Designing Augmented Reality Visualizations for Synchronized and Time-Dominant Human-Robot Teaming

Savannah Paul[1], Christopher Reardon[2], Tom Williams[1], and Hao Zhang[1]

[1]Colorado School of Mines, Golden, CO USA
[2]U.S. Army Research Laboratory, Adelphi, MD, USA

ABSTRACT

Team communication is crucial in multi-domain operations (MDOs) that require teammates to collaborate on complex tasks synchronously in dynamic unknown environments. In order to enable effective communication in human-robot teams, the human teammate must have an intuitive interface that supports and satisfies the time-sensitive nature of the task for communicating information to and from their robot teammate. Augmented Reality (AR) technologies can provide just such an interface by providing a medium for both active and passive robot communication. In this paper we propose a new Virtual Reality (VR) based framework for authoring AR visualizations, and demonstrate the use of this framework to produce AR visualizations that help facilitate high task performance in synchronized, time-dominant human-robot teaming. In this paper we propose a new framework that uses a virtual reality (VR) simulation environment for developing AR strategies as well as present a AR solution for maximizing task performance in synchronized, time-dominant human-robot teaming. The framework utilizes a Unity-based VR simulator that is run from the first person point of view of the human teammate and overlays AR features to virtually imitate the use of an AR headset in human-robot teaming scenarios. Then, we introduce novel AR visualizations that support strategic communication within teams by collecting information from each teammate and presenting it to the other in order to influence their decision making. Our proposed design framework and AR solution has the potential to impact any domain in which humans conduct synchronized multi-domain operations alongside autonomous robots in austere environments, including search and rescue, environmental monitoring, and homeland defense.

Keywords: Augmented Reality, Human-Robot Teaming, Cooperative Teaming, Time Dominant, Information Communication

1. INTRODUCTION

As a result of steady improvements in robotics, the integration of robots into the work space to operate alongside humans has become an increasingly popular solution to improving productivity and task efficiency across various domains.[1] However the ability for humans and robots to function as independent peers in the work space with coordinated actions has yet to be successfully implemented; often the effectiveness of human-robot teams and the consistency in their performance is not guaranteed.[2] Autonomous adaptation to perception changes, dynamic environments, and previously unseen events is an open problem; for this reason, manual teleportation or tightly supervised control is often used for complex tasks in the real world.[3] Human time is valuable, so supplementing teams or even replacing some human members with robots conserves resources and allows for humans to focus on the other tasks that cannot be completed by robots.[1] Now when considering human-robot teams in the context of MDOs, time dominance is a key requirement that needs to be factored into the teams behavior. In scenarios that require *time dominance* the exchange of information to support task performance is not only time sensitive but also competitive in the sense that each decision, successful or not, results in a notable consequence. For this reason it is crucial that both teammates operate synchronously with

Further author information: (Send correspondence to Hao Zhang)
Savannah Paul: savannahpaul@alumni.mines.edu
Christopher Reardon: christopher.m.reardon3.civ@mail.mil
Tom Williams: twilliams@mines.edu
Hao Zhang: hzhang@mines.edu

one another by sharing relevant information in an appropriate time frame and maintaining an awareness of each other's progress. Enabling this form of time dominant human-robot teaming would allow for robot teammates to operate as independent peers thereby increasing task efficiency of the team and reducing the amount of repetitive work done by the human.

Augmented reality has previously been proposed as a potential solution for improving human-robot communication as it can intuitively express information to the user while also enhancing their situational awareness.[4–6] Situational awareness plays a large role in a person's ability to make quick and informed decisions in dynamic environments,[7] so using AR to express the information gathered from the robots' perceptual data can the potential to improve the overall team's performance.[8]

Existing AR systems for enabling effective human-robot communication tend to focus on the fundamental aspects of interaction within teams such as explicit verbal understanding as well as implicit understanding of intentions from previous actions or gestures. In order to imitate natural human interaction, these AR systems have combined human speech and gestures in order to communicate intentions and commands to the robot teammates.One such multi-modal system[9] allows the human to reference objects and assign tasks to the robot in a natural manner by using AR to disambiguate the gestures and translate the information into an appropriate form for the robot to operate on (see also[10]). This system was able to disambiguate commands such as "go to this" or "go behind that" by using speech processing to parse the verbal commands in conjunction with gesture processing to recognize gestures and match them with words such as "that", "this", "here", or "there." Another approach focuses on another fundamental requirement for team members performing collaborative activities, the ability for each member to rapidly assess each others' actions and attitudes.[11] They developed a series of AR cues designed to communicate robots' intended movements in order to reduce the waiting time for shared resources. Using human experiments they found that each technique effected performance to a different extent, with techniques that used AR overlays placed throughout the scene resulting in better task completion than those summarizing the scene in a fixed location.

