Exploring Augmented Reality Visualizations for Communicating Robot Intent in Robotic Ultrasound

Josefine Schreiter*
Otto-von-Guericke University
Research Campus STIMULATE
Magdeburg

Minh Hoang Tan
Otto-von-Guericke University
Research Campus STIMULATE
Magdeburg

Robert Klank
Otto-von-Guericke University
Research Campus STIMULATE
Magdeburg

Christian Hansen
Otto-von-Guericke University
Research Campus STIMULATE
Magdeburg

Fabian Joeres
Otto-von-Guericke University
Research Campus STIMULATE
Magdeburg

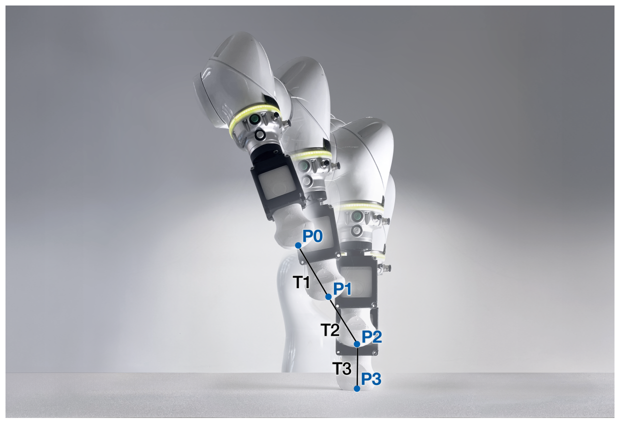


Figure 1: Initial positioning of a robotic ultrasound probe including predefined end effector poses P_0–P_3 and motion trajectories T_1–T_3 following an approach by Seitz et al. [9].

Abstract

Working safely in close proximity to robots, such as during robot-assisted medical procedures, requires communication of robot intent. We investigated how augmented reality (AR) visualizations can convey robot *poses*, *trajectories*, and *motion progress* by comparing several designs in a laboratory study with six participants. Results showed that *arrow-based* visualizations are preferred for representing robot *poses* and *trajectories*. *Explicit progress* displays were perceived as clearer and safer, but *trajectory-based* cues were preferred overall due to reduced distraction. Overall, visualizations that conveyed essential information while minimizing occlusion were rated most valuable. This work contributes an initial perspective towards structuring the AR design space for human-robot interaction with a robotic arm, illustrated through the example of robotic ultrasound.

Index Terms: Augmented Reality, Human-Robot Interaction, Robotic Ultrasound, Robot Intent, Mixed Reality.

1 Introduction

As robots increasingly transition from isolated industrial settings to everyday environments, human-robot collaboration is becoming more prominent. In medical contexts, for example, robotic systems assist clinicians during procedures. Robotic systems can generally be classified according to their level of autonomy, ranging from teleoperated systems to highly autonomous ones [13]. While full automation is expected in the more distant future, near-term systems will likely operate at relatively low levels of autonomy and rely on supervisory control [13]. As a result, clear communication of intended robot behavior is essential for successful collaboration [2]. Predicting such behavior is especially crucial for ensuring safety in shared work environments and establishing user trust [5]. These requirements are particularly relevant in robotic ultrasound applications, where the robot operates in close proximity to the patient and the operator. In this context, the tolerance for unexpected or poorly communicated robot motion is low, as such movements may compromise safety and reduce user trust. While physical safety mechanisms such as collision detection or compliance control are commonly applied to mitigate risks [10], the clear communication of intended robot behavior represents a complementary strategy that supports predictable and safe human-robot interaction. To address this challenge, prior work has explored the application of augmented reality (AR) to convey real-time robot behavior. AR enables spatial and immediate visualization of information directly within the work environment. In the context of robotic ultrasound, such approaches are particularly relevant for supporting users during the initial robot positioning and alignment of the ultrasound probe to the body surface. Seitz et al. [9] proposed a stepwise approach in which a robotic arm holding an ultrasound probe moved through four predefined intermediate poses leading to the final surface-contact pose. Their prototype further explored different autonomy levels for executing the overall trajectory, but did not incorporate visualizations to support understanding or monitoring the robot's motion. In contrast, several studies have examined AR-based interfaces for communicating robot intent. Tsamis et al. [11] introduced an AR interface to interact with a robotic arm mounted on a mobile platform, enabling both programming tasks and the visualization of cues about the robot's intended motion.

