Evaluating the Effects of Tracker Reliability and Field of View on a Target Following Task in Augmented Reality

Jonathan Ventura* Marcus Jang Tyler Crain Tobias Höllerer
University of California, Santa Barbara

Doug Bowman
Virginia Tech

Abstract

We examine the effect of varying levels of immersion on the performance of a target following task in augmented reality (AR) X-ray vision. We do this using virtual reality (VR) based simulation. We analyze participant performance while varying the field of view of the AR display, as well as the reliability of the head tracking sensor as our components of immersion. In low reliability conditions, we simulate sensor dropouts by disabling the augmented view of the scene for brief time periods. Our study gives insight into the effect of tracking sensor reliability, as well as the relationship between sensor reliability and field of view on user performance in a target following task in a simulated AR system.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Augmented Reality; H.1.2 [Models and Principles]: User/Machine Systems—Human Factors;

Keywords: augmented reality, immersion, user study, simulation

1 Introduction

Augmented Reality systems may be used to simulate X-ray vision. The ability to see through solid objects has potential applications in various fields ranging from construction [Webster et al. 1996] to medicine [Azuma 1997], military [Livingston et al. 2002] and search-and-rescue operations [Livingston 2005]. Unfortunately, implementing X-ray vision is not a trivial task. AR systems face certain unresolved issues, including those pertaining to tracking objects in the X-ray view. For example, GPS dropout errors [Reitmayr and Drummond 2006] may cause augmented position information to be temporarily lost. We hypothesize that sensor dropouts will decrease user performance, but this problem may be mitigated by increasing the augmented field of view (FOV).

A task common to many AR X-ray systems is to follow, using the augmented overlay, a moving target that is occluded by objects in the real world. Accuracy in AR system object tracking is an issue which is likely to have an effect on task performance. Since the augmented view is overlaid on top of the real view it is also interesting to study the effects of varying the augmented field of view as well as its interaction with tracking imperfections. We created a task that aimed to study these effects in a generic fashion in order to be generalizable while still being grounded in reality.

In this study, we have implemented an AR simulation in virtual reality to examine the effects of varying FOV and sensor dropouts in a target following task. We simulate video see-through AR using a VR head mounted display. This AR simulation gives us the ability to vary experiment conditions which may be difficult or even impossible to replicate on a real-world AR system. It also gives us the ability to keep unnecessary factors from generating noise in our data, such as illumination or visibility changes, weather problems, or environmental distractions.

We are considering FOV and the sensor dropouts, or persistence of the augmented imagery, as the varying components of immersion in our study. We adhere to Slater's definition [Slater 2003] of immersion as being the "objective level of fidelity of the sensory stimuli produced by a technological system," enabling us to view the components of immersion as the directly controllable aspects of the technological system.

The AR simulation is built as a target following task. The participant stands inside a virtual room and visually follows a person walking an unpredictable path outside the room. We wish to answer the subsequent questions. What effect does varying the augmented FOV have on user task performance? What effect do different dropout lengths have on performance? What are the interactions between these two variables?

2 Related Work

Livingston and Ai studied the effects of various sources of registration error with an AR simulation similar to our own [Livingston and Ai 2008]. Participants tracked a virtual car moving throughout a real environment, with a white box representing an augmented view of the car's location. A white box was continuously visible, even when the car itself was occluded by a building in the environment. Other virtual cars and their associated white boxes acted as distractors. At specific times during the experiment, the simulation would freeze and the participant identified the location of the correct car.

Bane and Höllerer propose a set of AR tools for virtual X-ray vision [Bane and Höllerer 2004]. Various augmented views are presented, each attempting to overcome "Superman's X-ray vision" problem of presenting too much AR information to the user. The toolset's usefulness is illustrated with an example of an outdoor user viewing the contents of a nearby building.

Ragan et al. developed an AR simulation using virtual reality to examine the effect of registration errors on an object manipulation tasks [Ragan et al. 2009]. The implemented task involved moving a virtual ring from one end of an irregularly shaped tube to the other, while avoiding collisions between the two objects. By using VR to simulate this task, they were able to independently control tracking jitter and latency variables in a pilot study.

