile rendering. Tactile rendering devices make use of the high tactile sensitivity of fingertips or hands [15]. Gloveone, AvatarVR, and Senso Glove [16] utilize glove-like designs. They incorporate hand-tracking functions and provide tactile feedback using electromagnetic actuators. Other wearable motion-tracking devices with tactile rendering have been introduced and summarised in [17]. However, few devices can provide both tactile and kinaesthetic rendering in MR applications. Although the HaptX Glove (HaptX) can provide both kinaesthetic and tactile rendering [18], the expensive price limits its wide applications. Moreover, it is quite heavy and may impose an additional physical burden on the operator. To this end, we aim to develop a low-cost solution to achieve the target of haptic rendering with both kinaesthetic and tactile rendering during bilateral teleoperation.

MR for Teleoperation: MR has been used for immersive and intuitive teleoperation of mobile robots [19], [20], a quadruped-manipulator robot [21], and has also been investigated for teleoperating a hyper-redundant robot for inspection tasks [22]. A new MR-based teleoperation system has been developed with maneuverability enhancement, where the virtual environment is augmented with real-time data obtained from the task space. A new interaction proxy is used for robot control while two control modes were developed for long-distance movements and fine movements respectively, which aims to decouple the operator from the control loop and reduce the cognitive burden [23]. However, neither kinaesthetic nor tactile rendering is available. A mixed-reality virtual fixture (MRVF) framework has been proposed for robotic welding, which ensures immersive and intuitive human-robot interaction with visual and haptic cues [24]. However, this framework can only provide kinaesthetic rendering to users when tools approximate the safety boundary but still lack tactile rendering. Though MR has been applied to different types of teleoperation applications, the attempts made to incorporate affordable tactile and kinaesthetic rendering devices into the MR-based immersive teleoperation system remain limited.

III. Methodology

A. Hardware System

In this paper, to provide an immersive teleoperation experience for human operators, MR will be investigated with the support of DT technology for teleoperation. Moreover, haptic rendering, including both kinaesthetic rendering and tactile rendering will be integrated to facilitate bilateral teleoperation with user perception enhancement. We conduct experiments using a low-cost desktop robot, the Dobot Magician (see Fig. 2 a), as an example. This robot has 4 degrees of freedom, while the precision for control is 0.2 mm. We employ a head-mounted Virtual Reality (VR) device, the Meta Quest 2, to construct the immersive teleoperation interface. Equipped with four infrared cameras, this VR device enables the user to simultaneously observe real-world objects and the virtual robot, facilitating the MR implementation. As for motion tracking and kinaesthetic rendering, Geomagic Touch (3D System) is used for teleoperation. Users can hold its stylus and generate motions, which can be mapped on the robot's end-effector for real-time control.

Force-sensitive resistors (FSRs) are a type of sensor widely used in robotic force/tactile sensing due to their low cost and ease of integration. They function by exhibiting a change in electrical resistance. As the applied force increases, the resistance decreases, facilitating measurements of contact pressure. We employed FSR sensors on the robot's end-effector to enable the detection of contact force. To achieve customizable mapping, the spatial resolution of the FSR sensor array and the tactile feedback array do not need to strictly coincide. The grasping force is relayed to the operator through the actuation of the tactile array. We can tailor the mapping relationship between the force detected by the robot and the intensity of the

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

tactile feedback delivered to the user via the tactile rendering device. For more complex scenarios, we can arrange FSR sensors in an array to improve the spatial resolution of tactile perception. This enhancement would allow a wider variety of tactile signals to be relayed to human operators via the tactile interfaces, thereby augmenting their haptic feedback experience during operation.

B. Digital Twin-driven Mixed Reality (DTMR) Framework

Unity, a game engine with powerful rendering capabilities, is utilized as our development software for constructing the MR environment. It ensures that the simulated robot exhibits high fidelity. Unity functions as a 3D simulator that mirrors a real robot, allowing the DT concept to be realized. The simulated robot (see Fig. 2 b) will be displayed when users wear the VR head-mounted display (HMD). To improve the ergonomics of teleoperation, the position and orientation of the virtual robot can be adjusted based on the users' preferences.

The DT transmits the control commands from the human operator to the real manipulator, while the physical and virtual scenes are displayed to the operator simultaneously, as shown in Fig. 2 c. The data transmission of the DT is implemented by a WebSocket. The commonly used polling protocol requires the client to check messages from the server all the time but a WebSocket allows the client to send and receive messages only when needed. This feature reduces the frequency of transmissions and improves data exchange efficiency significantly. The WebSocket client is set for the virtual robot arm and the client code is written in Unity. Meanwhile, the control commands are generated in the JSON data format and sent to the server for the physical robot control. Also, the physical robot states will be sent to the virtual robot to match each other in real-time. The integration of the virtual robot with the real-world scene observed by the cameras leads to the construction of the DTMR framework, aiming to provide operators with natural human-robot interaction and immersive visualization. Fig. 2 d illustrates the kinaesthetic rendering mechanism. Force feedback will be generated on the operator's hand once the virtual robot reaches the boundary of the customized workspace or interacts with a virtual object.

