Can Virtual Human Entourage Elements Facilitate Accurate Distance Judgments in VR?

Karla Paraiso[1] and Victoria Interrante[2]

[1] Arizona State University, Tempe AZ 85281, USA
[2] University of Minnesota, Minneapolis MN 55455, USA
kparaiso@asu.edu, interran@umn.edu

Abstract. Entourage elements are widely used in architectural renderings to provide a sense of scale and bring the drawings to life. We explore the potential of using a photorealistic, three-dimensional, exact-scale model of a known person as an entourage element to ameliorate the classical problem of distance underestimation in immersive virtual environments, for the purposes of enhancing spatial perception accuracy during architectural design reviews.

Keywords: Virtual environments, distance perception, virtual human entourage elements.

1 Introduction

Immersive virtual reality (VR) technology has tremendous potential as a tool for enabling architects and their clients to make important design decisions about features like room layout, ceiling height, etc. in a building project based on their first-person experience of the proposed structures before they are built. However, previous research has persistently found that people tend to systematically underestimate egocentric distances in immersive virtual environments presented via head-mounted display systems (HMDs) [22]. This can pose a particular problem for the use of VR in architectural design reviews, as the classically-reported average error rate of –25% [22] could make a 9-foot ceiling feel like an 8-foot ceiling when viewed from a typical female eye height of 5 feet off the ground. While the newest generation of lighter, more ergonomic and wider-field-of-view HMDs appear to afford more accurate distance judgments than HMDs of the past [e.g. 30, 3], recent studies have shown that their use does not eliminate the problem completely [12].

Previously proposed work-arounds to the distance underestimation phenomenon in HMD-based VR include: enforcing an acclimation period during which people can learn from their visual, proprioceptive, and possibly haptic feedback to adopt a more accurate interpretation of the mapping between what they see and how far away it is [23, 11]; immersing people in a photorealistic virtual replica of their concurrently-occupied physical environment before transitioning them into the virtual model of interest [26]; using an artificially exaggerated geometric field of view (gFOV) when rendering the virtual environment, effectively "minifying" the image contents [16]; artificially lowering the eye height used when rendering the virtual environment [15]; introducing a bright light into the periphery of the observer's visual field [7]; or embodying the user in a virtual self-avatar [24, 19]. Each of these potential interventions has various drawbacks as well as advantages. For instance: the visual adaptation obtained by physically walking over shorter distances does not extend to more accurate judgments of farther distances [11], and if adaptation is naturally occurring during immersion in a virtual environment, then the amount of artificial adjustment to eye height or gFOV necessary to evoke metrically accurate distance judgments for any particular user would likely vary over time. In addition, interventions that involve any sort of pre-process or extra encumbrance may be viewed as undesirable by the target population. By way of example, our architectural colleagues have found that clients can sometimes be hesitant even to want to wear an HMD, much less to "suit up" or even don simple wrist- and ankle-worn devices to facilitate full-body tracking. Hence, while many efforts have already been made to address the problem of distance underestimation in VR, and much progress has been achieved, further opportunity for advancement remains.

In that context, we report here our initial investigations into the potential for facilitating more accurate distance perception in HMD-based immersive virtual environments by introducing photorealistic, accurately-scaled, virtual human entourage elements into the modeled spaces.

2 Previous Work

There is a long history of including small human figures in architectural drawings, both as an indicator of scale and to convey a sense of how the space might be inhabited [1, 4]. The rationale for using human figures, as opposed to other, inanimate, objects to provide familiar size cues is, however, not immediately clear. There is considerable population variability in human height (e.g. 160-185cm for the average man, depending on his country of origin)[*], which suggests that the human body may have the potential to serve only as an inexact ruler for metric size; at the same time, however, psychologists have found that people are remarkably adept at judging the height of unknown persons from full-length photographs [10]. Furthermore, it is possible that, for evolutionary reasons, we afford privileged status to human figures when making perceptual judgments, or that, due to ecological constraints, size consistency in humans is simply more robust than size consistency in other objects (e.g. when one sees a door that is half as tall as a person standing next to it, it is more probable that the door is 3.5' tall than that the person is 14' tall).

