Virtual Shadow Rendering for Maintaining Situation Awareness in Proximal Human-Robot Teaming

Andrew Boateng and Yu Zhang
aoboaten,yzhan442@asu.edu
School of Computing, Informatics, and Decision Systems Engineering
Tempe, Arizona, USA

ABSTRACT

One focus of augmented reality (AR) in robotics has been on enriching the interface for human-robot interaction. While such an interface is often made intuitive to interact with, it invariably imposes *novel objects* into the environment. In situations where the human already has a focus, such as in a human-robot teaming task, these objects can potentially overload our senses and lead to degraded teaming performance. In this paper, we propose using AR objects to solely augment natural objects to avoid disrupting our natural senses while adding critical information about the current situation. In particular, our case study focuses on addressing the limited field of view of humans by incorporating persistent virtual shadows of robots for maintaining situation awareness in proximal human-robot teaming tasks. We designed a novel process to generate such shadows and verified that they worked better than other commonly adopted methodologies in teaming scenarios.

KEYWORDS

Augmented Reality, Virtual Shadow, Human-Robot Interaction

ACM Reference Format:
Andrew Boateng and Yu Zhang. 2021. Virtual Shadow Rendering for Maintaining Situation Awareness in Proximal Human-Robot Teaming. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion), March 8–11, 2021, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3434074.3447221

Figure 1: User view of persistent virtual shadow

1 INTRODUCTION

Over the past decade, there has been accelerated growth in robotic applications, making it no longer far-fetched to envision robots as part of our lives. What stands out to be most appealing are teaming applications where humans and robots complement each other to achieve complex tasks [15]. For effective collaboration, it is important for the human to maintain team situation awareness, which is known to benefit teaming performance [7]. Maintaining team situation awareness, requires the human to monitor the team status at any point of time. However, given our limited fields of view, it is easy to lose sight of our robotic teammates and hence the team situation awareness. This is especially true when we must focus on our own tasks at hand while interacting with robots intermittently. In such situations, loss of situation awareness can reduce team productivity and potentially cause serious safety risks.

1.1 Motivating Scenario

Consider Dan, a human worker, and his partner robot in a car assembly workshop. As a team, Dan is supposed to focus on his part of the job while the robot performs its tasks. In one scenario, Dan sends the robot to fetch a hot soldering rod. However, Dan may not know exactly when the robot will return. He would like to context-switch to his own tasks before the robot returns instead of idly waiting for the rod. In such situations, he would prefer to keep an eye on the robot at all times to avoid safety risks and to ensure receipt of the rod as soon as it arrives. One main problem we have here is that we are unable to perceive objects that are outside our field of view. When robots move behind our back, it is easy to lose awareness of the current situation.

While there are various ways to circumvent this problem, a general solution is unfortunately lacking. For example, one solution is to have the robot announce its status (e.g., approaching) using beeping sounds. Such a solution only works well when there is little noise. Another solution is to have the robot always approach from the front. It however, does not completely address the issue as the status of the robot while behind the human is lost.

To enable humans to see beyond the limitations of what our eyes offer, augmented reality (AR) devices are used to insert virtual objects into real-world to add useful information to the environment. AR has proven to be useful in industries such as education [2], health, security, manufacturing [8], and entertainment [20]. While AR provides an intuitive and general interface for communicating with humans, applying it to human-robot teaming tasks involves the following challenges:

(1) **Requiring persistent and dynamic information projection:** For mobile collaborative robots, dynamic information

HRI '21 Companion, March 8–11, 2021, Boulder, CO, USA

ACM ISBN 978-1-4503-8290-8/21/03…$15.00
https://doi.org/10.1145/3434074.3447221

must be constantly communicated to the human to maintain team situation awareness. Adding a meaningful dynamic holograms to the environment often requires the holograms and their dynamics to be custom-made. For example, consider holograms for indicating that a robot is picking up an object vs. moving forward.

(2) **Minimizing visual distractions:** Holograms are not part of our natural environment. So, adding such objects to the environment, especially when they are dynamic, will be distracting.

1.2 Contributions

Based on the above discussion, we see an imperative need to develop a fundamentally different principle of applying AR for robots to communicate with humans without overloading our senses with humans without overloading our senses. We propose ***Persistent Virtual Shadows (PVSs)*** as a general methodology for maintaining situation awareness without creating much distraction to humans. Since shadows are encountered every day, their addition to the environment will be perceived as natural and in no way overbearing for us. Furthermore, shadows are seamlessly integrated into our surroundings and provide rich and high-fidelity information about the actual objects.

