An Evaluation of Graphical Context When the Graphics are Outside of the Task Area

Cindy M. Robertson[1] and Blair MacIntyre[1] and Bruce N. Walker[2]

GVU Center, [1]School of Interactive Computing and [2]School of Psychology, Georgia Institute of Technology

ABSTRACT

An ongoing research problem in Augmented Reality (AR) is to improve tracking and display technology in order to minimize registration errors. However, perfect registration is not always necessary for users to understand the intent of an augmentation. This paper describes the results of an experiment to evaluate the effects of graphical context in a Lego block placement task when the graphics are located outside of the task area. Four conditions were compared: fully registered AR; non-registered AR; a heads-up display (HUD) with the graphics always visible in the field of view; and a HUD with the graphics not always visible in the field of view. The results of this experiment indicated that registered AR outperforms both non-registered AR and graphics displayed on a HUD. The results also indicated that non-registered AR does not offer any significant performance advantages over a HUD, but is rated as less intrusive and can keep non-registered graphics from cluttering the task space.

CR Categories and Subject Descriptors: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Evaluation/methodology

Additional Keywords: augmented reality, communicative intent, augmented environments, human-computer interaction.

1 INTRODUCTION

Augmented Reality (AR) has been shown to be a useful user-interface paradigm for various application domains. One main motivation for using AR is that the graphics are placed *in situ*, allowing AR systems to be designed to support hands-free interaction. However, in order to precisely align the graphics with physical objects, both the user and the objects must be accurately tracked (at least with respect to each other), and the whole system must be accurately calibrated. At this time, tracking and display technologies are not accurate enough to produce perfect registration between the graphical world and the physical world.

In our previous work, we discussed the AIBAS system, an adaptive intent-based augmentation system designed to use the communicative intent [11] of an augmentation to simplify the creation of AR applications that work in real-world situations with "good enough" tracking [9]. Our goal was to empower programmers by providing them with a framework to create augmentations that function in the presence of registration error. Our group has also modified an open source scene graph (OpenSceneGraph [7]) to estimate registration error at its transformation nodes (and thus the objects attached to those nodes) [1]. Using this estimate, we can design augmentations that adapt to changing registration error.

In a previous study, we added visual context (i.e., additional graphical cues) to an augmentation to help the user understand the intent of the augmentation when there was registration error present in the scene [10], and demonstrated that perfectly registered graphics are not necessary for users to accurately follow instructions in an AR system. Participants could understand and complete tasks when presented with potentially ambiguous augmentations (caused by registration error) when appropriate context was added to the augmentation.

We believe that a second reason AR is not widely used is the concern that the computer graphics may block a worker's view of the task space, thus interfering with their primary task. Such graphics could not only be annoying, but in certain tasks, could also be very dangerous. A desire to limit the amount of graphics in the task space guided our designs in AIBAS [9] and the previous experiment. However, when error, augmentations and context are present, the user's view of the task space may be unacceptably obscured.

Head-worn displays can still be useful in these situations, even if in-situ graphics are not used, because heads-up displays (HUDs) can also enable hands-free interaction. In this study, we evaluate the impact of situating graphics outside of the task area, comparing in-situ AR, graphics accurately positioned in 3D at the edge of the task space, and two alternative HUD designs that use orientation-only tracking to orient the graphical instructions based on the viewpoint of the worker. This study shows that, while the accurately registered AR presentation is most effective, graphical augmentations do not need to be located in the task area to be useful. This points the way to alternative display techniques when there are concerns about using AR in real world situations, such as when the task space must be unimpeded. Also, this paper will show that low-level orientation-only tracking can be used to provide enough information to create useful augmentations if a more sophisticated tracking system is not available.

