Omnisurface: Common Reality for Intuitive Human-Robot Collaboration

Chapter · February 2025

DOI: 10.1007/978-981-96-1151-5_34

Akhlak Uz Zaman
University of Alabama

Fujian Yan
Wichita State University

Omnisurface: Common Reality for Intuitive Human-Robot Collaboration

Akhlak Uz Zaman[1], Hui Li[1], Fujian Yan[2], Yinlong Zhang[3], and Hongsheng He[1(✉)]

[1] The University of Alabama, Tuscaloosa, AL 35487, USA
hongsheng.he@ua.edu
[2] School of Computing, Wichita State University, Wichita, KS 67208, USA
[3] Shenyang Institute of Automation, Shenyang 110016, China

Abstract. Effective communication and information projection are essential for human-robot teaming. The projection of images on non-planar surfaces using a conventional projector is challenging due to the inherent problem of distortion. The projection distortion occurs due to the variations in depth across the surface of the teaming workspace. As a result, the projected image, information, or symbols lose their original shape and create confusion during human-robot teaming. In this paper, we presented an innovative approach to perform distortion-free projections in the teaming workspace. A pre-warped image is constructed based on the surface geometry that the projector displays and accurately replicates the original projection image. Beyond the technical achievement, this research highlights the social acceptance of improved spatial augmented reality in human-robot teams. It fosters better teamwork, trust, and efficiency by enabling more intuitive and reliable interactions.

Keywords: Human-Robot Teaming · Distortion Correction · RGB-D sensor · Common Reality

1 Introduction

Robots have been used as a tool for humankind for a long time. Gradually, they become like a human partner in terms of performing a common task. The incorporation of robots into collaborative human-robot teaming has become more widespread in recent years. It has led to a transformation in numerous industries due to the combination of the distinct capabilities of humans and robots [1,2]. Visualization approaches utilizing a projector can play a crucial role in improving the intuitiveness of human-robot collaboration by presenting cooperative instructions in the workspace. The visualization of images using a projector on real objects, also known as spatial augmented reality [3], can enhance coordination and team performance by enabling humans to monitor the actions of robots. It assures proper teamwork to gain more job satisfaction [4]. Projectors

This research was funded by NSF grants #2427895 and #2420355.

H. Li et al. (Eds.): ICSR + InnoBiz 2024, LNAI 15170, pp. 1–12, 2025.
https://doi.org/10.1007/978-981-96-1151-5_34

can be utilized to arrange a mixed-reality canvas in the physical work environment through a virtual visualization of the work plan. It can provide intuitive feelings for humans and improve collaborative team performance [5]. Traditional projectors cannot understand surface geometry and adjust images to account for non-flat projection surfaces. Consequently, ordinary projectors' cannot handle the depth variation at different points of a non-flat surface and make the projected image distorted [6]. The use of RGB-D sensors has become popular nowadays for 3D surface reconstructions and their utilization in different robotics and computer vision applications. It can provide the pixel-level depth of a particular surface with a very high frame rate. The depth data of RGB-D sensors can be utilized to understand surface geometry [7].

The projection distortion problem causes a change in the original shape of the object to be projected. It can mislead the human participants to understand robots' activities properly and seriously hamper the overall performance and intuition of human-robot teaming. The main focus of this study is to solve the distortion problem to facilitate quality human-robot teaming. This paper has the following contributions,

1. This paper proposes a new visualization-based interface for human-robot teaming in real-world scenarios, enhancing intuitive communication and collaboration between humans and robots.
2. This research proposes a unique method for correcting projection distortion on non-planar surfaces using RGB-D sensor data, significantly improving the accuracy and clarity of visual instructions in human-robot collaboration tasks.

2 Human-Robot Teaming

A robot was a helping hand to its operator (a human) from the beginning. It has become a partner or co-worker with the advancement of artificial intelligence and robotics. Human-robot teaming creates interdependence between the two entities through communication, collaboration, and coordination. A dual collaboration concept was proposed that guided the design of the components of the collaboration [2]. The proposed architecture suggests some factors that are useful for proper human-robot teaming. Figure 1 shows the basic components of human-robot teaming, where both humans and robots have their specific abilities. They are interdependent on one another in terms of collaboration factors such as context, environment, and task.