Previous research on distance perception in HMD-based immersive virtual environments has found that sparsely adding individual items of furniture, such as a table in one spot and a chair in another, in an attempt to provide additional cues to familiar size in an otherwise large empty room, does not lead people to make more accurate distance judgments [6]. The importance of using representatively furnished models is however supported by real-world studies showing that the presence of furniture has a significant

[*] https://en.wikipedia.org/wiki/List_of_average_human_height_worldwide

impact on spaciousness judgments in an architectural context, with rooms judged (by architecture students) to be less spacious when empty or over-crowded than when moderately furnished [5].

With respect to the potential effectiveness of adding virtual human agents to a modeled interior space to facilitate distance perception accuracy in VR, prior work paints a mixed picture. Although several previous studies have found that participants make more accurate distance judgments when they are provided with an avatar self-embodiment, the underlying explanation for this improvement is unclear. On the one hand, research showing a significant impact of the self-avatar's geometric and motion fidelity on distance estimation accuracy suggests that embodiment might lead to improved performance by evoking a stronger sense of presence in the virtual environment or a clearer appreciation of the affordances for action in the virtual world [25]. On the other hand, studies showing that people tend to perceive gaps on the ground as wider when they are embodied in smaller feet [8] or generic objects on a table as being smaller when viewed in the context of self-embodiment with a larger virtual hand [17], suggest that people tend to use their own body as a ruler to calibrate their perception of the sizes of nearby objects in VR. It is not clear, however, to what extent one's own virtual body size, or the virtual body sizes of others, might influence judgments of mid-range distances in a virtual space [14]. Mohler et al. [19] found that distance judgments improved with self-embodiment in VR regardless of whether the embodiment was experienced via a first-person or third-person view. However, Linkenauger et al. [17] found that people do not scale their perception of the size of a generic object to the size of a non-embodied avatar's hands.

Most directly related to our current work, Ragan et al. [21], in a 2012 poster presentation, report an exploratory study using a desktop virtual environment in which the addition of static or well-animated non-realistic virtual characters led to more accurate spatial perception judgments than when no characters or badly-animated characters were used. However, contemporary work by McManus et al. [18] reported finding no impact of either a generic self-avatar embodiment or the presence vs. absence of a generic autonomous agent on peoples' distance perception accuracy in a realistic virtual room environment.

These results leave open the question of whether, and under what conditions, adding static or animated virtual human models to a virtual architectural environment might facilitate more accurate judgments of egocentric distance in those virtual interiors. It seems logical to expect that a static virtual human model of a known or assumed height might potentially serve as a reliable indicator of size or scale in an immersive virtual environment. Additionally, the presence of compellingly realistic dynamic virtual humans could potentially evoke the inference of similar affordances for one's own actions in the virtual environment as in the real world, facilitating more accurate distance judgments despite the lack of an embodiment of one's own. In our present reported work, we further add a novel contribution to the results of prior work through our decision to use a photorealistic virtual human model representing a faithful scale replica of an individual that the participant will have met in person just prior to their immersion in VR, an intervention designed to maximize the potential of the virtual human model to serve as a reliable metric for absolute size judgments in the virtual environment.

3 Our Experiment

3.1 Method

We used a mixed within-and-between subjects design to expose each participant to three different conditions of virtual human (VH) presence – no VH, static VH, or dynamic VH – in three different virtual hallway environments, prior to having them make action-based egocentric distance judgments in those environments by walking without sight to previously-viewed target locations indicated by a virtual white mark at one of five different pre-defined distances on the virtual floor. Different hallway models were used in the different VH conditions to avoid that an impression of the interior space derived under one VH condition would influence distance judgments queried under a different VH condition. The assignment of virtual human condition to hallway environment was randomized between participants.

