Analyzing the Effect of a Virtual Avatar's Geometric and Motion Fidelity on Ego-Centric Spatial Perception in Immersive Virtual Environments

Brian Ries*, Victoria Interrante, Michael Kaeding, and Lane Phillips
Department of Computer Science
University of Minnesota

Abstract

Previous work has shown that giving a user a first-person virtual avatar can increase the accuracy of their egocentric distance judgments in an immersive virtual environment (IVE). This result provides one of the rare examples of a manipulation that can enable improved spatial task performance in a virtual environment without potentially compromising the ability for accurate information transfer to the real world. However, many open questions about the scope and limitations of the effectiveness of IVE avatar self-embodiment remain. In this paper, we report the results of a series of four experiments, involving a total of 40 participants, that explore the importance, to the desired outcome of enabling enhanced spatial perception accuracy, of providing a high level of geometric and motion fidelity in the avatar representation. In these studies, we assess participants' abilities to estimate egocentric distances in a novel virtual environment under four different conditions of avatar self-embodiment: a) no avatar; b) a fully tracked, custom-fitted, high fidelity avatar, represented using a textured triangle mesh; c) the same avatar as in b) but implemented with single point rather than full body tracking; and d) a fully tracked but simplified avatar, represented by a collection of small spheres at the raw tracking marker locations. The goal of these investigations is to attain insight into what specific characteristics of a virtual avatar representation are most important to facilitating accurate spatial perception, and what cost-saving measures in the avatar implementation might be possible. Our results indicate that each of the simplified avatar implementations we tested is significantly less effective than the full avatar in facilitating accurate distance estimation; in fact, the participants who were given the simplified avatar representations performed only marginally (but not significantly) more accurately than the participants who were given no avatar at all. These findings suggest that the beneficial impact of providing users with a high fidelity avatar self-representation may stem less directly from the low-level size and motion cues that the avatar embodiment makes available to them than from the cognitive sense of presence that the self-embodiment supports.

CR Categories: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism – *virtual reality;* I.3.6 [**Computer Graphics**]: Methodology and Techniques – *interaction techniques*

Keywords: Spatial perception, presence, head mounted displays, immersive virtual environments, virtual avatars.

1 Introduction and Previous Work

Virtual reality technology is experiencing increasingly wide-spread use in many areas, including computer-aided architectural design [Anderson et al. 2003]. Through the use of head mounted displays (HMDs) and virtual environments, architects and their clients have the potential to preview alternative designs at full scale before they are constructed, and to experience these designs from an immersive, first-person perspective that facilitates the informed evaluation of design decisions such as choosing the heights of ceilings or defining the sizes of rooms. This use of VR during the design iteration process not only has the potential to aid architects in creating better designs but also to enable customers to more effectively weigh the cost/benefit tradeoffs of different design alternatives.

To achieve this potential, it is important to ensure that users' perception of 3D space in the virtual environment is equivalent to what their perception of that space would be in the real world. However, previous research has found that people typically tend to underestimate egocentric distances in virtual environments relative to in the real world [Henry and Furness 1993, Loomis and Knapp 2003]. This problem has to be overcome if we are to maximize the utility of virtual reality technology in areas such as architectural design and design evaluation.

Robustly enabling people to accurately perceive distances in virtual environments can be tricky. Waller and Richardson [2008] found that while providing people with visual feedback in a virtual environment can enable them to adapt to the apparent compression and thereby perform more accurately on distance judgment tests in the VE, it subsequently causes them to inaccurately overestimate distances in the real world. Likewise, while artificially scaling the images shown in the HMD has been shown to induce more accurate responses on distance estimation tasks in a virtual environment [Kuhl *et al.* 2006], neither the real world after-effects of this manipulation, nor the impacts of its long-term use in a VE, have yet been investigated. Our goal is to enable people to perceive distances accurately without inducing adaptation, by compelling them to cognitively interpret what they see via the HMD in the same way as they would interpret a corresponding view of the real world.

In recent investigations [Interrante *et al.* 2005, 2006], we have found that people tend *not* to significantly underestimate distances in a virtual environment that is a high-fidelity replica of a real environment that they have recently been in (a *virtual replica room*), and our subsequent studies exploring this phenomenon [Interrante *et al.* 2008] suggest that the accuracy of participants' distance judgments in a virtual environment may depend, at least in part, on cognitive factors such as their depth of presence in the virtual environment. Also supporting this premise is recent work by Steinicke *et al.* [2009], which finds that people who enter a novel virtual environment via a portal from a virtual replica room judge distances more accurately in the novel environment than do people who experience it immediately upon donning the HMD.

In the real world, preventing people from viewing their bodies does not interfere with their ability to accurately judge distances

e-mail: {riesbr, interran, phillips}@cs.umn.edu; mikleing@gmail.com

VRST '09, November 18-20, 2009, Kyoto, Japan.