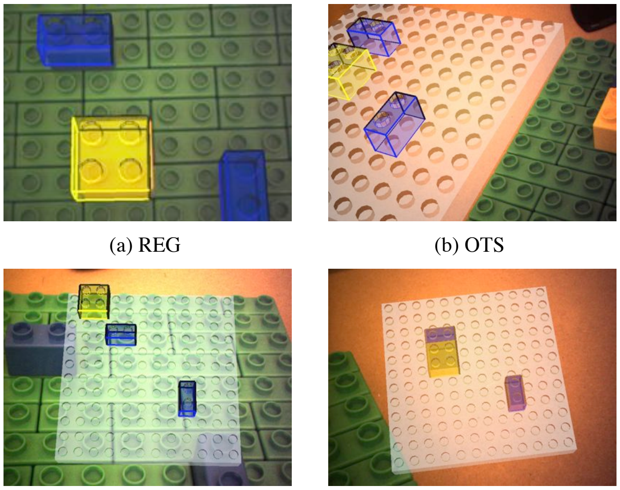
2 RELATED WORK

Tang et al. compared the effectiveness of augmented instructions in an assembly task [12]. That user study showed that the use of AR in the form of computer assisted instruction projected on a head-mounted display can improve task performance and can relieve mental workload as compared to a printed manual and computer assisted instruction using a monitor-based display. Two of our display cases are similar to one in this study, but the design of the study and the comparison with the other cases yields additional insight into the relationship between AR and HUDs.

Livingston et al. conducted a user study to determine which display attributes, including drawing style and opacity, best express occlusion relationships among far-field objects [4]. While

College of Computing, School of Interactive Computing
College of Sciences, School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332-0760
{leistner, blair, walkerb} @ cc.gatech.edu

IEEE International Symposium on Mixed and Augmented Reality 2008
15 -18 September, Cambridge, UK
978-1-4244-2859-5/08/$25.00 ©2008 IEEE



(a) REG (b) OTS

(c) HV (d) HS

Figure 1. The subject view of the four blocks of trials.

we are not explicitly concerned with occlusion in our studies, we are concerned about the most effective drawing style for representing our augmentations. That study provided insight when designing our augmentations. More recently, Livingston and Ai studied the specific impact of different kinds of registration errors on user performance in an outdoor AR system. Their results point the way toward more realistic error characterizations in future studies such as ours [5].

3 Orientation-only Tracking

One of the goals of this study is to evaluate the use of orientation-only tracking to create more usable HUDs as an alternative to *in-situ* AR. There are several benefits to this type of system, including cost and ease of design. Tracking systems, such as the InterSense IS-1200 tracker used in both of our experiments, are expensive and may be inappropriate for scenarios such as on-site equipment maintenance. However, if a HUD-based system using orientation-only tracking were close to the effectiveness of in-situ AR, the cost difference and difference in installation complexity would make such a system worth considering.

The initial design for the orientation-only HUD was to use three 3DOF orientation sensors (such as the InterSense InertiaCube series): one on the participant's body, one on their head, and one on the object being augmented (i.e., the Lego base plate in our experiment). If we assume the participant is facing the object, the angle between the body sensor and the object sensor would tell the system on which side of the object the participant is located. The HUD can therefore show the augmentation from roughly the viewpoint of the user. The angle between the head sensor and the body sensor allows us to implement a variation of the "always visible" HUD, where the graphics only appear when the user turns their head to one side.

In our experiment, we were already using a 6DOF InterSense IS-1200 head tracker, and knew where the base plate was located in the tracker coordinate system, obviating the need for a sensor on the base plate. Therefore, we attached one orientation sensor, an InterSense InertiaCube2, to the user's body. Together, these sensors allowed us to simulate the desired 3-sensor configuration.

In these two HUD cases, we wanted to see if performance in terms of error and time per block placement, as well as cognitive load, are comparable enough to the AR cases, to validate the usefulness of orientation-only tracking. If they are, the results could have a significant effect on application design.

4 Theory

This user study was designed to evaluate the effectiveness of graphical context in variations of AR where the graphics are intentionally not superimposed with the relevant objects in the world. HUD's have been shown to be useful, without the use of any tracking or registration. For example, the maintenance systems created with the Microvision Expert Technician System, a hands-free wearable display, have been used to provide mobile access to vehicle history and repair information. However, such systems have limitations compared to AR, such as not rendering the maintenance instructions from the viewpoint of the worker.

We have shown in our previous study that even a small amount of graphical context cues are useful in making "nearly-registered" AR usable, which made us wonder if graphical context would also be effective on a HUD, and how such HUD-based systems would compare to registered AR. The question we are asking could be phrased: When provided with enough context, how do AR and HUD-based systems compare? We have chosen to evaluate four cases.

Fully registered AR: When there is no visible misalignment between the graphics and the world, we can say there is no error. For our purposes, however, achieving absolutely no registration error in an AR system is impossible, so, we consider *no error* to be "negligible" error. In the case of the experimental setup, if the amount of registration error is less than half of the size of one of the Lego pegs on the base plate, there is no question as to where the block should be placed, and we consider this to represent no error. From here on, we will also refer to this case as the AR-registered case or REG.

