Opportunities and Challenges in Mixed-Reality for an Inclusive Human-Robot Collaboration Environment*

Extended Abstract[†]

Stephanie Arévalo Arboleda
Westphalian University of Applied Sciences
Germany
stephanie.arevalo@w-hs.de

Max Pascher
Westphalian University of Applied Sciences
Germany
max.pascher@w-hs.de

Jens Gerken
Westphalian University of Applied Sciences
Germany
jens.gerken@ w-hs.de

ABSTRACT

This paper presents an approach to enhance robot control using Mixed-Reality. It highlights the opportunities and challenges in the interaction design to achieve a Human-Robot Collaborative environment. In fact, Human-Robot Collaboration is the perfect space for social inclusion. It enables people, who suffer severe physical impairments, to interact with the environment by providing them movement control of an external robotic arm. Now, when discussing about robot control it is important to reduce the visual-split that different input and output modalities carry. Therefore, Mixed-Reality is of particular interest when trying to ease communication between humans and robotic systems.

KEYWORDS

Mixed-Reality, Robot Control, Human-Robot Collaboration, Severe Motor Impaired.

1 INTRODUCTION

Science fiction has encouraged people to imagine a world where technology enhances human capabilities. Technological advances in multiple areas, e.g. sensor technologies, artificial intelligence and display technologies allow researchers and practitioners alike to find new ways to amplify and augment human perception, physiology and cognition [1]. The field of Human-Robot Interaction (HRI) is on the forefront of these developments, as it is concerned with all these aspects.

We venture into HRI through a research project that aims to empower people with severe motor impairments, e.g. tetraplegics, to participate in activities of work and daily life in a self-determined way (see Fig. 1). Due to the limb movement constraint that our user group possesses, we have limited human input modalities available in such a HRI scenario. Furthermore, our users are unable to quickly move themselves around in physical space. The robot technology we are concerned with provides semi-autonomous functionality, thereby requiring the user to sometimes intervene with autonomous tasks (e.g. in case of failures due to the robot), but also completely overtake control for certain tasks (e.g. direct control yet beyond the capabilities of an autonomous system). Thereby, our research explores the continuum between teleoperation and peer-to-peer collaboration [2].

In this context, we find Mixed-Reality (MR) to be of particular interest, as prior research has often focused on enhancing the robot's perception with sensor technologies [4, 5], but left the human with traditional input and output devices. In contrast, MR has the capability to provide completely new ways of how humans perceive the interaction with the robot, particular in form of head-mounted Augmented-Reality (AR) glasses such as Microsoft HoloLens.

One of the main challenges we see in HRI, from a human perspective, is the general need of a mediatory interface. While our daily interaction with technology has gotten more direct in the last years to the extent that often input and output interaction space have merged (e.g. in the context of touch screens), this is not the case with HRI. Here, such mediatory interfaces often are disruptive in a sense that it is not possible for the user to keep both the interface and the robot conducting the task in focus. Commercial robot systems normally come with an external touch screen as operating device, requiring the user to shift attention between the robot and the device [6]. From our observations, this cannot only limit efficiency and effectiveness while performing tasks but also impose security risks in the context of human-robot collaborative work environments.

In this paper, we would like to present our understanding of the design opportunities and challenges for Mixed-Reality in Human-Robot Interaction. In doing so, we will keep a focus on our specific research scenario of supporting people with severe motor impairments.

*Produces the permission block, and copyright information
[†]The full version of the author's guide is available as acmart.pdf document

VAM-HRI, March 2018, Chicago, Illinois USA

Figure 1: Tetraplegic user operating a Universal Robots UR5 robot arm with MARG head-mounted sensors [3].

2 RELATED WORK

One of the first applications of AR in Human-Robot control relates to teleoperation for robot control. Milgram et al. emphasize the use of AR for “achieving spatial human-robot communication” which provides the following advantages: expanding the visualization of the environment, serving as a mediator between humans and a robotic system, enabling action optimization and warning the user of operational failures [7].

