Opportunities for In-Home Augmented Reality Guidance

Jaylin Herskovitz
University of Michigan
Ann Arbor, MI, USA
jayhersk@umich.edu

Eyal Ofek
Microsoft Research
Redmond, WA, USA
eyalofek@microsoft.com

Walter S. Lasecki
University of Michigan
Ann Arbor, MI, USA
wlasecki@umich.edu

Adam Fourney
Microsoft Research
Redmond, WA, USA
adamfo@microsoft.com

ABSTRACT

The use of Augmented Reality (AR) systems has been shown to be beneficial in guiding users through structured tasks when compared to traditional 2D instructions. In this work, we begin to examine the potential of such systems for home improvement tasks, which present some specific challenges (e.g., operating at both large and small scales, and coping with the diversity in home environments). Specifically, we investigate user performance of a common low-level task in this domain. We conducted a user study where participants mark points on a planar surface (as if to place a nail, or measure from) guided only by virtual cues. We observed that participants position themselves so as to minimize parallax by kneeling, leaning, or side-stepping, and when doing so, they are able to mark points with a high degree of accuracy. In cases where the targets fall above one's line of vision, participants are less able to compensate and make larger errors.

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3312933

KEYWORDS
Augmented reality; task guidance; computer-aided instruction; DIY

We discuss initial insights from these observations and participant feedback, and present the first set of challenges that we believe designers and developers will face in this domain.

INTRODUCTION AND BACKGROUND

AR has been used successfully in guiding users through a variety of structured, physical tasks, such as repairing equipment [2] and assembling objects [4, 7]. In this domain, AR systems show many clear benefits over traditional 2D paper or computer-delivered instructions. For example, AR guidance systems result in increased task efficiency by both reducing the time it takes to locate the next step of the task [6], and by reducing the cost of context switching between a workspace and a set of instructions [12]. AR can also decrease task error rates by reducing cognitive load, specifically in reducing the spatial cognition needed to transform 2D instructions to a 3D space [13]. Because of these advantages, years of effort have been spent improving the sensing and tracking technology at the heart of these guidance systems.

Although many of these systems have focused on industrial tasks [3, 6, 8], AR might also prove useful in guiding people through tasks in their homes, for example, in home improvement projects. However, while current industrial assembly guidance systems are successful, they are not readily deployable to the home due to several key differences between industrial and domestic settings. One major difference is in the complexity and scale of hardware that can realistically be deployed. Existing systems are mainly focused on being as accurate as possible, which requires specialized, external hardware to track the user's head and the location of relevant, predefined objects within a specified area. Not only is this cost prohibitive, but it also limits the mobility and scale of such systems. Recently, camera-based inside-out tracking has begun to address both problems, and high-quality consumer AR is becoming extremely close to availability. However, camera-based tracking is not a panacea, and there remain limitations on the accuracy of tracking. Any Do-It-Yourself (DIY) AR system targeted to consumers will need to deal with such inaccuracies for the time being.

A second difference between industrial and home settings is the scale and diversity of environments – DIY tasks can potentially span multiple rooms, and room architecture can vary widely from premises to premises. To overcome the variation of the environment, many AR guidance applications use only a small area of the room, such as a flat table, as a workspace [12, 14]. In industrial settings, applications may also rely on known, fixed geometry for tracking, such as when servicing a specific car engine [1, 6].

Here, we examine the effects of these differences on a low-level task common to many DIY projects: marking points on a planar surface (e.g., a wall) in a large physical workspace. DIY projects commonly include operations like adding fasteners (nails, screws, etc.) or cutting, and these operations typically involve the user measuring to a desired point and marking the point with a utensil. AR could potentially reduce the effort and error involved in this process.

Prior work has used AR to guide users to and instruct users to act on existing physical landmarks [6, 9], but in our case, users rely solely on virtual cues as guidance. Users may be misled by slight inaccuracy in the positioning of visual guides, unexpected parallax between the guides and work surface, and unexpected behavior related to occlusion (guides that should be occluded by the user's hands or tools but are instead rendered on top). To examine this, we create a task that asks users to view a series of virtual guidelines and mark a wall based on these features alone – a task that would involve tedious measuring if performed with traditional tools. We then measure how accurately this task can be performed with the Head Mounted Device's (HMD) built-in camera-based tracking system, and how issues such as occlusion and parallax may interfere with performance.