[Creem-Regehr *et al*. 2005]. However this does not directly imply that, in a virtual environment, allowing people to see their bodies will have no positive effect. Mel Slater and colleagues have long suggested that providing people with a virtual avatar can enhance their sense of presence in a virtual environment [Slater and Usoh 1994], and several other groups have found that even partial embodiment can enable enhanced performance on tasks within reaching distance [Linebarger and Kessler 2002; Lok et al. 2003; Salzmann and Froehlich 2008]. In recent research, we and others have found that when people are allowed to experience a fairly high quality, first-person avatar self-embodiment in a virtual environment, their ability to accurately assess egocentric distances in that environment [Ries et al. 2008], or in a different, subsequently experienced environment [Mohler et al. 2008], is improved. What remains unclear is exactly why and how these avatars are enabling this enhanced performance. The principal goal of the research presented in this paper is to investigate what particular aspects of a fully-tracked, high-fidelty avatar are critical to facilitating distance perception accuracy, and what simplifications in the avatar implementation might be feasible without compromising the avatar's effectiveness for this purpose.

In this paper, we explore two main elements of the avatar representation: geometric and motion fidelity. A geometrically faithful avatar provides users with robust familiar sizes cues that they could use to calibrate their perception of distances in a virtual environment. But constructing an exact-matching avatar for a particular individual can be a difficult task. For many applications it isn't practical to create a personalized avatar for each user based on a 3D scan of their body; in these cases we must begin with a generic avatar model and scale it to fit. In order to obtain a robust fit, we must not only match the overall height of the model to the height of the user, but also match the individual lengths and widths of each of the separately animated body parts, including the torso and limbs, and especially the feet. Ideally, this calibration process should be straightforward and only require a short amount of time to complete. Presently, rendering a well-fitting avatar requires considerably more effort than simply rendering the raw locations of the markers used for the body tracking. Similarly, a faithfully animated avatar has the potential to provide users with motion cues that enable them to perform a perceptual-motor calibration of the 3D virtual environment that facilitates making accurate judgments of distances in the VE. But achieving a low latency, high fidelity animation for the avatar requires considerable resources. Full-body tracking must be available to precisely capture fine details of the user's real-world motions. The tracked motion must then be mapped, in real time, onto an avatar skeleton that precisely fits the user's body. All of these requirements significantly raise the bar on the virtual reality equipment and software calculations needed. Our goal in this paper is to gain deeper insight into the level of fidelity in the avatar's geometry and motion that is required to enable people to more accurately judge distances in a high-fidelity, novel virtual environment when embodied. This analysis provides two benefits: it allows us to investigate possible simplifications in the procedures and expenses required for a successful virtual avatar implementation, and it sheds some light on the psychological effects that underlie the effectiveness of virtual avatars in facilitating accurate distance perception.

2 Experiment 0: No Avatar

In a small portion of an initial experiment, whose results were presented in a poster several years ago [Interrante *et al*. 2005], we investigated the accuracy with which disembodied users could estimate egocentric distances in a high fidelity virtual environment that was a faithful replica of a real place that they had never been in. The purpose of that experiment was to obtain a baseline measurement of participants' expected accuracy in a typical virtual environment distance estimation task whose conditions matched the conditions of similar previous studies by others. For our present paper, we extended this portion of that previous work by doubling the size of the participant group in the relevant condition.

2.1 Method

In this experiment, we used direct blind walking to assess the relative accuracy with which participants, in a within-subjects design, could estimate egocentric distances when immersed, without any avatar self-embodiment, a) in a novel, high fidelity virtual environment, versus b) in the corresponding real world environment. All of the participants whose results we analyzed in this study experienced the virtual environment before experiencing the corresponding real environment.

2.2 Participants

We recruited five new participants for this experiment, whose results we analyzed in conjunction with the results of the five participants in our 2005 experiment. The participants in the original experiment ranged in age from about 20 to 45 and included students from the Departments of Architecture and Computer Science recruited via email and classroom announcements, and members of the local community recruited via personal connections. The five new participants were randomly self-selected University of Minnesota students recruited via a hand-held placard from among passersby to the building housing our lab. Each participant was compensated with a $10 gift card for their time.

2.3 Apparatus

The virtual environment used in this study was a high-fidelity 3D model of a restricted-access hallway located on the 4th floor of the building housing our lab. Figure 1a shows a view of the virtual hallway and figure 1b shows a photograph of the actual hallway for comparison. The virtual model was constructed in SketchUp based on detailed measurements of the actual space, and the geometry of the model was texture-mapped with photographs taken in the real environment. Care was taken to accurately model not only the gross dimensions of the space but also its 3D geometric details, including the door handles, trim, and recessed light fixtures. The illumination of the virtual environment was captured from the lighting and shading present in the photographs, which were segmented and applied as decals over the 3D polygonal meshes comprising the corresponding elements in the model, with lighting disabled in the rendering pipeline.

The virtual environment was presented to participants in our lab, via an nVisorSX head mounted display. The lab space within which the experiment was conducted consisted of a large, open room, approximately 30' long and 25' wide in the center, tapering down to 16' wide at each end. The virtual hallway model was positioned within the lab space so that the length of the hallway was aligned with the long direction of the room. Although the ends of the virtual hallway extended beyond the ends of the lab in each direction, our experimental protocol utilized only the portion of the hallway model that fit within the room.

The HMD came equipped with foam blinders that blocked peripheral vision of the external environment, and two tiny displays that provided participants with separate 1280x1024 resolution images to each eye over a manufacturer-specified 60º diagonal field of view (~2.2 arc minutes per pixel) with 100%

Figure 1: Left: a screen shot of the virtual hallway environment used in our experiments; Right: a photograph of the real hallway from which the virtual model was derived.

stereo overlap. The device weighed approximately 1kg (2.2 lbs) and was attached by a 15' cable to a video controller box that was strapped onto to a small, wheeled cart. During the experiment, an assistant carefully managed the cables to keep their presence as transparent as possible to the user, holding them up at one end to provide relief to the participant's head from their weight and using the other end of the cables to tug the cart along as necessary. Figure 2 illustrates this setup. A desktop computer monitor in the room enabled the experimenters to see the right eye image being presented to the participant via the HMD.