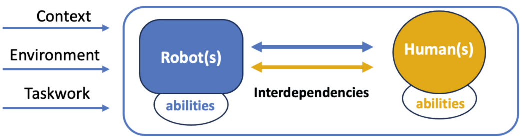


Fig. 1. Fundamental components responsible for Human-robot teaming.

2.1 Use of Visualization in Human-Robot Teaming

The concept of imaginary projection [4,8] in human-robot collaboration can potentially enhance team performance. Their main goal was to create a physical work environment using projectors and use this canvas as a friendly communication medium for collaborations. They proposed an object-tracking computer vision algorithm to recognize an object in the workspace and estimate the pose and state of the physical object. Their proposed method helps humans verify whether the alignment of the physical object with the robot's scope is perfect or not. They also initiated visual instructions to show the robot's intentions for the upcoming activities and warn humans if there were any safety issues. A projection-based communication channel was proposed that communicates between humans and robots to circulate necessary information using visual references [9]. A convolutional neural network architecture [10] was presented that transforms arbitrary surfaces into interactive touch interfaces. The architecture uses a CNN to detect fingertip touches on flat and curved surfaces with RGB-D sensor input, enabling intuitive human-robot interaction through touch gestures and reducing operator training time. The industrial scope of human-robot collaboration [11] was studied and showed some evidence that projector-based graphical visualization can improve industrial work efficiency and safety. The authors proposed a dynamic graphical user interface to interact with industrial robots and perform better human-robot interaction. A web-based interface [12] was designed to establish a dual communication system between robots and humans on the collaboration worktable. An augmented reality (AR) based collaboration framework [13] was proposed to establish better teamwork. They showed that AR-based visualization can have the potential for peer collaboration between humans and robots in unknown and instrument-less environments. The authors also proposed bidirectional communications between humans and autonomous robots to perform peer teamwork. Although image distortion can hamper the intuitiveness of the collaboration, none of the literature discussed above concentrated on image distortion problems.

2.2 Projector Distortion Correction

Image distortion occurs due to the improper geometry of the projection surfaces [14–16]. A novel system [14] was proposed to solve the image distortion problem by moving the projector direction from an uneven surface to an even surface. It does not work for our problem. Because, in most cases, human-robot teaming workplaces are non-plane. A distortion correction method [16] was suggested, utilizing the captured image from a camera attached to the projector. The system used an algorithm to correct the keystone distortion problem automatically. The authors considered the surface angle and computed a pre-warped trapezoid for keystone correction. The technique had no scope to capture non-plane surface geometry and correct distortion problems. A projector distortion correction-based system [17] was proposed for non-planner surfaces using a Kinect device as an RGB-D sensor. The authors implemented a 3D point cloud based on the non-planar surface data captured by Kinect. After that, the system performed geometric correction and generated a pre-warped image for distortion correction. The process utilized a projector to display the pre-warped image on the non-plane surface, minimizing the image distortion problem. The technique acknowledged that this method does not apply to real-time applications. This approach had a plan to reduce computation costs using GPUs so that their system could be used for real-time applications. According to the review, projector distortion correction on non-planar surfaces is a unique approach to initiating an intuitive interface for human-robot collaboration. None of this literature is directly aimed at this research goal. The literature mentioned above is really helpful in conceptualizing the background of our specific research goal.

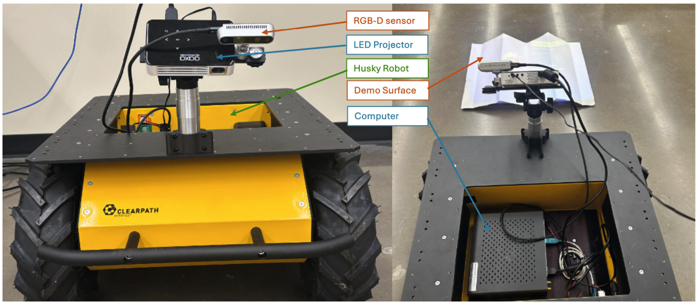


Fig. 2. Proposed distortion-free projection on omni-surface to facilitate human-robot teaming.