C. Soft Pneumatic Tactile Array (SPTA) for Tactile Feedback

Kinaesthetic and tactile renderings are both necessary to reconstruct the sensing of remote environments for users. Though the haptic controller (Geomagic Touch) can provide kinaesthetic rendering, it cannot rebuild the tactile information of the gripped objects. Therefore, as a supplement, the SPTA device was developed. Here, we describe the system integration and working mechanism of the SPTA. Initially, tactile sensors positioned on the robot's end-effector can convert contact pressure information into digital signals using a microcontroller. Subsequently, the tactile rendering system decodes these signals, transforming them into various pressures to actuate the tactile array, as shown in Fig. 3. The SPTA is characterized by its affordability since the overall system costs about £550, including £0.25 to £30 for the single array fabrication which depends on different lamination mold sizes, £40 for the electrical control system, and £500 for the pneumatic control system. The overall price can vary from £300 to £700 due to the selection of different pumps, valves, and electrical components.

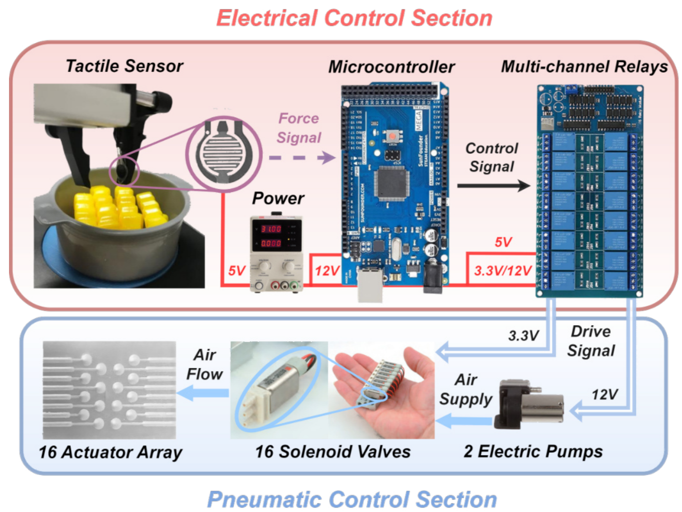

Fig. 3. The hardware structure of the SPTA device is designed such that, upon receiving commands from the microcontroller, the multi-channel relays power the pumps and valves. This mechanism facilitates the provision of regulated pressure to the SPTA's expandable air pockets via air tubes.

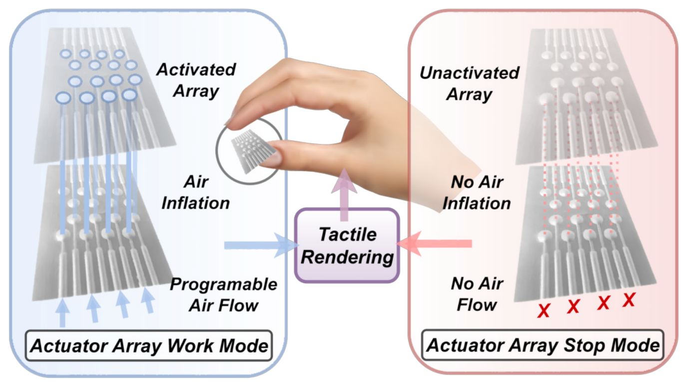

Fig. 4. SPTA actuator array working diagram. The 16 pneumatic actuators can be activated into various combinations, such as different numbers and shapes, aiming to provide the tactile rendering of texture features for users.

1) Soft Pneumatic Tactile Array: The Soft Pneumatic Tactile Array (SPTA) is constructed from Thermoplastic Polyurethanes (TPU), a material known for its wearable properties. As depicted in Fig. 4, two sheets of TPU film form a two-dimensional pattern of 16 tactile actuators by using thin film lamination technology. Each of these actuators serves as an expandable air pocket with a cylindrical chamber, the diameter of which is 3 mm. The overall dimension of SPTA is 30mm-by-20mm with a rectangular shape, well matching the size of humans' fingertips. When air flows into the actuator chamber, the elastic deformation of the deforming membrane can provide a tactile sensation. In addition, the compliant nature of SPTA provides conformity with the fingertip surface to ensure realistic sensation. The number of expandable air pockets can be tailored to meet the demands of various applications, and their chamber diameter can also be customized. By activating different numbers of actuators, we can convey varying levels of tactile rendering intensity, providing

