Assessing the Suitability and Effectiveness of Mixed Reality Interfaces for Accurate Robot Teleoperation

Francesco De Pace
Politecnico di Torino
Torino, Italy
francesco.depace@polito.it

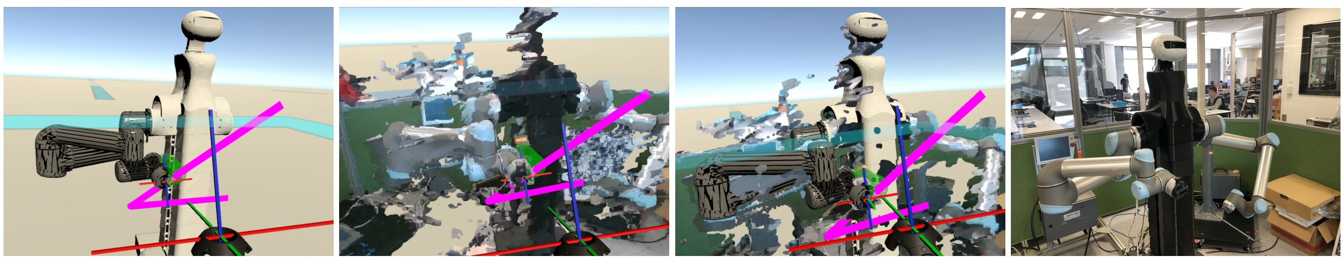
Gal Gorjup
The University of Auckland
Auckland, New Zealand
ggor290@aucklanduni.ac.nz

Huidong Bai
The University of Auckland
Auckland, New Zealand
huidong.bai@auckland.ac.nz

Andrea Sanna
Politecnico di Torino
Torino, Italy
andrea.sanna@polito.it

Minas Liarokapis
The University of Auckland
Auckland, New Zealand
minas.liarokapis@auckland.ac.nz

Mark Billinghurst
The University of Auckland
Auckland, New Zealand
mark.billinghurst@auckland.ac.nz



(a) (b) (c) (d)

Figure 1: Comparison of mixed and virtual reality interfaces: (a) the "pure" virtual interface (VR_S), (b) the "pure" point cloud interface (MR_S), (c) the point cloud and the virtual robot interface (MRR_S), and (d) the real robot in the laboratory space.

ABSTRACT

In this work, a Mixed Reality (MR) system is evaluated to assess whether it can be efficiently used in teleoperation tasks that require an accurate control of the robot end-effector. The robot and its local environment are captured using multiple RGB-D cameras, and a remote user controls the robot arm motion through Virtual Reality (VR) controllers. The captured data is streamed through the network and reconstructed in 3D, allowing the remote user to monitor the state of execution in real time through a VR headset. We compared our method with two other interfaces: i) teleoperation in pure VR, with the robot model rendered with the real joint states, and ii) teleoperation in MR, with the rendered model of the robot superimposed on the actual point cloud data. Preliminary results indicate that the virtual robot visualization is better than the pure point cloud for accurate teleoperation of a robot arm.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**.

VRST '20, November 1–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-7619-8/20/11.
https://doi.org/10.1145/3385956.3422092

KEYWORDS

Mixed Reality, Virtual Reality, Robot Teleoperation

ACM Reference Format:

Francesco De Pace, Gal Gorjup, Huidong Bai, Andrea Sanna, Minas Liarokapis, and Mark Billinghurst. 2020. Assessing the Suitability and Effectiveness of Mixed Reality Interfaces for Accurate Robot Teleoperation. In *26th ACM Symposium on Virtual Reality Software and Technology (VRST '20), November 1–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3385956.3422092

1 INTRODUCTION

There has been an increased research interest in developing methods that allow operators to use Virtual Reality (VR) and Mixed Reality (MR) technologies to remotely control [5, 7] and/or collaborate [6, 9] with robotic platforms. For example, Sun et al. [8] developed two types of control modes to tune the position, orientation, and force of an industrial manipulator in MR. Similarly, Whitney et al. described a remote teleoperation system [10, 11] to control a robotic arm in MR in a pick-and-place task. The results show that direct manipulation outperforms the MR teleoperation in terms of completion time and workload. To the best of our knowledge, no studies have been conducted to thoroughly analyze MR interfaces' effectiveness and accuracy in more complex path following tasks. In this work, we evaluate our MR robot teleoperation system for tasks that require highly accurate control of the end-effector position and velocity, such as remote surgery [6, 9] or welding [5, 7]. This is facilitated by the RGB-D sensors that allow for real-time 3D reconstruction of the physical surroundings.