3 Projection Distortion Correction on Omnisurface

The proposed research initiates a visualization interface according to the following setup. The projector and the RGB-D sensor are fixed on top of the collaboration workspace. The target surface of the collaboration workspace is assumed to be a non-planner surface. Initially, the calibration method established the relationship between the projector and the depth camera (the extrinsic parameters). The depth sensor uses calculated calibration parameters and allows the projection of surface area by the projector. It determines the optimal correction area, which is the maximum rectangular region within the projection surface to perform the distortion correction. The original projection image is transformed into a pre-warped image to fit the effective correction area, removing any geometric distortion on the target surface. Finally, the projector outputs the corrected image. Figure 2 shows the hardware setup of our visualization-based collaboration framework for human-robot teaming. A projector and RGB-D sensor are attached above the collaboration surface on top of a robotic vehicle. The proposed method captures the 3D surface depth using an RGB-D sensor and visualizes a virtual object model onto the real-life object. This spatial augmented visualization helps humans monitor robot activity intuitively.

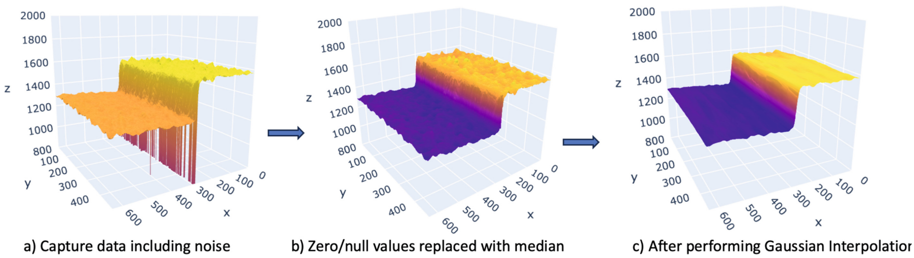


Fig. 3. 3D graphical representation of the Surface data a) captured using RGD-D sensor which includes noise, b) after replacing null/zero values with median, and c) after applying the Recursive Gaussian Interpolation.

3.1 Surface Data Acquisition and Refinement

It is imperative to extract noise from the depth data of the RGB-D sensor before initiating the distortion correction process. The presence of noise in the depth data can lead to a decrease in the surface smoothness and result in inaccurate distortion correction. At first, the system captures surface depth using an RGB-D sensor and then performs refinement. The proposed approach followed two steps to reduce noise: (i) replacing invalid null (practically zero) values with the median and (ii) performing the recursive Gaussian interpolation algorithm [18] to make a smoother surface. Figure 3 shows the progress of surface data refinement.

3.2 Inverse Projection Model

To project a 2D image onto a non-planar 3D surface using depth data, this approach must determine the inverse projection transformation. This involves first identifying the projection transformation equation, which generates the 2D image from 3D world coordinates using the pinhole camera model. Once this is established, the proposed technique can calculate the inverse projection transformation needed to project the 2D image onto the surface accurately as defined by its 3D world coordinates. The projection transformation equation from 3D world coordinates P_w to 2D image coordinates P_i, can be represented as

$$\mathbf{P}_\mathrm{i} = \mathbf{K} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \cdot \mathbf{P}_\mathrm{w} \tag{1}$$

where K, R, t, P_i, and P_w are denoted respectively as the camera intrinsic matrix, rotation matrix, translation vector, 2D image coordinates, and 3D world coordinates.

3.3 Pre-Warped Image Generation

To generate a pre-warped image, the proposed system needs to produce an equivalent opposite surface D'' of the given depth data D'. The proposed approach calculated the minimum depth $\mathbf{D}_{\min}$ and maximum depth $\mathbf{D}_{\max}$ for each of the depth values greater than zero. Then, prepare the equivalent opposite surface D'' using the following equation where every depth value $D''[i,j]$ is calculated from the original surface $D[i,j]$,

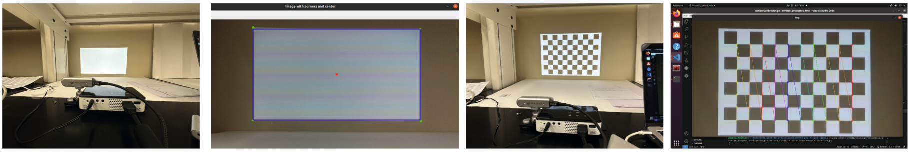
$$\mathbf{D}''[i,j] = \mathbf{D}_{\max} - (\mathbf{D}[i,j] - \mathbf{D}_{\min}) \tag{2}$$

To generate a pre-warped image, the opposite surface D'' is considered as coordinate depth data P_{w} and generates a 3D point-cloud using the pinhole camera model. The 2D view of this point-cloud is our expected pre-warped image P'_{i} which gives distortion-free projection on omni-surfaces.

