# A Discussion on Current Augmented Reality Concepts Which Help Users to Better Understand and Manipulate Robot Behavior

Kai Groetenhardt[1]

**Abstract:**

For a safer, more trustful, and more dynamic collaboration, humans should understand and be able to manipulate the behavior of robots they are interacting with. Therefore, a way for a meaningful communication has to be established that takes place in a common perceptual space. One way to accomplish that is to use augmented reality (AR) in which the robot is able to display information for the human in 3D space, and the human can send commands to the robot using interaction methods provided by AR devices. In this work, a brief overview of AR concepts is given and discussed. They are divided into three categories: (1) understanding the movement of robots, (2) understanding the internal states of robots, and (3) manipulating robot behavior. Whereas (1) and (2) already show a number of promising approaches, and (3) is still in need for more innovative ideas.

**Keywords:** robot behavior; human-robot interaction; augmented reality

## 1 Introduction

Looking at active research in robotics, it is imaginable that robots will increasingly find their way into private households. Therefore, it seems important that humans without technical background are able to understand the behavior of robots and can control them. Breazeal et al. [Br01] identified an overlapping perceptual space as a key requirement for effective human-robot interaction, and Collett et al. [CM06] further explains that the perceptual space differs in input and output, which is illustrated in Fig. 1. Consequently, one obstacle is that humans and robots have different perceptual spaces, which just partially overlap. Not every action a human can perform is within the perceptual space of a robot, and a robot cannot reach every form of perception that is available to a human. Furthermore, humans and robots have differing conceptual models of the world; robots use several sensors to collect data of their surroundings and use various routines to interpret it. Sometimes they also need knowledge about the real world, which is provided by a knowledge framework. Those different perceptual spaces and different conceptual models of the world are reasons why understanding robot behavior can be difficult for humans.

---

[1] RheinMain University of Applied Sciences, Faculty of Design - Computer Science - Media, Unter den Eichen 5, 65195 Wiesbaden, Germany, kai.groetenhardt@hs-rm.de

doi:10.18420/inf2020_90

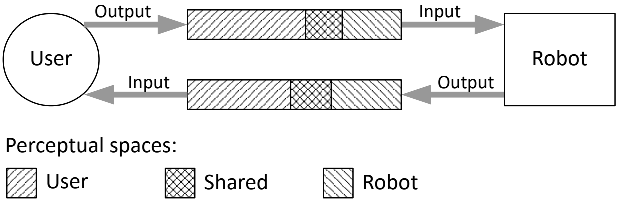


Fig. 1: An illustration showing the concept of the perceptual spaces of a robot and a human overlapping wherein a meaningful communication is possible. Source: Self-made rework of Figure 1 in [CM06].

Augmented reality (AR) can be a tool to widen the overlapping parts of the perceptual spaces of humans and robots, and is therefore able to make understanding robots easier for humans. Robots can send information to AR devices, which then can be visualized for the humans to perceive, and humans can use the interaction possibilities of those devices to send commands to the robot.

The aim of this paper is to give a brief overview of current work regarding human-robot interaction (HRI) with the focus on understanding and manipulating robot behavior using AR, and discuss possible future work. This paper starts with the categorization, presentation and discussion of the state of the art in Section 2. In Section 3 follows a conclusion summarizing what was learned, and suggesting ideas for future work to hopefully enhance the topic of using AR for HRI further.

## 2 Discussing State of the Art

AR in combination with robotics is not a new topic. In fact, the literature review of Geen et al. [Gr08] about human-robot collaboration AR approaches contains papers dating back to the early 2000s addressing that topic. In the past, efforts using AR were hampered by limitations in the available AR head-mounted displays (HMDs), which often were custom-built. With a new batch of AR HMDs like the *Microsoft HoloLens*[2] and the *Magic Leap 1*[3] some limitations of the past were reduced or eliminated and enabled researchers to conceptualize and implement new AR concepts for HRI.