Figure 1: An example of how varying the amount of AR usage impacts the user experience.

While these existing AR methods are sufficient for human-robot teams to accomplish straightforward tasks, there still exist many gaps that are left to be filled in order for AR to be an effective part of a complete solution to human-robot teaming in more complex scenarios. Previous approaches do not address the problem of selecting what information to communicate between teammates or when to communicate it. Varying the quantity and quality of communicated information enables robots to influence the behavior of the human teammates consuming said information to differing degrees of effectiveness. Research regarding the inclusion of different AR communication strategies on task performance has shown a correlation between task performance and various types of cognitive load imposed by the AR system, such as visual or auditory load, as well as correlations between task performance and different types of communication strategies, such as gestures and natural language.[12] While it is necessary to keep the human teammate informed with the most recent information, it is critical not to negatively affect decision making by overstimulating them as well. In Figure 1, we see three different designs for an AR overlay with varying amounts of information being displayed. The images on the right and left show how using too few visualizations is equally as detrimental as showing too many. The optimal solution lies somewhere in between these designs and avoids overwhelming the human's cognitive capacity with visuals while maintaining utility. In addition, several technical obstacles exist to communicating and visualizing information between robot and human teammates, including maintaining an accurate alignment between robot and human frames without the use of external instrumentation, communicating spatially accurate information between teammates in the context of that transform, and maintaining a manageable visualization of the information under the constraints imposed by the AR device.

In this paper, we propose a novel framework for designing new AR strategies, and an AR solution to time-dominant human-robot teaming. The proposed design framework utilizes a Unity-based virtual reality (VR) simulation environment that is run from the first person point of view of the human and overlays AR features to virtually imitate the use of an AR headset in human-robot teaming scenarios. Integrating the simulation into the AR design process allows for AR features to be iteratively prototyped, tested, and improved within a virtual environment to increase the productivity of real world testing. In addition, we introduce a new set of AR visualizations that displays information collected from both the human and robot teammates in order to support strategic communication and improve each team members decision making. The developed AR features reduces the density and complexity of the information displayed through the strategic use of colors and established military symbology.

The remainder of the paper is structured as follows. In Section 2, we introduce the proposed design framework. Section 3 presents visualization strategies designed for human-robot teaming in multi-domain operations. The paper is then concluded in Section 4.

2. PROPOSED DESIGN FRAMEWORK

2.1 Framework Overview

Our proposed design framework for developing new AR systems utilizes a Unity-based virtual reality (VR) simulation for creating and testing new AR strategies. One iteration of the design process is pictured in Figure 2. This framework begins with preliminary drafting of the appearance and behavior of the desired AR features and then enters a three stage cycle of (1) virtual prototyping of the AR design, (2) testing that virtual prototype within the VR simulation, and (3) refining the design based on the results from testing. This cycle is continued until the design is approved for real world AR testing. From here, the process may repeat or enter the iterative cycle again depending on the results from AR testing. The benefit of using this framework is that it maximizes the productivity of real world testing by allowing unsuccessful features to be phased out in the early stages of the design process before resource-intensive real-world testing is initiated.