3 AR Simulation

Our experiment involves a target following task which uses an augmented reality interface to allow people to be tracked behind walls. A participant stands in the middle of a square room with unique doors and/or windows on each wall. Virtual people walk unpredictable paths outside the building. The target to be followed is one of these people, visually distinguishable by a large black top hat. The augmented reality interface overlays a translucent red rectangle

*e-mail: jventura@cs.ucsb.edu

VRST 2009, Kyoto, Japan, November 18 – 20, 2009.
© 2009 ACM 978-1-60558-869-8/09/0011 $10.00

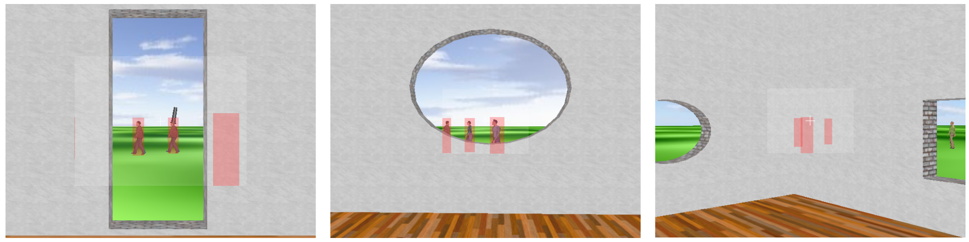


Figure 1: *Sample views of AR simulation. Transparent red rectangles overlay the people, indicating their location.* Left: *A user view of Top Hat (non-occluded) in the doorway.* Middle: *Three distractor people are visible through a large window.* Right: *A typical user view of a corner occluding several people.*

on each person, which is visible even when the person is occluded by a wall. However, the target person's overlay is identical to the other virtual people overlays; the target may only be distinguished when visible through a window or a door.

To actually run this experiment using a true augmented reality system would be extremely difficult. Many confederates would be needed to walk around the room, and their movements would need to be accurately reproduced for each new participant. Also, the confederates would need to be continously tracked for the augmented overlays to be displayed.

By simulating augmented reality, we are freed from the current restrictions of display technology. This simulated environment also affords easily controlled variables, and thus the experiment may be replicated.

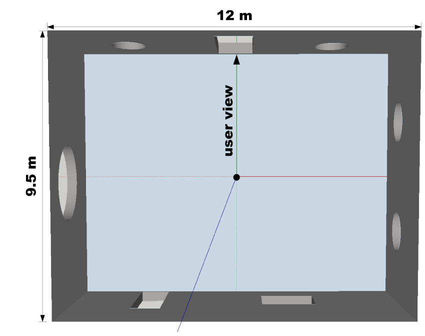


Figure 2: *Overhead view of virtual room (approx. measure). Participant is stationary in the center of the room, with initial orientation along "user view" arrow. Top-hat's initial position is outside the front doorway, directly in the participant's view.*

4 Experimental Design

The experiment was designed with several goals in mind. The task should have straightforward, quantitative results. Also, the task should be reasonably simple so as not to frustrate the participants. Furthermore, the task should be generic so as to be generizable, yet grounded in reality (i.e. not an abstract world), to enhance participant familiarity with the AR task.

We developed a target following task scenario, where the user is asked to visually track a virtual person as it moves throughout the scene. The participant is centered inside a room with various doors and windows giving a view to the outside world. They may change their orientation, but not change position (three degrees of freedom). The user's view to the virtual person may be occluded at any time by the walls of the building; at other times the person may be visible through doors or windows in the building. The augmented view element is composed of red marker rectangles indicating the location of the person to be followed and the distractor people.

The parameters of our experiment are divided between the "real" components of immersion, such as the field of view of the HMD, and the "augmented" components of immersion, such as the field of view of the simulated AR display, and the performance of the tracking sensor. We introduce periods of sensor dropouts, where the augmented overlays disappear, to simulate the effect of dropouts in a real tracking system. Such dropouts may occur with a magnetic sensor near interfering material. Different display options are available in this case; for example we could use the last known sensor reading, or predict future values. However, we determined through expert analysis that hiding the AR overlay during sensor dropouts would have the least effect on performance.