Participants. We recruited a total of 18 participants (10 m., 8 f., ages 19-29, $\mu = 21.5 \pm 2.8$) from our local University community via email lists and posted flyers. Participants were compensated with a $10 gift card to an online retailer. Our experiment was approved by our university's Institutional Review Board, and all participants gave written informed consent.

Materials. We used Autodesk Maya to create three different virtual hallway models, which we imported into Unreal Engine, where they were lit and populated with system-provided assets such as light fixtures, picture frames, plants, doors, and windows, plus various items of furniture obtained from Arbitrary Studio [2]. The images in the paintings were obtained using a Google search for commercially-free-licensed images. We constructed the hallways so that all three models were essentially structurally equivalent – each having the same basic length and width – but differing in appearance with respect to decorative details. Figure 1 shows what each hallway model looked like from the participant's starting position.

Fig. 1. The three hallways models used in our experiment.

We used the Skanect software by Occipital in conjunction with a Structure Sensor mounted on an ipad to capture a 3D model of the experimenter, which we imported into Mixamo to rig and animate. The experimenter wore the same outfit when conducting the study. The shoes were added to the model in a post-process.

Fig. 2. A photograph of the experimenter (left) and her corresponding 3D model (right).

In each of the static avatar conditions, the 3D virtual human model was placed in a visible position either beyond or to the side of the path over which the participant would need to walk in order to reach the targets used for the blind-walking distance judgments. In light of the findings of Jung et al. [9], suggesting that people's estimates of distances to forward-facing virtual humans may be affected by social influences, we made sure to orient the model so that she was facing away from the participant and to position her so that her attention was implied to be engaged by some other item in the virtual hallway. Figure 3 shows what this looked like in each case.

Fig. 3. A view of the static avatar in each of the three different virtual hallway environments.

For the animated avatar conditions, we programmed the virtual human to traverse a path either across the virtual hallway or down the hallway in a direction away from the participant. In each case, the participant first had a view of the empty hallway before the virtual human walked into view. The virtual human appeared either from a doorway at the far end of the hall, from a doorway at the near end of the hall, or from behind the participant and to their right, and it exited from view through another doorway. We imported basic walking and turning motions from Mixamo to animate the virtual model's limbs, and used Unreal Engine's Animation Blueprint to define the character's movement through the scene. Figure 4 shows a representative frame from each of the three different character animations.

Fig. 4. A representative view of the animated avatar in each of the three different virtual hallway environments.

The experiment was conducted in our virtual reality laboratory, which approximately spans a 30' x 29' space. The virtual environment was rendered using Unreal Engine, running on an ORIGIN PC with an Intel Core i7 6850K Hex-Core 3.6GHz processor, 32GB DDR4 SDRAM (2800MHz), and a single 8GB NVIDIA GeForce GTX 1080 Founders Edition graphics card. Participants viewed the virtual environment using an HTC Vive head-mounted-display, which presents two 1080 x 1200 pixel resolution images on OLED displays, one for each eye, over a combined field of view of approximately 100˚h x 110˚v. The device weighs approximately 1 lb. and attaches securely to the head via wide elastic straps. 6DOF head tracking was accomplished using Valve's Lighthouse Tracking system, which spanned an approximately 15' x 15' area at one edge of the open lab space.

Procedures. Participants were greeted at the door of our lab by the same experimenter who would be represented as a virtual human in the VR environment. Participants were scheduled by appointment so that each participant proceeded individually through the experiment, and no participant was exposed to any activity of any other participant. Each participant was first screened for adequate visual acuity, defined by the ability to successfully read lines of letters corresponding to 20/60 or above on a wall-mounted eye chart from a distance of 20' without wearing glasses. We chose the distance of 20' out of consideration of the focal distance of the optics in the HTC Vive, and participants were tested without glasses because they would not be able to wear their glasses in the HMD. None of the participants failed the visual acuity test. Next, participants were screened for stereo vision ability by asking them to identify two shapes of increasing depth complexity (a rectangle and a goldfish) presented as random dot stereograms on an Oculus Rift HMD. All of the participants passed the stereo vision test. After their eligibility to participate in the experiment had thus been verified, participants were given written instructions explaining the experiment procedure and were asked to sign an informed consent form. They then filled out a short survey providing basic demographic information (age, gender) and completed a baseline simulator sickness form (SSQ) [13].

