VR Collide! Comparing Collision-Avoidance Methods Between Co-located Virtual Reality Users

Anthony Scavarelli
Carleton University
1125 Colonel By Dr.
Ottawa, ON K1S5B6, CA
anthony.scavarelli@carleton.ca

Robert J. Teather
Carleton University
1125 Colonel By Dr.
Ottawa, ON K1S5B6, CA
rob.teather@carleton.ca

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053180

Abstract
We present a pilot study comparing visual feedback mechanisms for preventing physical collisions between co-located VR users. These include *Avatar* (a 3D avatar in co-located with the other user), *BoundingBox* (similar to HTC's "chaperone"), and *CameraOverlay* (live video feed overlaid on the virtual environment). Using a simulated second user, we found that *CameraOverlay* and *Avatar* had the fastest travel time around an obstacle, but *BoundingBox* had the fewest collisions at 0.07 collision/trial versus 0.2 collisions/trial for *Avatar* and 0.4 collisions/trial for *CameraOverlay*. However, subjective participant impressions strongly favoured *Avatar* and *CameraOverlay* over *BoundingBox*. Based on these results, we propose future studies on hybrid methods combining the best aspects of *Avatar* (speed, user preference) and *BoundingBox* (safety).

Author Keywords
VR; multi-user; co-location; collision avoidance.

ACM Classification Keywords
H.5.1. Multimedia Information Systems: Artificial, augmented and virtual realities; H.5.2. User Interfaces: Interaction Style; H.5.2. User Interfaces: Haptic I/O; I.3.7 3-D Graphics: Virtual Reality.

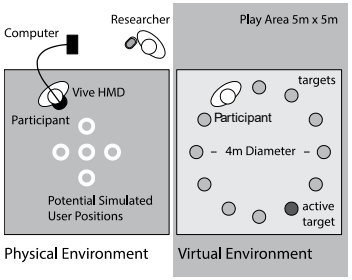

Figure 1. The physical and virtual layouts of the "play area", participants, and researcher controller.

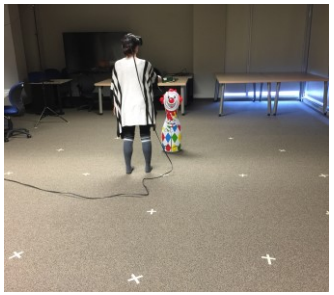

Figure 2. The physical layout of experiment space. "Clown" is used as collision obstacle during camera overlay mode. Tape marks virtual target positions.

Introduction

The recent resurgence of virtual reality (VR) has yielded several relatively low-cost head-mounted displays (HMDs). Notably, this includes the HTC *Vive* [11], a HMD that includes a room-scale tracking system allowing users to physically walk around in a modest sized virtual space. With greater mainstream access to this technology, opportunities arise for multi-user collaborative VR experiences. Many previous multi-user VR systems employed networking to connect multiple users to the same virtual environment (VE) [8, 9, 12]. In these cases, the physical space is not shared. However, as VR devices become more available we foresee that dedicated VR spaces will be shared by multiple users (e.g., one or more family members within a household). Since HMDs occlude the physical environment, this introduces the possibility of physical collisions between users, and hence the possibility of injury and/or equipment damage in multi-user co-situated VR scenarios. To address this problem, we propose methods to warn users of impending collisions with other VR users in physical proximity.

We present a pilot study evaluating three visual feedback modes (*Avatar*, *BoundingBox*, *CameraOverlay*, described below) to determine which best helps VR users avoid collision with a simulated "secondary user". Participants navigated a simple virtual environment, in situations eliciting either no collision, a glancing collision, or a head-on collision.

PRIOR WORK

Since VR systems have been traditionally expensive and inaccessible (outside of lab settings), there is relatively little work on co-situated VR users. Previous work on the topic used redirected walking to prevent user collisions within a large tracked VR space [1], or used avatars to help users identify and localize each other [5]. There are also lessons from commercial VR applications [8], [9], [12].