The following papers, describing the usage of AR for HRI, are divided into the three categories (1) understanding the movement of robots, (2) understanding the internal states of robots, and (3) manipulating robot behavior. Each category is limited to a maximum of two papers to not exceed the scope of this work. If a category contains two papers, they are chosen to be similar for a better comparability, but they are different regarding their goals and approaches. Those papers are not depicted to the full extent, instead their presentment is limited to the most relevant parts. To every paper the motivation is stated, followed by the

---

[2] https://www.microsoft.com/en-us/hololens

[3] https://www.magicleap.com/en-us

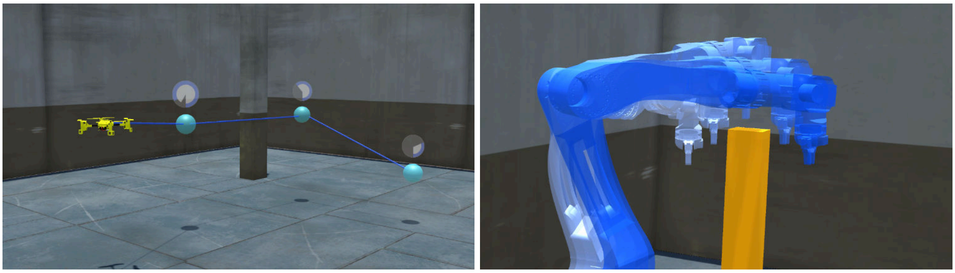
description of its AR concept, and concluded with the results of a user study, if available. At the end of each category, the concepts are being discussed.

## 2.1 Understanding the Movements of Robots

One defining characteristic of robots is their ability to move within the real world, sometimes in collaborative work spaces attended by humans. Unfortunately, robots do not necessarily have the ability to communicate their motion through gestures, gaze, or other social cues like humans. Here, two papers are presented showing possibilities to help the user understand the movement of robots with the help of AR.

Walker et al. [Wa18] argue that there are difficulties identifying when, where, and how a robot will move, which represents a primary challenge towards achieving safe and usable robotic systems. To tackle that problem, they introduce four concepts to indicate future movements of a flying robot using the AR HMD *Microsoft HoloLens*. The first concept, called *NavPoints* (shown in Fig. 2 (A)), adds virtual navigation points displayed as spheres into the 3D space. The spheres are connected through lines, which indicate in what order the robot will pass them. Above the spheres two radial timers are displayed, which show when the robot will arrive and when it will leave that position. The second concept, which is called *Arrow*, is similar, but a more minimal approach. An animated arrow shows the route the robot will take a few seconds in the future. As the arrow moves it leaves a line behind showing the path it was taking. The third concept is called *Gaze*, which augments the robot by a 3D model of an eye that is looking in the direction of travel. The fourth and last concept they presented is named *Utilities*. It is a 2D circular radar displayed at a corner of the user's perceptual space that shows the robots position relative to the user. Eventually, they compared the concepts by conducting a user study to see, among other things, how the displayed virtual imagery affected participant understanding of robot movement intent. The test showed that *NavPoints* ranked best followed by *Arrow*, *Gaze* and *Utilities*.

Rosen et al. [Ro20] indicate that a robot's movement intent can be shown on a 2D screen, but this requires the human to take their attention away from the robot's physical space to observe the display, which could be dangerous. Additionally, a 2D projection of a 3D motion can take time for a human to understand, requiring interaction to inspect different points of view. As a test scenario they chose a robot arm that performs a programmed movement with some objects nearby. The task for the human is to check if the robot will hit the objects before it even starts moving. To make that possible, a virtual 3D model of the robot arm is displayed multiple times along the planned path in 3D space visible through the AR HMD *Microsoft HoloLens*, shown in Fig. 2 (B). To have a reference, they implemented that same concept for a 2D screen with the possibility to move the virtual camera via mouse and keyboard. They compared both concepts and found that their AR system reduced the completion time of the task and increased the average accuracy of collision predictions.