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

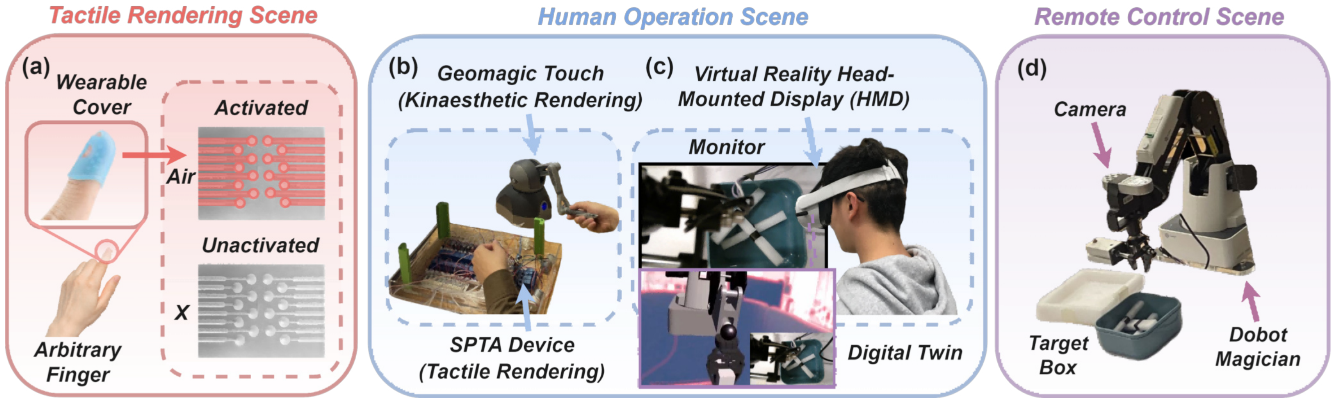

Fig. 5. The experimental setup is illustrated as follows: (a) shows the mode of operation for tactile rendering; (b) presents the Geomagic Touch and SPTA device; and (c) depicts the digital twin as seen during human operation. Experimental setups for user study are shown in (d). Given that our proposed SPTA device is wearable, users have the flexibility to attach the device to any finger they prefer.

information about the features of the object being grasped.

2) Electrical Control: The input signal modulation of the tactile/force sensor and the output signal modulation used to control the multi-channel relays are completed in a microcontroller unit (Arduino Mega 2560 R3). The 16-channel relays are used to control solenoid valves and extra two air pumps.

3) Pneumatic Control: Two DC mini air pumps ($12V$) act as the pneumatic power source for the actuation system, while 16 three-port solenoid valves ($3.3V$) control the direction of air flows. Each channel of the air tube is linked to an air pocket (tactile actuator) of the SPTA. When pressurized air is injected, the air pocket will be inflated. Such tactile actuators can rapidly inflate and deflate, with the tactile feedback response time falling within 200 ms. The speed of airflow, and thus the rate of air pocket inflation, can be adjusted by the working voltage of the pump. This effectively regulates the delay induced by the pumps. However, for situations that require high-frequency feedback, the pump delay could potentially constrain the response speed. This limitation can be mitigated by optimizing the pneumatic control algorithms.

IV. Experiments and Result Analysis

A. Experimental Setup and Participants

The experimental setup is depicted in Fig. 5. We initially assessed the key components of the proposed system and subsequently conducted user studies to quantify the improvements offered by the use of the SPTA device. A total of 10 participants, consisting of 2 females and 8 males (aged 26-36), were invited to participate in the user studies. Following their interaction with the system, all participants were asked to complete questionnaires that assessed their user experience regarding the system's performance. Additionally, a subset of 8 participants undertook a second user study. This additional study involved a quantitative analysis based on an object transportation task, allowing us to gain more detailed insights about the system. Regarding the participants' background, 6 individuals reported having prior experience with playing computer games, which may contribute to their adaptability to our system. Among them, 2 had experience with VR devices.

B. System Evaluation

We evaluated the key components of the proposed framework, including i) the DTMR teleoperation framework with two ablation studies, and ii) the SPTA device.

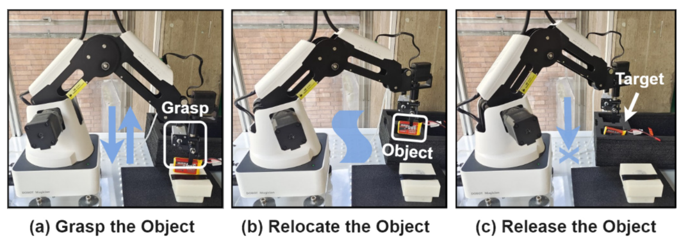

Fig. 6. The procedures of the pick-and-place task. (a) grasp the object; (b) relocate the object; (c) release the object.

1) Evaluation of the DTMR Teleoperation Framework: To evaluate the performance of the DTMR teleoperation framework, we conducted two ablation studies for the haptic controller and the DTMR respectively, which represent the fundamental elements of our bilateral teleoperation. Fig. 6 illustrates the procedures of the pick-and-place task employed in the two ablation studies. Before each test starts, the robot is set at an initial state. The operator first commands the robot end-effector to grasp the object, then relocates it above the target box. Finally, the object should be placed inside the box. In the first ablation study, participants were instructed to complete the pick-and-place task both with and without a haptic controller. In the second ablation study, the DT was alternately enabled and disabled for comparative purposes, with the incorporation of the haptic condition. Each participant received a demonstration of the teleoperation system before the user studies commenced. Subsequently, they were required to repeatedly perform the pick-and-place task to familiarize themselves with the DTMR system. Participants were also asked to describe their experience and complete a subjective feedback questionnaire, which addressed both ablation studies within a single survey. The questions for the subjective questionnaire are listed as follows:

- Q1: Which one is easier to use, Geomagic Touch (haptic controller) or VR controller (non-haptic controller)?
- Q2: Did the haptic controller improve your confidence in task completion?