In this paper, we introduce a shadow rendering system that enables humans to maintain critical information about the robot at all times. We make use of Hololens to generate the PVS projection experience (as shown in Fig 1). Instead of creating new objects, we only augment the existing objects (i.e., robots) via naturalistic projections to add the needed information to the environment. Our method is influenced less by environmental conditions, such as noise and lighting, and provides information that extends beyond our visual limitations in an intuitive way. Our contributions in this work include:

(1) Introduced and implemented a novel process for generating PVSs that involves modeling environment, virtual light placement, and virtual shadow rendering.
(2) Evaluated PVSs to demonstrate that they are able to communicate a rich set of information while being less distracting than the implemented baseline methods.

2 RELATED WORK

AR empowers us to visually perceive and interact with objects that are not present in the physical world [4]. Due to this ability, it is used in at least 12 distinct domains [12] that include military [19], marketing [17], geospatial [11] among others. The increasing use of this technology has drawn many researchers into the field of AR.

An immediate application of AR is to introduce novel objects that can be interacted with [10, 13, 20]. The focus has been on making the objects more interactive and realistic. For example, Wang et. al [21] make use of lightning and shading of real scenes to modify AR objects making them more lifelike. In robotics, these objects can be used as part of the interface to facilitate interaction [9].

Another use of AR is to provide information about objects or process (e.g., instructions) in the real-world. Such information is often superimposed onto physical objects or imposed into our views to display additional information. A closely related work to ours is the work on intention projection using projectors. In particular, Anderson et al. [3] projected virtual parts of an object onto another physical object to highlight the right place to insert the part. However, the humans in these prior works are required to pay full attention to the information provided, which may not be practical in human-robot teaming situations.

The increasing collaboration between humans and robots has paved way for investigating into effective ways of communication. Researchers have studied different modalities of communication other than vision, such as using sound [6, 14]. Others studied more subtle ways such as body language [16, 18]. However, these methods are not universal and can be easily affected by culture, demography, etc., in addition to environmental conditions, such as noise.

3 APPROACH

3.1 General Flow

The flow of our persistent shadow rendering process is illustrated in Fig. 2. As shown in the Fig 2, the robot first sends data about its orientation and position to the HoloLens. This information from the robot together with that of the low-energy Hololens scanning (Fig. 3) of the environment are combined to create a virtual environment in the Hololens akin to the real environment. Based on the robot's information and a pre-specified robot 3D model, we create the shadow of the virtual robot in this virtual environment and extract the shadow as a hologram. During the shadow generation process, our approach determines the right position to place our virtual lights to ensure that the shadow rendered is always in the view of the human. The necessary steps needed to create the persistent view of a robot are listed below, which will be detailed later:

(1) Environment Modeling
(2) Virtual light placement
(3) Virtual Shadow Rendering

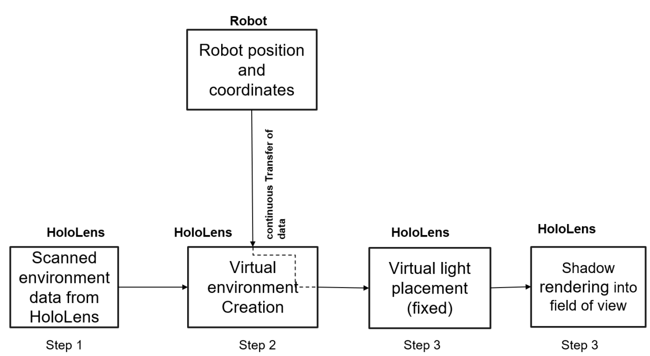

Figure 2: General flow of the shadow rendering technique

3.2 Environment Modeling

AR devices constantly scan their environment for anchoring and low-energy mapping. Our method combines the semi-autonomous nature of SLAM modeling, the accuracy of using pre-built maps and the easiness of building an environment in a virtual space.

Since the Hololens already possess a 3D scan of any environment, no extra work is required for mapping the environment. We extract this information for the use of our environment. The 3d Hololens generated model is converted into a transparent, cutout shadow-receiving game object in unity (see Fig. 4). This is done to prevent the model from interfering with the real-world environment when deployed. The cutout model is cast onto the surfaces of the real-world objects. The robot localizes in the real-world environment.

