InvisibleRobot: Facilitating Robot Manipulation Through Diminished Reality

Alexander Plopski,* *Member, IEEE*
Nara Institute of Science and Technology

Ada Taylor†
Carnegie Mellon University

Elizabeth Carter‡
Carnegie Mellon University

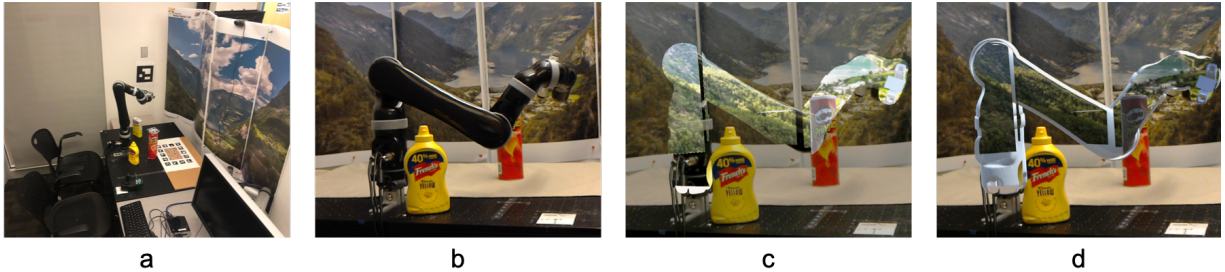Henny Admoni§ *Member, IEEE*
Carnegie Mellon University

Figure 1: (a) In our experiment environment, users sit to the left and right of the robot while grasping and moving objects on the table. (b) When users sit to the right, the robot's body occludes the user's view of the robot controller. (c) InvisibleRobot removes the robot's body from the user's view through diminished reality, thus allowing the user to manipulate the controller without adjusting their viewpoint or the robot's pose. (d) Alternatively, InvisibleRobot also shows the outline of the robot to help users maintain awareness of the robot's pose while controlling it in narrow or complex environments.

ABSTRACT

When controlling robots, users often face the issue of an operating area that is occluded by an element in the environment or the robot's body. To gain an unobstructed view of the scene, users have to either adjust the pose of the robot or their own viewpoint. This presents a problem, especially for users who rely on assistive robots as they can't easily change their point of view. We introduce InvisibleRobot, a diminished reality-based approach that overlays background information onto the robot in the user's view through an Optical See-Through Head-Mounted Display. We consider two visualization modes for InvisibleRobot: removing the robot body from the user's view entirely, or removing the interior of the robot while maintaining its outline. In a preliminary user study, we compare InvisibleRobot with traditional robot manipulation under different occlusion conditions. Our results suggest that InvisibleRobot can support manipulation in occluded conditions and could be an efficient method to simplify control in assistive robotics.

Keywords: Diminished Reality, Robot Manipulation, Optical See-Through Head-Mounted Display

Index Terms: Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

Robots can augment human capabilities by increasing an operator's capacity to perform tasks such as picking up heavy objects, reaching for items outside of their reach, and performing repetitive motions with unprecedented accuracy and precision. These strengths can be applied in environments ranging from a large-scale industrial setting to working within the home to expand an individual's physical capabilities. Particularly in close-quarters environments, operators often face occlusion problems when manipulating robots due to the layout of the environment or the robot's body being in the user's view [1, 2]. To overcome this problem, operators have to either re-position the robot or adjust their own viewpoint of the scene, which may not be always an option. In particular, persons with disabilities who use robots to extend their self-sufficiency cannot necessarily easily change their position to view the scene from a different location.

To address this limitation, we introduce InvisibleRobot (Fig. 1), a novel approach to support robot manipulation even when the robot's body occludes objects of interest. In particular, we apply diminished reality (DR) to remove the occluding robot from the user's view and allow direct manipulation of the object. DR has previously been considered in a variety of tasks, such as X-Ray vision [5], manufacturing [4], and teleoperation [6]. However, these systems were designed for video see-through systems and are not viable for in-situ robot control due to physical constraints of mobile devices and safety concerns of video see-through head-mounted displays. Although the robot could also be removed from the user's view through projectors and retroreflective material [3], this approach is impractical as it requires modification of the robot and estimation of the user's viewpoint by external cameras. To overcome these limitations, we use an optical see-through head-mounted display (OST-HMD) that presents the computer graphics (CG) to the user without occluding the scene.

2 INVISIBLEROBOT VIEW

We consider two ways of presenting users with a diminished view of the robot. The first diminished view overlays background scenery information onto the robot to remove it from the user's view, keeping only a digital representation of the gripper to support grasping and moving of objects occluded by the robot (Fig. 1c). This visualization method limits the presentation to the most essential elements and should help participants more easily operate the robot in occluded conditions.

*e-mail: plopski@is.naist.jp
†e-mail: adat@andrew.cmu.edu
‡e-mail: ejcarter@andrew.cmu.edu
§e-mail: henny@cmu.edu

A limitation of a purely diminished view of the robot is that it retains only subtle features of the robot's joints, making it potentially difficult for users to maintain a mental model of the robot's overall pose. This is especially important when the robot is operated in tight environments or its movement is constrained by its design, e.g., joint limits. To address this limitation, we developed a second visualization that augments InvisibleRobot with an outline of the robot's shape (Fig. 1d).