Non-registered AR: If we do not want graphics obscuring the task space, we can purposely locate the graphics at the edge of the space, at a fixed position and orientation with respect to the Lego base plate. By doing this, we still have useful orientation information, and a fixed position offset. From here on, we will also refer to this case as the AR-off-to-side case or OTS.

Heads-up display with the graphics always visible: A HUD shows graphical instructions in a fixed position on the head-mounted display. Orientation information can be used to ensure the graphics are oriented in roughly the same orientation as the user's view of the physical base plate, as discussed in Section 3. These graphics are always visible in the user's field of view (FOV). From here on, we will refer to this case as the HUD-visible case or HV.

Heads-up display with the graphics not always visible: With orientation-only sensing, we can create a HUD that only displays the graphical instructions when the user is looking to the side. The graphics are not always visible in the user's field of view. In this experiment, the angle between the user's head and the user's body must be 30 degrees or more in either direction in order for the graphics to appear on the head- mounted display. From here on, we will refer to this case as the HUD-side case or HS.

5 Methodology

A within-subjects experiment was conducted. The independent variable in this study was the method for displaying the graphics on the head-mounted display (including four different methods). The dependent variables included time to complete each task, the number of errors, and perceived mental workload. We had ten specific hypotheses, which we omit here for space reasons; a complete discussion can be found in [8].

5.1 Participants

Participants were recruited via email and word of mouth. The study included 28 participants (12 male, 16 female), ranging in age from 18 to 29 years.

5.2 The Setup

The experimental setup was similar to our previous experiment [9]. Participants stood next to the desk on which a Lego base plate was located in a fixed position relative to fiducial markers that were mounted on the wall in front of them. They wore the head-mounted display that contained an InterSense IS-1200 tracker, a 60 frames-per-second Point Grey Flea camera, and a Sony Glasstron video see-through optical display. The camera is mounted above a right angle prism, moving the optical center of projection of the camera closer to the participant's eyes than would otherwise be possible, with the intent of reducing the parallax offset of the video-mixed head-worn display. They also had an InterSense InertiaCube2 orientation-only sensor attached to their waist. This was used in two of the four blocks of trials.

5.3 Session Information

Participants were asked to complete an introductory demographic and AR/VR experience questionnaire. They were then asked to complete two tasks: an Edinburgh handedness test [6] (to ensure they used their dominant hand) and the Spatial Learning Ability Test [2] (to allow us evaluate the relationship between spatial abilities and task completion).

Participants were then asked to read a training document and complete a training exercise to familiarize them with the task and experimental setup. They were reminded that they would be evaluated based on the amount of time that it takes them to place each block as well as the number of errors they make while placing a block; therefore, it was important for them to work as quickly and as accurately as possible.

5.4 The Placement Task

The task consisted of the following: picking up the yellow block, pushing a button to start the trial, placing the block, and pushing the same button to end the trial. After each trial, participants were asked to rate their confidence about the block placement, using a 5-point Likert scale, from 1 (I think the block is in the wrong place) to 3 (I don't know if the block is placed correctly) to 5 (I think the block is in the correct place).

If the block was placed correctly, the participant was informed as such and advanced to placing the block in the next location. If the block was placed incorrectly, the participant was informed as such, and was instructed to attempt to place the block in the correct location again. These steps were repeated until the block was correctly placed. After all of the trials were completed, the participant answered a survey questionnaire about their experience, including portions of the NASA TLX rating questionnaire [3]. We did not do a complete NASA TLX evaluation as Hart and Staveland did in our analysis. Instead, we used their Likert scales for mental demand, physical demand, effort, frustration, and perceived performance as a model for evaluating subjective workload in our study.

In the design of this study, there were a total of four blocks of trials: perfectly registered graphical instructions (AR-registered or REG, shown in Figure 1(a)), graphics located to the left of the physical board (AR-off-to-side or OTS, shown in Figure 1(b)), heads-up display with the graphics always visible (HUD-visible or HV, shown in Figure 1(c)), and heads-up display where users have to turn their head to the right or left for the graphics to become visible (HUD-side or HS, shown in Figure 1(d)). There were context blocks shown in the graphics for all four blocks of trials. For the purposes of this study, the context took the form of two virtual blue Lego blocks, shown in Figure 1, that represented two physical blue Lego blocks that existed on the Lego base plate. In the three non-registered cases, the base plate was also drawn. For each of the blocks of trials, the participants were presented with 18 targets. In order to eliminate any order biasing, a 4x4 Latin Square was used to determine the order of presentation of the blocks of trials.