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

- Q3: How useful is the DTMR interface?
- Q4: Why do you feel the DTMR interface is helpful?

The first question asked users to select the more user-friendly interaction mode. 80% of the participants preferred the haptic controller, as it provides force feedback to prevent collisions and enables continuous control of the robot end-effector for dexterous manipulation. In contrast, the VR controller relies solely on the fingers to generate control commands, utilizing joysticks and buttons both of which have a low degree of freedom. This can make it challenging for users to generate the desired smooth control path in the 3D pick-and-place task.

The second question asked participants to select their preferred control mode based on their confidence in the task's success rate. 60% of the participants preferred the VR controller as a slower control rhythm is easier to master. However, when considering the penalty for touching boundaries and time constraints, all participants favored the haptic controller, which ensures more reliable performance through kinaesthetic rendering and manipulation efficiency.

The third question requires the user to indicate whether the DTMR framework is useful to improve performance or not. In traditional teleoperation systems, operators monitor the remote scene through video streaming. With the DTMR system, operators can observe the virtual robot overlapping with the real robot in the VR HMD. As a result, 80% of participants commented that the DTMR system gave an immersive manipulation experience, helping them build up confidence in complex manipulation tasks, while the last 20% found it hard to get used to the infrared cameras on HMD. These users preferred to view real-world scenarios through RGB cameras, which can be realized by other types of VR headsets such as Meta Quest Pro and Pico 4.

In the fourth question, participants were asked to explain why they felt the DTMR framework aided in improving their performance during teleoperation. Most of them pointed out that the DTMR framework provided a better 3D perception of the operating scene. The virtual robot can generate instant feedback when operators give control commands, acting as the reference of the real robot state in real time. The virtual robot's relative position can be adjusted based on user preference. A virtual target is also placed in the VR environment, enabling users to assess the relevant states between the physical robot and the real object more intuitively. Thus, the DTMR framework allows human operators to perform teleoperation with enhanced ergonomics, efficiency, and intuitiveness.

In summary, the haptic controller achieved a success rate of 40% with an average time cost of 19.6 seconds, while in contrast, the non-haptic controller required 35.4 seconds. With the assistance of the DTMR system, these two performances improved to 16.9 seconds and 32.5 seconds, respectively. However, we also observed that most grasping failures were due to the operator's inability to accurately judge the object's state during the grasping process, leading to unsuccessful grasps or weak grasps causing the object to slip. These results highlight the significance of our SPTA device, which provides tactile feedback to assist users in making such judgments.

2) Evaluation of the SPTA Device for Tactile Rendering: In an MR-based bilateral teleoperation system, it's important

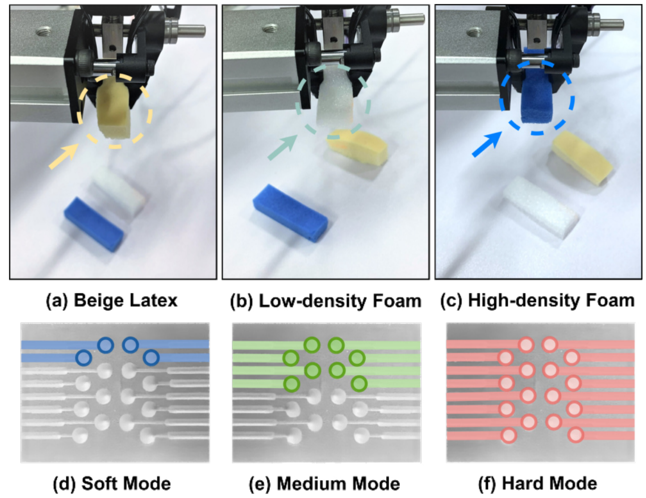

Fig. 7. Illustration of the samples with different stiffness. (a) beige latex; (b) low-density foam, which is stiffer than latex; (c) high-density foam, which is the stiffest among the three materials. To distinguish between different materials, the tactile array is configured with three activation modes, including (d) soft mode, with 4 arrays activated; (e) medium mode, with 8 arrays activated; and (f) hard mode, where all 16 arrays are activated.

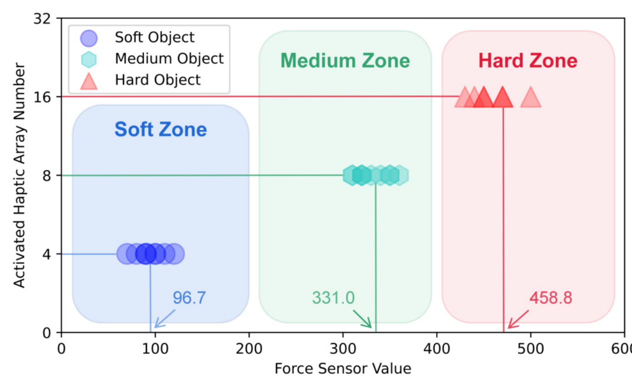

Fig. 8. The experimental results for stiffness estimation demonstrate that the values from the force sensor can be converted into measures of stiffness. We have established soft, medium, and hard classifications to aid in material identification, which in turn assists in tactile rendering via the SPTA device.

for operators to perceive the material properties (such as the stiffness) of the target object through a tactile feedback device. This can ensure the operators do not exert too much force on delicate objects. To this end, the main objective of this experiment is to evaluate the usability of the SPTA device and examine whether it can generate adequate tactile stimuli for users when grasping objects of varying stiffness.