3 PROTOTYPE SYSTEM

To evaluate the effects of InvisibleRobot on the user's ability to manipulate the robot, we created a simple scene shown in Fig. 1a. Although the ideal system would reconstruct the background in real-time and update the rendering accordingly, we opted to use a prepared model of the scene to account for the computational cost, network load, and rendering capabilities of our OST-HMD, the Microsoft HoloLens. To simplify the tracking of the environment, we attached fiducial markers to important elements, such as tracked objects, walls, and the table of our experiment setup (Fig. 1a). We calculated and recorded the transformations between all elements relative to the marker placed on the table.

Our robot is a Kinova Mico 2 arm that is connected to a computer running Ubuntu 16.4 and the ROS environment. We attached a camera to the robot's arm and used hand-eye calibration to determine its pose relative to the gripper. We then took an image of the table marker and computed its pose relative to the robot's base from the robot's kinematics and our hand-eye calibration. Finally, we manually adjusted the pose of the base and the robot joints to ensure visual alignment in the user's view. We also stored the location of the table marker relative to the HoloLens' coordinate system. The system recovers the stored pose on subsequent runs to simplify the operations.

During runtime, we load and render the models of the environment at the recorded poses. We share the robot's pose with the HoloLens over a local WiFi with a latency of about 50 ms. We found that although this delay is noticeable, it does not significantly affect the overall experience, as users tend to reduce the speed at which they control the robot when they need to perform minute adjustments.

4 PRELIMINARY EVALUATION

Our evaluation scenario simulates a common operation of the Kinova Arm where a user moves an object (a Pringles can) from one spot on the table to another. The environment has two additional distractor objects that users have to avoid while performing the task. Users can control the position and orientation of the robot's gripper through a 3DOF joystick. As the joystick provides only 3DOF of input to control the 6DOF robot arm, users can simultaneously control the position (x, y, z), rotation (pitch, yaw, roll), or the robot's gripper by switching between different joystick modes with the click of a button on the controller. Although it would be ideal to use a controller that provides full 6DOF control, we opted for the joystick as it is a common operation tool for this robot and is familiar to wheelchair users.

We conducted a preliminary experiment with three participants to gain feedback on the system's usability and their experiences using it. None of the participants had previous experience controlling the robot, while two participants had briefly used the HoloLens in the past. To study the impact of different occlusion conditions we asked participants to sit either to the left or to the right of the robot. When participants sat to the left of the robot, they could observe the operation area directly, without any occlusions. While sitting to the right, the robot's frame would occlude the initial and target locations, requiring users to move their head or take advantage of InvisibleRobot. We designed the two scenarios to determine whether participants liked using InvisibleRobot in general or if the ability to

switch this visualization on and off would be more suitable for practical applications. We randomized the order in which participants tried the different visualization techniques and location in which they sat.

5 RESULTS

Although participants practiced controlling the robot for five minutes before the experiment, they struggled with the controls, often failing to move the Pringles can to its target location within five minutes. Participants named confusion with the control modes as the main limitation, which is a common issue with this type of robot teleoperation.

Interestingly, participant feedback did not consistently indicate that one mode was more difficult than another. This could be because although we advised participants to focus only on the visualization, they judged the overall system where they struggled with the robot controls.

Nevertheless, two participants preferred using InvisibleRobot when the robot did not occlude the operation area while one participant preferred the standard view. When the robot was occluding the operation area, two participants preferred using InvisibleRobot with outline while one participant preferred using InvisibleRobot. This suggests that the outline helped the participants maintain a mental image of the robot to improve the controls, while the see-through option helped them grasp the object even when it was occluded by the robot.

6 FUTURE WORK

In the future, we plan to conduct a formal study to determine the applicability of our system to support assistive robot manipulation. We also plan to investigate how visualization misalignment that can be expected in practical situations affects the user's ability to control the robot, the user's satisfaction, and mental image. Finally, we plan to investigate how to update the background scene presented on the OST-HMD in real-time and identify other criteria for smart removal of the robot and the scene.

ACKNOWLEDGMENT

REFERENCES

[1] R. M. Aronson, T. Santini, T. C. Kübler, E. Kasneci, S. Srinivasa, and H. Admoni. Eye-Hand Behavior in Human-Robot Shared Manipulation. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, pp. 4–13, 2018.

[2] R. Holladay, L. Herlant, H. Admoni, and S. S. Srinivasa. Visibility Optimization in Manipulation Tasks for a Wheelchair-Mounted Robot Arm. In *Proceedings of the RO-MAN Workshop on Human-Oriented Approaches for Assistive and Rehabilitation Robotics*, 2016.

[3] M. Inami, N. Kawakami, and S. Tachi. Optical Camouflage Using Retro-Reflective Projection Technology. In *Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 348–349, 2003.

[4] S. Mori, M. Maezawa, and H. Saito. A Work Area Visualization by Multi-View Camera-based Diminished Reality. *Multimodal Technologies and Interaction*, 1(3):18, 2017.

[5] C. Sandor, A. Cunningham, A. Dey, and V.-V. Mattila. An Augmented Reality X-Ray System Based on Visual Saliency. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality*, pp. 27–36, 2010.

[6] K. Sugimoto, H. Fujii, A. Yamashita, and H. Asama. Half-Diminished Reality Image Using Three RGB-D Sensors for Remote Control Robots. In *Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robotics*, pp. 1–6, 2014.