The use of MR in a HRI context has become more popular during the last years. One of the most explored areas within mixed and augmented reality is remote robot control. For instance, enabling decision making based on the robot’s affordances by providing spatial information through a virtual environment augmented with real video information [8].

Kobashi et al. consider the use of MR in the development of a humanoid robot that plans its actions in a virtual environment and interacts with the real one. It provides the person who is controlling the robot an idea of possible results, facilitating then the controller’s decision making process [9]. In addition, trying to achieve better information exchange when using robots for object manipulation, Frank et al. propose mixed reality interfaces to enhance communication of spatial information with robots for object manipulation [10].

A different approach that could be discussed in the context of robot control is the use of drone technology. Drones have allowed people to explore areas that are hard to reach or present dangers to humans. In a way, it has enabled humans to see through a robotic system’s eyes. Artemiadis goes beyond the fact of just seeing the world and proposes the control of multiple devices, by the use of Brain-Machine Interfaces (BMIs), in cases where the use of joysticks is not feasible [11]. In this way, leaving room for exploring new interaction interfaces to control robotic systems.

2 DESIGN OPPORTUNITIES IN USING MIXED-REALITY FOR HRI

As we briefly discussed in the introduction, we see one of the main opportunities in using Mixed-Reality for HRI in the possibility to bridge the gap between input and output interaction spaces. To be more precise, it foremost allows us to have one unified output space, meaning the intermediary interface is merged with the real-world view of the robot & task. In a Mixed-Reality setting we are able to augment the real-world view of the environment, including the robot and task, with both feedback from and interaction controls towards the robot, thereby providing any digital visual information in the operator’s line of sight. Furthermore, the input space is also linked closely to this newly unified output space – for example typical hand gestures when using a Microsoft HoloLens must be performed in the line of sight. In our scenario, hand gestures are not applicable. Therefore, we are looking for example at gaze-based or head movement interaction – again, with a MR or AR head mounted device, those inputs are closely linked to the output space and provide a very direct interaction style.

We think that by bringing input and output space closer together, awareness of actions will be improved, decision taking should be facilitated and cognitive load caused by divided attention can be reduced.

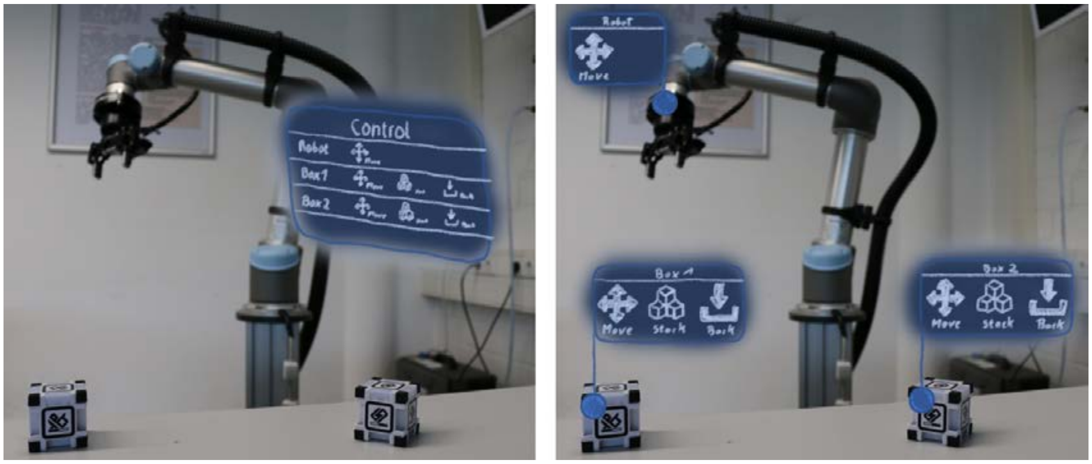

Figure 1: Variations of the interface (design mock-up): (left: 1) Centralized interface as a holographic window next to the objects (right: 2) Distributed interface with an option panel over the objects