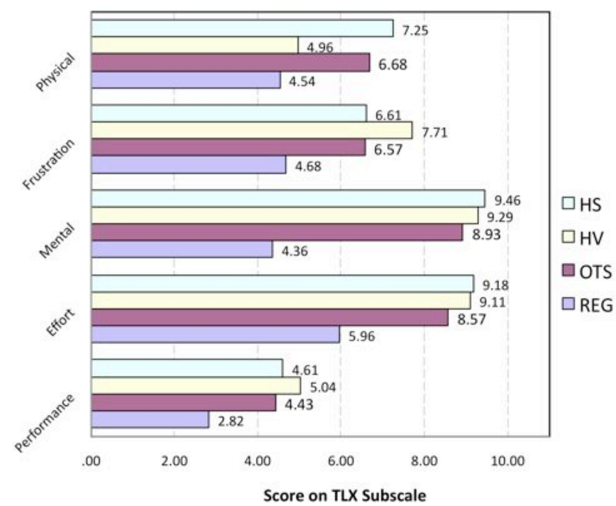


Figure 2. The subjective workload ratings for each condition.

5.5 Data Recorded

Several types of data were recorded during the experiment in addition to the questionnaires. First, trial data included block data (color, size), how many times the participant attempted to place each block, the time to complete each block placement, and the tracker data for each trial. Second, video data were collected, including a view of what the participant was seeing, a view of the participant from above to show where the participant's head was pointing, a frontal view of the participant to see where she was looking, and a view of the participant's hands to see how she placed the block.

6 Discussion

In this section we briefly discuss our findings. For a more detailed discussion, refer to Robertson's thesis [8].

As we expected, the participants completed the four Lego block placement tasks with few errors, but the times did vary among the conditions due to several factors such as head movement, need for memorization and screen clutter. The REG case always outperformed the other cases in terms of speed, errors and cognitive load; however the differences between the REG case and the other cases were not always significant.

We expected both AR conditions (REG and OTS) to outperform both HUD conditions. Contrary to our expectations, we found that the HV case outperformed the OTS case in terms of errors made and block placement times. We failed to anticipate the effects that head movement would have on placement times and the effects that memorization requirements would have on number of errors made and block placement times. These factors played a large role in shaping task performance in terms of error and time, as well as perceived performance, as shown in Figure 2. However, the subjects reported the AR conditions to be more natural than the HUD conditions, which was supported by their subjective workload ratings for effort, also shown in Figure 2.

In two of the cases we studied, the graphics were located outside of the participant's field of view, requiring them to look to the side to see the graphics (OTS and HS), while in the other two cases, the graphics were always visible (REG and HV). As expected, block placement times were slower in the two cases in which the users have to turn their heads to the side to see the graphics. There was a significant difference between both the REG and OTS cases and the REG and HS cases. In addition, the HV case took less time then both the OTS and HS cases; however, only the difference between the HV case and the HS case was significant (the HV visible case was approaching significantly less time than the OTS case, but was not actually significant). Similarly, the perceived physical demand increased when the graphics were not located in their field of view, although the values were not significantly different between all of the cases.

We expected the perceived mental workload and frustration of a participant to increase when the graphics were not located in their field of view, because they would need to remember the augmentations when the graphics were off-screen. This was true for REG versus OTS, but not for HV versus HS. While the graphics were visible in the HV case, it appears that the extra annoyance of having to switch focus between the physical Lego board and the non-registered graphics blocking the screen space imposed additional mental workload, as well as frustration, and caused the differences to be insignificant between HS and HV.

We had expected the HV case to be slow. Surprisingly, despite the fact that it was extremely annoying and frustrating to the users, subjects were still able to complete the tasks quickly in the HV case. These results actually show that graphics always visible via a HUD may be a good option performance wise, if registered AR is not possible. However, for ease of use, as evidenced by the subjective workload ratings for mental workload, effort, perceived performance and frustration shown in Figure 2, HV is either the worst or very close to the worst.