Holm [1] used redirected walking (subtle motion compression [6]) to allow multiple VR users to share a Huge Immersive Virtual Environment (HIVE)[7]. Results of multi-user studies reveal that redirected walking helped prevent collisions between users by either changing their velocities, or by stopping them entirely. However, it is unclear if users shared the same VE or not. For collaborative VR scenarios, multiple users must share the same virtual environment. Changing users' worldview (e.g., turning off the HMD[1]) or stopping them to avoid collision with another collaborating user is disruptive to collaboration, and likely would break presence. Moreover, redirected walking is impractical in scenarios like multi-user home VR, because of the large space required: Steinicke at al. [4] report that a circular arc with radius of 22m or greater is necessary to prevent users from detecting redirection. Unfortunately, the maximum tracking radius of the HTC Vive is ≈2.3m [11].

Streuber and Chatziastros [5] conducted an experiment where two participants carried a stretcher through a VR maze and the number of wall collisions noted [5]. They used an avatar to display the position of each user under several test conditions and report that user coordination was not improved by showing the other user's avatar. This may be related to the use of a stretcher for this "joint-task action", which minimizes direct interaction between users [5].

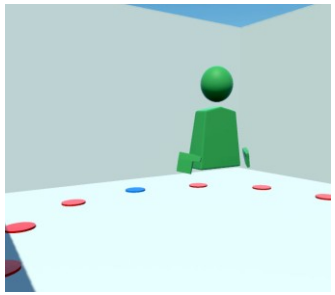

Figure 3. User POV of *Avatar* in-app.

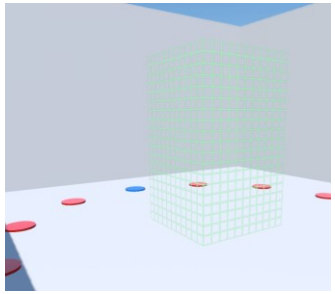

Figure 4. User POV of *BoundingBox* in-app.

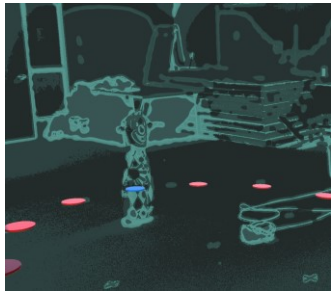

Figure 5. User POV of *CameraOverlay* in-app.

While no commercially available software currently supports multiple physically co-located VR users, some modern VR games include multiplayer functionality; but require separate tracked spaces. Tilt Brush [12], Rec Room [8], and Pool Nation VR [9] all use basic avatars (e.g., showing head, body, and two hands/controllers) to help users localize one another. The familiarity of avatars makes them an attractive candidate for avoiding collisions in physically co-located VR.

Finally, although *user* collision-prevention is not well researched, commercial systems include methods for preventing *environmental* collisions. Most well-known is Valve's "Chaperone" system, used with the HTC *Vive*. The Chaperone displays a subtle green grid when the user approaches the physical space boundary [11], becoming more noticeable as the user gets closer. Boundaries are mapped out in advance by the user during calibration. Since the Chaperone is only visible when the user is close to the space boundaries, it may maintain user presence better than avatars, which provide continuous reminders of the physical world.

METHODOLOGY

We conducted a pilot study on preventing user collisions, comparing three candidate techniques:

- *Avatar*: A green avatar consisting of a head, body, and two hands similar to the user representation depicting the position and orientations of the Vive HMD and controllers. See Figure 3.
- *BoundingBox*: A wire-frame grid, similar to Valve's Chaperone system [11], described above. It is invisible until the user comes within (10 cm)[1], at which point it appears. See Figure 4.
- *CameraOverlay*: The Vive's front-facing camera is used to produce an overlay of the physical environment displayed over the virtual space. We note that this condition may be impractical for many real-world use cases in VR, as it likely negatively impacts presence. However, we include it as a "control" condition, expecting it to offer best performance as the next closest thing to actually seeing the physical environment. See Figure 5.

Participants

We recruited 12 participants (8 male, 4 female, mean age 24, *SD* 5.8 years). All were VR novices, assessed via a pre-experiment survey. They were not screened for any particular traits, other than an ability and willingness to walk around in VR.

Apparatus and Test Environment

The experiment was conducted using an HTC Vive [11] HMD, connected to a PC with a Xeon 4-core CPU, 16GB RAM, and Nvidia GTX 970 GPU. The PC communicated with a mobile app on the researcher's smartphone over a local wireless network. The mobile app controlled the experiment conditions (see Figure 7). For the *CameraOverlay* condition, a children's punching bag was physically placed in the environment as a stand-in for a second user. The punching bag was co-located with the virtual position of the simulated second user. See Figures 2 and 5.