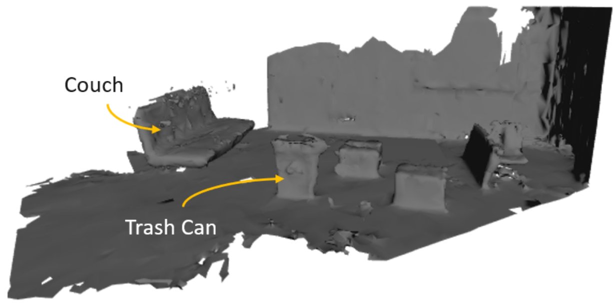

Figure 3: Low-Energy scan of our working scene by Hololens

To get virtual shadows of the robot onto the real world, we cast shadows from the robot unto the virtual environment. We then extract the shape, form and orientation of these shadows and project them unto the real surfaces corresponding to the cutout world. The design of unity's shadow casting algorithm enables the shadow to pass through objects. To address this issue, our method assigns the parts of the environment that are irrelevant the color of black, which makes them transparent in the real world therefore preventing them from receiving shadows that pass through objects.

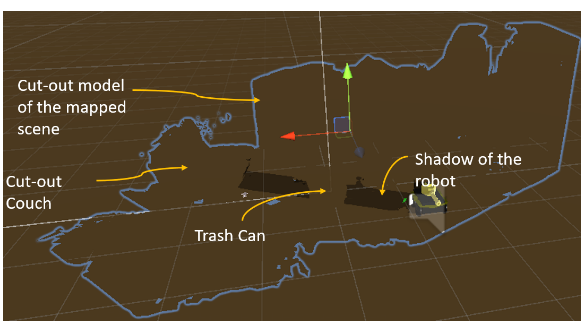

Figure 4: A transparent model of our working scene with robot shadows to be superimposed onto the real-world

3.3 Virtual Light Placement

We use directional light to enable humans see shadows whenever it is needed. The directional light is made a child of the main camera to keep it relatively constant. It is rotated such that the beams are directed towards the HoloLens' center. It is then strategically placed such that the robot is always between the Hololens position and the light source. This ensures the robot's shadow is always in the field of view. Due to the sensitivity of the Hololens, a slight tilt or rotation of the head changes the direction of the directional light. Unlike position, a relative change in the Euler angles between the directional light and the camera has significant effects on the direction of light rays and the intensity of the illumination.

Because the human head makes involuntary movement, the rotation transforms of the Hololens change a lot. It is therefore necessary to keep the relative orientation of the light source constant. We scripted the directional light so that the Euler angular difference between it and the camera (it's parent) is zero. This enables the directional light to have constant rotational angles relative to the camera. This is important in ensuring that we have a constant feed of the shadow in the human's field of view.

3.4 Shadow Rendering

Unity uses a depth buffer system to keep track of all surfaces that are close to the light. If any surface comes in direct line with light source, the surface will be illuminated. The other surfaces will not be illuminated thereby being in the shadow [1]. We model our shadows of the robot using the unity shadow rendering approach. After the shadows are created, and rendered onto the virtual objects, it takes the form of the object (just like any shadow). To prevent projecting unnecessary information, we disable shadow casting for all objects except for the robot. We then model these shadows as game objects and render them via HoloLens (just like any hologram) into the real world. Since the virtual world is a replica of the real world, a mere scaling of the shadow by the right amount will align them perfectly with real world objects giving the real shadow effect.

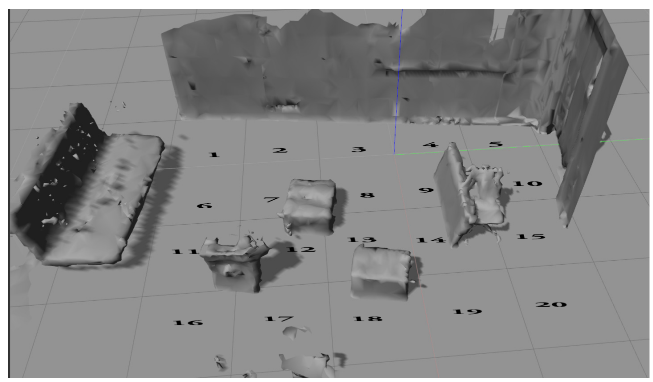

Figure 5: Discretized environment viewed in ROS

4 EXPERIMENTAL DESIGN

We evaluated three different methods that can be used to maintain a robot's status in proximal human-robot teaming. They included our persistent virtual shadow method (PVS), a method that uses an arrow to indicate the robot's status, and a minimap method. The two baseline methods were chosen since they are commonly used methods for providing situation awareness [5]

For fair assessment, we kept the same scenarios and environment for all methods. The survey was a between-subject design. For experimental purposes, we mapped and modeled the environment into a ROS world (Fig. 5). The modeled world was discretized to enable participants specify a particular location by mentioning the number assigned to it.