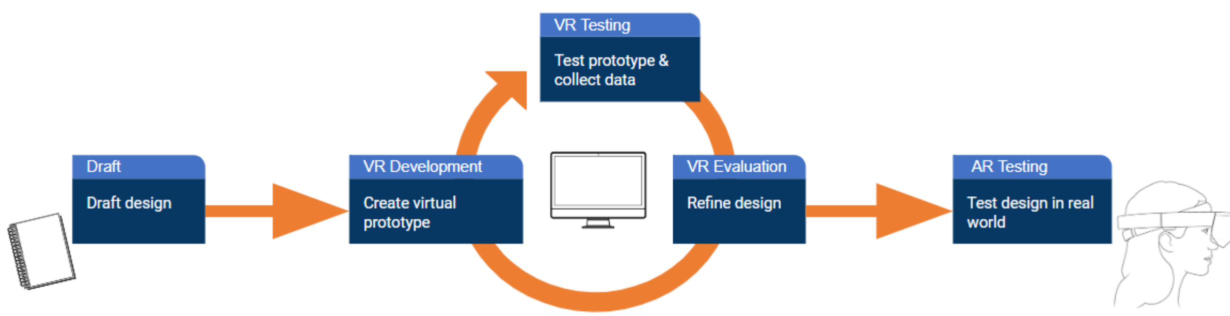


Figure 2: Design Framework Flowchart

The Unity based simulation environment created for DCIST* interfaces with ROS allowing for a range of robots and devices to be loaded into a variety of detailed virtual Unity scenes and interact with the scene in the same way a robot would in the real world. This simulation tool is utilized during the three stage cycle of the framework and serves as the basis for the virtual environment. The interfacing between Unity and ROS is supported through the use of ROS#, a set of software libraries providing web socket connections for Unity-ROS intercommunication. The Unity and ROS nodes during the virtual stages of the framework later develop into the application run on AR headset worn by the human and physical robot during the real world testing stage of the framework. In other words, ROS# is essential for sharing information between the human and robot during both the VR and AR stages of the framework. ROS# is used to support Unity-ROS communication for both the VR and AR implementations of the design since both AR and VR applications can be developed on the

*Distributed and Collaborative Intelligent Systems and Technology Collaborative Research Alliance http://dcist.org/

Unity platform. This results in a seamless transition from virtual to real world testing. To this end, the VR design of the AR application must conform to real world restrictions by maintaining a relatively light load on the Unity node in order to produce high quality visualizations. As a result, all calculations and data processing are computed on the ROS node and used to generate and send different events to the Unity node in order to trigger or alter various visualizations. ROS topics are the channels that are responsible for sending these events between Unity and ROS.

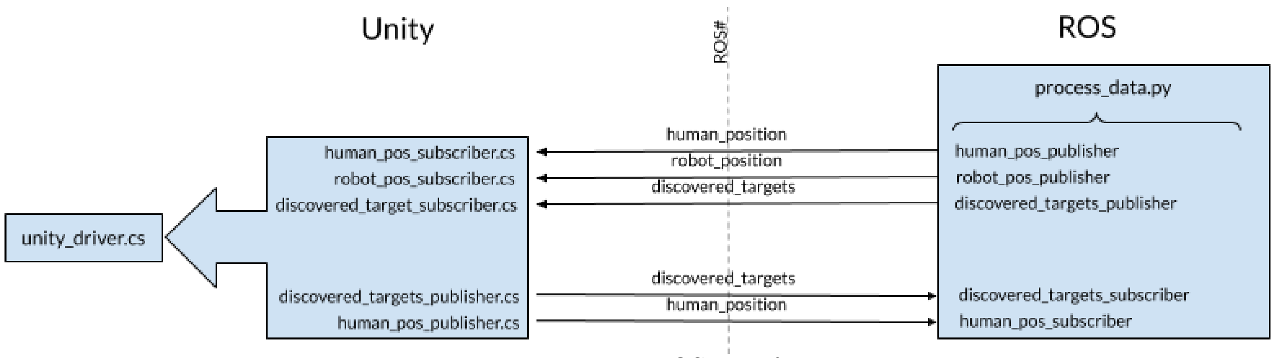


Figure 3: Unity-ROS Interface

Figure 3 shows the relationship between all these components and how the collected data travels throughout the system. On the ROS node exists a python script that processes and summarizes the scene by performing periodic calculations using information such as the elapsed time, target and teammate positions, and meeting point position. These calculations are used to gather and manage information about the scene that can be useful to the human teammate, when a new piece of information arrives or is updated an event is sent as a message to the Unity node. The Unity node will then generate or modify an AR visualization to be displayed to the human based on the type of message received. While some messages are event-based, other messages such as teammate locations or the current time are sent at a constant rate to remain time synchronous and ensure that the human and robot and both kept informed with the most accurate information. The events sent from ROS to Unity encapsulate all the information required to construct the AR visuals allowing for the Unity node to remain dedicated to solely rendering visuals. Keeping the Unity node dedicated to visualizations enables the three stage cycle of the framework to be more modular and design centered by providing the ability to revise strictly the AR visuals without needing to modify the ROS node.