We measured out a 4m radius circle on the lab floor, marking off locations on the physical floor where goal markers appeared on the virtual floor. See Figures 1 and 2. Participants physically walked across the circle of targets, from one target to the opposite target on the other side. The experiment employed two custom

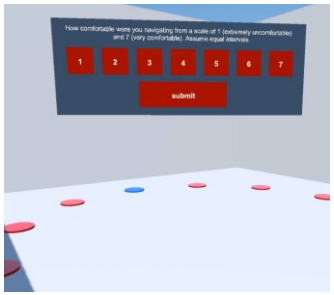

Figure 6. User POV, in experimental VE, of Survey to be completed after each trial in-app.

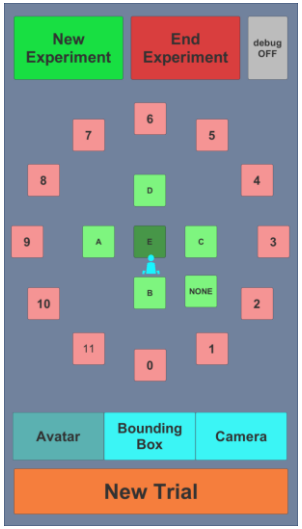

Figure 7. Mobile app view that the researcher uses to control the experiment wirelessly. Also shows current status and real-time position of participant within the VE.

applications developed in Unity [10]. The first displayed the VE to participants via a desktop VR app, see Figures 3 through 6. The second was a mobile app for controlling the state of the VR system, see Figure 7. When prompted by the experimenter (via the mobile app), the target location – a disc on the floor of the virtual environment – would turn blue, indicating the participant should walk to it. The desktop VR application logged participant movement times across the circle, as well as their motion trail. The desktop VR application also presented a 7-point Likert scale questionnaire to assess user comfort. The questionnaire was completed after each trial, and is seen in Figure 6. To make selections on the questionnaire, the user remotely pointed a Vive controller (i.e., employing ray-casting) and pressed the analog trigger button to select a response. The mobile app enabled the researcher to follow participants while simultaneously controlling the experiment. This helped prevent any actual collisions or tripping over cables.

Procedure

Upon arrival, participants read a written explanation of the experiment purpose and procedure. They then provided informed consent, and completed the pre-experiment questionnaire. Next, they were outfitted in the HTC *Vive* HMD, and were given roughly 5 minutes to practice walking around the virtual environment. After this practice period, the actual experiment began. Participants were instructed to stand at a specified target location. When they were ready, the experimenter activated another target location via the mobile app. The disc for that target location would turn blue, and participants would walk to that location. Each "crossing" of the circle was considered one trial.

To enhance external validity, we used 3 types of trials: 1) those with no simulated user (i.e., no possibility of collision); 2) those with a simulated user in a glancing collision position (i.e., obstacle *near* the motion path); and 3) those with a simulated user in a head-on collision position (i.e., obstacle right in the motion path). In the *BoundingBox* and *Avatar* conditions, the obstacle was completely simulated. In the *CameraOverlay* condition, the experimenter placed a physical prop (a blow-up clown) in place of a virtual obstacle. See Figures 2 and 5. In all cases, participants were instructed to avoid the obstacle to the best of their ability.

Upon completion of a trial, the participant answered a survey (Figure 6) about their comfort levels with the current condition. In total, participants completed 18 trials with each condition. Upon completion of the experiment, participants completed a short questionnaire asking them their subjective assessment of each collision avoidance method in terms of Efficiency, Safety, Pleasantness, and Suitability (each ranked on a 7-point Likert scale). In total, the experiment took about 30-40 minutes per participant.

Design

The experiment employed a within-subjects design with one independent variable, collision avoidance method, with three levels: *BoundingBox*, *Avatar*, and *CameraOverlay.* The collision avoidance method order was counterbalanced per a Latin square. To further avoid learning effects, the starting and ending positions of each trial were randomized without replacement.

There were three dependent variables - movement time, collision count, and comfort level. Movement time

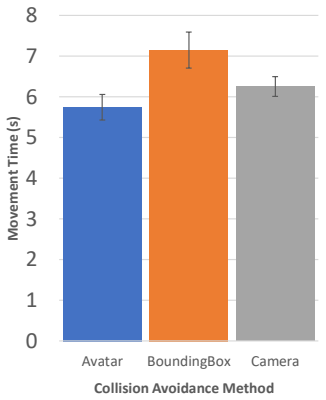

Figure 8. Movement Time averages graphed against method. Standard Error used.