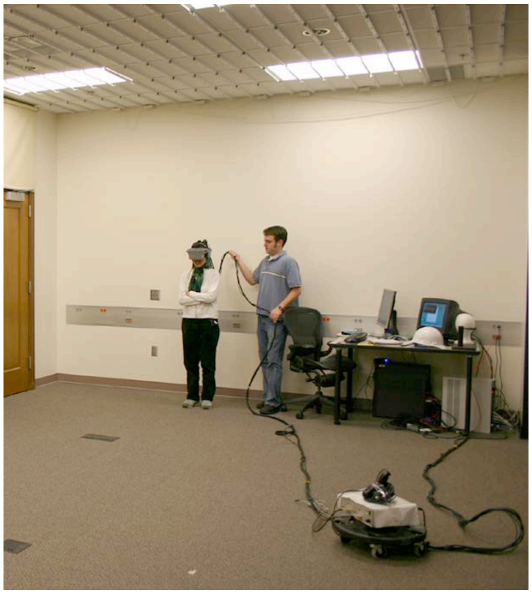

Figure 2: A photograph from 2005 showing the lab space, HMD and HiBall tracker, and illustrating the role of the assistant in managing the cables.

For this experiment, head tracking was provided by a HiBall 3000 optical ceiling tracking system manufactured by 3rd Tech. This system enables extremely high accuracy, low latency tracking with 6 degrees of freedom for up to four tethered sensors. Because we only needed to track the head, we used a single sensor attached to the back of the HMD.

The virtual environment was rendered using OpenGL on a desktop computer with a Xeon 2.83GHz processor with 2.0GB of RAM and nVidia Quadro 4500 graphics card.

Throughout all phases of the experiment, including both real world and virtual world trials, participants wore a small portable radio that was configured to provide static noise to each of their ears via tiny headphones. The purpose of the radio was to mask out subtle ambient sounds, such as the hum of the computers in the lab, to prevent participants from using any form of auditory information, including echolocation, to infer their proximity to any point in the lab at any time during the experiment.

2.4 Procedure

Each participant began by entering our laboratory, reading written instructions describing the experimental protocol, and signing the consent form. Then, the experimenters verbally repeated the instructions to the participant while assisting them in putting on and adjusting the HMD and radio.

The experiment began with 20 trials of blind walking in the virtual hallway environment. At the start of each trial, participants stood near one of the ends of the room, looking down the virtual hallway, and our computer software generated a virtual tape mark centered within the width of the hallway at a random distance 8'-25' in front of them. Participants took visual aim at the mark, and then, when they were ready to begin, closed their eyes, said 'ready', and walked with their eyes closed until they thought they reached the marked location. They then stopped and said 'done'. Upon hearing the 'ready' signal, the experimenter pressed a key to record the participant's starting location and clear the display to black, which prevented participants from inadvertently acquiring any visual input while walking. Upon hearing 'done', the experimenter pressed a different key to record the participant's ending location while leaving the display black. Participants were then guided to walk, with their eyes still closed, until they reached an arbitrary spot near the opposite end of the room, at which point they were told to turn around 180º to face down the hallway in the other direction. The display was then turned back on and a new target location was presented. Participants were never given any feedback about their performance at any time.

After completing the virtual world trials, participants were taken upstairs to the actual hallway, where they performed 10 trials of blind walking in the real environment as a control. In the real world trials, participants wore a blindfold instead of the HMD, and the starting and ending locations of each trial were marked by pieces of cloth laid down at arbitrary locations by the experimenter while the subject was not looking. Because the hallway was not tracked, both the walked distance (between the participant's starting and ending locations) and the presented distance (between the starting and ending cloth markers) had to be measured manually by the experimenters.

2.5 Results

Figure 3 shows the average relative errors in the virtual world and real world distance estimates obtained from each of the ten participants in this experiment. The position of each point along the horizontal axis is defined by the average relative error in the participant's distance judgments in the real world, computed as: (walked_distance – presented_distance)/(presented_distance) and the position of each point along the vertical axis is defined by the average relative error in the distance judgments made in the virtual environment. Points will be close to the plotted diagonal line when a participant's virtual and real world errors are nearly equivalent, and will fall below this line when participants underestimate distances in the virtual world relative to in the real world.

We performed separate ANOVAs on the data from each participant to assess the significance of the differences in the magnitude and direction of the errors in their distance estimates across all trials under the real world and virtual world presentation conditions. For each participant's data, the point is either rendered as hollow (white) if the virtual and real world errors were not significantly different ($p > 0.05$), unsaturated (grey) if they were only slightly significantly different ($0.01 \le p \le 0.05$), or filled with color if they were strongly significantly different ($p \le 0.01$). The horizontal and vertical error bars around each point bound the 95% confidence intervals of the means.