2 THE MR SYSTEM

The system has been designed to connect two different environments, the local robotic environment and the remote user/operator environment (LE and RE, respectively). The LE contains a Universal Robot UR5 manipulator along with its controller, a personal computer (LPC_1) running Ubuntu 18.04 and the Robot Operating System (ROS) Melodic, and two Intel RealSense D415 cameras. LPC_1 is used to exchange data with the robot controller and send and receive data over the Local Area Network (LAN). The two depth cameras are connected to LPC_1, acquire, compress, and share camera frames based on the User Datagram Protocol (UDP) through the network. The RE setup also contains a computer (RPC_1) running Windows 10 and an immersive VIVE Pro VR headset, with one controller and two tracking stations. RPC_1 receives frames from LPC_1 over the LAN and runs the user interface developed in Unity3D. In order to properly visualize the point cloud and to detect the robot position and orientation in the virtual environment, two different calibration procedures are utilized. The depth cameras are extrinsically calibrated using the approach proposed in [1], ensuring high fidelity visualization. The alignment of the robot relies on the use of an Aruco marker [3] placed at a known position with respect to the real manipulator, making it possible to recognize its position and orientation in the virtual space. The operator is able to remotely control the robot arm by pressing the side button of the Vive controller, mapping the relative position and orientation of the controller to the robot end-effector. The operator can also use the teleporting interface to move around in the virtual environment by pressing the touchpad of the Vive controller. Virtual reference system indicators are rendered on both the controller and the robotic end-effector to highlight the axes of translation and rotation.

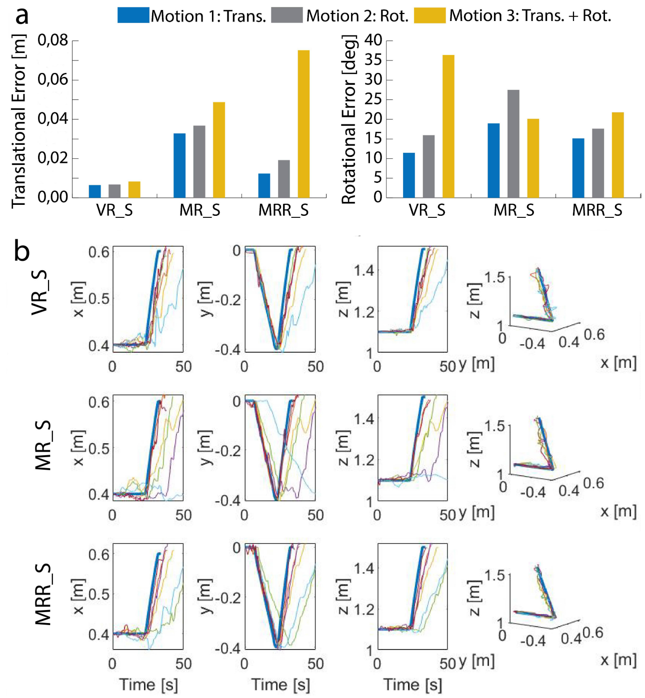


Figure 2: a) The translational (left chart) and rotational (right chart) errors for PT. b) Performance for the speed task. Blue line denotes the baseline. The first three columns present the end-effector position. The last column presents the teleoperated end-effector trajectories in 3D space.

3 MR INTERFACE EVALUATION

In order to investigate the effectiveness of teleoperation in MR, the proposed system (henceforth called MR_S) has been compared with two other interfaces (Figure 1): a "pure" VR version of the system (VR_S) and an MR version of the system with the virtual representation of the robot superimposed on the real one (MRR_S). Because of the calibration errors and the low resolution of the point cloud, the VR_S and the MRR_S interfaces were introduced and compared so as to investigate how these affect the MR teleoperation. Six users (aged between 25 and 31 years old) were asked to accomplish four different tasks with the above three interfaces and to complete a comprehensive survey about their experience with robotic systems. The survey showed that participants had a moderate exposure to robotics technologies (on average 3.6 out of 5).

