Designing Indicators to Show a Robot's Physical Vision Capability

Hong Wang*   Tam Do†   Zhao Han‡

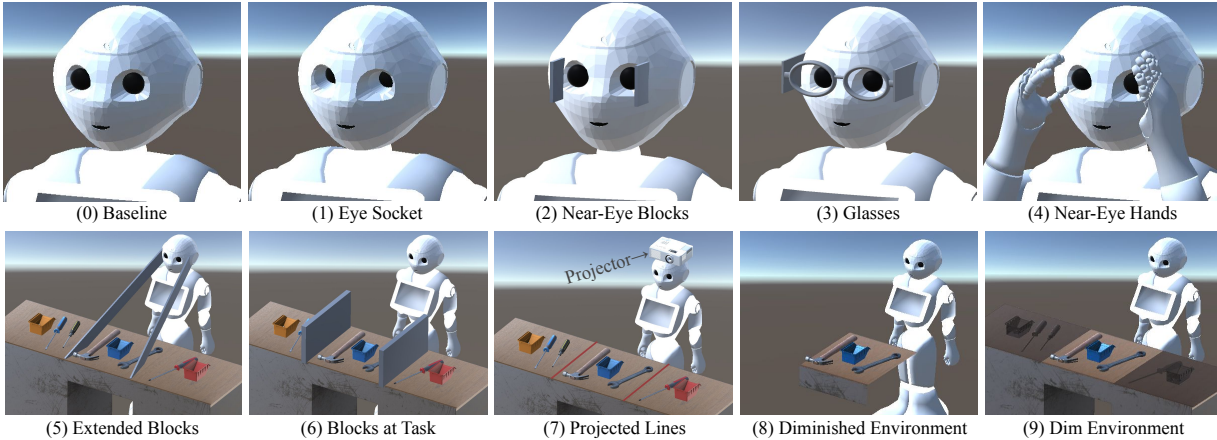RARE Lab, Department of Computer Science and Engineering, University of South Florida

(0) Baseline	(1) Eye Socket	(2) Near-Eye Blocks	(3) Glasses	(4) Near-Eye Hands
(5) Extended Blocks	(6) Blocks at Task	(7) Projected Lines	(8) Diminished Environment	(9) Dim Environment

Figure 1: Nine indicator designs with a baseline in the beginning, to be registered to the real robot. See next page for details.

ABSTRACT

In human-robot interaction (HRI), studies show humans can mistakenly assume that robots and humans have the same field of view, possessing an inaccurate mental model of a robot. This misperception is problematic during collaborative HRI tasks where robots might be asked to complete impossible tasks about out-of-view objects. In this initial work, we aim to align humans' mental models of robots by exploring the design of field-of-view indicators in augmented reality (AR). Specifically, we rendered nine such indicators from the head to the task space, and plan to register them onto the real robot and conduct human-subjects studies.

Index Terms: Computer systems organization—Robotics; Human-centered computing—Mixed / augmented reality

1 INTRODUCTION

Mental models are structured knowledge systems that enable people to engage with their surroundings [1]. During human-human interactions, a shared mental model was shown to improve team performance with a mutual understanding of the task, which is true not only among human teams but also in human-agent teams [2], applicable to physically embodied agents like robots.

However, in human-robot teaming and collaboration scenarios, because robots more or less resemble humans, humans can form an inaccurate mental model of robots' capabilities, leading to mental model misalignment. Frijns et al. [3] noticed this problem and proposed an asymmetric interaction model: Unlike symmetric interaction models where roles and capabilities are mirrored between humans and robots, asymmetric interaction models emphasize the distinct strengths and limitations of humans and robots.

---

*e-mail: hongw@usf.edu
†e-mail: tamdo@usf.edu
‡e-mail: zhaohan@usf.edu

One mental model misalignment case is the assumption that robots possess the same field of view (FoV) as humans. Although humans have over $180°$ FoV, a robot's camera typically has $60°$ horizontal FoV. This discrepancy can result in inefficiencies during interactions. Specifically, Han et al. [4] investigated how a robot can convey its incapability of handing a cup that is both out of reach and out of view. Yet, in the study context, participants assumed the cup on the table was within the robot's FoV, and expected the robot to successfully hand it to them. In those cases, robots benefit from a more accurate mental model, leading to fewer explanations and clearer instructions, e.g., "the cup on the right" rather than "the cup".