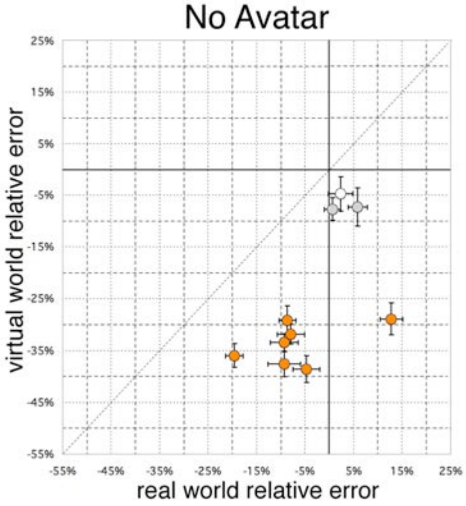

Figure 3: A plot of the average relative errors in distance judgments made by participants with no provided avatar.

We also ran an overall ANOVA, across participants, on the average relative errors made by each participant in the virtual and real worlds. We found that, as a group, the ten participants in this study made significantly greater errors in the virtual environment than in the real world $\{F(1,18) = 17.84, p < 0.001\}$.

2.6 Discussion

Of the ten participants in this experiment, one performed equivalently well in the real and virtual worlds, two performed only slightly worse in the virtual world and seven performed significantly worse in the virtual than in the real world. Overall, as confirmed by the ANOVA analysis, these results indicate that when participants are directly immersed, with no self-embodiment, in a high quality virtual environment that represents a real place they have never been to before, they will typically tend to underestimate egocentric distances in the virtual environment relative to in the real world. This finding replicates the general conditions and results of similar previous experiments by other research groups and lays the foundation for subsequent experiments in which the conditions of embodiment are varied.

3 Experiment 1: High Fidelity Avatar

The goal of our next experiment was to assess the extent to which providing participants with a high-fidelity, fully tracked, first-person avatar self-embodiment might enable them to make distance judgments in a novel virtual environment that are closer to the judgments that they would make in the real world. A preliminary version of this experiment was presented in a short paper last year [Ries *et al.* 2008]; for our present study we extend that work by increasing the number of participants by over 60%.

3.1 Method

In this experiment, as in experiment 0, we use direct blind walking to assess the relative accuracy with which participants can estimate egocentric distances in a novel, high-fidelity virtual environment versus in the real world. However this time we provide the participants with a high quality, fully tracked, first person avatar self-embodiment in the virtual environment.

3.2 Participants

We recruited four new participants for this experiment, whose results we analyzed in conjunction with the results of the six participants in our 2008 experiment. All ten participants were recruited in the same fashion, by standing on the sidewalk in front of our building advertising the experiment to passing pedestrians using a hand-held placard. As before, each participant was compensated with a $10 gift card.

3.3 Apparatus

The apparatus used in this experiment is similar to that used in experiment 0, with a few changes. Specifically, the virtual environment, computer hardware and head-mounted display used in this experiment are the same as those used previously; but instead of using the Hiball system to track the position and orientation of the head only, we switched over to using a Vicon motion capture system to track both the head mounted display and the participant's entire body. The Vicon system consisted of 12 MX40 cameras, mounted on poles close to the ceiling and positioned around the room to optimize the coverage of the portion of the room where the participant would be walking. Body tracking was accomplished by having participants wear a tightly fitting, black, non-reflective body suit covered with retroreflective markers; additional markers were also attached to participants' hands and feet. The head mounted display was defined as a separate object and tracked independently from the body via six retroreflective markers that were attached to it in an asymmetrical arrangement.

Each participant was embodied using the same generic avatar, shown in figure 4. The avatar was represented by a texture-mapped mesh, purchased from TurboSquid, that was manually skinned to the default Vicon skeleton in an offline pre-process.

We used a modified version of the OGRE 3D rendering API for the virtual environment and animated avatar rendering, which was extended to enable proper stereo rendering in OpenGL. The appearance of the virtual environment was identical in the OGRE-based renderer as in the renderer that we used in experiment 0.

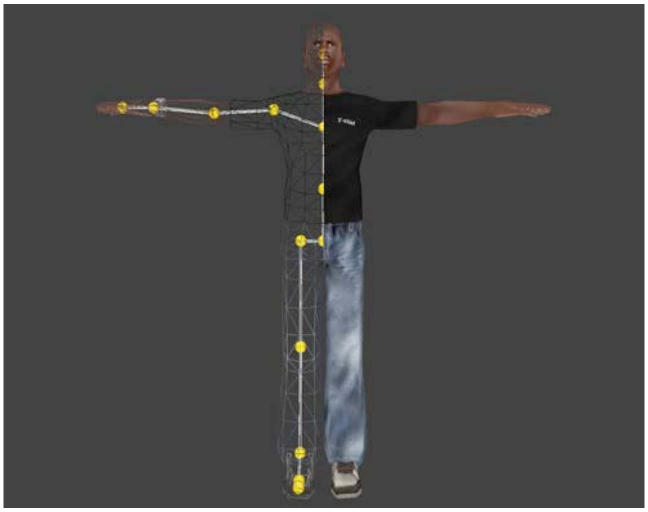

Figure 4: The virtual avatar used in experiment 1.

Despite the changes in apparatus, we did not notice any significant difference in the overall end-to-end latency of the tracking and rendering systems between experiments 0 and 1, though we did not attempt to robustly quantify the end-to-end system latency in either case and it is likely that subtle differences existed.