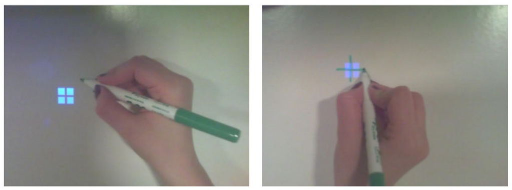

Figure 1: Point marking procedure. Left: A point before marking. Right: Desired outcome.

PROCEDURE

Fifteen participants were recruited (4 female, 11 male; 1 left handed) from a large technology company. The average age was 29 (min 17, max 44). Three participants were frequent users of AR technology (i.e., they used an AR headset at least once a month over the past year). However, the majority of participants had very limited AR experience, including four participants who had never before used an AR device. Upon arriving to the study, participants were asked to complete a survey which captured demographic information and previous AR/Virtual Reality (VR) experience. Participants were then familiarized with the Microsoft HoloLens HMD.

The task was completed in a workspace approximately 2.25 meters long and 1.5 meters wide. Along the longer side was a wall upon which a large magnetic whiteboard was mounted. Room lighting was set to ensure that the whiteboard was well lit so that participants could see the board and their markings through the dark lenses of the HoloLens. Since the wall was white and devoid of visual features, we affixed a calibration marker (a removable sticker) at the approximate center of the whiteboard – this position ensured that the marker was visible to the HoloLens cameras from most viewing angles. We used Vuforia Engine's Unity package to track the marker from the HoloLens [11]. The marker provides visual features for the tracking, and establishes a reference point in 3D space from which we independently measure the accuracy of each participant's actions.

Table 1: Correlation coefficients

	Median Error	Median Horizontal Error	Median Vertical Error
X-Axis Position	0.01225	0.10875	-0.08060
Y-Axis Position	**0.49686**	0.14331	**0.50632**

Participants were then asked to use a dry-erase marker to draw a mark as accurately as possible at 16 predetermined locations on the physical whiteboard. The location of each point was indicated by the AR HMD as a one centimeter by one centimeter blue square with a red cross in the center (see Figure 1). Participants were told they should trace the cross onto the whiteboard with the marker.

Users were shown one point at a time, and performed the HoloLens's Air-Tap gesture (a built-in system gesture analogous to a click on a traditional mouse) to move to the next point. The points were selected pseudo-randomly within a one meter by one meter square centered vertically approximately 1.35 meters from floor height. The horizontal range of the points was 0.83 meters, and the vertical range was 0.86 meters (minimum 0.90 meters from floor height, maximum 1.75 meters from floor

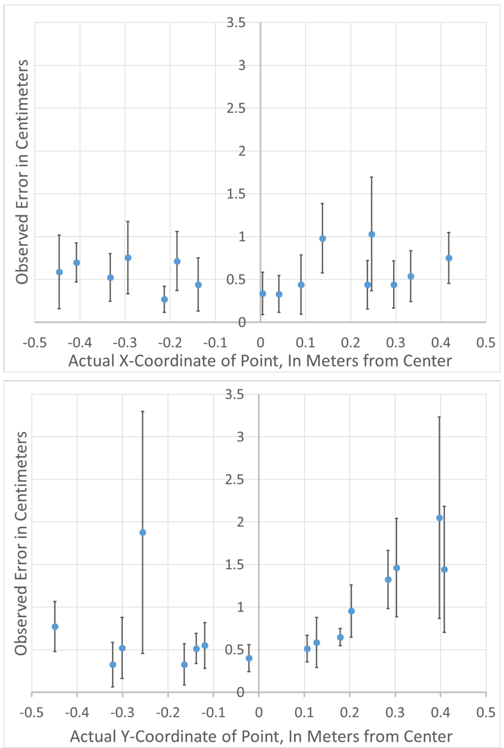

Figure 2: Average error for a point versus its horizontal position (top) and vertical position (bottom) on the wall. Position here is shown relative to the marker's position. As the points get higher than the marker, the accuracy of the marking deteriorates.

height). The closest pair of points was 10.9 centimeters apart. Each participant was shown the same points in the same sequence.