(A) (B)

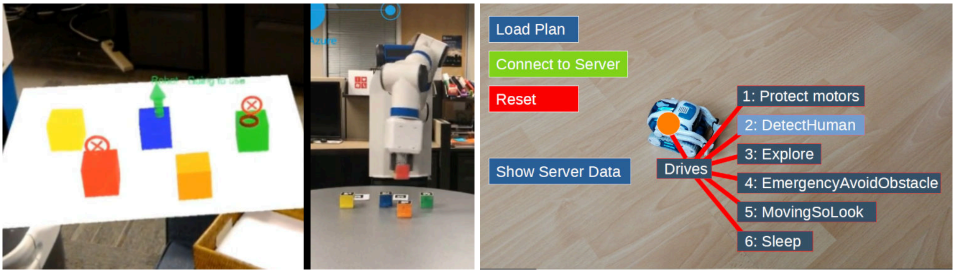
Fig. 2: Two different concepts communicating robot movement intent to humans. (A) shows the concept of view through the *Microsoft HoloLens* displaying the *NavPoints* concept from Walker et al. [Wa18] in which connected waypoints, the arrival, and the departure time of a robot can be seen. Source: Self-made rework of Figure 1A in [Wa18]. In (B) the concept of Rosen et al. [Ro20] is shown in which several steps of the planned movement are displayed in full size. Source: Self-made a rework of Figure 1 in [Ro20].

Comparing the concepts of Walker et al. and Rosen et al., it is apparent that they both show the robots motion intents, but target different scenarios. Walker et al. show where the robot will be located for users to adapt their own behavior towards the robot. Rosen et al. show the robot's future movement for the human to be able to intervene in the robot's behavior. It is imaginable to combine both concepts, but divide them into a planning and execution mode, which can be switched by the user. For the planning mode, the concept of Rosen et al. could be used to see detailed movements and to identify collisions. In execution mode, the concept of Walker et al. would show the path of the robot and when it will reach waypoints.

## 2.2 Understanding Internal States of Robots

To not only understand the movement of robots, but also the robots' decision-making process that leads to movements or other actions, an interface to the humans' perceptual space needs to be established. In this section, two papers are discussed showing a robot's plan of action via AR.

Chakraborti et al. [Ch18] cite the *Roadmap for U.S. Robotics* [Ch09] by saying "humans must be able to read and recognize robot activities in order to interpret the robot's understanding". They argue that attempts were made to accomplish the idea with natural language, but the state of the art limits the scope of such interactions, especially where precise instructions are required. To show an alternative, they communicate the intentions of a robot using AR to a collaborating human. Their setup consists of a robot that is tasked to stack colored boxes and a human who is equipped with a *Microsoft HoloLens* and has the ability to claim boxes through an AR interface. A virtual 3D model of boxes mirroring the boxes that are positioned in front of the robot is displayed for the human in 3D space, as can be seen in



(A) (B)

Fig. 3: Two different concepts to communicate a robot's intent to a human. (A) shows the concept of Chakraborti et al. [Ch18]. On the left hand side, the view through the *Microsoft HoloLens* and the mirrored model of the boxes in front of the robot can be seen. The virtual boxes are annotated with symbols that indicate the robot's plan. On the right hand side, the robot is displayed while executing its task of stacking boxes. Source: A frame taken from a video linked in [CSKK18] with kind permission of Tathagata Chakraborti. In (B), the concept of the *Android* AR app of Rotsidis et al. [Ro19] can be seen that shows the robot annotated with its current plan represented as a hierarchical graph in which the active task "DetectHuman" is highlighted. Source: Self-made rework of Figure 2 in [Ro19].

Fig. 3 (A). Those virtual boxes can be annotated by the robot to show what its intentions regarding those boxes are. The robot marks a box with a green upward pointing arrow to communicate that this box is the next one the robot is going to pick up. Also, boxes the robot intends to use in the future are marked with a circled red cross. The human on the other hand, has the ability to claim boxes for themselves even if the robot already has indicated to use them. In that case, the robot removes the mark at the corresponding box and chooses another one to complete its task. Unfortunately, there is not a user study yet, but Chakraborti et al. announced their intention to conduct one.

Rotsidis et al. [Ro19] state that it's important for end-users to have a mental model of their robot that contains the capabilities and awareness of its limitations in order to trust it. Subsequently, through transparent decision-making of the robot it is possible for the users to adjust their expectations and forecast certain actions of the robot. The authors' attempt to tackle that challenge is based on an AR application running on an *Android* handheld device that shows the plan of the robot in form of an hierarchical graph. If the app detects the robot, it displays the graph next to it. The graph shows tasks the robot is able to perform and highlights the task the robot is currently executing, which can be seen in Fig. 3 (B). The user has the possibility to interact with the graph to see more or less information. They conducted a user study that showed the robot is perceived more alive, livelier, and friendlier with the app than without it.