4 Experiment

4.1 Experiment Setup

An "Intel RealSense D435" RGB-D sensor is used to capture non-planar surface depth. This sensor is small, lightweight, and low-cost. The "AAXM P300 Neo" is a small and lightweight LED projector. The proposed approach utilizes this projector to display the corrected image on the human-robot workspace (Fig. 4).



a)White rectangle Projection b) Projection surface detection c) Chessboard Projection d) Corner detction

Fig. 4. Projection Surface Calculation and Calibration.

4.2 Projector and RGB-D Sensor Calibration

To perform the calibration between the projector and the RGB-D sensor, we first positioned both devices in a fixed setup with overlapping fields of view, aimed at a calibration surface. Using a checkerboard pattern, we captured multiple images from the RGB-D sensor in various orientations. We then calibrated the RGB-D sensor's RGB camera by detecting the corners of the checkerboard pattern and utilizing the OpenCV library [19] to estimate its intrinsic parameters. For the projector, we treated it as an inverse camera, projecting a known pattern and capturing it using the RGB-D sensor. This allowed us to calibrate the projector and determine its intrinsic parameters. Subsequently, we computed the extrinsic parameters by matching corresponding points from both the projector and the RGB-D sensor. This process enabled us to determine the spatial relationship between the two devices and align their coordinate systems. Finally, we validated the calibration by projecting patterns and confirming their alignment with the 3D depth data captured by the RGB-D sensor.

4.3 Projection Surface Calculation

The field of view (FOV) of the RGB-D sensor and the projector is not the same. In practice, the RGB-D sensor covers more surface area than the projector display. To ensure precise projection, we calculated the specific projection area from the entire surface and used the depth array of this specific area to generate a pre-warped image. We then placed a white rectangle on the surface, covering the entire area of the projector display, and detected the coordinates of the four corners of the rectangle.

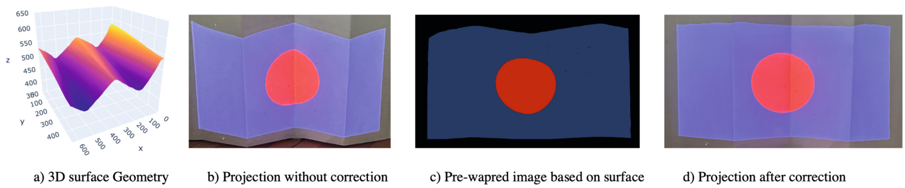


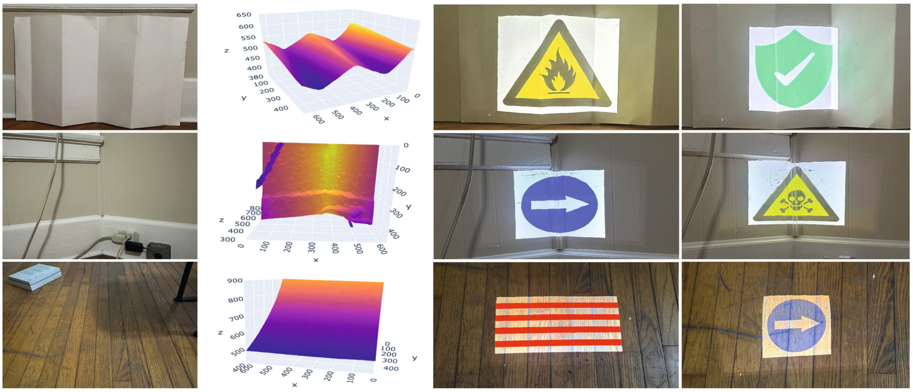
Fig. 5. Major steps for distortion correction on a non-planner surface using RGB-D sensor.

4.4 Distortion Correction and Projection

The proposed method requires to build a pre-warped image from the original image. This pre-warped image is prepared based on the non-planner surface data captured by the RGB-D sensor. Figure 5 presents the major steps for distortion correction on a non-planner surface using an RGB-D sensor by our proposed method. Here, Fig. 5 a) represents the 3D surface data captured using the RGB-D sensor, (b) shows the projection without any correction, (c) represents the pre-warped image generated using our proposed method, and (d) represents the projection based on the proposed method. In order to test our distortion correction method, we chose to project various symbols onto uneven surfaces such as room corners, wavy walls, and wooden floors. These surfaces are not typically suitable for projection using a regular projector due to their significant depth and color variations. First, we captured the 3D depth of the surface and refined the surface data to eliminate any noise. We then created pre-warped images using the noise-free depth data based on our proposed distortion correction method. Finally, we projected these pre-warped images onto the surfaces one by one. The result we observed was extremely promising. Figure 6 shows three surfaces, refined 3D depth data, and two corrected projection examples for each surface.