*e-mail: josefine.schreiter@ovgu.de

Their visualization approach comprised three modules for depicting *navigation paths* and *arm movements*, including sphere-shaped virtual objects and *safety zone* rendering realized through a semi-transparent sphere. Their results showed reduced task completion times and robot idle times, fewer workflow interruptions, and increased safety and user trust compared to standard operation. Beyond such application-specific interfaces, other work has focused more explicitly on the design and comparison of visualization concepts for communicating robot intent. Gruenefeld et al. [3] investigated three different progress visualizations for industrial robot motion to enhance spatial awareness and reduce collision risk. They distinguished between *path* visualizations depicting virtual end effector trajectories, *preview* showing a virtual replica of the robot arm, and a *volume* representation using a cylindrical shape. Their results indicated that participants reported the highest confidence and perceived safety with the *volume* visualizations and were most accurate in anticipating subsequent robot motion, while requiring the least head movement compared to the other visualization approaches. In this context, Walker et al. [12] introduced an AR-based design framework including four visualization approaches for conveying the intended motion of an aerial robot. They compared the approaches in a standardized manner in a user study with 60 participants including subjective and objective metrics. Their results showed that AR-based visualizations significantly improved objective task efficiency and users' comprehension of robot intent. Additionally, the evaluated designs revealed trade-offs between intent clarity and user perception, with more explicit visualizations generally leading to higher perceived communication quality and usability.

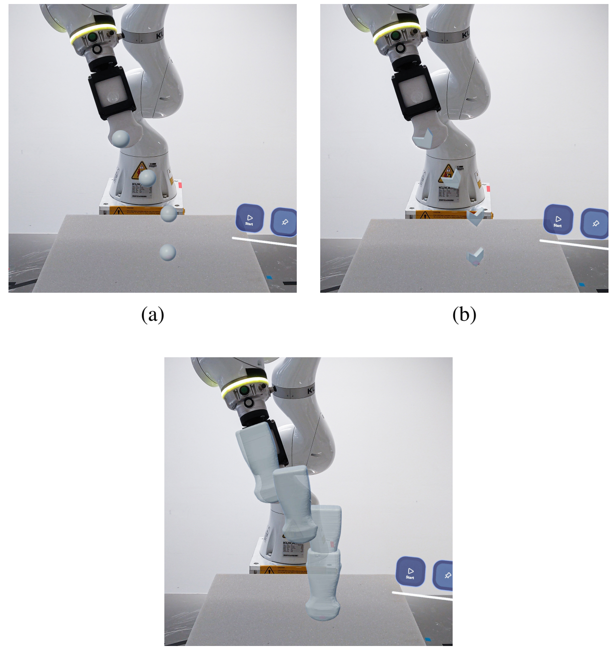
The present work adopts an exploratory approach to investigate visualization concepts for communicating intended poses and motions of a robotic ultrasound system. An initial robot-positioning strategy for surface alignment was adopted from a work of Seitz et al. [9], and a set of visualization concepts was systematically identified and implemented. These concepts explored different representations of end effector *poses*, *trajectories*, and *progress*, which were assessed in a user study with regard to comprehensibility, predictability, and perceived safety. Rather than proposing a finalized interface design, the goal of this work is to compare these visualization elements. In addition, a touchless control was implemented, including a graphical user interface (GUI), to facilitate human-robot interaction.

2 VISUALIZATION

Different representations were used across the categories *pose*, *trajectory*, and *progress*. The designs follow prior work in robot-related AR visualization, as referenced in the corresponding sections, and were specifically adapted to the requirements of robotic ultrasound. The conceptual principles of each visualization approach are described in the following subsections, while the corresponding implementation details are provided in Section 3.2.

Pose Representations of this category visualize the current or future Cartesian poses (x,y,z,a,b,c) of the robotic end effector along the path. Representations include *sphere*, *arrow*, and *virtual replica* (see Fig. 2). The *sphere* representation uses simple three-dimensional spherical objects [6, 12] to indicate the Cartesian position. *Arrow* elements are displayed as V-shaped arrowheads, encoding both Cartesian position and the movement direction toward the next target pose [4, 12]. *Virtual replicas* depict the Cartesian pose, including both position (x,y,z) and orientation (a,b,c), and mimic the proportions of the real tool [8].