In the first tasks (pose tasks, PT), users had to move the end-effector to three different positions and orientations in 3D space, highlighted by a red virtual asset (called *ghost*). During the last task (speed task, ST), users had to follow the ghost's movement along a specific trajectory (pure translation)[1]. Both PT and ST had been pre-recorded with the real robot to obtain the base-line data. The parameters collected were: i) the end-effector pose (PT), ii) the end-effector trajectory with respect to time (ST), iii) the usability based on the SUS questionnaire [2], iv) the workload based on the NASA-TLX questionnaire [4]. Regarding the SUS scores (*S*), both VR_S (*S*=80) and MRR_S (*S*=71) proved to be valuable solutions, whereas MR_S provided unsatisfactory results (*S*=58). These outcomes appear to be confirmed by the workload scores (VR_S (*S*=34), MRR_S (*S*=39), MR_S (*S*=60)), suggesting that the pure point cloud seems to be inadequate to teleoperate a robot. In contrast, a virtual representation of the robot greatly improves usability. Regarding PT, it is evident that translational errors are minimal for VR_S, followed by MRR_S and MR_S (Figure 2a). On the other hand, rotational errors appear to be quite high, independently of the employed interface. ST results show similar trends in speed tracking (columns 1-3 in Figure 2b). However, trajectories obtained through the MRR_S interface seem to match the baseline more closely than others (column 4 in Figure 2b).

4 CONCLUSIONS AND DISCUSSION

Preliminary results suggest that a pure point cloud interface seems less efficient than interfaces that also render the virtual representation of the robot. Future work will involve more users, considering the operator's body motions in 3D space and evaluating the operator's appreciation of the interfaces. Moreover, rotation tasks will also be considered. The results will provide useful insights to understand to what extent human operators can effectively and accurately control and teleoperate robotic platforms using MR interfaces.

[1] A video demonstrating the interfaces employed and the four different tasks can be found at the following URL: https://youtu.be/qgY5OKUMrg0

REFERENCES

[1] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[2] John Brooke. 1996. SUS: A quick and dirty usability scale. *Usability evaluation in industry* (1996), 189.

[3] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280 – 2292.

[4] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[5] Y. Liu and Y. Zhang. 2015. Toward Welding Robot With Human Knowledge: A Remotely-Controlled Approach. *IEEE Transactions on Automation Science and Engineering* 12, 2 (2015), 769–774.

[6] Cai Meng, Tianmiao Wang, Wusheng Chou, Sheng Luan, Yuru Zhang, and Zengmin Tian. 2004. Remote surgery case: robot-assisted teleneurosurgery. In *IEEE International Conference on Robotics and Automation (ICRA)*, Vol. 1. IEEE, 819—823.

[7] Amruta Rout, BBVL Deepak, and BB Biswal. 2019. Advances in weld seam tracking techniques for robotic welding: A review. *robotics and computer-integrated manufacturing* 56 (2019), 12–37.

[8] Da Sun, Andrey Kiselev, Qianfang Liao, Todor Stoyanov, and Amy Loutfi. 2020. A new mixed-reality-based teleoperation system for telepresence and maneuverability enhancement. *IEEE Transactions on Human-Machine Systems* 50, 1 (2020), 55–67.

[9] Andreas Tobergte, Rainer Konietschke, and Gerd Hirzinger. 2009. Planning and control of a teleoperation system for research in minimally invasive robotic surgery. In *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 4225–4232.

[10] David Whitney, Eric Rosen, Elizabeth Phillips, George Konidaris, and Stefanie Tellex. 2020. Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In *Robotics Research*. Springer, 335–350.

[11] David Whitney, Eric Rosen, Daniel Ullman, Elizabeth Phillips, and Stefanie Tellex. 2018. Ros reality: A virtual reality framework using consumer-grade hardware for ros-enabled robots. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1–9.