The experiment has two independent variables: the field of view of the AR interface and the length of sensor dropout periods. Each trial lasts 60 seconds, and includes seven sensor dropouts. The total vertical field of view of the HMD is 36 degrees, while the three possible values for the augmented field of view are 10, 20, and 34 degrees. The length of dropouts vary between two seconds (highest), one second (medium) and zero seconds (lowest). We experimented with longer dropout periods but deemed them too long. There is a total of nine conditions. For each participant we tested each condition three times giving a total of 27 trials per participant.

During each trial there are 20 people in addition to a virtual man *Top Hat* wearing a tall black top hat; each walking an unpredictable path outside of the room. The paths for each virtual person were randomly generated from a set of coordinates exterior to the room. Paths were constrained to stay within the rectangular $50 \times 50\ m^2$ area surrounding the room. At each path point, the virtual person randomly changes its walking speed within the range of four to seven meters per second. Top Hat has the additional constraints as follows. He starts each trial just outside the front door of the building, standing in the participant's plain view. In order to make paths of similar difficulty Top Hat walks around the entire building at least once, and is visible through the windows for at least 10 seconds. These constraints were met by randomly generating paths until all were satisfied. We generated 27 sets of paths in this fashion, each corresponding to a specific trial and used a latin squares ordering to discourage learning effects.

4.1 Implementation

We implemented this experiment with Python and the Vizard virtual reality toolkit. Head-tracking was handled with a InterSense InertiaCube3 orientation tracker with a refresh rate of 180 Hz. We used a Pro-View 60 head mounted display (HMD) displaying 640×480 video at 60 Hz, and 36 degrees vertical field of view. As we were not concerned with the effects of stereo, we disabled the stereoscopic feature. The experiment ran on a 2.4 GHz Intel Core 2 machine with two gigabytes of RAM, a NVIDIA GeForce 9800 GX2 graphics card and Windows XP SP3.

4.2 Participants

We ran this experiment with 19 participants; eleven male and eight female. The participants ranged in age from 23 to 59, 15 were in the range 23 to 30, and the remaining four in the range 51 to 49. None were colorblind. Most users were heavy computer users with at least some familiarity with virtual reality and augmented reality. Only one participant reported regularly playing 3D video games. After every nine trials, the participants took a mandatory five minute break. Some experienced light fatigue, but none experienced strong effects of motion sickness.

5 Analysis

In our experiment we define performance as effectiveness in following Top Hat. The dependent variable we measure is the angular distance (in yaw) between the the user's viewing direction and the direction towards the target character. This measure ranges between zero and 180 degrees (we used the smaller of the two options, clockwise and counter-clockwise). We record one measurement per video frame, at about sixty frames per second, resulting in approximately 3600 measurements for a one minute trial.

Because the video does not always run precisely at 60 Hz, each trial might have a slightly different number of measurements. We also record a timestamp for each measurement, so that we can determine their exact frequency. As a pre-processing step, we linearly re-sample the data for each trial so that each has exactly 3600 measurements at 60 Hz.

We observed that participants tend to switch between two different states during the experiment: a target following state, and a lost state, where they are searching for the target. The two states are easily visible in the data. In the target following state, the angular error is generally low. We had asked participants to keep the cross hairs on the target as closely as possible, but different participants achieved different levels of accuracy. In the lost state, the error starts to steadily climb, and may fluctuate depending on where the target moves. The error may even return to near-zero during a lost state, if the participant unwittingly crosses the target's path.

The simplest metric is the average error over an entire trial. This may not accurately represent sensor dropouts, because the error does not necessarily stay high during the lost state, and also different participants may generally keep the cross hair further from the target even in the target following state.

We may detect the participant's state more explicitly by applying a threshold τ to the data. The threshold specifies the maximum angular error that represents successful following of the target. When conducting the trials we noticed that the participants spent a majority of the time correctly following Top Hat. By computing the median angular error, we found that the participants successfully tracked Top Hat within an error margin of four degrees. We also noticed that during trials as Top Hat would change direction the participant's angular error would increase by a small amount for a short period of time, without losing the target. As a result we doubled the four degrees, giving us a threshold of $\tau = 8$ degrees.