As shown in Fig.5 a, we incorporated the tactile array into an adjustable wearable cover, which can be conveniently fitted to the user's most sensitive finger. This significantly enhances the adaptability of the SPTA device, allowing for easy extension to multiple fingers on any hand or other controller. Furthermore, the layout of the tactile array and wearable cover is designed to match the shape of the human finger, adhering to ergonomic principles.

A crucial aspect of our investigation is to evaluate whether a human operator, guided by the tactile rendering device, could effectively discern differences in stiffness among various objects during the process of grasping. We selected three test samples, each demonstrating unique stiffness properties:

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

beige latex, low-density foam, and high-density foam, aiming to represent a broad range of stiffness characteristics and to provide a comprehensive evaluation of both the robot's sensory responses and the human operator's ability to differentiate these variations in stiffness.

When contact occurs, the sample starts to deform and meanwhile, the force sensor generates digital signals. The beige latex is soft, with the least stiffness. Low-density foam is stiffer than latex since it is made of two layers of white foam sheets. The high-density foam consists of three blue-colored layers and thus is the stiffest one among them. To ensure fair comparisons, all samples have the same size of $1cm \times 1cm \times 5cm$, as shown in Fig. 7 a-c. Once the robotic gripper successfully grasps the target object, the force sensors mounted on the tips can help estimate the object's stiffness.

TABLE I. Relationship between force sensor value U_{Analog} and stiffness k

U_{Analog}	100	200	300	400	500
$k(N/m)$	80	210	600	2000	4000

Three threshold areas have been established to meet these identification requirements: the soft, medium, and hard zones. These settings were determined by the experimenters based on preliminary trials. These areas then help generate corresponding tactile feedback via the SPTA device. More specifically, each region corresponds to a specific number of air pockets activated in the SPTA, namely 4, 8, and 16. Consequently, users can directly perceive the stiffness of the object they are grasping through tactile feedback.

During user studies, the robot picked up different objects, and users were required to classify the stiffness based on the feedback from the SPTA device. The force sensor values U_{Analog} and perceived stiffness k were derived from preliminary trials and statistical analysis to map the U_{Analog} to k. We use the output curve of the force sensor (which describes the relationship between the analog voltage output U_{Analog} and force value) in combination with the measurements of the dimensions of the object's deformation to determine the stiffness range k for each material. We collected ten sets of data for a quantitative analysis of SPTA performance, as summarized in Fig. 8. The results indicate that when grasping samples of different material types, U_{Analog} values fall within the three threshold areas. These areas' mean values and variances are 96.7±16, 331.0±21, and 458.8±41. According to Table. I, the stiffness k of the three materials is approximately $0-600N/m$, $600-2000N/m$, and $2000-4000N/m$ respectively. The stiffer the material of the object, the greater the number of expandable air pockets in the SPTA device that are required to be activated. The experimental results validate that the SPTA device can discern the relative stiffness of objects.

C. User study based on an Object Transportation Task

1) Experiment Design: Here, quantitative comparisons were made between scenarios with and without tactile rendering, which is known to be the most critical factor for enhancing teleoperation performance. Given that the advantages of using the haptic controller were verified in the DTMR framework's performance evaluation, we assessed the SPTA function in this user study by comparing the teleoperation performance when the SPTA device was either disabled or enabled. Participants were asked to conduct a multiple-object transportation task that involved 3 stages:

- Stage 1: In this initial stage, the operator should identify the object's location. The robot should then be manoeuvered to gradually descend towards the object, ensuring that it doesn't come into contact with the plate.
- Stage 2: Once the object is within reach, the operator should close the gripper and remain attentive to the real-time sensory feedback. If the SPTA fails to activate, it means the grip has not been successful, and the operator should revert to Stage 1 to retry the process.
- Stage 3: Upon successful activation of the SPTA, indicating a valid grip, the operator should guide the robot to transfer the target objects from the white plate into the designated box one at a time, repeating until all objects have been relocated.

An affordable desktop robot (Dobot Magician) was used in this experiment. The distance between the object's initial position and the empty box is set to 20 cm. Although similar to the previous pick-and-place task, it presents additional challenges due to view occlusions and the low color contrast between objects and plates. To ensure each user was familiar with the tactile rendering function in this task, all participants were required to be involved in a training session before user studies began. More specifically, they would familiarize themselves with both the SPTA hardware and software.