In this paper, we aim to address the FoV discrepancy by first exploring the design space. Specifically, we propose nine situated augmented reality (AR) indicators (Figure 1 except for the baseline in Figure 1.0) to align human expectations with the real vision capability of robots. We are particularly interested in AR as the robot's hardware is hard to modify, and AR allows fast prototyping to narrow down the design space. Moreover, the situatedness allows placing indicators near the eyes, which have the FoV property, and the task objects, which are negatively affected by FoV misperception, rather than, e.g., on a robot's FoV-irrelevant chest screen. With those indicators rendered, we plan to register them onto the robot and conduct user studies to narrow down and evaluate the designs.

2 RELATED WORK

AR for Robotics: Robotics researchers have integrated AR in multiple domains. Examples include AR systems for fault visualization in industrial robots [5], integration in robotic surgical tools [6], and AR-enhanced robotics education for interaction [7]. These studies showed the potential of AR in enhancing human-robot interactions.

AR Design Elements for HRI: From the literature, Walker et al. [8] developed a taxonomy and proposed four categories of design elements: Virtual Entities, Virtual Alterations, Robot Status Visualizations, and Robot Comprehension Visualizations. Our designs fall under "Virtual Alterations – Morphological" (Figure 1.1 to 1.3) and "Virtual Entities - Environmental" (Figure 1.5 to 1.9).

AR for Robot Communication: One particular line of research related to our work is to leverage AR to better understand robot

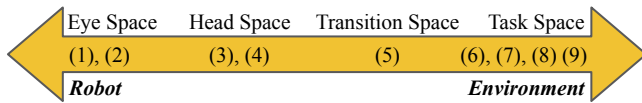| Eye Space	Head Space	Transition Space	Task Space
(1), (2)	(3), (4)	(5)	(6), (7), (8) (9)
Robot			*Environment*

Figure 2: Our design spectrum: robot to environment, and vice versa.

actions and intentions to improve transparency, thereby gaining trust and acceptance. For example, Han et al. [9] added an AR arm to a mobile robot to enable gesturing capabilities for communication. Walker et al. [10] demonstrated how AR can visually convey the intent of aerial robots that lack physical morphology for communication. In our work, we want the robot to implicitly communicate its FoV capability, so that interactants can form a correct mental model of the robot and provide clearer requests.

Mental State Attribution to Robots: Besides, researchers have also studied humans' automatic attribution of mental states *to* robots [11]. This line of work differs significantly from our work, which focuses on helping humans form a more accurate mental model *of* robots.

3 DESIGN TAXONOMY AND SPECTRUM

As robots are physically situated in our physical world, we grouped our designs into four connected areas between the robot and its environment, as shown in Figure 2. *Eye space* designs focus on the modifications at the robot's eyes, which possess the property of FoV. Examples include design (1) and (2) in Figure 1. *Near head space* closely ties to the robot's head around eyes. Examples include design (3) and (4) in Figure 1. *Transition space* includes designs that extend from the robot into its operating environment, such as design (5) in Figure 1. As the indicator moves closer to the task setting, we hypothesize that those designs will better help people to identify the performance effects of FoV. Finally, the designs grouped in *task space*, design (6) to (9) in Figure 1, are those not attached to the robot but rather placed in its working environment. Spectrum in Figure 2 offers a visual breakdown of our indicator designs, emphasizing the continuum from the robot space to the environment space.

4 DESIGNS: STRATEGIES TO INDICATE ROBOT'S FOV

To indicate the robot's FoV, we designed the following nine strategies. The number prefixes are the same as in Figure 1.

(1) Eye Socket: We deepen the robot's eye sockets using an AR overlay over the existing eye sockets. It creates a shadowing effect at the robot's eyes. As the sockets deepen, they physically limit what angle the eyes can see, thus matching the cameras' FoV. Note this design is only possible with AR, as physical alteration is difficult.

(2) Near-Eye Blocks: We add blocks directly to the sides of the robot's eyes to functionally block those outside of the camera's horizontal FoV. This design is possible both physically and in AR.

(3) Glasses: Similar to the near-eye block but with aesthetic purposes and more familiar in daily life, we add a pair of glasses with solid temples to obscure those outside of the field of view. Note that this design can also be added both physically or with AR.

