A Hands-Free Virtual-Reality Teleoperation Interface for Wizard-of-Oz Control

Nhan Tran
MIRRORLab
Colorado School of Mines
Golden, CO
nttran@mines.edu

Josh Rands
MIRRORLab
Colorado School of Mines
Golden, CO
jmrands@mines.edu

Tom Williams
MIRRORLab
Colorado School of Mines
Golden, CO
twilliams@mines.edu

ABSTRACT

Effective robotic teleoperation techniques are critical to Wizard-of-Oz (WoZ) style HRI experimentation. Current WoZ teleoperation methods such as GUIs, joysticks, and wearable sensors, may result in a loss of naturalistic interaction, impeding experimental effectiveness and validity. Recent work has investigated the use of Virtual Reality technologies to provide a WoZ teleoperator with their robot's perspective. This technique enhances visual feedback and enables more natural interaction, but introduces its own challenges. In this paper we propose a WoZ teleoperation interface that pairs a VR display with technologies for hands-free robot control in order to address those challenges while providing an immersive VR experience for robot teleoperators.

KEYWORDS

Virtual Reality, Wizard of Oz, Teleoperation, Interface Design

ACM Reference Format:
Nhan Tran, Josh Rands, and Tom Williams. 2018. A Hands-Free Virtual-Reality Teleoperation Interface for Wizard-of-Oz Control. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for Human-Robot Interaction (VAM-HRI'18).* ACM, New York, NY, USA, 4 pages.

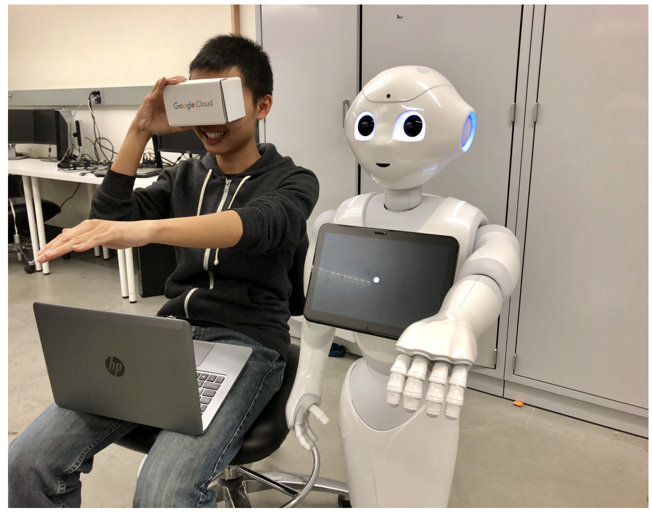

Figure 1: The teleoperation interface, and its integration with the Softbank Pepper robot.

1 INTRODUCTION

Human-Robot Interaction (HRI) researchers need to be able to effectively evaluate how humans will interact with autonomous robots. This is particularly challenging as researchers often need to be able to evaluate autonomous capabilities that do not exist yet, and/or need to evaluate autonomous capabilities within a carefully experimentally controlled environment. One common method of accomplishing this is the Wizard of Oz (WoZ) paradigm.

In the WoZ paradigm, an experimenter remotely controls a robot, triggering some subset of the capabilities it would ideally execute autonomously, such as movement, speech, or cognition [21]. While this paradigm allows autonomous capabilities to be evaluated without requiring those capabilities to actually be implemented, it comes with its own set of implementation challenges. Specifically, experimenters are typically required to design point-and-click robot control interfaces, a practice which can be repetitive and time consuming. Moreover, such interfaces are not always effective, as the time necessary for an experimenter to decide to issue a command, click the appropriate button, and have that command take effect on the robot is typically too long to facilitate natural interaction.

What is more, the WoZ paradigm is often implemented from a third-person point of view and/or displayed on a screen. This type of remote teleoperation risks decreasing the situational awareness of the teleoperator and can harm the effectiveness of the experiment [5]. And while this method is superior to joystick or gamepad-based teleoperation, the teleoperator is unable to view the robot's perspective. This can be both ineffective and dangerous; operating a robotic arm from varying visual perspectives, for example, can cause serious performance challenges for the teleoperator [15]. Virtual Reality (VR) teleoperation provides one possible solution to this problem. VR puts the teleoperator in the robot's perspective, which improves depth perception and enhances visual feedback, resulting in an overall more immersive experience [14]. One issue that arises with this type of immersive VR teleoperation, however, is that the teleoperator can no longer see their teleoperation interface, often requiring the use of two teleoperators rather than one. We propose a novel teleoperation interface which provides hands-free WoZ control of a robot while providing the teleoperator with an immersive VR experience from the robot's point of view.