Trajectory Representations of this category visualize the straight-line paths between the defined end effector poses. This includes *continuous ribbon*, *sphere segments*, *box segments* and *arrow segments* (see Fig. 3). *Continuous ribbon* visualizes the trajectory as a continuous, viewer-oriented surface [3]. *Sphere segments* depict the trajectory using regularly spaced spheres [12, 6]. Similarly, *box segments* depict the trajectory using regularly spaced boxes, implicitly indicating direction through axis alignment [3]. *Arrow segments* depict the trajectory using regularly spaced V-shaped arrowheads, explicitly encoding the movement direction [4, 12].



(a) (b) (c)

Figure 2: Representations of included *pose* visualizations including (a) spheres, (b) arrows and (c) virtual replicas.

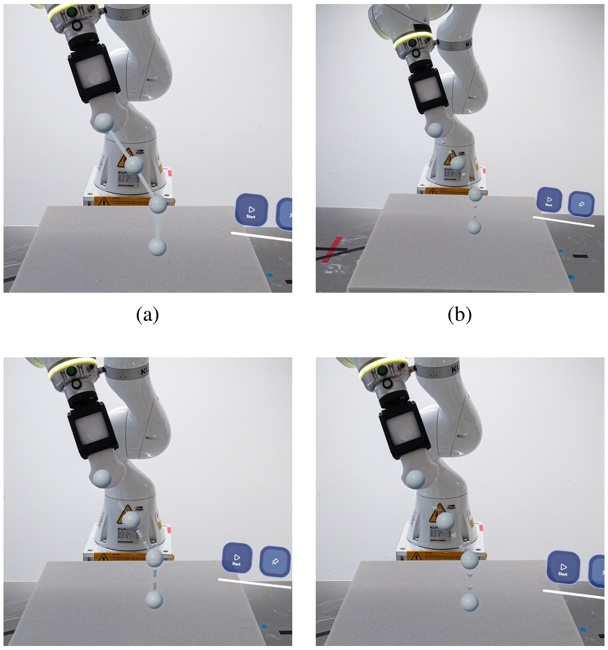
Progress Representations of this category focus on different approaches to visualize the end effector's progress along the path between two target poses. It is implemented through one *trajectory-based* approach and two approaches including *external indicators* (see Fig. 4). *Trajectory-based* indicators include a trajectory gradient where trajectory segments are colored ranging from black near the current pose to white for distant poses [1, 3]. *External indicators* include *percentage display* (exact percentage of completed trajectory) [14], and *circular display* (progress indicated by clockwise filling) [12]. Both are positioned above the next target pose.

3 IMPLEMENTATION

3.1 Apparatus

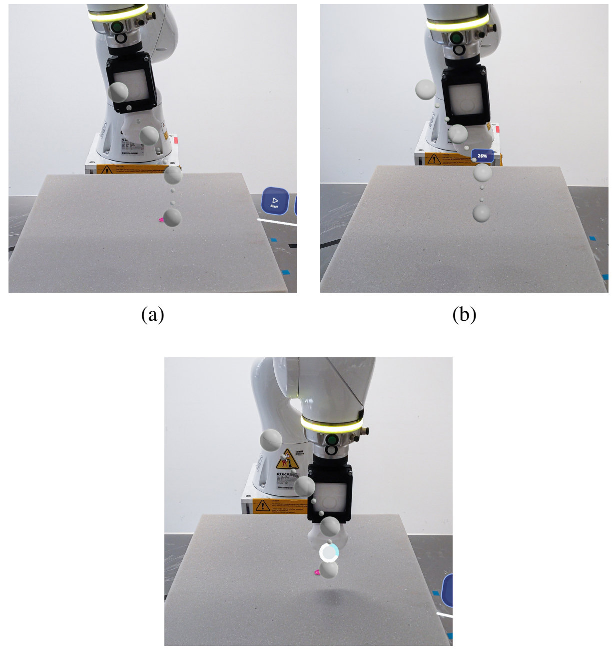
The experimental setup consisted of a robotic arm of type *KUKA LBR iiwa 16*[1] using *Sunrise.OR 16* operated in *SmartServoLIN* to enable continuous updating of target coordinates. A mock-up ultrasound probe was mounted on the robot's end effector. All motion commands were referenced to the tool center point, corresponding to the tip of the ultrasound probe. AR content was presented through a head-mounted display of type *HoloLens 2*[2]. All AR representations were implemented using *Unity*[3]. The system comprised the physical robot, a desktop *Unity* application, and a *HoloLens Unity* application (see Fig.5). The robot continuously streamed motion status and Cartesian pose data to the *HoloLens*,

[1]KUKA AG, Germany
[2]Microsoft Corporation, USA
[3]Unity Software Inc., San Francisco, USA



(a) (b) (c) (d)

Figure 3: Representations of included *trajectory* visualizations including (a) continuous ribbon, (b) sphere segments, (c) box segments and (d) arrow segments.