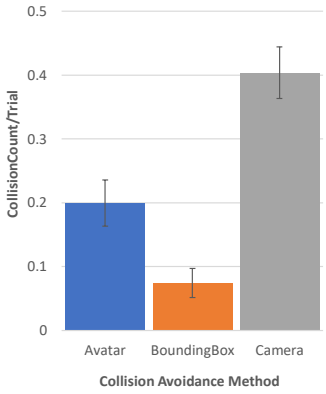

Figure 9. Number of Collisions' averages graphed against method. Standard Error used.

(in seconds) was the time from leaving a starting point to reaching the target disc. A collision was defined as the participant coming within 10 cm of the obstacle (consistent with personal space boundaries per Sambo and Iannetti [2]). Comfort was assessed via the post-trial questionnaire on a 7-point Likert scale.

Participants completed 18 trials with each collision avoidance method, for a total of 18 * 3 = 54 trials per participant. Thus, over all 12 participants, the experiment consisted of 648 total trials.

RESULTS AND DISCUSSION

One-way ANOVA on the ratio data revealed a significant difference in movement time ($F_{1,12} = 14.21$, $p < .05$) – see Figure 8. *Avatar* offered the fastest movement time, while the *BoundingBox* was slowest. A Scheffé posthoc test revealed that only *Avatar* and *BoundingBox* collision were significantly different ($p < .05$).

One-way ANOVA also revealed a significant difference in collision count between the three collision avoidance methods ($F_{1,12} = 28.21$, $p < .05$) – see Figure 9. *BoundingBox* offered the lowest collision count, at an average of 0.07 collisions per trial (roughly 1 collision per 12 or 13 trials). *CameraOverlay* fared the worst, with an average collision count of 0.4. In other words, nearly half of the CameraOverlay trials included a collision. This is surprising, as we had expected a live camera feed to offer *better* performance than the simulated aids; this may also reveal the difference in physical perception seen through a single lens (no depth) camera.

Post-trial comfort levels were compared using the non-parametric Friedman test on the ordinal data. They were found to be non-significant ($p > .05$).

Friedman tests on the post-experiment ordinal questionnaire data show significant differences in Efficiency, Safety, Pleasantness, and Suitability – see Figure 10.

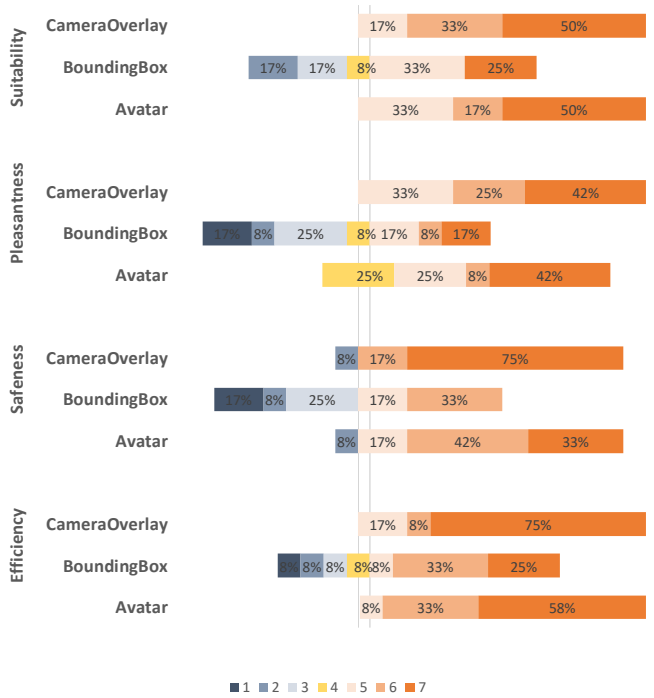

Figure 10. Post-experiment survey response averages visualized across all methods and questions pertaining to Suitability, Pleasantness, Safeness, and Efficiency.

Post-hoc analysis using Conover's *F* generally revealed no significant differences between *CameraOverlay* and *Avatar*, but significant differences between these two collision avoidance methods and *BoundingBox*.

