

A Replication Study Testing the Validity of AR Simulation in VR for Controlled Experiments

Cha Lee*

University of California, Santa Barbara

Scott Bonebrake†

University of California, Santa Barbara

Tobias Höllerer‡

University of California, Santa Barbara

Doug A. Bowman§

Virginia Tech

ABSTRACT

It is extremely challenging to run controlled studies comparing multiple Augmented Reality (AR) systems. We use an “AR simulation” approach, in which a Virtual Reality (VR) system is used to simulate multiple AR systems. In order to validate this approach, we carefully replicated a well-known study by Ellis et al. using our simulator, obtaining comparable results.

Index Terms: I.3.7 [Three-Dimensional Graphics and Realism]: Virtual Reality—AR Simulation; I.3.6 [Methodology and Techniques]: Device independence—Replication

1 INTRODUCTION

The inherent components of AR pose challenging issues when trying to conduct controlled experiments between systems. Hardware for AR systems can differ greatly and finding compatible systems may prove to be extremely hard. AR applications also typically rely on the real world as a backdrop to virtual objects. This makes replicating the exact real world conditions impossible in outdoor applications and difficult for indoor applications. We use an “AR simulation” approach, in which a VR system is used to simulate a range of AR systems. AR simulation, as proposed by Gabbard et al. [2] and Ragan et al. [3], may prove to be a viable method of conducting controlled experiments in AR. By using VR, we can simulate real world conditions and hardware configurations (display, tracking, processing power) of an AR system and provide repeatable conditions for experiments. But are the results of experiments using AR simulation valid for real-world AR systems? Our research begins to address this question by attempting to replicate a well-known AR experiment.

2 DESIGN

The experiment we chose to replicate was the second study in Ellis et al. [1]. Our VR system consisted of a Kaiser Proview 60 HMD, WorldViz Precision Position Tracking (PPT) system, two wired Intersense InertiaCube2, and a VRML model based on the ReCWEB lab at UCSB. To simulate the original HMD used, we rendered a transparent region equivalent to the 19 deg field of view of the original HMD. All augmentations were only viewable within this region while the simulated real objects could be seen in the entire view of the user, as shown in Figure 1. The two Intersense InertiaCube2 in conjunction with three markers (two for head, one for hand) from the PPT system were used to track the head and hand of the user in 6 degrees of freedom. A *simulated* real hand was then used to represent the actual hand of the participant.

*e-mail: chalee21@cs.ucsb.edu

†e-mail: egscott2004@gmail.com

‡e-mail: holl@cs.ucsb.edu

§e-mail: bowman@vt.edu

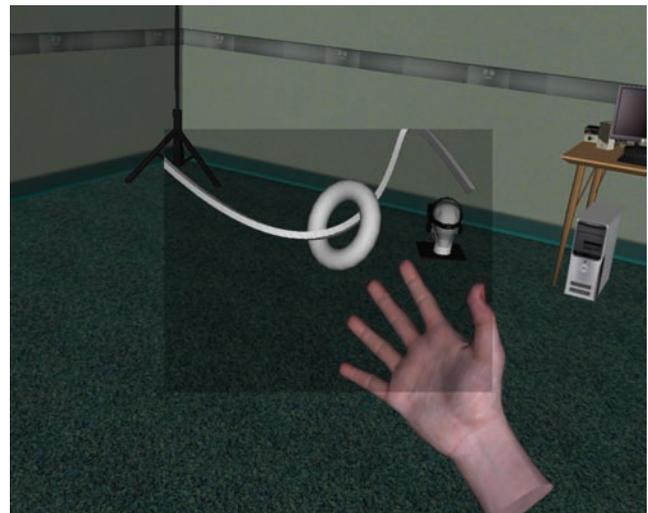


Figure 1: Screenshot of participant's view

For the experiment we had 14 users, ages 21 - 33, nine male and five female. All users were comfortable around computers. All users were able to perceive stereo as verified by a random dot stereogram. The HMD was then automatically calibrated and registered with each user with a constant error of less than two cm. Previous work [4] has found evidence that users are able to adapt to small perturbations in head registration for motor performance in VR tasks. Participants were asked to trace a virtual 3D path with a virtual ring rigidly attached to the participant's hand. Independent variables were ring size, path type (angular and smooth), and total system latency (on the virtual objects). There were five different latencies and six different paths for a total of 30 different conditions per ring type. The conditions were randomly ordered in a block of trials and each participant performed three blocks. Seven participants used the small ring, and seven the large ring. The dependent measure in this experiment was the number of collisions between the ring and the path. For more details on the experimental design, please refer to [1].

3 RESULTS

Preliminary analysis of the results showed that users reached asymptotic performance after the first block of trials; thus only the second and third blocks were used in the following analysis. All significant effects on tracing performance are shown in Table 1. These results are comparable to the results from [1] with our results also showing a significant interaction between path and latency.