As one might expect, the spatial relationship between the context blocks and the target location of the virtual block affected a participant's ability to place a block. When the context block was adjacent to the target location of the block, the participants performed best (in terms of errors and placement time). When the target block location was lined up horizontally or vertically with a context block, the participant performed better (in terms of errors and placement time) than if the target location has no other simple spatial relationship to the context block. These results have implications for designing AR systems. When designing an AR system that uses context, it is important to leverage as many contextual cues and spatial relationships as possible. It would behoove a designer to put careful thought into the relationship between the context and the task because well-placed context should improve task performance.

7 Conclusions

This study has shown that registered AR outperforms both non-registered AR and graphics displayed on a HUD. We have also shown that non-registered (off the side) AR does not offer any significant performance advantages over a HUD, but is rated as less intrusive and can keep non-registered graphics from cluttering the task space.

We found that HUD-based displays can work well when the HUD graphics are oriented with the user's view of the workspace (as in our experimental conditions). One would expect this from VR-based implementations of correctly oriented HUD-style graphics, such as Stoakley et al's WIM [12]. However, our implementation achieved this orientation alignment in the real, not virtual, world using orientation-only trackers. The fact that the OTS case was not significantly different than either of the HUD cases is interesting, because the HUD cases are easier and cheaper to implement. This has significant implications for the designs of future mobile HMD-based systems.

8 Acknowledgements

We would like to thank Lisa Mauney for her help with the data analysis. We acknowledge the many students of the Augmented Environments Lab who helped keep this project running, especially Enylton Coelho and Maribeth Gandy. This work was supported by NSF CAREER Grant 0347712, the Office of Naval Research under Contract N000140010361, and a grant from the Georgia Tech GVU Center.

References

[1] Coelho, E. M., MacIntyre, B., and Julier, S. (2004). (osgAR): A Scenegraph with Uncertain Transformations. In Proc. *International Symposium on Mixed and Augmented Reality*, Arlington, VA, November 2-5, pp. 6-15.

[2] Embretson, S.E. (1997). The factorial validity of a cognitively designed test: The Spatial Learning Ability Test. Educational and Psychological Measurement, 57, 99- 107.

[3] Hart, S. and Staveland, L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P. and Meshkati, N. (Ed.), Human Mental Workload. Elsevier, North Holland, Amsterdam (1988), pp. 139-183.

[4] Livingston, M.A., and Zhuming Ai. (2008). The Effect of Registration Error on Tracking Distant Augmented Objects. To appear in Proc. *International Symposium on Mixed and Augmented Reality*, Cambridge, England.

[5] Livingston, M. A., Swan, J. E. I., Gabbard, J. L., Höllerer, T. H., Hix, D., Julier, S. J., Baillot, Y., and Brown, D. (2003). Resolving Multiple Occluded Layers in Augmented Reality. In Proc *International Symposium on Mixed and Augmented Reality*, Tokyo, Japan, October 7-10, pp. 56-65.

[6] Oldfield, R. 1971. The assessment and analysis of handedness: The Edinburg Inventory. *Neuropsychologia* **9:** 97–113.

[7] Open Scene Graph. from http://www.openscenegraph.org

[8] Robertson, C.M., Using Graphical Context to Reduce the Effects of Registration Error in Augmented Reality, PH.D. Thesis, Georgia Institute of Technology, 2007.

[9] Robertson, C., and MacIntyre, B. (2003). Adapting to Registration Error in an Intent-Based Augmentation System. In A. Y. C. Nee & S. K. Ong (Eds.), Virtual and Augmented Reality Applications in Manufacturing (pp. 143-163). London: Springer Verlag.

[10] Robertson, C., and MacIntyre, B. (2007). An Evaluation of Graphical Context as a Means for Ameliorating the Effects of Registration Error. In Proc. *International Symposium on Mixed and Augmented Reality*, Nara, Japan, November 13-16, pp. 99-110.

[11] Seligmann D., and Feiner S. (1991) Automated Generation of Intent-Based 3D Illustrations. Computer Graphics, 25(4) Proc. ACM SIGGRAPH, Las Vegas, Nevada, pp. 123-132.

[12] Stoakley, R., Conway, M., Pausch, R., Virtual Reality on a WIM: Interactive Worlds in Miniature, CHI'95, pp. 265-272.

[13] Tang, A., Owen, C., Biocca, F., & Mou, W. (2003). Comparative Effectiveness of Augmented Reality in Object Assembly. In Proc. *SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, FL, April 5-10, pp. 73-80.