3. DESIGN OF TIME-DOMINANT AR VISUALIZATIONS

We use the proposed framework and simulation environment to develop an AR solution for maximizing task performance in synchronized, time-dominant human-robot teaming. Our design is intended for cooperative human-robot search and exploration tasks. In order to test and evaluate the effectiveness of our design we needed to model this type of scenario as an exercise for the human and robot to complete in the VR simulation. To accomplish this, we employ target-search scenario. Using a simulated real-world environment populated with numerous buildings, we distribute targets throughout the scene for the human and robot to find. To emulate the time-dominant nature of real world MDOs the human and robot are given a time limit and a location to reach by that time. Performance of the exercise is measured by the amount of targets the team is able to interact with. Targets in this case represent tasks of varying priority that the robot teammate can identify and localize but not interact with, for example searching files or checking for signs of life, so the robot must alert the human to the targets' locations for the human to address. Two types of targets are scattered throughout the scene as seen in Figure 4 and each take a fixed amount of time for the human to interact with. The time limit imposed on the simulation prevents the human from being able to address all the targets, making leveraging the information provided by the AR visuals essential in order to maximize the number of target interactions.

The is a challenging scenario for human-robot teams because it requires each teammate to be perpetually aware of each others' positions as well as have a shared understanding of each others' progress and future intent. In order to maximize the team's performance, the robot and human need to cover as much ground as possible while avoiding redundant searching. This requires both the robot and human to be aware of not only where the other is in the scene but also where they have both previously searched. As previously stated, AR has high potential for expressing spatial related data, like keeping the robot visible through walls, however this information needs to be accurately sized and placed in the world frame in order to be functional. Further, keeping the human's point of view unobscured is essential in search-related tasks, so an extra emphasis is placed on making sparse but effective AR visuals. We believe that providing AR visualizations to improve the human teammates situational awareness will provide a significant advantage in this scenario. For the team to be most successful, the human must be kept informed of the robot's location, the location of the meeting place, and the locations of the targets discovered by both the human and robot. As the robot moves through the scene it is collecting information on the location and type of the targets encountered as well as the time spent in each area of the map. This information can be useful to the human when attempting to make critical choices regarding different areas and the time-value tradeoff for searching them as well as mentally building a concrete understanding of the scene.

In the following sections, we describe how the proposed iterative design framework was used to develop AR visualizations that we expect to improve the team task performance in this time dominance scenario.

Figure 4: VR Simulation Map and Key

3.1 Phase line

We adapted the military phase line symbol into an AR visual to function as the next meeting place for the human and robot. Traditionally a phase line is a control measure used by the military in order to coordinate team movement.[13] Phase lines are displayed on maps, typically drawn along geographical features as seen in Figure 5. Initially our design was similar to the traditional phase line, however after two iterations of virtual testing we determined that it could be enhanced to improve situational awareness by extending its design into a three dimensional wall tall enough for the human to view from far away and behind buildings as seen below in Figure 6. The location of the phase line is initially stored on the robot and is sent to the human with the initialization event and displayed at the start of the exercise. The phase line visual can be leveraged to keep team movements synchronous by placing more phase lines throughout the scene and altering the times to reach them. This would not only help coordinate team movement but could also serve to allow the robot to determine what information is valid in the current context. Since this visual is sufficiently tall, it appears above buildings and can be seen from most locations