This interface integrates a VR headset with a Leap Motion Controller. With these tools, we have created an environment that gives the experimenter hands-free, immersive control of a robot's social interactions. Within this framework, the VR headset is interfaced directly with the robot's camera, allowing the experimenter to see exactly what the robot sees. Furthermore, the Leap Motion sensor reads input from the experimenter's gestures and automatically generates these gestures on the robot in real time, without requiring the teleoperator to see this teleoperation interface.

2 PREVIOUS WORK

Many different robot teleoperation interfaces have previously been proposed. For example, Hashimoto et al. [9], Quintero et al. [20] created interfaces in which the teleoperator viewed the robot's environment from a first- and third-person points of view, respectively, and controlled movement by touching the camera stream of the robot. In another teleoperation system Xi et al. [26] utilized computer vision techniques to capture the operators movements, and replicated said movements on a remote robot. This technique required the teleoperator to hold a 4-ball frame, and through computer vision and inverse kinematics they calculated the position and orientation of their hand, further using this data to control the robot. WoZ robot teleoperation interfaces range from immersive VR to gesture control via analog joysticks. Further implementations utilize hands-free teleoperation. Our work extends the capabilities of past teleoperation interfaces by integrating VR and hands-free operation.

Virtual and augmented reality robotic teleoperation interfaces are effective in creating an immersive environment for the operator. Liu et al. [13] state that VR robot interaction significantly improves performance over on-screen HRI. In situations where perception is challenging, VR enhances performance fundamentally due to stereo cues. VR teleoperation enhances field of view and allows the operator to simulate a more naturalistic interaction between subject and robot [14]. The following interfaces implement robot teleoperation via virtual/augmented reality.

Miner and Stansfield [14] describe a teleoperation interface that uses voice control and gesture commands in a simulated VR environment. The setup utilized task-level orientation in which participants issued voice commands to be executed by the robot. Gong et al. [7] and Hedayati et al. [10] created more immersive teleoperating experiences by integrating trackpads and joysticks with Augmented Reality (AR) glasses, which allowed the user to see from the robot's perspective while augmenting this perspective with information of relevance to the robot's teleoperator. Recent work has also investigated the use of suit-based teleoperation for controlling humanoid robots. This teleoperation interface requires the operator to wear a full exo-suit equipped with arm, shoulder, chest, and head sensors [4]. Immersive VR teleoperation interfaces such as this enable teleoperators to effectively replicate natural human motions, enabling more natural WoZ studies – however the need for such a suit introduces its own complications. Another robotic teleoperation interface was created by Pereira et al. [17]. Their interface utilized a VR headset and joysticks that enabled the operator to see from the robot's perspective, control the gestures of the robot, and control the direction of the robot's gaze. This setup contained both a "Solo" setup as well as a "Pair" setup to optimize the control of their telepresent robot. In the Solo setup, the gestures of the robot were controlled via a joystick while the operator also controlled verbal interactions. In the Pair setup, the joysticks were replaced by a game controller with two analog joysticks of similar size. One operator controlled the non-verbal interactions while immersed in VR, and another operator was responsible for the verbal interactions while watching a computer monitor display of the robot's camera stream. All of the above interfaces utilize VR teleoperation to create an immersive environment for the operator, resulting in more naturalistic social capabilities for the teleoperated robot and increasing the effectiveness of the WoZ technique.