4.5 Performance and Accuracy

In order to evaluate the performance of our proposed model for point projections accurately, we use the Root Mean Square Error, which is a standard statistical



a) Rough/uneven surface images b) 3D Surface data captured by RGB-sensor c) Corrected projection examples

Fig. 6. Examples of Corrected projection on rough surfaces such as wavy walls, room corners, and floors.

metric for measuring average error magnitudes. It involves a series of mathematical operations to calculate, and here are the detailed equations used to represent error, $\mathbf{E} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$. In this equation, the squared difference $(y_i - \hat{y}_i)^2$ represents the squared discrepancy between each target point, denoted as y_i , and its corresponding projected point, $\hat{y}_i$, for the i-th observation in a dataset where n represents the total number of pairs of points. To calculate the RMSE, we took 40 random (x,y) coordinates of target points, found the coordinates of the projected points, and calculated the errors between the target and projected points. We found the Root Mean Square Error value is as good as 3.34 mm, the average error is 3.13 mm, and the percentage of the error is 0.40%. Figure 7 presents the scatter plot illustrating the target points (in green) and projected points (in red) for each pair, with gray dashed lines connecting each target to its projected point. This visual helps you see the dispersion of the points in the 400 mm x 250 mm space and the error represented by the distance between corresponding points. Figure 8 shows a few examples of the projections, visually comparing ordinary and corrected projections using our proposed method.

4.6 Social Acceptance

In order to assess the social acceptance of our correction method for projection distortion on omni-surfaces to enhance human-robot collaboration, we have designed a survey with three evaluation criteria: originality, visibility, and understandability. We have assigned a weight to each criterion to calculate the final acceptance score for each perspective. The criteria are as follows: 1) Originality measures how closely the projection resembles the original image; 2) Visibility assesses whether the user can clearly see the robot's instructions; 3) Understandability gauges the extent to which the user can understand the robot's intention after viewing the projection. The final score across all criteria is 4.3

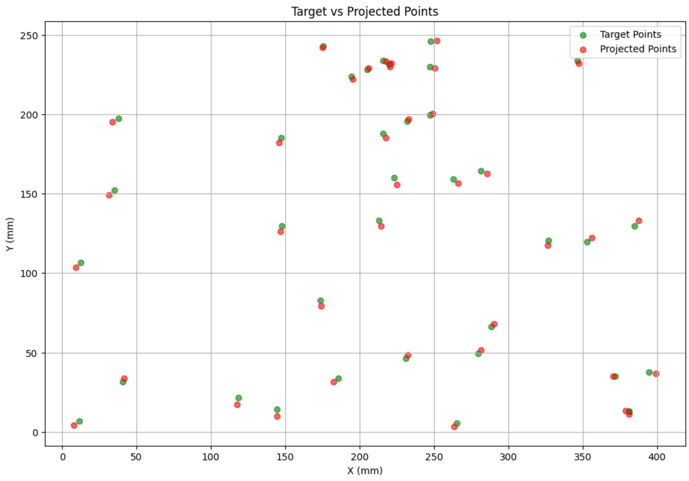


Fig. 7. Scatter Plot of Target and Projected Points with Error Visualization.