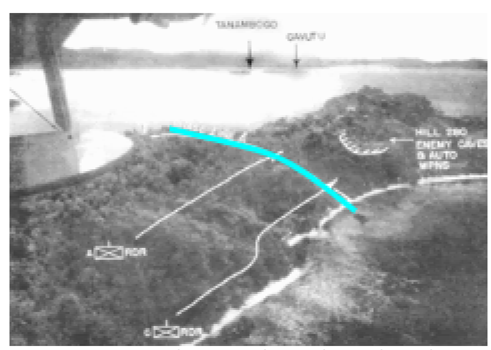
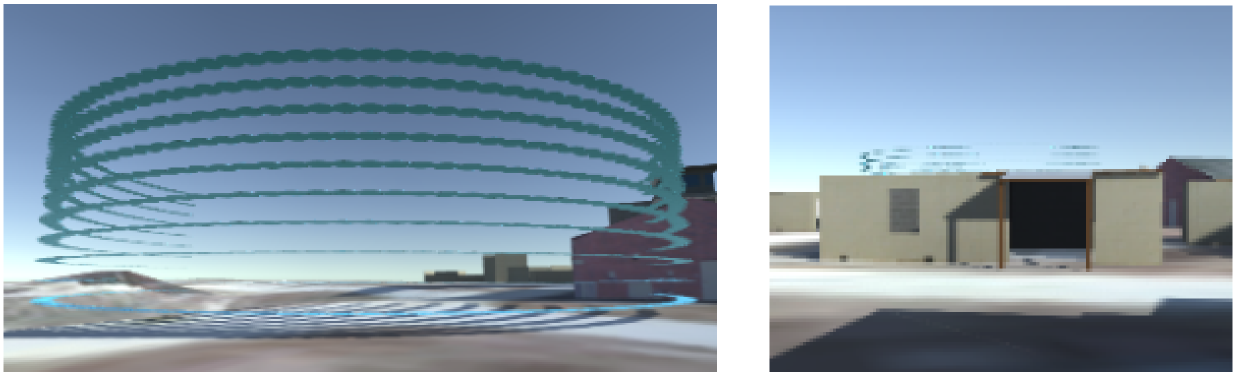


Figure 5: Military phase line drawling

throughout the environment. This helps the human maintain situational awareness and orientation within the environment, similar to how landmarks function to situate people in unfamiliar environments.



(a) View from up close (b) View from far away

Figure 6: AR phase line visual

3.2 Human timebar

This AR visualization is fixed to the human teammates view, so this visual follows them in the left corner of their view as they move throughout the simulation. The timebar, seen in Figure 7, numerically displays the time remaining for the team to reach the next phase line at the top of the bar and below displays a visual representation of the remaining time with respect to the total allotted time for the exercise as well as the portion of that remaining time required to reach the phase line based on the humans current location. On the ROS side, the total and spent time is managed as well as the human's phase line exit time that is calculated from the human and phase line locations both are sent to the human on the Unity node at a constant interval keeping the timebar up to date at all times. Displaying the remaining time relative to the total time that was given for the exercise allows the human to visually assess the appropriate pace to reach the meeting point by evaluating the rate at which the remaining time, represented by the green portion of the bar, shrinks. The blue portion represents the time to reach the phase line based on the humans current location, this value updates dynamically as the human moves throughout the environment so as they travel further from the phase line the blue portion will consume a larger amount of the green free time and will consume less as they move closer. This is intended to give the human a quick understanding of the time they have left for exclusively searching and interacting thereby enabling them to make faster decisions related to target searching without having to jeopardize reaching the exit on time.

Figure 7: Human timebar

3.3 Target timebars

The target timebars are displayed above each target as they are discovered and are analogous to the human timebar with similar colors corresponding to the same values as in the human timebar. These values are calculated on the ROS node and sent to Unity at constant intervals just as the human timebar is. The target timebar however also includes in orange the amount of time that represents the cost for the human to interact with the target, based on the type of target. The green free time portion on the target timebar therefore shows what the human's free time will be after interacting with the target. For teams to be successful in time-dominant scenarios they must make decisions in terms of the trade-offs between the time cost and value of an action.

By using similar color schemes in the design on the human and target timebars the human is able to make fast visual assessments of the aforementioned factors. Figure 8 below shows how the human can use the two timebars to determine the time related value of interacting with a target. Because the blue portion represents the exit time to reach the next phase line, which is dynamically calculated based on the human's position, it can additionally be used by the human from further away to determine if they can even afford to travel to a given target. Using the human's own timebar in conjunction with the target timebars makes them explicitly aware if they are running out of time well before they are actually in danger of not reaching the phase line. The simplicity of this design makes it easier for the human to rapidly compare and make competitive decisions between targets without being overburdened by mental calculation. Target timebars rotate to follow the human teammate so they can be seen unobstructed from a further distance.

