Enhancing Telepresence in Human-Robot Interaction with Depth Sensing and Augmented Reality

HiMER Lab — Human-in-Mind Engineering Research

Shyam Ganatra, MS Robotics and Autonomous Systems
(Mentor: Heejin Jeong, PhD, Assistant Professor)
School of Manufacturing Systems and Networks | The Polytechnic School

Introduction

Teleoperation enables remote robot control in fields such as manufacturing, healthcare, and disaster response, where direct human presence is difficult or risky. [1]

Traditional teleoperation setups lack immersive depth perception, which can impair task accuracy in cluttered or complex environments. [2]

This project aims to enhance teleoperation by integrating depth-sensing and augmented reality (AR), improving spatial awareness, user experience, and operational efficiency.

Findings and Progress Thus Far

Enhanced Depth Perception: The integration of depth perception has proven successful. The system now enables users to perceive depth information, providing a more immersive and accurate representation of object locations. This improvement reduces the difficulty of tasks requiring precise distance estimations, such as grasping or positioning objects.

Challenges: One of the primary challenge was integrating and synchronizing the depth camera data with the robot's inbuilt sensors. Since the teleoperation requires immediate feedback to the user, any delays could affect the accuracy of system

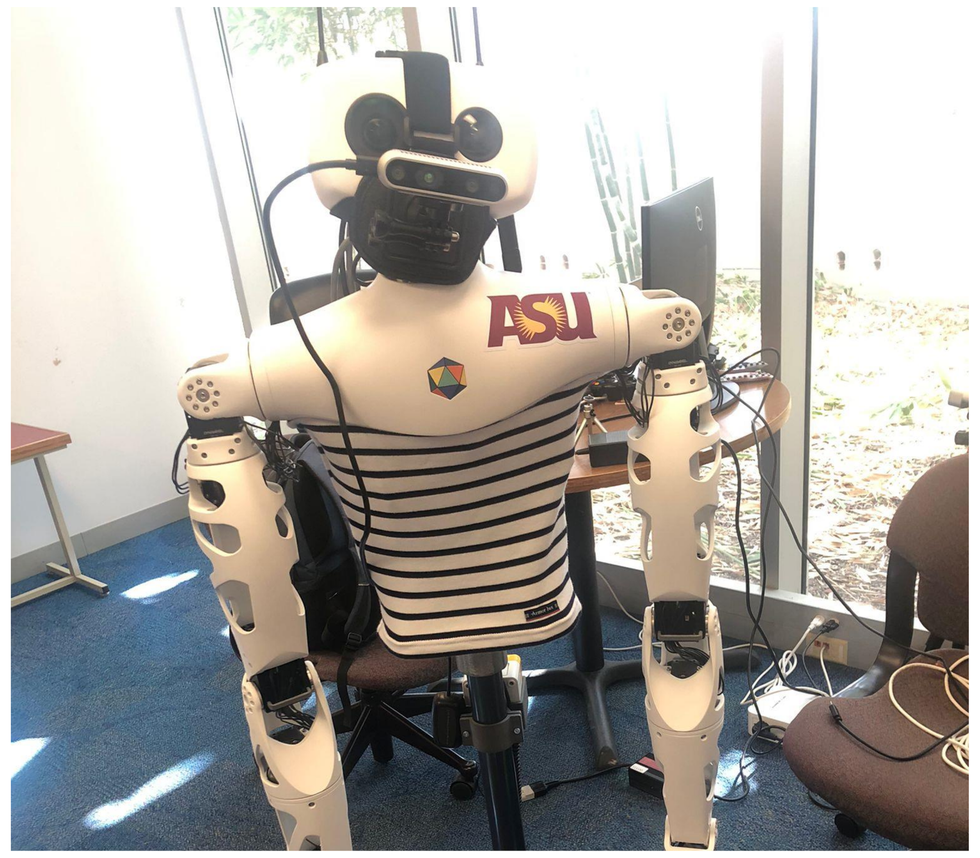

Fig 1: Reachy robot with an added camera

Research Methods

Hardware Integration:

- Integrated an **Intel RealSense D345i** depth camera into the **Reachy robot** to capture real-time depth data using stereo vision and active IR, delivering spatial information that surpasses traditional 2D cameras. [3]
- Configured the **RealSense SDK** and **ROS** nodes to synchronize depth data with Reachy's motor control, enabling responsive, real-time feedback for enhanced teleoperation.