Using this threshold we measure a participant's *time to failure*, the time until the a lost state is encountered. However, this may not be a very descriptive measure, since a participant may reach the lost state sooner or later depending on the movement of the avatars, which varies between trials.

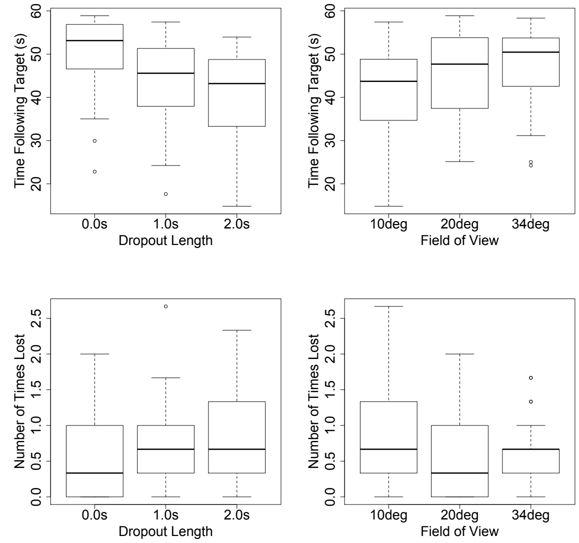


Figure 3: *Box plots for our two independent variables and two metrics. In general, performance increases when the length of sensor dropout periods is lessened, and when the AR field of view increases.*

We also consider a measurement we call *time following target*, which is the amount of time spent successfully following the target. This metric seems to most generally represent how well a participant performed. Trials with either more or longer lost periods will result in a lower total time following the target.

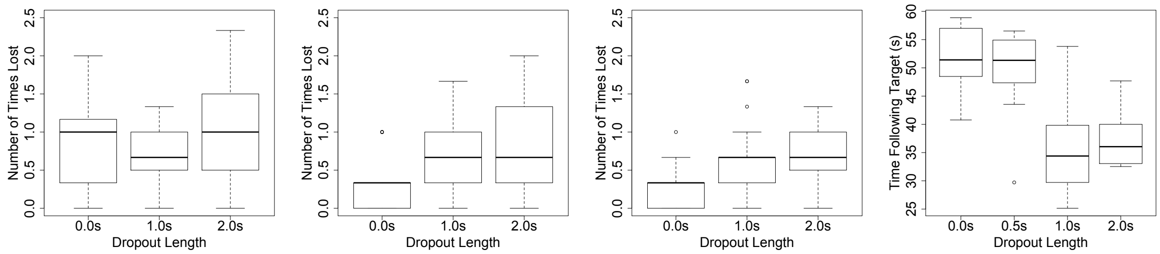
For a more specific measure of performance, we count the number of times a participant reaches the lost state in a trial, or *number of times lost*. We only record a lost state when the error stays above the threshold τ for a minimum of four seconds. A four second threshold was used because it is double the maximum dropout length, allowing the user a few seconds to resume successful tracking after the dropout before they are considered to have lost the target.

6 Results

Figure 3 shows box plots of two metrics versus our two independent variables.. We performed two-way ANOVA for the time following target metric as an overall analysis of the effects of our two factors. Both the failure length and the field of view have a significant effect on performance with this metric with $p < 0.001$. We also used a two-way ANOVA to examine the number of times the participant loses the target. Both field of view and dropout length have a significant effect with $p < 0.001$. There also is an interaction between the two factors at the $p = 0.05$ level. Figures 4(a)-(c) show box plots for these conditions.

From this plot we can see that with a small field of view, performance is equally bad with no dropouts or one second dropouts. When the field of view is increased to the medium level, performance dramatically increases when there are no dropouts, but only gets slightly better when there are dropouts. Finally, there is not a significant difference in performance between the medium and large field of view ($p = 0.575$), which suggests that the medium field of view is enough for good performance.