Recorded videos of real-life human collaboration with robots (Fig. 6) were given to the participants in the survey. After taking the survey, participants were provided the NASA Task Load standard

Figure 6: The human-robot teaming task setting and user view of the three methods: arrow, map, shadow, respectively

questionnaire (TLX) for subjective studies and evaluation of the effectiveness of the methods used.

For each method, the evaluation is divided into 2 separate parts: 1) position and orientation and 2) movement. This is done to gauge and show how effective our method is for projecting a rich set of information about the robot's status. We showed various pictures and videos of the robot shadow in different positions and orientations. The robot also performed different activities while behind the human. These same scenarios were used for the map and arrow methods. The participants were tasked to identify, e.g., where the robot was and what it was doing. Fig. 6 shows sampled screenshots of the three methods presented to the participants in the survey.

5 RESULTS AND ANALYSIS

We recruited 70 human subjects to take the survey on Mturk. 18 out of the total responses were discarded for failing the sanity checks.

5.1 Position and Orientation

From Fig. 7, it is observed that both arrow and map methods performed badly in one of the categories. Arrow did not perform well for position. This is mainly because arrows are able to correctly point to a direction of an object but not able to provide much depth information. The lack of depth information made it difficult for participants to estimate the position of the robot. The map method did better because we are used to reading maps and obtain spatial awareness when reading them. It was therefore easy to estimate the robot's position. PVS performed comparable to map if not better. However, the map did worse in orientation. This is due to the fact that the representation of the robot on the map is circular.

PVS outperformed the other methods for orientation. This is because neither the map nor arrow methods are able to communicate information about the object's orientation or shape. Shadows on the other hand is a projection of the real object so it provides information about both. The participants in the map and arrow methods however can tell orientation based on the robot's movement.

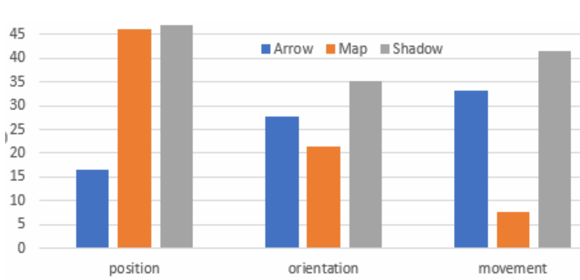

Figure 7: Results for position, orientation and movement

5.2 Movement

In general, PVS performed better than the two baselines. The reason has to do with the details presented by the shadow method. Shadows provide high-fidelity dynamic information about the actual objects. The map method however performed poorly.

A student t-test is performed between each pair of methods based on the overall accuracy. It yielded a p-value of 0.001 between the shadow and the map method. Similarly, we got a p-value of 0.007 between the shadow and the arrow method.

Fig 8 presents the result of subjective metrics. Even though we did not observe a significant difference between methods (which may be due to sample size), it is observed that the shadow appeared to have outperformed the other methods in most of the metrics. One main metric worth discussion is "*Frustration*". When people are not able to solve a problem or are not able to figure out something, their frustration level increases. Having the lowest frustration score for the shadow is an indication that, the shadow method did not distract the users much. Surprisingly, most people thought the shadow was less physically demanding. This could be due to the fact that it is easy to correlate the position of the actual robot with its shadow. In which case, participants did not have to scroll up and down the pages to figure out the status of the robot.

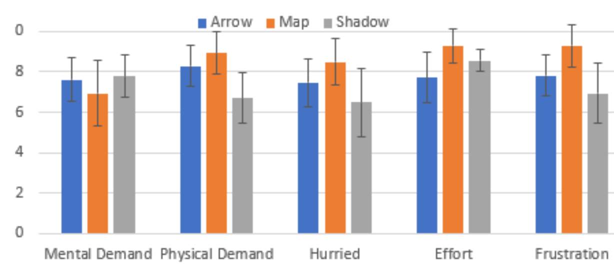

Figure 8: NASA TLX results with standard deviation bars

6 CONCLUSIONS

In this paper, we address the problem of maintaining situation awareness in proximal human-robot teaming. We apply AR technology to augment the existing objects rather than create new ones to minimize distractions while providing useful information. We propose persistent virtual shadows to enable the human maintain sight of the robot during interaction tasks. Our study has shown that our method is effective and introduces less distraction.

ACKNOWLEDGMENTS

We thank the reviewers for their helpful comments. This research is supported in part by the NSF grant IIS-1844524, the NASA grant NNX17AD06G, and the AFOSR grant FA9550-18-1-0067.

REFERENCES

[1] 2015. Unity - Manual: Shadows. (2015).

[2] 2018. Trends in Educational Augmented Reality Studies: A Systematic Review. *Malaysia online journal of educational technology* 6, 2 (2018), 60–74.