(a) (b) (c)

Figure 4: Representations of included *progress* visualizations including (a) trajectory-based gradient, (b) external percentage indicator and (c) external circular indicator.

while user input from the AR interface was sent back to the robot controller. Real-time adjustments of the visualization parameters were performed within the desktop application. All components communicated via bidirectional *UDP*. Spatial registration between the robot and the *HoloLens* coordinate system was achieved through a manual alignment procedure. Users placed a virtual robot base model onto the physical robot base using hand gestures and confirmed the alignment through a virtual button. All AR visualizations - except GUI elements - were subsequently aligned to the registered robot base frame.

3.2 Visualization Approaches

The robot poses P_0–P_3 were predefined (see Fig.1) following an approach of initial robot positioning by Seitz et al. [9]. In this setup, P_0 represented an arbitrary starting point, P_2 a pose 10 cm above the surface-contact pose P_3, and P_1 an interpolation between P_0 and P_2. All representations were aligned with the computed Cartesian coordinates (x,y,z) and, where applicable, to the corresponding orientation (a,b,c).

For the *arrow* pose representation, an additional logic was implemented to encode subsequent motion direction. *Trajectory* visualizations were instantiated by placing segments at regular intervals along each sub-path. A dedicated module stored central parameters such as trajectory length or motion direction. *Progress* was visualized using a global progress value, with the *trajectory gradient* derived by mapping colors to each segment based on its relative position along the path. The *percentage display* and the *circular display* followed an analogous logic.

3.3 Human-Robot Interaction

The system provided motion control of the robot through a GUI anchored within the user's field of view (see Fig. 2). The GUI consisted of a motion-control button with three automatic switching states (*start*, *stop*, *continue*) and a pinning button to fix the interface in place. After each target pose was reached, the motion-control button returned to the *start* state. Pressing the *start* button triggered additional auditory feedback.

4 Evaluation

The study aimed to compare the different visualization approaches with respect to comprehensibility, predictability, and perceived safety. Participants were recruited without requiring specific prior knowledge or expertise. A table positioned in front of the robot simulated a patient table (see Fig.1). Participants stood at a predefined distance of 85 cm from this table to ensure consistent viewing conditions and full visibility of both the robot and AR visualizations. An experimental supervisor operated the desktop application and intervened if necessary using the robot's smart pad. Visualization approaches were evaluated within their respective categories, with elements presented in a randomized order. At the beginning, participants were briefed on the study aims and procedure, provided informed consent and completed a demographic questionnaire. Prior to each category, participants received a short explanation of the respective visualization's function. Each representation was presented by having the robot execute all four target poses once before returning to the initial pose. Preferred *pose* and *trajectory* visualizations were used when assessing the *progress* visualizations. After each representation in the *pose* and *trajectory* category, participants verbally completed a custom-designed questionnaire. It consisted of three Likert-scale items (1: very bad, 5: very good) assessing comprehensibility, predictability, and perceived safety (one item each), a ranking of the elements to capture individual preferences, and three open-ended questions for general feedback. The *progress* questionnaire was likewise developed specifically for this study. It comprised six questions, including two Likert-scale items (1: very bad, 5: very good) assessing comprehensibility and perceived safety (one item each), one ranking, and two open-ended questions for general feedback.

Figure 5: System architecture illustrating the three application components and their bidirectional data exchange.