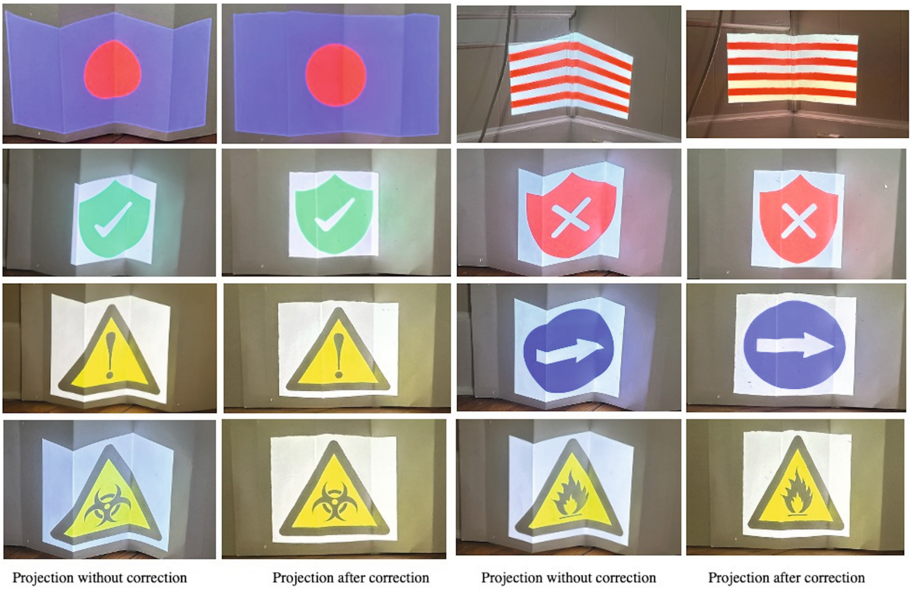


Fig. 8. Comparison between ordinary projection and corrected projection.

out of 5, and the visual representation of Fig. 9 allows us to compare the different aspects of evaluation. The result demonstrates how solving a technical issue can significantly improve the social acceptance of human-robot collaboration, thus advancing social robotics.

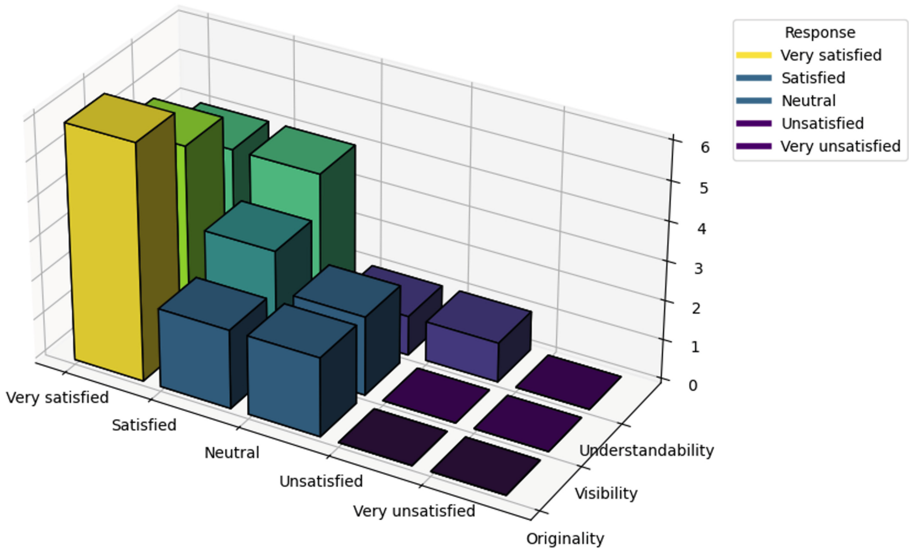


Fig. 9. User feedback distribution across criteria in evaluating projection distortion correction on omni-surfaces.

5 Conclusion

This research introduced an innovative approach to mitigate projection distortion on non-planar surfaces and improved human-robot collaboration. By leveraging an RGB-D sensor to capture surface geometry, the suggested system successfully corrects distortions in real time, thereby improving the accuracy and clarity of visual instructions in collaborative contexts. This innovation increased team trust and productivity while also enhancing the technical aspects of human-robot interaction. The technology has the potential to be widely applied in businesses that depend on human-robot cooperation, as demonstrated by its capacity to produce distortion-free projections on complicated surfaces. Further optimization for real-time processing and the application of this method to dynamic surfaces will be explored in future work.