After such preparation, each participant conducted four trials one by one, two with tactile rendering and two without tactile rendering. To circumvent any potential order effects, the sequence of the four trials was determined randomly. Each trial included 3 instances of pick-and-place operations. Therefore, every operator was required to complete 12 operations in total. Consequently with 8 operators, our data collection comprised 32 trials which were split 16/16 between with and without SPTA, encompassing a total of 96 trials of the object transportation task. The values of the six predefined evaluation metrics were recorded for quantitative evaluation: i) time cost, ii) grip times, iii) empty times, iv) slipped times, v) bottom times, and vi) success rate. Participants were required to follow two rules when conducting the task:

- Rule 1: Each participant needs to transport all objects into the target box within three minutes (180s).
- Rule 2: Each participant should prevent the gripper from touching the bottom plate during experiments.

2) Results Analysis: After the formal test was finished, all six evaluation metrics for each user of their four experiment trials were summarised and averaged for the following analysis. To quantify the effectiveness of the SPTA tactile rendering function, we evaluate the level of improvement between performances with and without SPTA, whose details and comparisons are summarised in Table. II and Fig. 9.

The evaluation metrics are defined as follows:

(1) **Time Cost (TC)**: This metric represents the duration taken by a participant to complete all stages of a multi-object

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

transportation task. If this exceeds 3 minutes, the participants cannot continue the task (Rule 1). According to the experimental results, with tactile rendering, the average time cost for participants to finish the task was 110.5s, and all people finished the task within the time limitation. However, without SPTA, the average time cost increases to 143.1s. Among all the trials, three users exceeded the 180s limit.

(2) **Grip Times (GT)**: This metric is used to describe the total number of gripping operations performed by each participant in a single trial. Ideally, the participant would need only three gripping operations to transport three objects. However, without the SPTA device, participants had to attempt more gripping operations before successfully grasping the target. To compare, the operators enjoy higher confidence in terms of their judgment of successful gripping operations when using the SPTA device. The average attempts were reduced from 9 to 7 when applying SPTA.

(3) **Empty Times (ET)**: This metric describes the situation when the participants control the robot to lift up its end-effector but without successfully picking up the target object. Without haptic rendering, the participants failed to lift the object with an average of 5 times. This number is reduced when using tactile rendering since the operators receive a better perception of the target object state through the contact information provided by the SPTA device.

(4) **Slipped Times (ST)**: This metric indicates the number of slipped times during the task. If the object is grasped unstably, it may drop out of the gripper in the transportation process. From the experimental results, the use of tactile rendering helps to reduce the slipped times by two-thirds.

(5) **Bottom Times (BT)**: This metric represents the number of times that the end-effector touches the plate during experiments. To keep the test safe, participants were required not to touch the plate bottom during operation (Rule 2), but carefully judge the distance between the end-effector and the plate. Results show that users were almost twice as likely to break this rule without using SPTA than with the tactile device.

(6) **Success Rate (SR)**: This metric indicates the ratio of the final number of successful object transportation divided by the total number of gripping operations for each trial. With the assistance of the SPTA tactile rendering, the success rate was nearly one in two, but without the SPTA device, the success rate dropped to two in five.

TABLE II. Experiment Results in Comparision for the Second User Study (Average across 8 Participants)

	TC	GT	ET	ST	BT	SR
Without	143.1s	8.43	5.19	0.56	1.44	39.59%
With	110.5s	6.88	1.38	0.19	0.81	48.54%
p-value	**0.0052**	0.2079	**0.0010**	0.1091	0.1654	0.0992

As shown in Fig. 9, based on the evaluation metric of enhancement, the performance of task execution is much improved when using the SPTA device. The average 'time cost' and the number of 'grip times' were both reduced by about one-fifth. Also, the 'empty times' and 'slipped times' can be reduced by 73.49% and 66.67% respectively, indicating the benefits to the robustness and stability. Moreover, the metric of 'bottom times' was reduced by nearly 50%, which demonstrates that tactile rendering can enhance safety during teleoperation. Using the SPTA device, the success rate for all participants is improved by about 9%. This is due to the fact that tactile rendering is necessary for contact-rich tasks. It can bring extra sensing information to the operators and serves as a significant supplementary feature for visual feedback.

Prior to undertaking any further statistical analyses, we performed Shapiro-Wilk tests to assess normality, setting the significance level at 0.05. The data collected for two evaluation metrics (time cost and success rate demonstrated non-parametric characteristics, with p-values standing at 0.0749 and 0.6006 respectively. The results for the remaining metrics conformed to the assumption of a normal distribution. The user study used a within-subject design, with all participants completing two rounds of control modes. As a result, we performed Wilcoxon signed-rank tests for a non-parametric comparison between the variables (time cost and success rate), and T-tests for the other metrics to establish the statistical differences between the two control modes. We considered a p-value of less than 0.05 to be significant. The p-values for the different evaluation metrics are displayed in Table II. According to these results, the differences in time cost and the number of empty times metrics between with and without SPTA are statistically significant.