5 Results

Six participants took part in the study (5 male, 1 female), aged between 22 and 31 (M=25.00, SD=3.16). Two had no visual impairment, one was farsighted, and three nearsighted, all with corrected vision. Prior expertise with AR was assessed using a five-point Likert scale (1 = no prior experience, 5 = very high prior experience), with ratings ranging from 2 to 5 (M=4.17, MAD=1.33). Prior expertise with robots ranged from 1 to 3 (M=2.00, MAD=0.89).
Participants generally found *pose* visualizations useful for understanding the robot's current position and its intended motion. Comprehensibility was rated highest for *spheres* (M=4.67, SD=0.52) followed by *arrows* (M=4.33, SD=1.21) and *replicas* (M=3.83, SD=0.75) (see Fig. 6). With regard to predictability, *arrows* received the highest ratings (M=4.33, SD=0.82). Perceived safety was comparable for *arrows* (M=3.83, SD=0.75) and *replicas* (M=3.50, SD=0.55), and lowest for *spheres* (M=2.50, SD=0.55). Overall, *arrows* were ranked highest, followed by *replicas* and *spheres*. Participants emphasized that *arrows* clearly conveyed motion direction, perceived as the most critical information, whereas *replicas* were criticized for partially occluding the environment.
Trajectory visualizations were generally perceived as valuable for understanding both direction and progress of the intended robot motion. Comprehensibility ratings were comparable across approaches, however, *ribbon segments* received lower scores for predicting subsequent robot motion (M=2.50, SD=1.52) and were criticized for occluding the real scene (see Fig. 7). In contrast, *sphere segments* (M=3.0, SD=1.26), *box segments* (M=3.33, SD=1.37), and *arrow segments* (M=3.17, SD=1.60) showed similar high predictability ratings, with *sphere segments* particularly noted for minimal occlusion. Perceived safety was considered comparable across elements, however, *arrow segments* were evaluated most positively (M=4.00, SD=0.89). Participants ranked *arrow segments* highest due to their clear indication of motion direction and intuitive interpretation, followed by *sphere segments*, *box segments*, and *ribbon segments*. Half of the participants considered the progress visualizations helpful, whereas the other half regarded them as unnecessary. Comprehensibility was evaluated similarly for the *percentage display* (M=3.67, SD=1.03) and the *circular display* (M=3.67, SD=0.82), and less positively for the *trajectory gradient* (M=3.17, SD=1.17) (see Fig. 8). Perceived safety ratings showed moderate variations, with the *percentage display* (M=3.83, SD=1.60) rated highest, followed by the *circular display* (M=3.5, SD=0.55) and the *trajectory gradient* (M=3.17, SD=0.75). The *trajectory gradient*, however, received the best overall ranking, as participants emphasized its advantage of not introducing an additional display and therefore avoiding distraction. The *percentage display* was valued for providing explicit numerical information, while the *circular display* was generally considered intuitive but less precise.

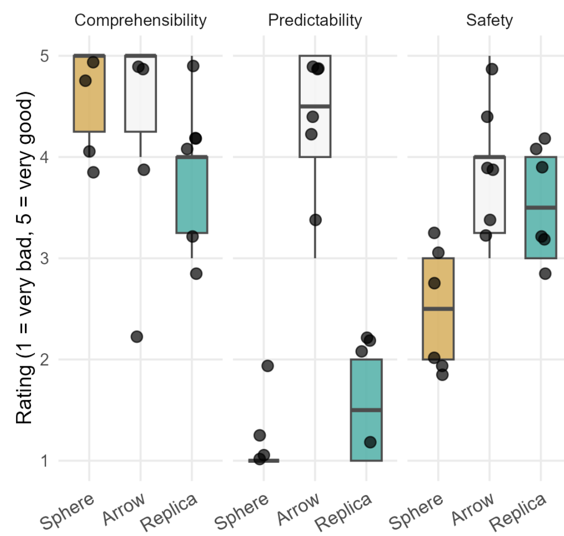


Figure 6: Participant ratings of comprehensibility, predictability, and perceived safety for the three *pose* visualizations.

6 Discussion & Conclusion

The study results indicate that *pose* and *trajectory* visualizations are most effective when they clearly convey directional intent and motion progress without introducing visual clutter. Visualizations that communicate essential information - such as end effector orientation or subsequent motion direction - while minimizing occlusion were rated most valuable, with *arrow-based* representations consistently preferred. *Replica-based* pose visualizations may offer benefits for conveying end effector orientation but require a reduction in visual magnitude to avoid obstructing the real scene. *Trajectory* visualizations need further refinement, particularly regarding segment density. The effectiveness of some *progress* indicators, such as *trajectory-based* gradients, appeared to depend on the underlying trajectory representation and might be more perceivable in continuous forms (e.g. ribbons) than discrete segmentations (e.g. spheres). In the context of study objectives, *progress* visualizations were perceived as supplementary rather than essential for understanding or predicting robot behavior, particularly for short, straight-line robot motion. Additional displays risk overwhelming users and should therefore be applied sparingly.
In a clinical context, preferences and requirements may differ from those observed under laboratory conditions. Clinicians are likely to prioritize visualizations that minimize distraction and integrate seamlessly into the clinical workflow [7]. For example, explicit vi-