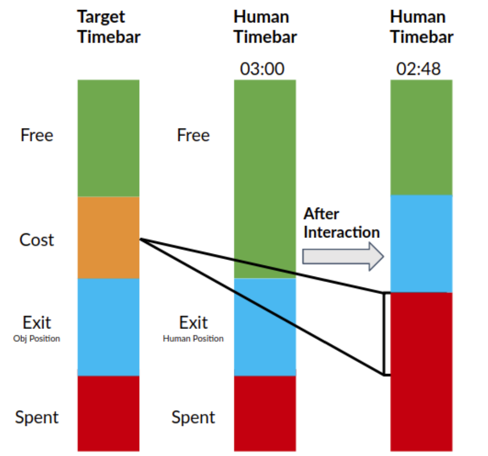


Figure 8: Human timebar before and after interacting with a target

3.4 Target marker

Target markers appear above a target after being found by either the human or robot. The color of the marker represents the type of target and whether or not it has been interacted with yet as seen in Figures 9, 10, and 11 below. This information is stored with the other target information on the ROS node of the robot, and sends events to Unity to update the color of the marker. These visualizations aid in orienting the human within the environment by functioning as virtual breadcrumbs, showing the human where they have searched as well as where the robot has searched and what it has discovered. These visuals can be seen above buildings to give the human an idea of a building's target density. This not only assists in choosing between buildings but also alerts the human of nearby targets that do not require extensive searching to reach.

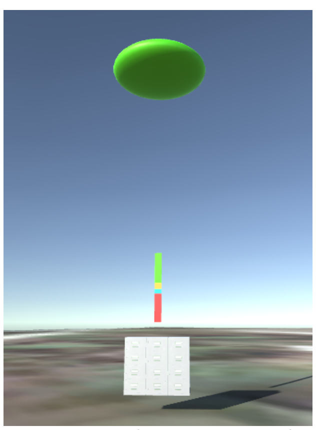


Figure 9: File target marker

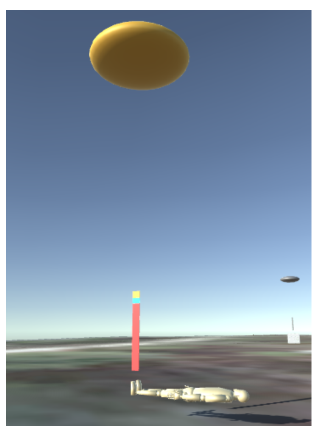


Figure 10: Dummy target marker

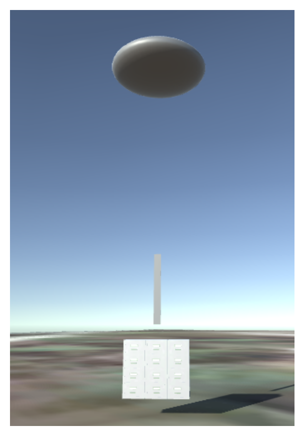


Figure 11: Post-interaction target marker

3.5 Summary timebar

The summary timebars represent the cost for the human to search the entire building and interact with all the targets inside. They are displayed at the entrance of a building after the robot has entered and exited the building, as seen in Figure 12, and are visually similar to the other timebars other than being thicker in width to provide visual distinction. The human can strategically use these visuals in combination with the target marker density to optimally choose between searching different buildings. This visual was introduced after initial drafting to add to the effectiveness of the target markers by identifying decision points for the human through consolidating the cost of the known targets in a building. The addition of this feature proved to aid in the humans performance and understanding of imminent time constraints.