[3] R. S. Andersen, O. Madsen, T. B. Moeslund, and H. B. Amor. 2016. Projecting robot intentions into human environments. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 294–301.

[4] Igor Baranovski, Stevan Stankovski, Gordana Ostojić, Sabolč Horvat, and Tegeltija Srđan. [n.d.]. Augmented reality support for self-service automated systems. *Journal of Graphic Engineering and Design (JGED)* 11, 1 ([n. d.]), 6.

[5] F. Bork, U. Eck, and N. Navab. 2019. Birds vs. Fish: Visualizing Out-of-View Objects in Augmented Reality using 3D Minimaps. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 285–286. https://doi.org/10.1109/ISMAR-Adjunct.2019.00-28

[6] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. 2017. A Survey on Dialogue Systems. *ACM SIGKDD Explorations Newsletter* 19, 2 (Nov 2017), 25–35. https://doi.org/10.1145/3166054.3166058

[7] Nancy J. Cooke, Jamie C. Gorman, Christopher W. Myers, and Jasmine L. Duran. 2013. Interactive Team Cognition. *Cognitive Science* 37, 2 (2013), 255–285. https://doi.org/10.1111/cogs.12009 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12009

[8] S.J Henderson and S Feiner. 2009. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. , 135-144 pages.

[9] Ayoung Hong, Burak Zeydan, Samuel Charreyron, Olgac Ergeneman, Salvador Pane, M. Toy, Andrew Petruska, and Brad Nelson. 2017. Real-Time Holographic Tracking and Control of Microrobots. *IEEE Robotics and Automation Letters* 2 (01 2017), 1–1. https://doi.org/10.1109/LRA.2016.2579739

[10] Carolien Kamphuis, Esther Barsom, Marlies Schijven, and Noor Christoph. 2014. Augmented reality in medical education? *Perspectives on medical education* 3, 4 (September 2014), 300—311. https://doi.org/10.1007/s40037-013-0107-7

[11] Mark Livingston, Simon Julier, Dennis Brown, J. II, Joseph Gabbard, and Deborah Hix. 2008. An augmented reality system for military operations in urban terrain. (12 2008).

[12] M. Mekni and A. Lemieux. 2014. Augmented Reality : Applications , Challenges and Future Trends.

[13] George Michalos, Panagiotis Karagiannis, Sotiris Makris, Önder Tokçalar, and George Chryssolouris. 2016. Augmented Reality (AR) Applications for Supporting Human-robot Interactive Cooperation. *Procedia CIRP* 41 (2016), 370 – 375. https://doi.org/10.1016/j.procir.2015.12.005 Research and Innovation in Manufacturing: Key Enabling Technologies for the Factories of the Future - Proceedings of the 48th CIRP Conference on Manufacturing Systems.

[14] Dylan Moore and Wendy Ju. 2018. Sound as Implicit Influence on Human-Robot Interactions. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. Association for Computing Machinery, New York, NY, USA, 311–312. https://doi.org/10.1145/3173386.3176918

[15] Harley Oliff, Ying Liu, Maneesh Kumar, Michael Williams, and Michael Ryan. [n.d.]. Reinforcement learning for facilitating human-robot-interaction in manufacturing. *Journal of Manufacturing Systems", volume = 56, pages = 326 - 340, year = 2020, issn = 0278-6125, doi = https://doi.org/10.1016/j.jmsy.2020.06.018, url = http://www.sciencedirect.com/science/article/pii/S0278612520301084* ([n. d.]).

[16] Raul Benites Paradeda, Mojgan Hashemian, Rafael Afonso Rodrigues, and Ana Paiva. 2016. How Facial Expressions and Small Talk May Influence Trust in a Robot. In *Social Robotics*. Springer International Publishing.

[17] J. Pilet, V. Lepetit, and P.Fua. 2008. Fast Non-Rigid Surface Detection, Registration and Realistic Augmentation. *Internationa Journal in Computer Vission*, 109–122.

[18] Darja Stoeva and Margrit Gelautz. 2020. Body Language in Affective Human-Robot Interaction. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*. Association for Computing Machinery, 606–608. https://doi.org/10.1145/3371382.3377432

[19] E. C. Urban. 1999. AThe information warrior.

[20] Yolanda Vazquez-Alvarez, Ian Oakley, and Stephen A Brewster. 2012. Auditory display design for exploration in mobile audio-augmented reality. , 987-999 pages.

[21] Yang Wang and Dimitris Samaras. 2003. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graphical Models* 65, 4 (2003), 185 – 205.