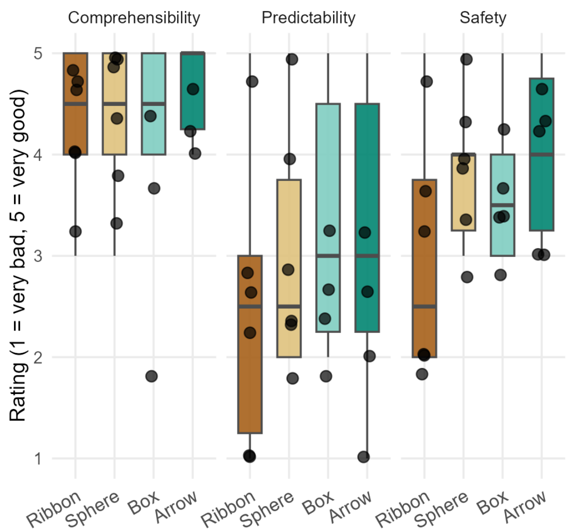


Figure 7: Participant ratings of comprehensibility, predictability, and perceived safety for the four *trajectory* visualizations.

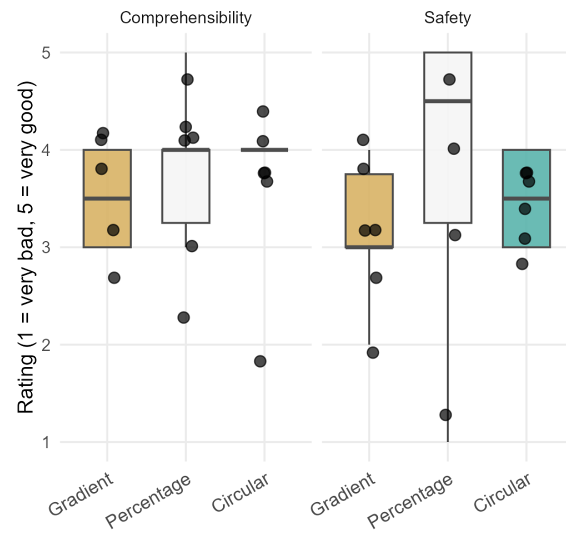


Figure 8: Participant ratings of comprehensibility and perceived safety for the three *progress* visualizations.

sualizations that were found useful in prior work [12] may be perceived as intrusive or cognitively demanding in clinical settings or by specific user groups. This highlights the importance of considering domain-specific constraints and user preferences when transferring visualization concepts from laboratory settings to clinical practice.

The study is subject to several limitations. First, the sample size was small, and participants showed comparably high prior familiarity with AR, limiting generalizability. In addition, the study participants did not have a clinical background. Second, spatial registration was performed manually, introducing potential alignment inaccuracies and inter-participant variability. Third, the abstract and simulated study environment do not fully reflect the complexity and constraints of the realistic clinical setting. In future iterations, the surface-contact pose will be integrated using marker-based tracking, with an optical marker placed on the surface to define the intended probe pose. Moreover, subsequent evaluations should be conducted with radiologists to assess the applicability and effectiveness of the visualization approaches in realistic clinical contexts including standardized questionnaires.

This work presents an exploratory investigation of AR-based visualization elements to support human-robot interaction in a robotic ultrasound application by facilitating spatial understanding and motion prediction. By systematically identifying and comparing key visualization concepts, including representations of *poses*, *trajectories*, and *progress*, the study provides an initial structuring of the design space for communicating intended robot behavior. The findings do not constitute a finalized visualization design but rather highlight visualization elements that appear particularly promising for the considered use case based on user feedback. As such, the results serve as a foundation for further refinement and validation. Future work should include evaluations with clinicians and more complex, realistic scenarios to assess the applicability of the identified design directions and to guide their integration into practical robotic systems.

Acknowledgments

This work was funded by the Federal Ministry of Education and Research within the research campus *STIMULATE* (Grant No. 13GW0473A), as well as by the European Regional Development Fund within the framework of the initiative *Saxony-Anhalt Research Focus* (Grant No. ZS/2023/12/182010).