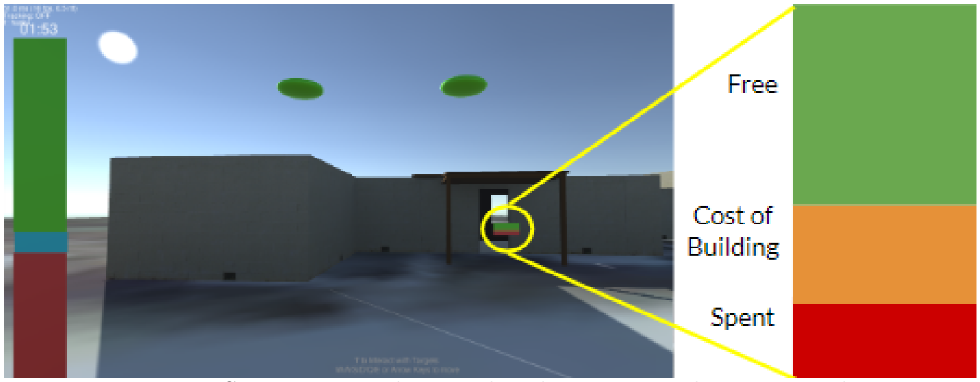


Figure 12: Summary timebar in placed in scene with target markers

4. CONCLUSION

In this paper we have presented a novel framework for designing and testing new AR strategies and have applied this framework to develop an AR solution to time-dominant, synchronized human-robot teaming. Through the development of our AR solution we demonstrated the impact of the iterative framework on the progression of the AR design and its potential to enhance future AR development. Additionally the specific AR strategy we developed explored the effects of various AR strategies on situational awareness and task performance in human-robot teams conducting MDOs. The time-dominant nature of our scenario and specific application to military operations had a major influence on the final AR design and exercise for evaluating it. The system we created collected information from both human and robot teammates in order to produce AR visualizations to the human that shape strategic team communication. This was done by reducing the density and complexity of the information collected through strategic use of color and establishes military symbology as seen with the phase line and consistent color scheme.

Acknowledgments.

This work was partially sponsored by the CCDC ARL DCIST CRA W911NF-17-2-0181.

REFERENCES

[1] Harriott, C. E., Buford, G. L., Zhang, T., and Adams, J. A., "Human-human vs. human-robot teamed investigation," in [*ACM/IEEE International Conference on Human-Robot Interaction (HRI)*], (2012).

[2] Scott A. Green, J. Geoffrey Chase, X. C. M. B., "Evaluating the augmented reality human-robot collaboration system," in [*International Journal of Intelligent Systems Technologies and Applications*], (2008).

[3] Hoffman, G. and Breazeal, C. (2004).

[4] Reardon, C., Lee, K., Rogers, J. G., and Fink, J., "Augmented reality for human-robot teaming in field environments," in [*International Conference on Human-Computer Interaction*], (2019).

[5] Williams, T., Szafir, D., Chakraborti, T., and Amor, H. B., "Virtual, augmented, and mixed reality for human-robot interaction," *ACM/IEEE International Conference on Human-Robot Interaction* (2018).

[6] Williams, T., Szafir, D., and Chakraborti, T., "The reality-virtuality interaction cube," *ACM/IEEE International Conference on Human-Robot Interaction* (2019).

[7] Endsley, M. R., "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society* **37**(1), 32–64 (1995).

[8] Martha Grabowskiab, Aaron Rowenb, J.-P. R., "Evaluation of wearable immersive augmented reality technology in safety-critical systems," *Safety Science* **103**, 23–32 (2018).

[9] Scott A. Green, Xioa QiChen, M. B. J. G. C., "Collaborating with a mobile robot: An augmented reality multimodal interface," *IFAC World Congress* **41**, 15595–15600 (2008).

[10] Sibirtseva, E., Kontogiorgos, D., Nykvist, O., Karaoguz, H., Leite, I., Gustafson, J., and Kragic, D., "A comparison of visualisation methods for disambiguating verbal requests in human-robot interaction," in [*IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*], (2018).

[11] Walker, M., Hedayati, H., Lee, J., and Szafir, D., "Communicating robot motion intent with augmented reality," in [*ACM/IEEE International Conference on Human-Robot Interaction (HRI)*], (2018).

[12] Tran, N., Phung, T., Mizuno, K., Hirshfield, L., Grant, T., and Williams, T., "Exploring mixed reality robot communication under different types of mental workload," *ACM/IEEE International Conference on Human-Robot Interaction* (2020).

[13] [*Terms and military symbols*], Headquarters, Departments of the Army (2013).