In addition to these teleoperation techniques, there have been approaches that use VR hand-controllers [6, 8, 12, 16, 22, 24], approaches that use glove controllers [1, 2], as well as hands-free adaptations. Hands free teleoperation creates a more natural human-computer interaction that reduces training time for operators by having a more user-friendly operating interface. In addition, hands-free teleoperation techniques can reduce costs by not requiring extensive hardware and sensors to gather motion data. The Leap Motion Controller is a hands free teleoperation which uses optical 3D sensors to map the users hand and finger positions [23]. Bassily et al. [3] created an algorithm that mapped human hand motion to a 6-DOF Jaco robotic arm using a Leap Motion Controller and inverse kinematics calculations. In traditional joystick or gamepad control, understanding the complex movements of a robotic arm with 6-DOF is time consuming and difficult, but the Leap Motion Controller, coupled with inverse kinematics mathematics, can simplify this process by solely requiring the user to replicate the gesture he/she desires of the robot, making it a powerful hands-free teleoperation device [18]. This previous research suggests that technologies like the Leap Motion Controller can provide effective and precise control for teleoperators, decreasing complexity while increasing usability from a user perspective. We would also draw the reader's attention to the work of Gaurav et al. [6], presented concurrently with our own at VAM-HRI 2018 [25], which uses Kinect data in a similar way to our own use of Leap Motion data.

To the best of our knowledge, however, no previous WoZ implementation has combined Virtual Reality display with Leap Motion input. We believe that the Leap Motion Controller combined with VR headset will provide teleoperators a more usable, hands-free, and immersive interface for effective remote robotic control.

3 INTEGRATED APPROACH

We propose an interface in which WoZ experimenters experience the robot's surrounding environment using a VR headset and control the robot utilizing a Leap Motion sensor. Images are streamed from a robot's camera to a VR headset, mimicking human stereoscopic vision using two lenses positioned above a display screen. Figure 3 shows the streaming images within our proposed interface, separated into left and right views in the display screen. After putting the display into the VR headset, experimenters would see in simulated 3D stereoscopic vision as the lenses help reshape the separated views.

In order to teleoperate the robot using hand gestures, Leap Motion data is manipulated to extract features such as hand position and orientation (e.g., pitch, roll, yaw). Figure 4 shows the visualization of the tracking data produced by the Leap Motion. Each arrow represents a finger, and each trail represents the corresponding movement of that finger. We then use the following equations to transform these features into a gesture command which can be sent to the robot.

$$robotGesturePitch = \begin{cases} low & \tau_{p_1} < humanGesturePitch < \tau_{p_2} \\ high & \tau_{p_2} < humanGesturePitch < \tau_{p_3} \end{cases}$$

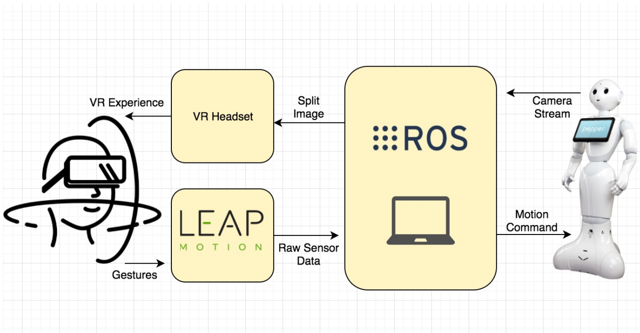

Figure 2: Architecture diagram: The user interacts directly with a VR headset (e.g., Google Cardboard) and a Leap Motion gesture sensor. These devices send data to and receive data from a humanoid robot (e.g., the Softbank Pepper) using an instance of the ROS architecture whose Master node is run on a standard Linux laptop.

$$robotGestureRoll = \begin{cases} low & \tau_{r_1} < humanGestureRoll < \tau_{r_2} \\ high & \tau_{r_2} < humanGestureRoll < \tau_{r_3} \end{cases}$$

Here, parameters $\tau_{p_1} < \tau_{p_2} < \tau_{p_3}$ and $\tau_{r_1} < \tau_{r_2} < \tau_{r_3}$ are manually defined pitch and raw thresholds. While in this work our initial prototype makes use of these simple inequalities, in future work we aim to examine more sophisticated geometric and approximate methods for precisely mapping human gestures to robot gestures.

All components of the proposed interface are integrated using the Robot Operating System (ROS) [19]. As shown in Figure 2, the Leap Motion publishes raw sensor data to the controller computer. This data is then converted to Twist data and then into motion commands which are sent to the robot via the Softbank NaoQi API. The controller computer also utilizes the NaoQui API to obtain a remote camera stream from the Softbank Pepper robot[1], publishing the resulting image data to a topic subscribed to by the Android VR app. This app then uses this image data to display what the robot sees in the user's split view. Figure 3 shows the split view that a user would see in the application. In addition, images can be streamed to this VR app from a ROS simulator (e.g., Gazebo [11]) or some other source, using the same process. In this work, we use a single camera, as Pepper has a single RGB camera rather than stereo cameras. In the future, we hope to use stereoscopic vision as input for a more immersive VR experience. This will, however, come with its own set of computational challenges.