In addition to comparable statistically significant effects, all of the effects were in the same direction as in the original study: the number of collisions increased with angular paths, with the small

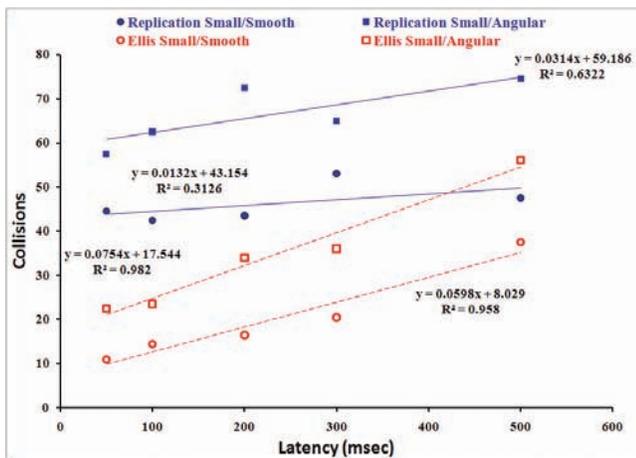


Figure 2: Comparison of trend lines for tracing performance of the small ring. Our replication study results are represented by the solid blue data points.

ring size, and as latency increased. However, there were some interesting differences in the absolute performance data, as shown in Figures 2 and 3. The participants' absolute tracing performance is much worse in our experiment. We believe this is mostly due to the differences in the collision algorithms used. It was unclear how the original experiment determined collisions during sustained contact of the ring and path. Initially we only counted this type of collision once, but this benefited the careless participants in the pilot study we ran. As a result, we decided to add collisions for every 200 msec the ring stayed in contact with the path which increased collisions overall but penalized the careless participants. Ellis et al.'s experiment also showed a visible increase in collisions as system latency increased for the small ring. This effect was less observable and in fact starts to level off as if a ceiling effect was occurring in our experiment. Due to the lower clearance for the small ring, users moved their head extremely close to their hand to get a better view which occasionally caused small jitter issues, because PPT is a vision tracking system which depends on line of sight. Adding our rigorous collision criteria, the small ring task may have been too difficult within our system.

4 DISCUSSION ON EXPERIMENTAL REPLICATION

During the course of this work, we learned some valuable lessons with regards to replicating and simulating previous experiments. One important reason for choosing the experiment in [1] was the very detailed description of the design and analysis the authors provided in the paper. In addition, Dr. Ellis was kind enough to be available for questions. Even with all this, understanding the original setup and design of the experiment was extremely challenging. We did not have access to the original models used, lacked information on the environment except for the video provided, lacked

Table 1: Significant Effects on Tracing Performance

Effect	df	F level	P level
Ring	1, 12	9.075	P < 0.011
Path	1, 12	25.638	P < 0.001
Latency	4, 48	14.245	P < 0.001
Path x Latency	4, 48	7.484	P < 0.001
Path x Ring x Latency	4, 48	3.348	P < 0.017

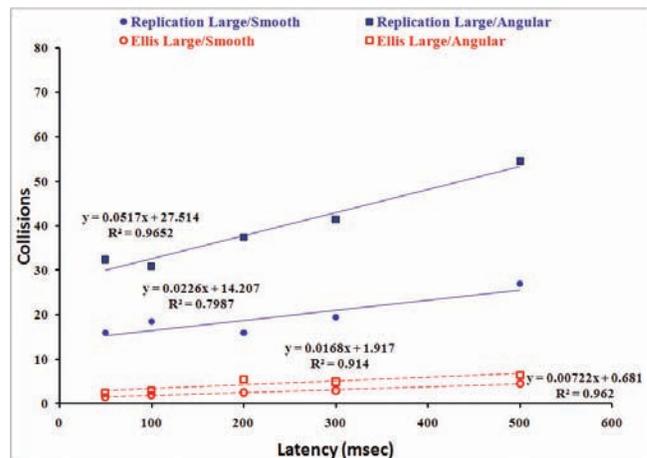


Figure 3: Comparison of trend lines for tracing performance of the large ring. Our replication study results are represented by the solid blue data points.

access to the original collision algorithm, and lacked the raw results from the previous study. Replicating the original models was impossible, so our models were only approximations. In summary, it would be extremely hard if not impossible to repeat any experiment without very detailed notes or the guidance of the original authors. This highlights the importance and need for even more detailed reports on experiments in the community. Although this may not be feasible within a conference or journal paper format, this information is invaluable to the repeatability of these experiments.

5 CONCLUSION AND FUTURE WORK

We have repeated a well known AR experiment by Ellis and colleagues and obtained comparable results. We believe this is a step toward validation of the idea that AR simulation is a viable and effective method of performing controlled experiments. For future work, we are interested in experimenting with other factors in AR systems such as display size, field of view, and resolution within our simulation environment.

ACKNOWLEDGEMENTS

We would like to thank Dr. Stephen R. Ellis for his time and invaluable guidance in replicating the original study. We would also like to acknowledge Masaki Miyanojara and the ReCVEB lab at UCSB for providing the virtual model of their lab.

REFERENCES

- [1] S. R. Ellis, F. Breant, B. Manges, R. Jacoby, and B. D. Adelstein. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In *VRAIS '97: Proceedings of the 1997 Virtual Reality Annual International Symposium (VRAIS '97)*, page 138, Washington, DC, USA, 1997. IEEE Computer Society.
- [2] J. L. Gabbard, J. E. Swan, II, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence: Teleoper. Virtual Environ.*, 15(1):16–32, 2006.
- [3] E. Ragan, C. Wilkes, D. A. Bowman, and T. Höllerer. Simulation of augmented reality systems in purely virtual environments. *Virtual Reality Conference, IEEE*, 0:287–288, 2009.
- [4] D. W. Sprague, B. A. Po, and K. S. Booth. The importance of accurate vr head registration on skilled motor performance. In *GI '06: Proceedings of Graphics Interface 2006*, pages 131–137, Toronto, Ont., Canada, Canada, 2006. Canadian Information Processing Society.