Improving Fluidity Through Action: A Proposal for a Virtual Reality Platform for Improving Real-World HRI

Carlos Baptista De Lima
c.v.baptistadelima@swansea.ac.uk
Swansea University
Swansea, United Kingdom

Julian Hough
julian.hough@swansea.ac.uk
Swansea University
Swansea, United Kingdom

Frank Förster
f.foerster@herts.ac.uk
University of Hertfordshire
Hatfield, United Kingdom

Patrick Holthaus
p.holthaus@herts.ac.uk
University of Hertfordshire
Hatfield, United Kingdom

Yongjun Zheng
y.zheng20@herts.ac.uk
University of Hertfordshire
Hatfield, United Kingdom

ABSTRACT

Achieving truly fluid interaction with robots with speech interfaces remains a hard problem. Despite technical advances in sensors, processors and actuators, the experience of Human-Robot Interaction (HRI) remains laboured and frustrating. Some of the barriers to this stem from a lack of a suitable development platform for HRI to improve the interaction, particularly for mobile manipulator robots. In this paper we briefly overview some existing systems and propose a high-fidelity Virtual Reality (VR) HRI simulation environment with Wizard-of-Oz (WoZ) cabability applicable to multiple robots including mobile manipulators and social robots.

KEYWORDS

Human-Robot Interaction, Fluidity, Virtual Reality Environments, Wizard-of-Oz

ACM Reference Format:

Carlos Baptista De Lima, Julian Hough, Frank Förster, Patrick Holthaus, and Yongjun Zheng. 2024. Improving Fluidity Through Action: A Proposal for a Virtual Reality Platform for Improving Real-World HRI. In *International Conference on Human-Agent Interaction (HAI '24), November 24–27, 2024, Swansea, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3687272.3690881

1 INTRODUCTION

The challenge of achieving fluid interaction while maintaining users' trust remains prevalent within the field of human-robot interaction (HRI). Despite advances in robot vision, motion, manipulation and automatic speech recognition, state-of-the-art human-robot interaction (HRI) is slow, laboured, and fragile. Improving the feeling of fluidity is a difficult task to achieve, with a variety of challenges from the robot control algorithms to the physical constraints of the robots [5, 6].

To allow for rapid development of HRI models, virtual reality (VR) is often used to create a development and testing environment, which largely isolates hardware constraints that robots may have concerning factors such as robot vision and motion. The use of VR environments with simulations of real-world constraints is not new, however VR in robotics has mainly been created for automatic optimisation of robotic motion planning, often within the self-reinforcement learning paradigm. On the other hand, there are VR and mixed reality simulations of HRI for social robots such as Flobi [8], however, these VR environments are largely specific to each robot and without a focus on real-world manipulator robots that achieve physical tasks.

VR is a powerful tool that can replicate the testing conditions required for Wizard-of-Oz (WoZ) experiments with experimenters tele-operating the robot, and in fully autonomous systems. In fact, there is a significant amount of research within this field looking to explore the benefits of using VR for the purpose of testing and using robots [2, 4, 12]. However, a question remains as to whether the virtual environment provides a comparable test-bed in relation to real-world testing for interactive robots.

2 BRIEF OVERVIEW OF EXISTING VR SYSTEMS FOR HRI SIMULATION

Surveying 33 HRI studies that compared the differences in the effects of physically co-present robots, telepresent robots, and virtual agents on human interactors, Li found that in the majority of cases human responses differed in a way that was favourable towards physical co-presence [7]. The effects under investigation of the surveyed studies included subjective ratings of trustworthiness, enjoyableness, or social presence as well as behavioural measures such as time on task, response times, or task completion times. Notably, only one study investigated the effect of embodiment on the interactors' linguistic behaviour: Fischer, Lohan & Foth [3] found that it was mainly interactional linguistic features such as the use of copulas, vocatives, understanding checks where participants' speech differed between the different embodiments, indicating a stronger common ground and interpersonal relationship between their participants and the physical robot. Importantly, none of the surveyed studies employed immersive virtual environments and is unclear whether the effects in favour of physically co-present robots would extend to these.

There are existing reviews of VR tools being used as test-beds for robots. Wonsick and Padir [12] explore various VR interfaces that were being used for controlling and interacting with robots. They

HAI '24, November 24–27, 2024, Swansea, United Kingdom

ACM ISBN 979-8-4007-1178-7/24/11
https://doi.org/10.1145/3687272.3690881

found that the use of a VR interface on the whole reduces the task completion time and increases operator performance. There was an absence of studies on interactive manipulator robots, suggesting a need to further explore this area.

Simpson, Richardson and Richards explored the usage of a Wizard-of-Oz (WoZ) experience within a collaborative multi-user VR environment [9]. They showed that although the system worked for the most part, their WoZ setup was still cumbersome, and some participants were unsure if they were completing the task correctly. This system did not involve the usage of any real-world robots.

Arntz et al [1] present a VR lab with the virtual model of the Pepper robot which interfaces with a WoZ system in the Unity game engine. The system also gathers data in real time, including eye-tracking. Such information is very important to gather when understanding how a user looks around and interacts within this space and this is particularly true for social robots. Furthermore, their WoZ system allows for changes to the environment to be made during run-time which allows for more flexibility in the testing process.