4 CONCLUSIONS AND FUTURE WORK

In this paper we have proposed an immersive, hands-free interface for WoZ teleoperation. The system consists of a VR headset, which connects to the robot's camera, and a Leap Motion Controller to enable natural real-time interaction in WoZ style experimentation. The immersive experience provided by the VR headset and the hands-free nature of the Leap Motion Controller combine to create a user-friendly and effective interface for robot teleoperation.

[1] While in this work we use the Softbank Pepper robot, our general framework is not necessarily specific to this particular robot.

Figure 3: VR Stream

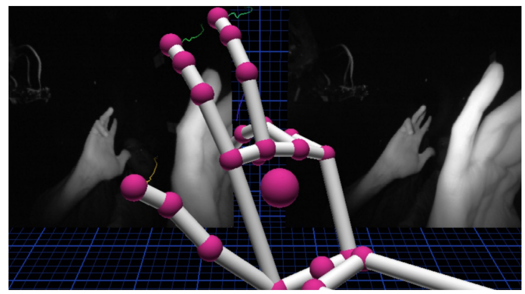

Figure 4: View of gesture tracking data in the Leap Motion SDK Diagnostic Visualizer

In future work, we plan to extend this preliminary interface in several ways. First, we aim to evaluate this interface in comparison to other teleoperation techniques such as traditional WoZ GUIS and third-person VR setups [7, 9, 14, 20]. Second, we note that the proposed approach, which passes the robot's camera feed directly into the teleoperator's VR display, provides the opportunity for Remote Augmented Reality. As such, we would like to integrate WoZ-relevant augmented reality cues into this VR display, to provide further information to the teleoperator, similar to what Gong et al. [7] accomplished with their teleoperation interface. Third, We would like to integrate our interface into an entirely virtual environment, for experiments in which participants interact with *simulated* robots within a virtual environment. Finally, we would like to integrate speech-recognition and text-to-speech and/or vocoder functionality to this teleoperator interface. This would allow teleoperators to either further teleoperate the robot using verbal commands, or to provide the robot with speech input which can immediately be uttered by the robot which they are teleoperating.

REFERENCES

[1] Jordan Allspaw, Jonathan Roche, Nicholas Lemiesz, Michael Yannuzzi, and Holly A Yanco. 2018. Remotely Teleoperating a Humanoid Robot to Perform Fine Motor Tasks with Virtual Reality–. In *Proceedings of the 2018 Conference on Waste Management.*

[2] Jordan Allspaw, Jonathan Roche, Adam Norton, and Holly A Yanco. 2018. Remotely Teleoperating a Humanoid Robot to Perform Fine Motor Tasks with Virtual Reality–. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI).*

[3] D. Bassily, C. Georgoulas, J. Guettler, T. Linner, and T. Bock. 2014. Intuitive and Adaptive Robotic Arm Manipulation using the Leap Motion Controller. In *ISR/Robotik 2014; 41st International Symposium on Robotics*. 1–7.

[4] Maxwell Bennett, Tom Williams, Daria Thames, and Matthias Scheutz. 2017. Differences in Interaction Patterns and Perception for Teleoperated and Autonomous Humanoid Robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[5] J. Y. C. Chen, E. C. Haas, and M. J. Barnes. 2007. Human Performance Issues and User Interface Design for Teleoperated Robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 6 (Nov 2007), 1231–1245. https://doi.org/10.1109/TSMCC.2007.905819

[6] Sanket Gaurav, Zainab Al-Qurashi, Amey Barapatre, and Brian Ziebart. 2018. Enabling Effective Robotic Teleoperation using Virtual Reality and Correspondence Learning via Neural Network. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[7] L. Gong, C. Gong, Z. Ma, L. Zhao, Z. Wang, X. Li, X. Jing, H. Yang, and C. Liu. 2017. Real-time human-in-the-loop remote control for a life-size traffic police robot with multiple augmented reality aided display terminals. In *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*. 420–425. https://doi.org/10.1109/ICARM.2017.8273199