Algorithm Development:

Specialized algorithms were designed to process the depth data using **Point Cloud Library (PCL)** and **OpenCV** for filtering, segmentation, and spatial analysis.

- **Distance Estimation:** A **Euclidean clustering algorithm** segments objects in Reachy's environment, with object distances calculated based on point cloud centroids. [4]
- **Bounding Box Generation:** Detected objects are enclosed in 3D bounding boxes using the **Minimum Volume Bounding Box** algorithm, ensuring consistent spatial representation. [5]

VR Interface Enhancement:

The Meta Quest Pro VR headset was adapted to include augmented reality (AR) overlays that display depth and object data in real-time, using OpenXR and Unity3D for smooth integration. [6]

- **Overlay Elements:** Objects in Reachy's environment are highlighted with bounding boxes and distance labels, with color-coded proximity indicators (e.g., green for safe, red for close). [7]
- **User Interface Enhancements:** Integrated with Unity's Canvas system for low-latency rendering, the interface offers adjustable transparency and a "distance awareness" mode that highlights objects by proximity. [8]

User Experience Testing: To validate the system, user studies will be conducted comparing the traditional teleoperation with he enhanced system with depth perception and AR overlays.

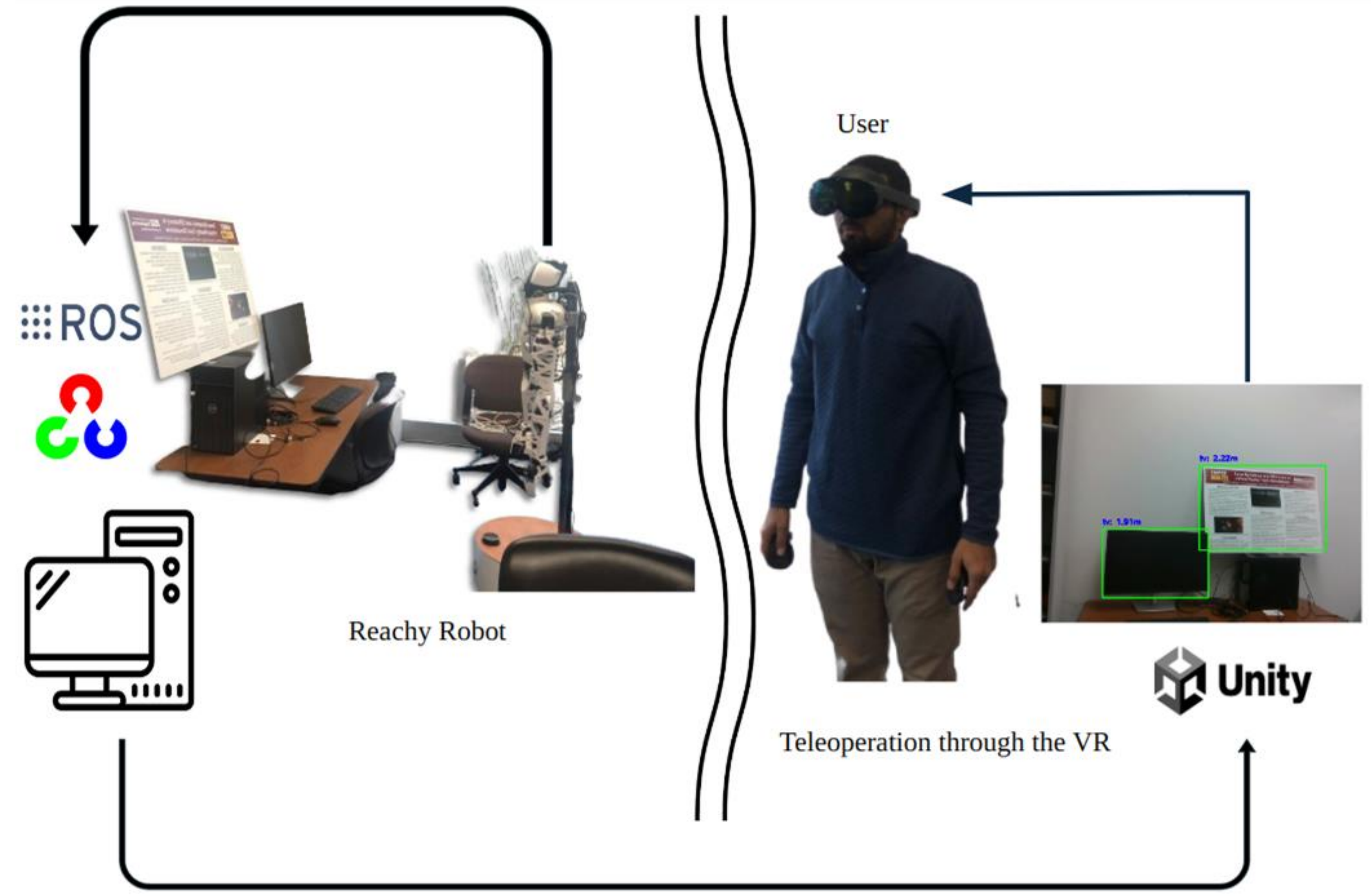

Fig 2: Concept image of VR Teleoperation with depth information

References

1. Aschenbrenner, D., Fritscher, M., Sittner, F., Krauß, M., & Schilling, K. (2015). *Teleoperation of an industrial robot in an active production line.* IFAC-PapersOnLine, 48(10), 159-164. DOI: https://doi.org/10.1016/j.ifacol.2015.08.125