Considering the time following target metric, performance decreases between zero and one seconds, but seems to level off be-



(a) (b) (c) (d)

Figure 4: *Box plots (a)-(c) relate sensor dropout length to the number of times the participant loses the target, with (a) low, (b) medium and (c) upper field of view. Box plot (d) shows the total time the participant is following the target versus the length of the sensor dropout periods, including the half second dropout condition. We collected the data for box plot (d) from seven participants, using the medium field of view.*

tween one and two. We ran a post-hoc analysis to determine which conditions were significantly different. Using Tukey's HSD, we found that the one second and two second conditions were significantly different than zero ($p < 0.001$ for both), but not significantly different from each other ($p = 0.239$). This suggests that we may be seeing a thresholding effect above one second sensor dropouts.

To investigate this threshold, we had seven participants complete an extra three trials with the medium field of view and sensor dropout periods of 0.5 seconds. Figure 4(d) shows a box plot of time following target versus sensor dropout length for these seven participants. Again using Tukey's HSD to compare the means, we did not find a significant difference between the zero length and half second length conditions, but did find a significant difference between those and the higher length conditions. With further experimentation, we could further delineate the performance curve as the dropout length increases from zero to one second.

7 Conclusions and Future Work

We designed a AR experiment to study the effects of tracker reliability and field of view on a target following task. We developed a virtual reality based AR simulation to enable control over these independent variables.

We found that for our target following task, good performance could not be achieved with a low field of view, even with perfect tracker reliability. With a reliable tracker, a medium field of view is enough for good performance. We did not see a significant performance increase when increasing to a high field of view. These results suggest that a reasonable AR field of view is crucial for performance in target following tasks, but the benefit does not continue to improve with increasing FOV.

Decreased tracker reliability, due to periods of sensor dropout, adversely affects task performance. However, the performance decrease levels off at one second dropouts. Preliminary analysis also shows no significant difference between having no dropouts and half second dropouts. This suggests that below a certain dropout length, good performance may be maintained, given a sufficiently large FOV. Our findings suggest that AR interface designers should consider ways to mediate the effect of sensor dropouts longer than one second, although dropouts less than one half second long do not harm performance.

Future work lies in further investigating the effect of the dropout length below the one second level. Our work also suggests new experiments designed to test the interaction between reliability of the head-based rendering and various other immersion components.

Acknowledgements

We would like to thank Cha Lee and Steffen Gauglitz for help in designing and running our experiments, and our study participants. This work was partially supported by NSF CAREER grant IIS-0747520, and NSF IGERT grant DGE-0221713.

References

AZUMA, R. T. 1997. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments 6* (August), 355–385.

BANE, R., AND HÖLLERER, T. 2004. Interactive tools for virtual x-ray vision in mobile augmented reality. In *Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on*, 231–239.

LIVINGSTON, M., AND AI, Z. 2008. The effect of registration error on tracking distant augmented objects. In *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*.

LIVINGSTON, M. A., ROSENBLUM, L. J., JULIER, S. J., BROWN, D., BAILLOT, Y., EDWARD, J., II, S., GABBARD, J. L., AND HIX, D. 2002. An augmented reality system for military operations in urban terrain. In *In Interservice/Industry Training, Simulation, and Education Conference*, 89.

LIVINGSTON, M. 2005. Evaluating human factors in augmented reality systems. *Computer Graphics and Applications, IEEE 25*, 6 (Nov.-Dec.), 6–9.

RAGAN, E., WILKES, C., BOWMAN, D., AND HOLLERER, T. 2009. Simulation of augmented reality systems in purely virtual environments. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, 287–288.

REITMAYR, G., AND DRUMMOND, T. 2006. Going out: robust model-based tracking for outdoor augmented reality. In *Mixed and Augmented Reality, 2006. ISMAR 2006. IEEE/ACM International Symposium on*, 109–118.

SLATER, M. 2003. A note on presence terminology. *Presence Connect 3*.

WEBSTER, A., FEINER, S., MACINTYRE, B., MASSIE, W., AND KRUEGER, T. 1996. Augmented reality in architectural construction, inspection, and renovation. In *In Proc. ASCE Third Congress on Computing in Civil Engineering*, 913–919.