A second opportunity arises from a certain problem, we came across in our prior research in this field, which is concerned with decision making. People have a hard time to understand and correctly predict the movement of industrial robot arms, as those do not suffer the same limitations as humans' joints. In a supervisory control setup, this may lead to situations in which the operator decides to take action although the robot is following a certain task correctly - or even worse, not taking action when one is needed. In a Mixed-Reality setting, we think that such situations could be reduced by providing visual movement predictions in real time. For instance, displaying superimposed images showing the trajectory to reach an object could be one worthwhile approach. Again, it is crucial that such information is presented in the line of sight of the operator, as they can thereby judge if the information is correct and the chances of missing them because of a split attention situation are minimized.

3 DESIGN CHALLENGES IN USING MIXED-REALITY FOR HRI

For all these opportunities, we also identified several design challenges which arise by using MR in the context of robot control.

A first important challenge is the specific design of the user interface in a MR or AR approach. As already mentioned, MR enables to present information and possible actions in the line of sight of the user. However, how these are organized and laid out and how that impacts for example the Usability is not yet clear. We see at least three possible approaches for a User Interface (UI) design. 1. A centralized "flat" interface could provide feedback information and controls, similar to current external mediatory interfaces. 2. By making use of the depth tracking capabilities of modern AR glasses such as HoloLens, interface controls and feedback components could be augmenting the real-world objects in 3D space. We call this a distributed interface (See 1 and 2 in Fig. 2) 3. Another option would be to take the "drone control" approach and putting the user in the perspective of the robot arm. Here, a MR approach closer to the VR spectrum of the continuum might be necessary. We call this an egocentric interface.

The centralized interface is closest to the kind of interfaces people are used to from interacting with today's technologies but is limited in the way it capitalizes on the possibilities of Mixed-Reality. Instead, the distributed interface makes especially use of these, while raising the challenge of not being less efficient as actions and feedback might be spread out and outside of our current attention focus. The egocentric interface allows a very natural and direct way of interaction, as the robot becomes much more an extension of ones' body. However, providing a perfect egocentric view of the environment out of the "robots' eyes" is challenging and might not always be effective, when e.g. an observer perspective provides more context information and thereby may allow more precise control.

Another challenge, especially in our specific scenario with severe motor impaired users, is the integration of input techniques. As already mentioned, typical hand gestures are not applicable. Three approaches have shown potential for future applications in our research. 1. MARG or IMU sensors which make use of the head position and thereby allow head movements as input commands [3]. The HoloLens already has this type of input integrated. 2. Speech interaction as an input modality, which has reached new levels with the success of speech assistants such as Apple Siri or Google Assistant [12, 13]. 3. Gaze-based interaction, making use of precise eye-gaze movements and fixations. The latter has been used more and more to allow people with motor impairments to interact with computers.

However, it cannot be generalized that the use of one modality suits all use cases. That is why, we need to consider the use of different modalities and the combination of these, thereby designing multi-modal approaches.

When combining input and output as presented, another challenge arises: Should we only interact with digital UI elements or can some interaction be done implicitly by tracking the "interaction" with the real-world? The reason why this is important is that a robot arm such as the UR5 has 7 degrees of freedom (DoF), including the grip. With head movements, we are only able to address 3 DoF directly. The challenge gets event bigger if we have multiple robot arms in more complex manufacturing scenarios. This then means that even for just moving the robot arm around in a remote- control scenario, we would have to switch between different operating modes. Therefore, we think it is important to consider and explore what we call implicit mode switching, where for example by simply looking at one specific robot arm, the interaction is started with no need to explicitly switch between robot arms via a centralized UI.