2. Pang, G., Yang, G., & Pang, Z. (2021). *Review of robot skin: A potential enabler for safe collaboration, immersive teleoperation, and affective interaction of future collaborative robots.* IEEE Transactions on Medical Robotics and Bionics, 3(3), 681-700. DOI: https://doi.org/10.1109/tmrb.2021.3097252
3. Rusu, R. B., & Cousins, S. (2011). *3D is here: Point cloud library (PCL).* In 2011 IEEE International Conference on Robotics and Automation (pp. 1-4). IEEE. DOI: https://doi.org/10.1109/ICRA.2011.5980567
4. Gonzalez, R. C., & Woods, R. E. (2008). *Digital Image Processing* (3rd ed.). Pearson. (For foundational methods in clustering, bounding boxes, and segmentation).
5. Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. International Journal of Computer Vision, 13(2), 119-152. DOI: https://doi.org/10.1007/BF01427149
6. Omarali, B., Denoun, B., Althoefer, K., Jamone, L., Valle, M., & Farkhatdinov, I. (2020). Virtual reality-based telerobotics framework with depth cameras. In 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 1217-1222). IEEE. DOI: https://doi.org/10.1109/RO-MAN47096.2020.9223502
7. Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient Graph-Based Image Segmentation. International Journal of Computer Vision, 59(2), 167–181. https://doi.org/10.1023/b:visi.0000022288.19776.77
8. Chen, J. Y. C., Haas, E. C., & Barnes, M. J. (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), 1231–1245. https://doi.org/10.1109/tsmcc.2007.905819

Acknowledgements

This project was made possible with mentorship of **Dr. Heejin Jeong** at the **HiMER Lab**, ASU. I extend my gratitude to the to the **MORE** program for funding essential research components. Special thanks to fellow lab members and collaborators who provided valuable feedback throughout the development and testing phases.

MORE — Masters Opportunity for Research in Engineering

ASU Ira A. Fulton Schools of Engineering — Arizona State University