3.4 Procedure

After verbally indicating their willingness to participate in the experiment, participants entered our laboratory and were instructed to put on a two-piece body suit over the clothes they were wearing. To expedite the suiting up process, the suit already had retroreflective markers attached to it, which we re-positioned as necessary, after the participant put on the suit, to ensure optimal placement. We then attached additional markers to the participant's hands and feet, using toupee tape. After the participant was properly suited up, we had them perform a range of motion to define the correspondence between the marker locations while the person was moving and segments of the animated skeleton in the Vicon calibration system. Participants were guided to properly perform the range of motion, which involved moving each of their limbs in a wide arc and rotating each of their joints, by mimicking the movements of one of the experimenters.

After completing the range of motion, the participant was invited to read the written instructions describing the experimental protocol, and to sign the consent form for the experiment. During this time, the manual portion of the Vicon calibration process was completed by a second experimenter. At the end of this process, participants were given the opportunity to view their skeleton on the desktop monitor of the computer running the Vicon system and to verify that they controlled it.

Finally, the virtual avatar mesh was scaled to match the height of the user and, using the markers on the ends of the user's big toes as a guide, the virtual avatar's feet were scaled to match the length of the participant's. When the avatar model was ready, the participant was instructed to move to a location at the far end of the room and put on the radio and head-mounted display.

The remainder of the procedure is identical to that of experiment 0, except that at the beginning of each trial in the virtual environment, the starting point as well as the ending point was indicated by a virtual tape mark. This was done to surreptitiously ensure that each participant was equivalently and repeatedly exposed to their avatar representation during the experiment. Participants were instructed to begin each trial by looking down and lining their toes up to the starting piece of tape, which, at a key press from the experimenter, was generated at a random spot approximately 8-12" in front of the participant's current location. This required them to shuffle their virtual (and real) feet forward a short distance to get them aligned with the location of the tape mark, ensuring that motion information as well as graphical information was provided by this action.

As in experiment 0, each participant performed 20 trials of blind walking in the virtual hall, followed by 10 trials of blind walking in the corresponding real hall. Participants were directed to take off the body suit before performing the real world trials.

3.5 Results

The results of experiment 1 are shown in figure 5. As in experiment 0, we performed separate ANOVAs on the data from each participant and color-coded the points according to the statistical significance of the differences in the participant's errors between the real and virtual world conditions. As before, the error bars bound the 95% confidence intervals of the means.

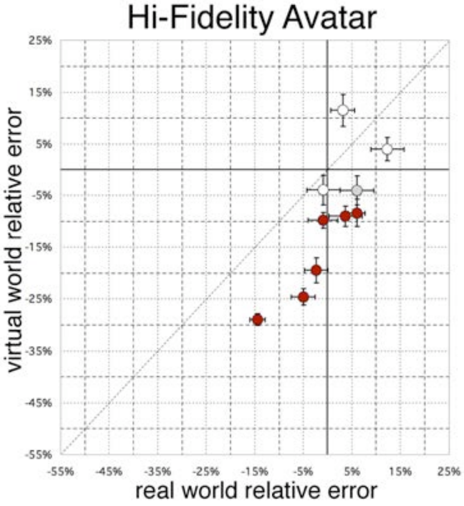

Figure 5: A plot of the average relative errors in distance judgments made by participants given a high fidelity, first person, avatar self-embodiment.

We also performed an overall ANOVA on the average results from each participant, and found that, as a group, the ten participants in experiment 1 made significantly greater errors in the VE than in the real world {$F(1,18) = 4.934$, $p = 0.0394$}.

3.6 Discussion

Of the ten participants in this experiment, three appeared to judge distances with approximately equivalent accuracy in the real and virtual worlds, one performed only slightly worse in the virtual world, and six underestimated distances in the virtual world relative to in the real world to a statistically significant extent. Individual performance also varied in accuracy within each environment condition, with, for example, some people tending to walk slightly long in the real world on average and others tending to walk slightly short. Such individual variations are to be expected and are generally not noteworthy. Overall, as indicated by the ANOVA analysis, what these results mainly tell us is that people will typically still tend to underestimate egocentric distances in a novel virtual environment relative to in the real world even when they are provided with a fully tracked, high fidelity, first-person avatar. However, the data also suggest that the magnitude of the errors that participants make when thus embodied is reduced relative to the magnitude of the errors that are made when participants are immersed in a virtual world without any self-embodiment. An ANOVA analysis of the virtual minus real world errors made by participants across experiments 0 and 1 finds a significant difference between the no avatar and full avatar conditions {F(1,18) = 7.387, p = 0.0141}. This indicates that while providing the virtual avatar doesn't enable people to perceive distances in a novel virtual environment as accurately as in the real world, it does provide a significant improvement over what is possible when only wearing an HMD with head tracking.

4 Experiment 2: Stiff Avatar

Having established that providing people with a custom-fitted, fully tracked avatar embodiment can make a difference, we next set out to explore the potential of using various types of simplified avatar representations. Achieving a fully tracked avatar requires a time-consuming set-up procedure and expensive tracking equipment. Might providing people with an avatar of the same geometric fidelity as in experiment 1, but with more rudimentary animation capabilities, that could be provided by a simpler and more economical tracking system, work just as well? If the increased accuracy we observed in experiment 1 was a result of enabling participants to calibrate their perception of egocentric distances in the virtual world using the familiar size cues provided by a virtual body that was correctly scaled to match their own, then providing them with that same avatar without the motion capabilities should have the same positive effect.