While there are substantial contributions to real-time VR HRI environments, there is still a lack of a system that incorporates all the requirements for a single VR environment for testing and improving fluidity of interaction. For example, the test lab by Arntz et al [1] was exclusively designed for social robots. Furthermore, while eye-tracking is important, it is unclear if such systems also capture gestures made by the user and audio, and integrating all these data streams in a single logging system. In some cases it is not clear what type of VR system was being used (e.g. [10]). There has also been little research exploring these test labs with manipulator robots which can cause technical difficulties in terms of the precision of movement that would be needed to pick up and move objects around in a space. Furthermore, despite papers showing that they record gesture or eye tracking, they do not do both. Thus, there is a need for creating a VR environment which, in similar to the work done by Arntz et al [1] and Whitney et al [11] is a testing bed for various ROS-enabled robots and that can be supported on a variety of headsets and track a variety of data streams in such a way that vriables can be adjusted in real-time within a WoZ interface. This would allow for the ability to adjust and be adaptable to existing models.

3 PROPOSAL FOR A VIRTUAL REALITY HRI PLATFORM

Based on the gaps in the literature identified above, we propose a highly realistic, high-fidelity, VR platform that supports the development of multiple ROS-enabled robots with both WoZ and fully automated HRI. In particular, we are interested in exploring the fluidity of such interactions with those who are either elderly or less-able. This is important as robots such as Fetch could significantly improve their quality of life within their home if they can trust the robot to correctly complete the actions required of them in a fast and efficient manner.

The first test case is to create an environment and space to work with the mobile manipulator robot, Fetch, along with the base initial environment that we would use to replicate the space of a standard home in the VR space as in Figure 1. As we are concerned with assistive robotics, [12] identified that the average age range of the participants from their review were in the 20-30 age bracket, so the scenario is designed to reflect the needs of populations who might benefit from assistive robotics in the home.

Figure 1: An image of the environment which would then be replicated into the Virtual Reality (VR) space.

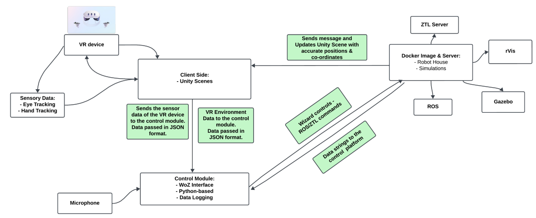

Figure 2: A diagram showing the system architecture of the proposed VR-HRI system.

The system will collect face and eye-tracking, hand tracking and audio logging data. Furthermore, the system will allow the isolation of certain capabilities such as audio and speech processing and legibility of motion. For the WoZ mode of the system, different capabilities would be either automated or Wizarded depending on the focus of the experiment or data collection. For example, for collecting participant vocal utterances in a Wizarded mode, the audio from microphones would be streamed to a Wizard teleoperating the virtual robot's motion and recorded, rather than go through a dialogue system. The architecture of the VR-based system can be seen in Figure 2, using Unity for the VR simulation environment and ROS as the middleware.

The VR environment will simulate the real-world environment and robot closely for final comparison of the two modalities, and the efficacy of building in-robot models via VR testing.

ACKNOWLEDGMENTS

Baptista de Lima, Hough and Förster's work is funded by the FLUIDITY project (EPSRC grant EP/X009343/1).

REFERENCES

[1] Alexander Arntz, André Helgert, Carolin Straßmann, and Sabrina C Eimler. 2024. Enhancing Human-Robot Interaction Research by Using a Virtual Reality Lab Approach. In *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*. IEEE, 340–344.

[2] Luca Crosato, Chongfeng Wei, Edmond SL Ho, Hubert PH Shum, and Yuzhu Sun. 2024. A Virtual Reality Framework for Human-Driver Interaction Research: Safe and Cost-Effective Data Collection. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 167–174.

[3] Kerstin Fischer, Katrin S. Lohan, and Kilian Foth. 2012. Levels of embodiment: linguistic analyses of factors influencing hri. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (Boston, Massachusetts, USA) *(HRI '12)*. Association for Computing Machinery, New York, NY, USA, 463–470. https://doi.org/10.1145/2157689.2157839

[4] Elisabeth Ganal, Lenny Siol, and Birgit Lugrin. 2023. PePUT: A Unity Toolkit for the Social Robot Pepper. In *2023 32nd IEEE International Conference on Robot and*

Human Interactive Communication (RO-MAN). IEEE, 1012–1019.

[5] Patrick Holthaus, Trenton Schulz, Gabriella Lakatos, and Rebekka Soma. 2023. Communicative Robot Signals: Presenting a New Typology for Human-Robot Interaction. In *International Conference on Human-Robot Interaction (HRI 2023)*. ACM/IEEE, Stockholm, Sweden, 132–141. https://doi.org/10.1145/3568162.3578631

[6] Julian Hough and David Schlangen. 2016. Investigating fluidity for human-robot interaction with real-time, real-world grounding strategies. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 288–298.

[7] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (2015), 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

[8] Ingo Lütkebohle, Frank Hegel, Simon Schulz, Matthias Hackel, Britta Wrede, Sven Wachsmuth, and Gerhard Sagerer. 2010. The bielefeld anthropomorphic robot head "Flobi". In *2010 IEEE international conference on robotics and automation*. IEEE, 3384–3391.

[9] James Simpson, Michael Richardson, and Deborah Richards. 2022. A Wizard or a Fool? Initial Assessment of a Wizard of Oz Agent Supporting Collaborative Virtual Environments. In *Proceedings of the 10th International Conference on Human-Agent Interaction*. 299–301.

[10] 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 HRI (VAM-HRI)*.

[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.

[12] Murphy Wonsick and Taskin Padir. 2020. A systematic review of virtual reality interfaces for controlling and interacting with robots. *Applied Sciences* 10, 24 (2020), 9051.