Interestingly, Chakraborti et al. chose to show a virtual copy of the real objects and annotated them instead of annotating the real objects directly in AR. The authors did not disclose why the went this way, but it would be interesting to find out if direct annotations could improve the usability. Comparing the concepts of Chakraborti et al. and Rotsidis et al., it is clear that both show the robots' plan, but, like in the previous Section 2.2, one is more detailed in its approach. Rotsidis et al. only show what task is being executed, whereas Chakraborti et al. also show how the current task is being executed. Additionally, Chakraborti et al. developed a specific vocabulary to communicate the robots plan in form of annotations. In contrast, Rotsidis et al. communicate the plan via text arranged in a graph. Of course, a more detailed approach is not always the better choice since too much information could lead to problems of its own, for example by overloading the user or showing unwanted information. It needs to be determined in what scenario one concept is more suited than the other, or if a scalable solution combining the two concepts would be the better approach.

## 2.3 Manipulating Robot Behavior

If a robot needs to be taught how to execute a new task or change its behavior, a typical way to achieve this is to reprogram it by using text-based or even visual programming languages. But there are also other approaches like programming by demonstration (PbD), which is a field of research of its own that also includes AR solutions (e. g. [OK18], [Qu18]), or rather unconventional approaches like knowledge patching used in the following paper.

Liu et al. [Li18] point out that machine learning methods have reached a remarkable level of effectiveness in specific tasks, but still have their limitations. For example, they lack interpretability of the knowledge representation, especially about how and why a decision is made, which plays a vital role in the scenarios where robots work alongside humans. In their system, they use interpretable knowledge represented by an And-Or-Graph (AOG) instead. Their setup consists of a robot with two arms and a *Microsoft HoloLens*, which can, among other things, display a 2D interface for the AOG in 3D space in front of the robot. The task to be solved is for the robot to open a medicine bottle with a lid that does not only have to be twisted but also pushed. The user starts with an AOG that describes how to open a normal bottle. The robot needs to be taught a push movement using PbD within the AR environment for it to be patched into the AOG by the human. To do the patching, the human can interact with the graph using hand gestures, comparable to mouse clicks and mouse drags, to remove and add nodes. The interface can be seen in Fig. 4, also a video [In] showing the whole process is available.

One could argue that the described part of Liu's et al. concept is a movement snippet manager that allows the user to combine little movements into a more complex motion, which is an interesting approach to not overburden the user. In their paper, they changed the process of opening a bottle globally, which means even non-medical bottles get opened using that push and twist movement. It would be interesting to see this concept combined with some sort of a teachable object recognition system to be able to chose the opening process in a more targeted manner. Consequently, that would need to be implemented into the AR interface, which would present a new challenge. Researching robot behavior modeling via AR, unfortunately, revealed very little approaches outside the PbD field, which

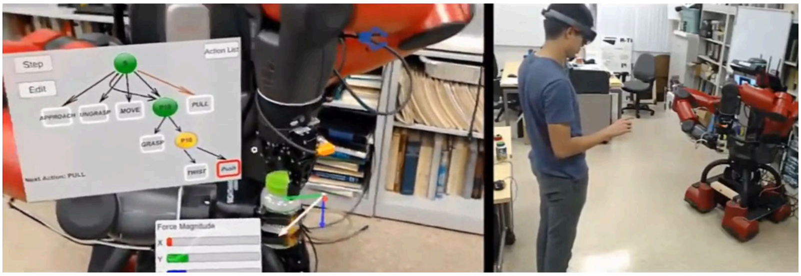


Fig. 4: The concept of Liu et al. [Li18] from two different perspectives. On the left hand side, the view trough the *Microsoft HoloLens* is displayed in which the human can see the AOG representing the knowledge to open a medicine bottle. On the right hand side, a third person's view without virtual elements is displayed. Source: A frame taken from a video [In] linked in [Li18] with kind permission of Hangxin Liu.

resulted in only one paper in this section without something comparable, which leaves that topic open for more innovative ideas.