4.1 Participants

We recruited ten new participants for this experiment, in the same manner as for experiments 0 and 1. Each participant was compensated with a $10 gift card.

4.2 Apparatus

The hardware apparatus (computer, HMD and tracker) and rendering software used in this experiment is identical to that used in experiment 1. The only difference is that, in order to emulate the availability of limited tracking information, instead of animating the full range of motion of every part of the body, as in experiment 1, we limited the animation of the avatar to root translation and rotation, keeping the rest of the body in a fixed pose. To acquire this 'neutral pose', we asked each participant, after completing the range of motion calibration, to stand straight with their hands at their side and we locked that pose manually.

Figure 6: A third-person view of the stiff-avatar condition.

We then kept track of the position and orientation of the root node of the participant's skeleton as they moved, and used that information only to update the position and orientation of the stiff avatar model. Figure 6 shows what this looked like.

4.3 Procedure

The procedure for this experiment was identical to the procedure in experiment 1 with the only difference being that when participants shuffled forward to line up their virtual feet with the virtual tape mark indicating the starting position of a virtual world trial, the movement of the virtual avatar did not exactly match their actual movement – rather the avatar model simply translated forward as a rigid entity.

4.4 Results

Figure 7 shows the results of experiment 2, using the same layout and color coding scheme as in experiments 0 and 1. As in experiments 0 and 1, an ANOVA on the aggregate data from experiment 2 revealed significant differences in the virtual world versus real world errors, overall {F(1,18) = 11.15, p = 0.00366}.

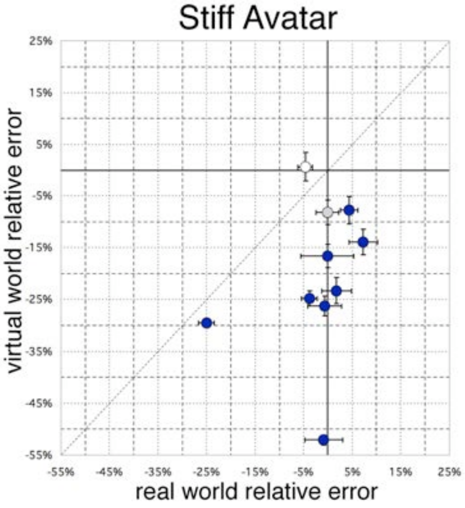

Figure 7: Average relative errors in distance judgments made by participants given a stiff-avatar self-representation.

4.5 Discussion

The magnitude of the average distance estimation errors made by the different participants in experiment 2 spanned an exceptionally large range, from being nearly equivalent in the virtual and real worlds to being underestimated in the virtual world by over 50% while being accurate in the real world to within 5%. These large individual differences complicate interpretation of the results. An ANOVA comparison of the virtual minus real world errors made by participants across experiments 1 and 2 does not reveal a significant difference between the full avatar and stiff avatar conditions {F(1,18) = 2.148, p = 0.1599}, but this should not be interpreted as a robust indication that the stiff avatar is as effective as the full avatar. A similar comparison across experiments 0 and 2 also fails, and to a greater extent, to reveal a significant difference in the virtual world minus real world errors between the stiff avatar and no avatar conditions {F(1,18) = 0.395, p = 0.538}.

5 Experiment 3: Dot Avatar

Our last experiment sought to investigate the efficacy of providing participants with a visually simplified avatar representation. There are several reasons why it could be advantageous to sidestep the process of animating and rendering a full avatar model. First of all, because the motion tracking process is sensitive to occlusion, under conditions of unrestricted movement the system often fails to successfully locate all of the tracking markers. In extreme cases, this can lead to errors in the definition of the skeleton, causing the avatar mesh to become hideously deformed. This doesn't happen often, but when it does, it is obvious and disturbing. If we bypass the character animation process and just render the raw marker locations captured by the cameras, the effect of isolated marker visibility failures is much less catastrophic. Secondly, obtaining an avatar model that bears a reasonable resemblance to any individual participant is very difficult, and it is unclear what psychological effects might result from various types of mis-matches in the avatar representation, from gender, to race, to age or level of physical fitness. Using a simplified (abstracted) avatar is not only easier, but also permits side-stepping these potentially complicated issues. If participants in experiment 1 were primarily using the perception-action correspondence provided by the motion of their virtual body to calibrate their perception of distances in the virtual world, or if participants are equally willing to accept a faithfully animated but inexactly rendered avatar as 'themselves' for the purposes of establishing presence in a VE, then we could find that providing participants with a dot avatar representation is just as effective as giving them a fully rendered embodiment.

5.1 Participants

Ten new participants were recruited for, and compensated for their participation in, this study in the same way as before.

5.2 Apparatus

The hardware apparatus and rendering software used in experiment 3 was identical to that used in experiments 1 and 2. The only difference from experiment 1 is that we did not render the full avatar mesh. Rather, each real physical marker attached to the user's body was represented by a virtual white sphere of the same size whose position was updated in real time to match the location of the physical marker. Figure 8 illustrates the result.

5.3 Procedure

The procedure for experiment 3 was identical to the procedure in experiment 1 with the only difference being that instead of lining up their virtual feet with the starting piece of tape, participants were asked to line up the virtual markers on each of their big toes.