The evident advantage of SPTA is avoiding both empty and unstable grip operations. Participants do not need to waste extra time lifting up the robotic gripper if it fails to grasp the target object. In addition to Table. II, we separately recorded the number when the SPTA helped identify the failure of grasping with an average of 2.48 times during each trial. This is the main reason behind the significant differences in time cost (TC) and the number of empty times (ET) in the performance with and without enabling the SPTA device. Moreover, the unstable tactile signal can be an early warning for the operators to adjust the grasp pose, which can help operators increase the success rate of object transportation.

In summary, results in this user study indicated that human operators could finish the task with higher efficiency and better performance when using tactile rendering since it provides touch information with the target object as the supplementary feature beyond the visual display.

V. Conclusions and Future Work

In this paper, we constructed a DTMR framework for immersive teleoperation. A tactile rendering device, SPTA, is developed to cooperate with a commercial haptic controller, providing tactile and kinaesthetic rendering respectively. The combination of the DTMR and haptic rendering facilitates immersive bilateral teleoperating. The framework performance was evaluated through user studies, focusing on the effectiveness and usability of the DTMR and SPTA systems. Future work will focus on exploring advanced haptic rendering technologies to better represent the texture of target objects. We also plan to apply other consumer-grade VR/MR systems, such as HTC VIVE, Microsoft Hololens, and Sony PlayStation VR, and evaluate the effectiveness of our system for more complicated contact-rich dexterous manipulation.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

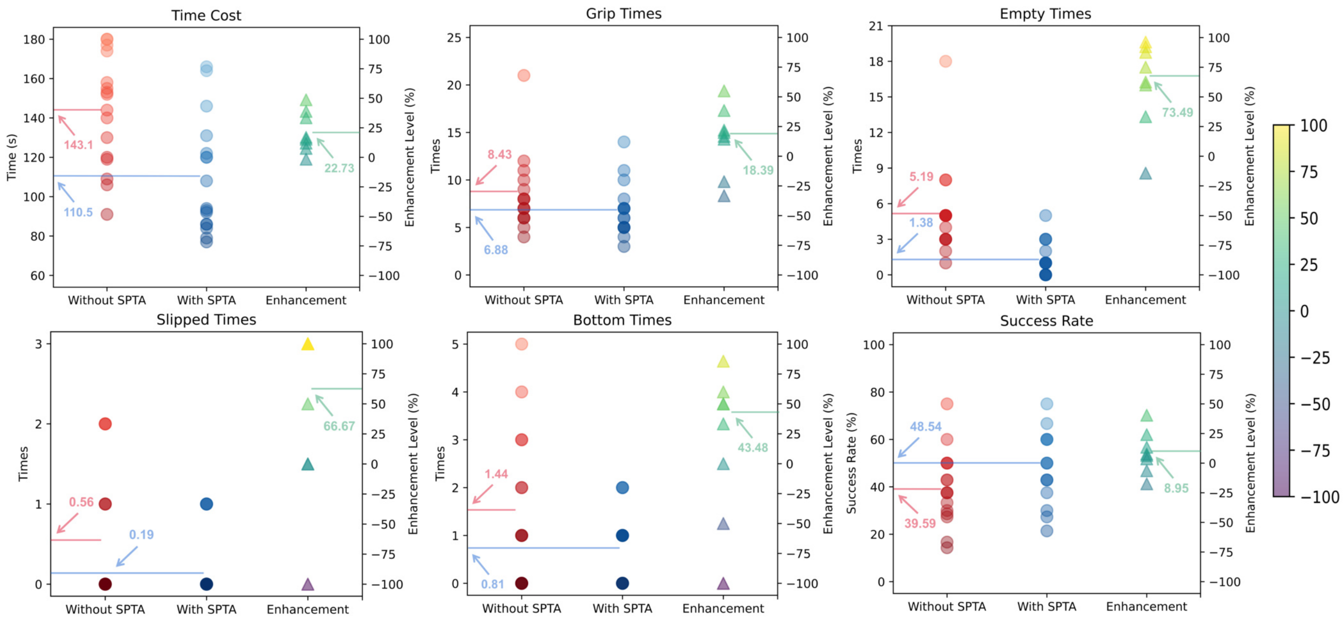

Fig. 9. Experiment results of the user study. The average performance for each participant under different conditions. The red points and blue points represent the data value without using and with using the SPTA tactile rendering device respectively, while the triangle markers indicate the enhanced percentage of the operator's behavior after introducing haptic rendering. Also, the average performances across all participants are illustrated by lines on each sub-figure.

References

[1] D. Zhang, J. Liu, L. Zhang, and G. Z. Yang, "Hamlyn crm: a compact master manipulator for surgical robot remote control," *International Journal of Computer Assisted Radiology & Surgery*, vol. 15, no. 3, pp. 503–514, 2020.

[2] A. Handa, K. Van Wyk, W. Yang, J. Liang, Y.-W. Chao, Q. Wan, S. Birchfield, N. Ratliff, and D. Fox, "Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 9164–9170.

[3] A. Gao, R. R. Murphy, W. Chen, G. Dagnino, P. Fischer, M. G. Gutierrez, D. Kundrat, B. J. Nelson, N. Shamsudhin, H. Su *et al.*, "Progress in robotics for combating infectious diseases," *Science Robotics*, vol. 6, no. 52, p. eabf1462, 2021.