References

1. Matheson, E., Minto, R., Zampieri, E.G., Faccio, M., Rosati, G.: Human-robot collaboration in manufacturing applications: a review. Robotics **8**(4), 100 (2019). https://doi.org/10.3390/robotics8040100
2. Mingyue Ma, L., Fong, T., Micire, M.J., Kim, Y.K., Feigh, K.: Human-robot teaming: concepts and components for design. In: Hutter, M., Siegwart, R. (eds.) Field and Service Robotics. SPAR, vol. 5, pp. 649–663. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-67361-5_42
3. Casas, S., Gimeno, J., Casanova-Salas, P., Riera, J.V., Portalés, C.: Virtual and augmented reality for the visualization of summarized information in smart cities:

a use case for the city of Dubai. In: Smart Systems Design, Applications, and Challenges, pp. 299–325. IGI Global (2020). https://doi.org/10.4018/978-1-7998-2112-0.ch015
4. Andersen, R.S., Madsen, O., Moeslund, T.B., Amor, H.B.: Projecting robot intentions into human environments. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 294–301. IEEE (2016). https://doi.org/10.1109/ROMAN.2016.7745145
5. Ganesan, R.K.: Mediating human-robot collaboration through mixed reality cues (Master's thesis, Arizona State University) (2017)
6. Boroomand, A., Sekkati, H., Lamm, M., Clausi, D.A., Wong, A.: Saliency-guided projection geometric correction using a projector-camera system. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 2951–2955. IEEE (2016). https://doi.org/10.1109/ICIP.2016.7532900
7. Jing, C., Potgieter, J., Noble, F., Wang, R.: A comparison and analysis of RGB-D cameras' depth performance for robotics application. In: 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), pp. 1–6. IEEE (2017). https://doi.org/10.1109/M2VIP.2017.8211432
8. Ganesan, R.K., Rathore, Y.K., Ross, H.M., Amor, H.B.: Better teaming through visual cues: how projecting imagery in a workspace can improve human-robot collaboration. IEEE Robot. Autom. Mag. **25**(2), 59–71 (2018). https://doi.org/10.1109/MRA.2018.2815655
9. Weng, T., Perlmutter, L., Nikolaidis, S., Srinivasa, S., Cakmak, M.: Robot object referencing through legible situated projections. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 8004–8010. IEEE (2019). https://doi.org/10.1109/ICRA.2019.8793638
10. Yan, F., Chavez, E., Yihun, Y.S., He, H.: Touch detection in augmented Omnisurface for human-robot teaming. J. Manag. Eng. Integr. **15**(2), 92–99 (2022). https://doi.org/10.62704/10057/24828
11. Hietanen, A., et al.: Proof of concept of a projection-based safety system for human-robot collaborative engine assembly. In: 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 1–7. IEEE (2019). https://doi.org/10.1109/RO-MAN46459.2019.8956446
12. David, J., Järvenpää, E., Lobov, A.: A web-based mixed reality interface facilitating explicit agent-oriented interactions for human-robot collaboration. In: 2022 8th International Conference on Mechatronics and Robotics Engineering (ICMRE), pp. 174–181. IEEE (2022). https://doi.org/10.1109/ICMRE54455.2022.9734094
13. Reardon, C., Lee, K., Rogers, J.G., Fink, J.: Communicating via augmented reality for human-robot teaming in field environments. In: 2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), pp. 94–101. IEEE (2019). https://doi.org/10.1109/SSRR.2019.8848971
14. Nakamura, N., Hiraike, R.: Active projector: image correction for moving image over uneven screens. In: Companion of the 15th Annual ACM Symposium on User Interface Software and Technology, pp. 1-2 (2002)
15. Han, D.: Real-time digital image warping for display distortion correction. In: Kamel, M., Campilho, A. (eds.) ICIAR 2005. LNCS, vol. 3656, pp. 1258–1265. Springer, Heidelberg (2005). https://doi.org/10.1007/11559573_152
16. Li, B., Sezan, I.: Automatic keystone correction for smart projectors with embedded camera. In: 2004 International Conference on Image Processing, 2004. ICIP 2004, vol. 4, pp. 2829–2832. IEEE (2004). https://doi.org/10.1109/ICIP.2004.1421693

17. Manevarthe, B., Kalpathi, R.: Geometric correction for projection on non planar surfaces using point clouds. In: Proceedings of the 12th International Conference on Distributed Smart Cameras, pp. 1–6 (2018). https://doi.org/10.1145/3243394.3243694
18. Wang, X., Liang, Y., Pan, Q., Yang, F.: A Gaussian approximation recursive filter for nonlinear systems with correlated noises. Automatica **48**(9), 2290–2297 (2012). https://doi.org/10.1016/j.automatica.2012.06.035
19. Camera Calibration and 3D Reconstruction (2024). OpenCV.org. https://docs.opencv.org/4.x/d9/d0c/group__calib3d.html. Accessed 25 July 2024