Once the participant marked the 16 points, a picture of the whiteboard was taken with a high-resolution digital camera. Based on the location of the pen marks in the image, and the transformation defined by the image of the calibration marker, we measured the participant's accuracy for each point. Immediately following the task, participants completed the NASA Task Load Index (TLX) questionnaire [5] and provided written answers to a series of open-ended questions about their experience.

RESULTS

Over all participants, the median error was 1.011 centimeters (Min: 0.186 cm., Q1: 0.708 cm., Q3: 1.719 cm., Max: 3.230 cm.). For each point, we then analyzed the correlation between its actual position on the wall and the median error over all participants (both in each component direction and overall). The correlation coefficients are shown in Table 1. We observe that the height of a point on the wall is strongly correlated with the absolute error of the observed markings. This effect is also shown in Figure 2, which shows the errors in the markings of each point relative to their distance from the center of the whiteboard (and the position of the calibration marker, the most predominant feature on the wall).

The average TLX Score over all 15 participants with all six categories weighted equally was 79.44 (out of 100) (SD = 12.54). Details of the scores are shown in Figure 3. We observed that there was no correlation between a participant's reported TLX Score (i.e., perceived difficulty of the task) and their average error (r = 0.07029). Some of the variation in these scores can be attributed to the design of the task, for example, participants mentioned that the limited device field of view combined with the fact that the points were randomly distributed made searching for the next point difficult (P1, P2, P5, P7, P8). Others mentioned that points became slightly out of focus when viewing them from a close range (P6, P8, P10, P14, P15).

However, participants also noted other visual factors that affect performance and perceived difficulty of the task. P15 noted that "It was hard to be precise because of the difference in depth. My eyes had to focus simultaneously on the mark and where I want to draw the mark." Other participants also confirmed that parallax was an issue, and that it was difficult to focus on the wall and point at once – a perceptual issue potentially caused by the annotation partially occluding the dry-erase marker. Some users noted that the calibration marker drifted from the correct position (possibly when out of the field of view of the device's cameras), resulting in a difference in depth between the point and the wall (P3, P7, P12, P14). One possible improvement to this can be to add more textures along the working area. Both of these factors have more of an effect on points higher than eye level, as the depth differences become more apparent then. We observed more inaccurate markings for higher points, see Figure 2 and Table 1. For points lower than eye level, participants would bend down or

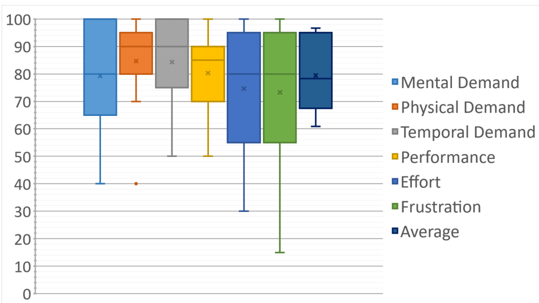

Figure 3: Participant reported NASA TLX Scores.

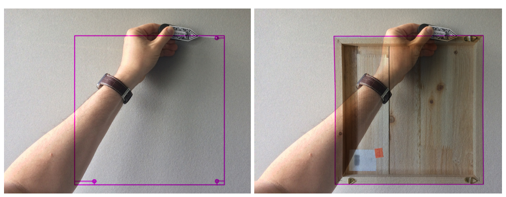

Figure 4: AR can potentially be used for many effort saving shortcuts in this domain. For example, this image shows a mock-up of an AR application for hanging a frame; instead of measuring the distance between two hooks at the back of a frame (left), the device could take a picture of the back of the frame and display it on the wall, letting the user easily see where nails are needed (right).

kneel in order minimize parallax and view the point as accurately as possible. This strategy could not be applied for targets that were above eye level.

DISCUSSION AND FUTURE WORK

Over time, we are likely to see changes in hardware improving device field of view, tracking accuracy, and computing power for standalone AR headsets, and a large amount of effort has already been put into making this possible. It might be some time until high quality, large field of view, low latency, light, and affordable headsets are widely available. Until that point, existing systems are likely to become cheaper and more common, thus, work for the immediate future should investigate what useful applications will be possible with the hardware available. From our observations, we highlight some considerations.