[8] Kerstin S. Haring, Victor Finomore, Dylan Muramato, Nathan L. Tenhundfeld, Mormon Redd, James Wen, and Brian Tidball. 2018. Analysis of Using Virtual Reality (VR) for Command and Control Applications of Multi-Robot Systems. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[9] Sunao Hashimoto, Akihiko Ishida, Masahiko Inami, and Takeo Igarashi. 2011. Touchme: An augmented reality based remote robot manipulation. In *The 21st International Conference on Artificial Reality and Telexistence, Proceedings of ICAT2011*, Vol. 2.

[10] Hooman Hedayati, Michael Walker, and Daniel Szafir. 2018. Improving Collocated Robot Teleoperation with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 78–86.

[11] Nathan Koenig and Andrew Howard. 2004. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2149–2154.

[12] Jeffrey I Lipton, Aidan J Fay, and Daniela Rus. 2018. Baxter's Homunculus: Virtual Reality Spaces for Teleoperation in Manufacturing. *IEEE Robotics and Automation Letters* 3, 1 (2018), 179–186.

[13] O. Liu, D. Rakita, B. Mutlu, and M. Gleicher. 2017. Understanding human-robot interaction in virtual reality. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 751–757. https://doi.org/10.1109/ROMAN.2017.8172387

[14] N. E. Miner and S. A. Stansfield. 1994. An interactive virtual reality simulation system for robot control and operator training. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. 1428–1435 vol.2. https://doi.org/10.1109/ROBOT.1994.351289

[15] A. Nawab, K. Chintamani, D. Ellis, G. Auner, and A. Pandya. 2007. Joystick mapped Augmented Reality Cues for End-Effector controlled Tele-operated Robots. In *2007 IEEE Virtual Reality Conference*. 263–266. https://doi.org/10.1109/VR.2007.352496

[16] Yeonju Oh, Ramviyas Parasuraman, Tim McGraw, and Byung-Cheol Min. 2018. 360 VR Based Robot Teleoperation Interface for Virtual Tour. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[17] André Pereira, Elizabeth J Carter, Iolanda Leite, John Mars, and Jill Fain Lehman. 2017. Augmented Reality Dialog Interface for Multimodal Teleoperation. In *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 764–771.

[18] Y. Pititeeraphab, P. Choitkunnan, N. Thongpance, K. Kullathum, and C. Pintavirooj. 2016. Robot-arm control system using LEAP motion controller. In *2016 International Conference on Biomedical Engineering (BME-HUST)*. 109–112. https://doi.org/10.1109/BME-HUST.2016.7782091

[19] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. 2009. ROS: an open-source Robot Operating System. In *ICRA workshop on open source software*, Vol. 3. Kobe, Japan, 5.

[20] Camilo Perez Quintero, Oscar A Ramirez, and Martin Jägersand. 2015. VIBI: Assistive vision-based interface for robot manipulation.. In *ICRA*. 4458–4463.

[21] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012).

[22] Eric Rosen, David Whitney, Elizabeth Phillips, Daniel Ullman, and Stefanie Tellex. 2018. Testing Robot Teleoperation using a Virtual Reality Interface with ROS Reality. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[23] Frank Weichert, Daniel Bachmann, Bartholomäus Rudak, and Denis Fisseler. 2013. Analysis of the accuracy and robustness of the leap motion controller. *Sensors* 13, 5 (2013), 6380–6393.

[24] David Whitney, Eric Rosen, Elizabeth Phillips, George Konidaris, and Stefanie Tellex. 2017. Comparing Robot Grasping Teleoperation across Desktop and Virtual Reality with ROS Reality. In *Proceedings of the International Symposium on Robotics Research*.

[25] Tom Williams, Daniel Szafir, Tathagata Chakraborti, and Heni Ben Amor. 2018. Virtual, Augmented, and Mixed Reality for Human-Robot Interaction. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 403–404.

[26] B. Xi, N. Liu, T. Lu, J. Lu, Y. Cai, and S. Wang. 2017. Robot teleoperation system based on SVDD. In *2017 Chinese Automation Congress (CAC)*. 2830–2835. https://doi.org/10.1109/CAC.2017.8243258