[4] D. Zhang, J. Chen, W. Li, D. B. Salinas, and G.-Z. Yang, "A microsurgical robot research platform for robot-assisted microsurgery research and training," *International journal of computer assisted radiology and surgery*, vol. 15, no. 1, pp. 15–25, 2020.

[5] D. Zhang, W. Si, W. Fan, Y. Guan, and C. Yang, "From teleoperation to autonomous robot-assisted microsurgery: A survey," *Machine Intelligence Research*, pp. 1–19, 2022.

[6] J. Lin, X. Guo, W. Fan, W. Li, Y. Wang, J. Liang, W. Liu, L. Wei, and D. Zhang, "Tims: A tactile internet-based micromanipulation system with haptic guidance for surgical training," *arXiv preprint arXiv:2303.03566*, 2023.

[7] D. Zhang, Y. Guo, J. Chen, J. Liu, and G.-Z. Yang, "A handheld master controller for robot-assisted microsurgery," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 394–400.

[8] X. Gao, J. Silvério, S. Calinon, M. Li, and X. Xiao, "Bilateral teleoperation with object-adaptive mapping," *Complex & Intelligent Systems*, vol. 8, no. 4, pp. 2983–2990, 2022.

[9] D. Zhang, J. Liu, A. Gao, and G.-Z. Yang, "An ergonomic shared workspace analysis framework for the optimal placement of a compact master control console," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2995–3002, 2020.

[10] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S.-M. Choi, "A review on mixed reality: Current trends, challenges and prospects," *Applied Sciences*, vol. 10, no. 2, p. 636, 2020.

[11] Y. Yin, P. Zheng, C. Li, and L. Wang, "A state-of-the-art survey on augmented reality-assisted digital twin for futuristic human-centric industry transformation," *Robotics and Computer-Integrated Manufacturing*, vol. 81, p. 102515, 2023.

[12] D. Zhang, J. Liu, L. Zhang, and G.-Z. Yang, "Design and verification of a portable master manipulator based on an effective workspace analysis framework," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 417–424.

[13] G. D. Kessler, L. F. Hodges, and N. Walker, "Evaluation of the cyberglove as a whole-hand input device," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 2, no. 4, pp. 263–283, 1995.

[14] C. Pacchierotti, S. Sinclair, M. Solazzi, A. Frisoli, V. Hayward, and D. Prattichizzo, "Wearable haptic systems for the fingertip and the hand: taxonomy, review, and perspectives," *IEEE transactions on haptics*, vol. 10, no. 4, pp. 580–600, 2017.

[15] H. Zhao, A. M. Hussain, A. Israr, D. M. Vogt, M. Duduta, D. R. Clarke, and R. J. Wood, "A wearable soft haptic communicator based on dielectric elastomer actuators," *Soft robotics*, vol. 7, no. 4, pp. 451–461, 2020.

[16] C. Anthes, R. J. García-Hernández, M. Wiedemann, and D. Kranzlmüller, "State of the art of virtual reality technology," in *2016 IEEE aerospace conference*. IEEE, 2016, pp. 1–19.

[17] J. Yin, R. Hinchet, H. Shea, and C. Majidi, "Wearable soft technologies for haptic sensing and feedback," *Advanced Functional Materials*, vol. 31, no. 39, p. 2007428, 2021.

[18] J. Perret and E. Vander Poorten, "Commercial haptic gloves," in *Proceedings of the 15th Annual EuroVR Conference*. VTT Technology, 2018, pp. 39–48.

[19] K. Li, R. Bacher, W. Leemans, and F. Steinicke, "Towards robust exocentric mobile robot tele-operation in mixed reality," in *5th International Workshop on Virtual, Augmented, and Mixed Reality for HRI*, 2022.

[20] Y. Su, X. Chen, T. Zhou, C. Pretty, and G. Chase, "Mixed reality-integrated 3d/2d vision mapping for intuitive teleoperation of mobile manipulator," *Robotics and Computer-Integrated Manufacturing*, vol. 77, p. 102332, 2022.

[21] C. C. Ulloa, D. Domínguez, A. Barrientos, and J. del Cerro, "Design and mixed-reality teleoperation of a quadruped-manipulator robot for sar tasks," in *Climbing and Walking Robots Conference*. Springer, 2023, pp. 181–194.

[22] A. Martín-Barrio, J. J. Roldán-Gómez, I. Rodríguez, J. Del Cerro, and A. Barrientos, "Design of a hyper-redundant robot and teleoperation using mixed reality for inspection tasks," *Sensors*, vol. 20, no. 8, p. 2181, 2020.

[23] 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.

[24] Y.-P. Su, X.-Q. Chen, T. Zhou, C. Pretty, and G. Chase, "Mixed reality-enhanced intuitive teleoperation with hybrid virtual fixtures for intelligent robotic welding," *Applied Sciences*, vol. 11, no. 23, p. 11280, 2021.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.