Given the perceptual issues noted above, an initial challenge in system design will be working around these factors. Some have straightforward solutions, for example, limited field of view has previously been addressed by systems that point the user in the direction of a target object if it is out of their line of sight. Work has also addressed the effect that scale of virtual objects can have on expected gestural interaction [10]. The effects of other issues on accuracy, such as occlusion, can be diminished through careful object placement, and future work should look to address these in full.

However, despite these limitations, AR guidance still has opportunities and advantages that we observed. For example, in our experiment, virtual guidelines for marking points are much likely faster than having to measure each point with physical tools, such as a ruler and level. Any modification to our setup, such as moving to a different wall, also only requires the visual marker to be moved.

Additionally, given the level of accuracy that we observed, there are opportunities for AR targeting to be used in real-world projects. While projects that require a high degree of accuracy may be out of the question for now, projects that can leverage relative placement or measurement are well-suited to current device capabilities. For example, when building a table, a user could utilize AR to measure and cut the length of one leg, and then use that leg as a reference to cut the others.

There are also many opportunities for future work investigating another advantage to AR in this space: its function as an external visual reference. For example, when hanging a picture, a user might need to measure all hooks on the back of a frame relative to its center so she can plan where to install nails. An AR system could behave as a memory to the user in this case, storing not only prior measurements but also displaying an image of the back of the frame directly on the workspace, making planing easy (Figure 4). Taking occasional snapshots of the workspace can also help the user recreate the status of devices or circuits prior to any changes she may have made. Data that was measured in one project may be recovered in the future for other projects in the same space, such as locations of studs, wires, or pipes in the walls. Advanced systems could even detect mistakes using the device's camera and help the user correct them.

CONCLUSION

In this paper, we have begun to examine the potential of commercial AR devices to aid in home improvement tasks. We did so through a user study which evaluated performance of a common low-level operation in this space – marking points on a planar surface – guided by AR cues. We found that users were able to mark points with a high degree of accuracy under optimal conditions, and discuss visual factors that affect performance, such as parallax and occlusion. We then present considerations and realistic opportunities for AR guidance systems in the home.

REFERENCES

[1] Hugo Alvarez, Iker Aguinaga, and Diego Borro. 2011. Providing guidance for maintenance operations using automatic markerless augmented reality system. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. IEEE, 181–190.

[2] Steven Feiner, Blair Macintyre, and Dorée Seligmann. 1993. Knowledge-based augmented reality. *Commun. ACM* 36, 7 (1993), 53–62.

[3] Pierre Fite-Georgel. 2011. Is there a reality in industrial augmented reality?. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. IEEE, 201–210.

[4] Mika Hakkarainen, Charles Woodward, and Mark Billinghurst. 2008. Augmented assembly using a mobile phone. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 167–168.

[5] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[6] Steven Henderson and Steven Feiner. 2011. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE transactions on visualization and computer graphics* 17, 10 (2011), 1355–1368.

[7] Steven J Henderson and Steven K Feiner. 2011. Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. IEEE, 191–200.

[8] Oliver Korn, Albrecht Schmidt, and Thomas Hörz. 2013. The potentials of in-situ-projection for augmented workplaces in production: a study with impaired persons. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 979–984.

[9] Michael R Marner, Andrew Irlitti, and Bruce H Thomas. 2013. Improving procedural task performance with Augmented Reality annotations. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*. IEEE, 39–48.

[10] Tran Pham, Jo Vermeulen, Anthony Tang, and Lindsay MacDonald Vermeulen. 2018. Scale Impacts Elicited Gestures for Manipulating Holograms: Implications for AR Gesture Design. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 227–240. https://doi.org/10.1145/3196709.3196719

[11] PTC Inc. 2018. Vuforia Engine. https://www.vuforia.com/

[12] Stephanie Rosenthal, Shaun K Kane, Jacob O Wobbrock, and Daniel Avrahami. 2010. Augmenting on-screen instructions with micro-projected guides: when it works, and when it fails. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 203–212.

[13] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 73–80.

[14] Giles Westerfield, Antonija Mitrovic, and Mark Billinghurst. 2015. Intelligent augmented reality training for motherboard assembly. *International Journal of Artificial Intelligence in Education* 25, 1 (2015), 157–172.