Open-ended comments provided by participants suggest that the *CameraOverlay* was deemed the "favourite" collision avoidance method, and *BoundingBox* was least preferred. This is likely because the speed with which it became visible was too quick; participants reported being unpleasantly surprised by it. Interestingly, despite this, *BoundingBox* had the fewest collisions, suggesting that it produces safer movement through the space. This may be because it provides an "exaggeration" of personal space boundaries compared to the other collision avoidance methods due to its use of a solid shape, and also because participants walked much slower, anxiously watching for when it would appear next.

Avatar offered quicker movement than BoundingBox, but at the cost of more collisions (~270% more). Additionally, participant comments revealed that they personified the avatar, projecting annoyances at "being in the way" or, more positively, "being cute". This may suggest that avatars could help users better understand their appearance as "real people" in future experiments. *Avatar* was also more preferred than *BoundingBox* as it could be seen from a distance (i.e., it did not suddenly appear, surprising participants like *BoundingBox*). While this seems ideal for avoiding collisions, participants also noted within post-questionnaires that if multiple users are sharing the same space in *unrelated* VR experiences, it may break presence to see other user avatars.

It is also interesting that in only 2 of 648 trials did a participant walk *through* the simulated user. In both cases the participants noted how strange it felt to do so. Realistic behaviors in virtual environments have long been taken as evidence of presence responses [3]. This suggests that presence was quite strong, despite the simplicity of our experimental VR environment. This may also help explain why participants tended to personify the avatar in the *Avatar* method.

CONCLUSION

Our pilot study revealed that each collision avoidance performed best in one dependent variable. This, in addition to participants' comments that suggest there may be room for hybrid methods that allow greater safety. For example, a new hybrid approach might be based on the slower but potentially safer *BoundingBox*, combined with a more easily personified, faster, and more visible *Avatar* method.

We note here a limitation of our results: participants may behave differently when they know that no collision is actually possible, as the obstacle was simulated. Future work will involve conducting a similar study, but using multiple participants instead of a simulated stand-in, exploring hybrid methods, moving users, and various timings for the *BoundingBox* fade-in to find a better balance between immersion (the box being invisible) and comfort (adjusting distance threshold at which it fades in). Upcoming wireless HMDs may also help decrease user anxiety about movement in co-located VR, as there is no risk of tripping over the other users' cables.

References

1. Jeannette E Holm. 2012. Collision Prediction and Prevention in a Simultaneous Multi-User Immersive Virtual Environment. .
2. Chiara F Sambo and Gian Domenico Iannetti. 2013. Better safe than sorry? The safety margin surrounding the body is increased by anxiety. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 33, 35: 14225–30.
3. Martijn J. Schuemie, Peter van der Straaten, Merel Krijn, and Charles A.P.G. van der Mast. 2001. Research on Presence in Virtual Reality: A Survey. *CyberPsychology & Behavior* 4, 2: 183–201.
4. Frank Steinicke, Gerd Bruder, Student Member, Jason Jerald, Harald Frenz, and Markus Lappe. 2010. Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics* 16, 1: 17–27.
5. Stephan Streuber and Astros Chatziastros. 2007. Human interaction in multi-user virtual reality. *Proceedings of the 10th International Conference on Humans and Computers. HC 2007*, 1–6.
6. Jianbo Su. 2014. Motion Compression for Telepresence Locomotion. *Presence: Teleoperators and Virtual Environments* 16, 4: 385–398.
7. David Waller, Eric Bachmann, Eric Hodgson, and Andrew C. Beall. 2007. The HIVE: A huge immersive virtual environment for research in spatial cognition. *Behavior Research* 39, 4: 835–843.
8. Rec Room on Steam. *1 Jun, 2016*. Retrieved November 30, 2016 from http://store.steampowered.com/app/471710/.
9. Pool Nation on Steam. *18 Oct, 2013*. Retrieved November 30, 2016 from http://store.steampowered.com/app/254440/.
10. Unity - Game Engine. *2016*. Retrieved November 30, 2016 from https://unity3d.com/.
11. 2016. Vive | Discover Virtual Reality Beyond Imagination. Retrieved November 15, 2016 from https://www.vive.com/ca/.
12. 2016. Tilt Brush. Retrieved November 20, 2016 from https://www.tiltbrush.com/.