This manuscript benefited from language editing and grammar suggestions provided by *ChatGPT* (OpenAI, San Francisco, CA, USA).

References

[1] O. Ashtiani, H.-J. Guo, and B. Prabhakaran. Impact of motion cues, color, and luminance on depth perception in optical see-through ar displays. *Frontiers in Virtual Reality*, 4:11, 2023. doi: 10.3389/frvir.2023.1243956 2

[2] R. T. Chadalavada, H. Andreasson, R. Krug, and A. J. Lilienthal. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR)*, pp. 1–6. IEEE, 02.09.2015 - 04.09.2015. doi: 10.1109/ECMR.2015.7403771 1

[3] U. Gruenefeld, L. Prädel, J. Illing, T. Stratmann, S. Drolshagen, and M. Pfingsthorn. Mind the arm. In B. Preim, A. Nürnberger, and C. Hansen, eds., *Proceedings of the Conference on Mensch und Computer*, pp. 259–266. ACM, New York, NY, USA, 09062020. doi: 10.1145/3404983.3405509 2

[4] M. Gu, A. Cosgun, W. P. Chan, T. Drummond, and E. Croft. Seeing thru walls: Visualizing mobile robots in augmented reality. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, pp. 406–411. IEEE, 08.08.2021 - 12.08.2021. doi: 10.1109/RO-MAN50785.2021.9515322 2

[5] J. M. Haney and C.-J. Liang. A literature review on safety perception and trust during human-robot interaction with autonomous mobile robots that apply to industrial environments. *IISE transactions on occupational ergonomics and human factors*, 12(1-2):6–27, 2024. doi: 10.1080/24725838.2023.2283537 1

[6] L. Kastner and J. Lambrecht. Augmented-reality-based visualization of navigation data of mobile robots on the microsoft hololens - possibilities and limitations. In *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pp. 344–349. IEEE, 18.11.2019 - 20.11.2019. doi: 10.1109/CIS-RAM47153.2019.9095836 2

[7] J. Lenz, I. Richter, and S. Meister. Automated filtering and visualization of patient-centered data from electronic health records in emergency care: A scoping review. *Journal of multidisciplinary healthcare*, 18:6503–6517, 2025. doi: 10.2147/JMDH.S555000 4

[8] T. Mielke, F. Heinrich, and C. Hansen. Gesturing towards efficient robot control: Exploring sensor placement and control modes for

mid-air human-robot interaction. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7801–7807. IEEE, 19.05.2025 - 23.05.2025. doi: 10.1109/ICRA55743.2025.11127519 2

[9] P. K. Seitz. *Implementation of safe human robot collaboration for ultrasound guided radiation therapy*. PhD thesis, Heidelberg University Library, 2023. doi: 10.11588/heidok.00033606 1, 2, 3

[10] A.-N. Sharkawy and P. N. Koustoumpardis. Human–robot interaction: A review and analysis on variable admittance control, safety, and perspectives. *Machines*, 10(7):591, 2022. doi: 10.3390/machines10070591 1

[11] G. Tsamis, G. Chantziaras, D. Giakoumis, I. Kostavelis, A. Kargakos, A. Tsakiris, and D. Tzovaras. Intuitive and safe interaction in multi-user human robot collaboration environments through augmented reality displays. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, pp. 520–526. IEEE, 08.08.2021 - 12.08.2021. doi: 10.1109/RO-MAN50785.2021.9515474 1

[12] M. Walker, H. Hedayati, J. Lee, and D. Szafir. Communicating robot motion intent with augmented reality. In T. Kanda, S. Ŝabanović, G. Hoffman, and A. Tapus, eds., *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 316–324. ACM, New York, NY, USA, 02262018. doi: 10.1145/3171221.3171253 2, 5

[13] G.-Z. Yang, J. Cambias, K. Cleary, E. Daimler, J. Drake, P. E. Dupont, N. Hata, P. Kazanzides, S. Martel, R. V. Patel, V. J. Santos, and R. H. Taylor. Medical robotics-regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science robotics*, 2(4), 2017. doi: 10.1126/scirobotics.aam8638 1

[14] H. Yu, Q. Fang, S. Fang, R. M. Aronson, and E. S. Short. How much progress did i make? an unexplored human feedback signal for teaching robots. 2