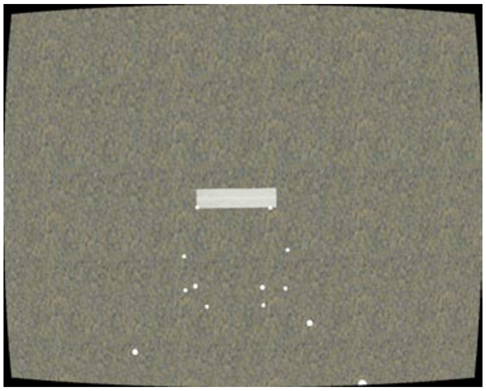

Figure 8: A first-person view of the dot avatar representation. The white spheres are drawn at the locations of the reflective markers worn by the tracked participant.

5.4 Results

Figure 9 shows the results in the dot-avatar condition. Overall, participants again underestimated distances in the virtual world relative to in the real world {F(1,18) = 19.65, p < 0.001}.

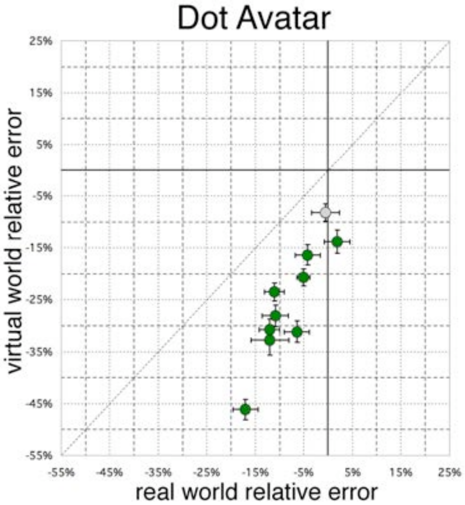

Figure 9: Average relative errors in distance judgments made by participants embodied as a dot avatar.

5.5 Discussion

A comparison of the results between experiments 1 and 3 shows that participants who were embodied using a dot avatar rather than a full avatar underestimated virtual verus real world distances to a significantly greater extent{F(1,18) = 5.221, p = 0.0346}. We did not find a significant difference in the magnitude of participants' virtual-minus-real world errors in the dot avatar and no avatar conditions {F(1,18) = 1.18, p = 0.295}. We also found no significant difference between the dot avatar and stiff avatar results {F(1,18) = 0.014, p = 0.907}.

6 General Discussion

Figure 9 shows a summary comparison of the results of the four experiments we conducted. The fully-tracked, full-body virtual avatar condition is the only one in which there is a significant improvement over the baseline no-avatar condition. This result suggests that the enhanced spatial accuracy enabled by the full-avatar embodiment is not derived solely from the low-level cues to familiar size or motion that the avatar embodiment provides, as these cues were also available, albeit separately, with each of the reduced fidelity avatar implementations. Rather, our results suggest that the enhanced performance that the full avatar enables may depend on higher level, cognitive factors that influence the interpretation of these lower level cues.

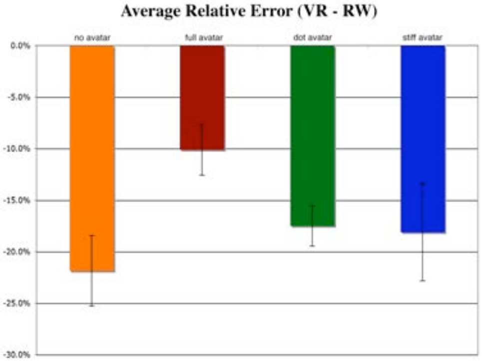

Figure 10: Overall comparison of the difference in participants' average relative errors in the virtual and real worlds, for each of the four different avatar types tested. Error bars bound the 95% confidence intervals of the means.

It may be that there is a minimal level of avatar fidelity that is required to enable participants to interpret their embodiment as a sufficiently plausible self-representation to allow them to accept what they see in the VE through the avatar's eyes as equivalent to what they would see in the real world through their own.

When we perform an ANOVA on the aggregate data from all 40 participants in all 4 experiments, interpreting the results of each experiment as indicative of the effects of a variation of the general condition *avatar type*, we do not find a significant main effect of 'avatar type' overall $\{F(3,36) = 2.101; p = 0.117\}$. This result is not surprising in light of the wide range of individual differences, and the fact that participants' errors were very similar, on average, in three of the four avatar conditions tested (stiff avatar, dot avatar, and no avatar). What our results suggest is *not* that distance perception accuracy varies with avatar type *in general*, but rather that it may be selectively facilitated by an appropriate confluence of avatar configuration characteristics.

In interpreting our results, it is also important to clarify that the fact that our experiments did not find a statistically significant difference in the errors in participants' distance judgments in the no avatar vs dot avatar and no avatar vs. stiff avatar conditions does *not* mean that performance under these conditions has been proven to be equivalent. It is possible that in a larger study, with more participants, the trends towards moderately increased accuracy that our results suggest could be found to be significant. It is also possible that in future studies employing different types of reduced-fidelity avatars, significant performance differences between simplified avatar and no avatar conditions will be found. Nevertheless, our results are somewhat discouraging from an equipment-purchasing standpoint, as they do not immediately reveal any notable shortcuts or cost savings that can be used to enable the same level of spatial judgment accuracy as can be obtained with a fully tracked, full-body avatar.