## 3 Conclusion and Future Work

In this work, a brief overview of AR concepts dealing with robot behavior was given and discussed. The state of the art shows that there are already several concepts proven to be helpful in understanding robot behavior. Others look promising, but their effectiveness needs to be tested. The presented papers differ in their aims and level of detail in a way that makes them prone to be combined. Combinations of the described concepts could lead to improvements that would be interesting to see in future work. All things considered, using AR to understand robots seems to be a viable approach to further pursue.

In contrast, more accessible interfaces to manipulate robot behavior in AR seem to be a difficult endeavor. After a thorough research, only one paper could be found that chooses an approach (at least partly) deviant to PbD. More ideas need to be developed and tested to see if AR is the right tool to manipulate robot behavior.

During the discussion, some suggestions for improvements were made, which could be conceptualized in more detail in future work. Especially the concept of Chakraborti et al. [Ch18] is an interesting candidate to pursue further to see if annotating real objects instead of the virtual copies of them feels more natural for the users.

## Bibliography

[Br01] Breazeal, Cynthia; Edsinger, Aaron; Fitzpatrick, Paul; Scassellati, Brian: Active vision for sociable robots. IEEE Transactions on systems, man, and cybernetics-part A: Systems and Humans, 31(5):443–453, 2001.

[Ch09] Christensen, Henrik I; Batzinger, T; Bekris, K; Bohringer, K; Bordogna, J; Bradski, G; Brock, O; Burnstein, J; Fuhlbrigge, T; Eastman, R et al.: A roadmap for US robotics: from internet to robotics. Computing Community Consortium, 44, 2009.

[Ch18] Chakraborti, Tathagata; Sreedharan, Sarath; Kulkarni, Anagha; Kambhampati, Subbarao: Projection-aware task planning and execution for human-in-the-loop operation of robots in a mixed-reality workspace. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 4476–4482, 2018.

[CM06] Collett, T. H. J.; MacDonald, B. A.: Developer Oriented Visualisation of a Robot Program. In: Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction. HRI ’06, ACM, New York, NY, USA, pp. 49–56, 2006.

[Gr08] Green, Scott A; Billinghurst, Mark; Chen, XiaoQi; Chase, J Geoffrey: Human-robot collaboration: A literature review and augmented reality approach in design. International journal of advanced robotic systems, 5(1):1, 2008.

[In] Interactive Robot Knowledge Patching Using Augmented Reality - YouTube. `https://www.youtube.com/watch?v=AqjmIhKGGus`. accessed: 07/04/2020.

[Li18] Liu, Hangxin; Zhang, Yaofang; Si, Wenwen; Xie, Xu; Zhu, Yixin; Zhu, Song-Chun: Interactive robot knowledge patching using augmented reality. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 1947–1954, 2018.

[OK18] Ostanin, M; Klimchik, A: Interactive robot programing using mixed reality. IFAC-PapersOnLine, 51(22):50–55, 2018.

[Qu18] Quintero, Camilo Perez; Li, Sarah; Pan, Matthew KXJ; Chan, Wesley P; Van der Loos, HF Machiel; Croft, Elizabeth: Robot programming through augmented trajectories in augmented reality. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 1838–1844, 2018.

[Ro19] Rotsidis, Alexandros; Theodorou, Andreas; Bryson, Joanna J; Wortham, Robert H: Improving robot transparency: An investigation with mobile augmented reality. In: 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). IEEE, pp. 1–8, 2019.

[Ro20] Rosen, Eric; Whitney, David; Phillips, Elizabeth; Chien, Gary; Tompkin, James; Konidaris, George; Tellex, Stefanie: Communicating robot arm motion intent through mixed reality head-mounted displays. In: Robotics Research, pp. 301–316. Springer, 2020.

[Wa18] Walker, Michael; Hedayati, Hooman; Lee, Jennifer; Szafir, Daniel: Communicating Robot Motion Intent with Augmented Reality. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. HRI ’18, Association for Computing Machinery, New York, NY, USA, p. 316–324, 2018.