(4) Near-Eye Hands: A straightforward approach that does not require any modifications or additions is for the robot to raise its hands directly to the sides of its eyes to act as visual indicators of the extent of its field of view. Practically, it can be used in a quick-start guide after the robot is shipped to the users without them wearing AR headsets. This design does not need AR or physical alteration.

(5) Extended Blocks: To more accurately show the range of the robot's field of view (e.g., which objects that the robot cannot see), we connect the blocks from the robot's head to the task environment, so that people know exactly how wide the robot can see. Note that this design can only be practically made possible with AR.

(6) Blocks at Task: A more task-centric design is to place the blocks to demonstrate the extent of the robot's field of view only in the robot's task environment, e.g., table. Unlike the last *Extended*

Blocks design, this one is in the environment rather than connected to the robot. Note that this design can also only be placed with AR.

(7) Projected Lines: Rather than near-eye AR displays or physical alteration, this design uses an overhead projector to project lines onto the robot's operating environment to indicate the robot's FoV. This projected AR technology frees interactants from wearing head-mounted displays or holding phones or tablets, thus making it ergonomic and scalable to a crowd, beneficial in group settings.

(8) Diminished Environment: Diminished reality is a technique to remove real objects from a real scene. We thus propose to delete everything that the robot cannot see, leaving only what is within the robot's field of view. Note that this design requires AR.

(9) Dim Environment: Rather than removing all the content that the robot cannot see, this design reduces the brightness of those content. Compared to diminished reality, we believe this design will help people maintain the awareness of the task environment. Similarly, this design also requires AR.

5 CONCLUSION AND FUTURE WORK

In this paper, we proposed nine designs to visually indicate a robot's physical vision capability, aiming to help align human expectations with a robot's actual capabilities for humans to form an accurate mental model of robots. We are planning to conduct user studies to narrow down as well as subjectively and objectively evaluate our designs. Tentative metrics include understandability of indicated FoV as well as accuracy and efficiency by measuring whether participants can identify out-of-view objects and how many are there.

REFERENCES

[1] J. R. Wilson and A. Rutherford, "Mental models: Theory and application in human factors," *Human Factors*, vol. 31, no. 6, pp. 617–634, 1989.

[2] C. M. Jonker, M. B. Van Riemsdijk, and B. Vermeulen, "Shared mental models: A conceptual analysis," in *Int'l Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, 2010, pp. 132–151.

[3] H. A. Frijns, O. Schürer, and S. T. Koeszegi, "Communication models in human–robot interaction: an asymmetric model of alterity in human–robot interaction (amodal-hri)," *Int'l Journal of Social Robotics*, vol. 15, no. 3, pp. 473–500, 2023.

[4] Z. Han, E. Phillips, and H. A. Yanco, "The need for verbal robot explanations and how people would like a robot to explain itself," *ACM THRI*, vol. 10, no. 4, 2021.

[5] G. Avalle, F. De Pace, C. Fornaro, F. Manuri, and A. Sanna, "An augmented reality system to support fault visualization in industrial robotic tasks," *IEEE Access*, vol. 7, 2019.

[6] S. Das and S. Vyas, "The utilization of AR/VR in robotic surgery: A study," in *Proceedings of the 4th Int'l Conf. on Information Management & Machine Intelligence*, 2022.

[7] M. Pozzi, U. Radhakrishnan, A. Rojo Agustí, K. Koumaditis, F. Chinello, J. C. Moreno, and M. Malvezzi, "Exploiting vr and ar technologies in education and training to inclusive robotics," in *Educational Robotics Int'l Conference*, 2021, pp. 115–126.

[8] M. Walker, T. Phung, T. Chakraborti, T. Williams, and D. Szafir, "Virtual, augmented, and mixed reality for HRI: A survey and virtual design element taxonomy," *ACM THRI*, vol. 12, no. 4, pp. 1–39, 2023.

[9] Z. Han, Y. Zhu, A. Phan, F. S. Garza, A. Castro, and T. Williams, "Crossing reality: Comparing physical and virtual robot deixis," in *2023 ACM/IEEE HRI*, 2023.

[10] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *Proceedings of the 2018 ACM/IEEE HRI*, 2018, pp. 316–324.

[11] S. Thellman, M. de Graaf, and T. Ziemke, "Mental state attribution to robots: A systematic review of conceptions, methods, and findings," *ACM THRI*, vol. 11, no. 4, pp. 1–51, 2022.