7 Future Work

While the highest fidelity avatars we have studied so far help, they still do not enable fully accurate distance perception. However these avatars still lacked realism, and the extent of users' exposure to them was minimal. In future work it would be interesting to see if performance could be further improved by providing people with an avatar that is a higher fidelity match, or by allowing them to have greater exposure to their avatar, *e.g.* via a virtual mirror.

It would also be interesting to explore the consequences of various types of deliberate mis-matches in the avatar embodiment, such as providing people with an avatar of the opposite gender. An intriguing potential of VR self-embodiment is enabling 'perspective-taking' [Yee and Bailenson 2006]. Could we use avatar embodiment to enable people to accurately perceive the virtual world from the perspective of someone else of a different size? This particular situation would be particularly useful in ergonomic studies of auto interior design.

Acknowledgments

This research was supported by the National Science Foundation through grants IIS-0313226 and IIS-0713587, and by the University of Minnesota through a Digital Technology Center seed grant, and by the Linda and Ted Johnson Digital Design Consortium Endowment and Lab Setup Funds.

References

ANDERSON, L., ESSER, J., AND INTERRANTE, V. 2003. A virtual environment for conceptual design in architecture. In *Proc. 9th Eurographics Workshop on Virtual Environments/7th Intl. Workshop on Immersive Projection Technology* (Zurich, Switzerland, May 22–23, 2003), 57–63.

CREEM-REGEHR, S., WILLEMSEN, P., GOOCH, A., AND THOMPSON, W. 2005. The influence of restricted viewing conditions on egocentric distance perception: implications for real and virtual indoor environments. *Perception*, 34, 2 (February 2005), 191–204.

HENRY, D., AND FURNESS, T. 1993. Spatial perception in virtual environments: evaluating an architectural application, In *Proceedings of the Virtual Reality Annual International Symposium,* (Seattle, Washington, September 18–22, 1993), 33–40.

INTERRANTE, V., ANDERSON, L., AND RIES, B. 2005. Lack of 'presence' may be a factor in the underestimation of egocentric distances in immersive virtual environments [Abstract]. Journal of Vision 5, 8 (May 2005), 527.

INTERRANTE, V., RIES, B., AND ANDERSON, L. 2006. Distance perception in immersive virtual environments, revisited. In *Proceedings of the IEEE Conference on Virtual Reality,* (Alexandria, Virginia, March 25–29, 2006), 3–10.

INTERRANTE, V., RIES, B., LINDQUIST, J., KAEDING, M., AND ANDERSON, L. 2008. Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators and Virtual Environments* 17, 2 (Apr. 2008), 176–198.

KUHL, S. A., THOMPSON, W. B., AND CREEM-REGEHR, S. H. 2006. Minification influences spatial judgments in virtual environments. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization* (Boston, Massachusetts, July 28–29, 2006), 15–19.

LINEBARGER, J. M., AND KESSLER, G. D., 2002. The effect of avatar connectedness on task performance. *Technical Report LU-CSE-02-007, Lehigh University, Department of Computer Science and Engineering.*

LOK, B., NAIK, S., WHITTON, M. C., AND BROOKS JR., F. P. 2003. Effects of handling real objects and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. In *Proceedings of the IEEE Conference on Virtual Reality* (Los Angeles, California, March 22–26, 2003), 125–132.

LOOMIS, J. M., AND KNAPP, J. M. 2003. Visual perception of egocentric distance in real and virtual environments. In Virtual and Adaptive Environments, L. J. Hettinger & M. W. Haas (Eds.), 21–46.

MOHLER, B. J., BÜLTHOFF, H. H., THOMPSON, W. B., AND CREEM-REGEHR, S. H. 2008. A full-body avatar improves egocentric distance judgments in an immersive virtual environment. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (Los Angeles, California, August 9–10, 2008), 194.

RIES, B., INTERRANTE, V., KAEDING, M., AND ANDERSON, L. 2008. The effect of self-embodiment on distance perception in immersive virtual environments. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*, (Bordeaux, France, October 27-29, 2008), 167–170.

SALZMANN, H., AND FROEHLICH, B. 2008. The two-user seating buck: Enabling face-to-face discussions of novel car interface concepts. In *Proceedings of the IEEE Conference on Virtual Reality* (Reno, Nevada, March 8–12, 2008), 75–82.

SLATER, M., AND USOH, M. 1994. Body centred interaction in immersive virtual environments. In Artificial Life and Virtual Reality, N. Magnenat-Thalmann and D. Thalmann (Eds.), 125–148.

SLATER, M., STEED, A., MCCARTHY, J., AND MARINGELLI, F. 1998. The influence of body movement on subjective presence in virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40, 3 (September 1998), 469–477.

STEINICKE, F., BRUDER, G., RIES, B., HINRICHS, K. H., LAPPE, M., AND INTERRANTE, V. 2009. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization* (Chania, Crete, September 30–October 2, 2009), *to appear*.

WALLER, D., AND RICHARDSON, A. R. 2008. Correcting distance estimates by interacting with immersive virtual environments: effects of task and available sensory information. *Journal of Experimental Psychology: Applied*, 14, 1 (March 2008), 61–72.

YEE, N. AND BAILENSON, J. 2006. Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. In *Proceedings of the ISPR 8th Annual International Workshop on Presence* (Cleveland, Ohio, August 24-26, 2006), 147–156.