In addition, there are two types of gaze-based interactions as input modalities which may be of use: fixation and smooth-pursuit [14]. By the use of fixation, the user can interact and control elements by dwelling at an object for a specific time. While this is generalizable to most standard user interface controls, it requires a high precision tracking and asks the user to explicitly "hold still", which is not a natural viewing behavior. Through smooth-pursuit the user instead follows a small target with their gaze in order to invoke certain controls. Each such control must have their own target or stimuli element with a unique movement pattern. To identify the desired action, the system only needs to correlate the user's gaze with the trajectory of potential stimuli, with no need for precise and calibrated gaze tracking.

3 CONCLUSIONS

In summary, in our research we explore the opportunities and challenges of using MR in human-robot collaboration. Our focus is to bring input and output modalities closer together, for presenting information and feedback in the line of sight and even context-specific by augmenting the real world environment in 3D space. We seek for interaction opportunities for people with severe motor impairments. In this context, by using MR technology, we propose 3 different UIs that merge the I/O modalities: centralized, distributed and egocentric, which can be used on different scenarios. This, leads us to some challenges regarding the input techniques, we consider integrating gaze, speech and MARG sensors to achieve a wider spectrum of use cases.

This research will serve as a base for future empirical studies and gather feedback about our design criteria. Beyond this, we think that many of our observations might be generalizable to the general context of MR and HRI.

ACKNOWLEDGMENTS

This research is supported by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

[1] A. Schmidt, S. Schneegass, K. Kunze, J. Rekimoto, and W. Woo, "Workshop on Amplification and Augmentation of Human Perception," 2017, pp. 668–673.

[2] M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," Found. Trends Hum.-Comput. Interact., vol. 1, no. 3, pp. 203–275, 2007.

[3] N. Rudigkeit, M. Gebhard, and A. Gräser, "Evaluation of control modes for head motion-based control with motion sensors," in 2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings, 2015, pp. 135–140.

[4] T. Matsuda, Robot Vision : New Research. New York: Nova Science Publishers, Inc, 2009.

[5] T. Peynot, S. Monteiro, A. Kelly, and M. Devy, "Editorial: Special issue on Alternative Sensing Techniques for Robot Perception," J. Field Robot., vol. 32, no. 1, pp. 1–2, Jan. 2015.

[6] "KUKA smartPAD," KUKA AG. [Online]. Available: https://www.kuka.com/en-gb/products/robotics-systems/robot-controllers/smartpad. [Accessed: 07-Feb-2018].

[7] P. Milgram, S. Zhai, D. Drascic, and J. Grodski, "Applications of augmented reality for human-robot communication," 1993, vol. 3, pp. 1467–1472.

[8] C. W. Nielsen, M. A. Goodrich, and R. W. Ricks, "Ecological Interfaces for Improving Mobile Robot Teleoperation," IEEE Trans. Robot., vol. 23, no. 5, pp. 927–941, Oct. 2007.

[9] K. Kobayashi, K. Nishiwaki, S. Uchiyama, H. Yamamoto, S. Kagami, and T. Kanade, "Overlay What Humanoid Robot Perceives and Thinks to the Real-world by Mixed Reality System," in Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, Washington, DC, USA, 2007, pp. 1–2.

[10] J. A. Frank, M. Moorhead, and V. Kapila, "Mobile Mixed-Reality Interfaces That Enhance Human–Robot Interaction in Shared Spaces," Front. Robot. AI, vol. 4, 2017.

[11] P. Artemiadis, "Brain-Swarm Control Interfaces: The Transition from Controlling One Robot to a Swarm of Robots," Adv. Robot. Autom., vol. 06, no. 01, 2017.

[12] "iOS - Siri," Apple. [Online]. Available: https://www.apple.com/ios/siri/. [Accessed: 08-Feb-2018].

[13] "Google Assistant - Your own personal Google," Google Assistant - Your own personal Google. [Online]. Available: https://assistant.google.com/. [Accessed: 08-Feb-2018].

[14] A. Esteves, D. Verweij, L. Suraiya, R. Islam, Y. Lee, and I. Oakley, "SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality," 2017, pp. 167–178.