Each participant was randomly assigned to a different set of three blocks of hallway/avatar combinations, defined so as to ensure that each participant would be exposed to each hallway and each avatar, and so that over the 18 total participants, each different hallway would be seen in combination with each different avatar, in each different possible order. Specifically, six participants were assigned to the six different possible presentation orders of the combinations {(H1,A1), (H2,A2), (H3,A3)}, six to different presentation orders of {(H1,A2), (H2,A3), (H3,A1)}, and six to different presentation orders of {(H1,A3), (H2,A1), (H3,A2)}. Due to an unfortunate oversight during the manual execution of the experiment, however, we noted when compiling the data that two participants had inadvertently been presented with incorrect stimuli: participant 11 was immersed in the condition (H1,A1) in the second block instead of (H3,A1), meaning that they saw H1 twice and didn't see H3, and participant 12 was presented with {(H2,A3), (H3,A2), (H1,A1)} instead of {(H2,A1), (H3,A2), (H1,A3)}, meaning that overall, some avatar/hallway combinations were seen more frequently than others.

During each block of trials, the participant began by standing at a pre-defined location in the registered real and virtual environments, marked by some tape on the floor of our lab. Figure 1 shows what each of the hallway environments looked like from the "home base" position. After the participant had had a chance to briefly look around (without moving from the home base position), their starting position was recorded using a keypress, a small white target was presented at a pre-defined location on the floor of the virtual hallway, and participants were asked to fixate on the target and when ready, to close their eyes, say "ready", and walk to where they thought the target was. Upon hearing the word "ready", the experimenter pressed a key to turn the display to black so that participants would not be able to see anything even if they did open their eyes. When the participant stopped, the experimenter used another key press to record their stopping location. As we had discovered, during pilot testing, that positional recording at the farthest distances could sometimes fail when the participant was facing away from their starting position because of the limitations in the range of the tracking system, we asked participants on each trial to turn around in place so that we could make a second recording of their ending position. When analyzing the data, we used information derived from the second recordings to infer the participant's ending position in the rare cases where the first recording failed. More details on this procedure are provided in the results section. With their eyes still closed, participants were then led on a circuitous path back to the home base position to start the next trial. Each block consisted of a total of 5 trials, in which targets were presented at distances of 8', 10', 12', 14' and 16', in randomized order.

Following each block of trials, participants removed the HMD, filled out a SSQ survey, and enjoyed some water and a package of pretzels, cookies, or crackers. After the final block of trials, participants completed a brief presence questionnaire, in which they provided numeric answers on a scale from 1 to 7 to a total of 13 different questions intended to assess various aspects of their sense of presence in the virtual environment. These questions, provided in the Appendix, were drawn from a combination of the Witmer-Singer IPQ [28] and the Slater-Usoh-Steed presence questionnaire [27]. Finally, participants were asked four additional exit survey questions related to their impression of the relative realism of the different hallway environments, the realism of the virtual human, the strategy they used to arrive at the target square, and any suggestions they had for improving the virtual environment experience.

3.2 Results

On eight of the 270 total trials, the outward facing ending position was recorded as (0,0,0) due to the inability of the HMD's sensors to see the light pulses emanating from the Lighthouse tracking stations. These trials affected a total of four participants. For each of these participants, we computed the median offset between the positions successfully recorded in the outward-facing and inward-facing directions on all of their other trials to derive an average "correction vector" that we then added to the inward-facing direction recorded at each point where the outward-facing position was unresolved, in order to infer the missing value(s). This procedure was necessary because the tracked position of the HMD moved in a systematic way when participants rotated in place, due to the HMD being located in front of the face while the axis of rotation was closer to being through the middle of the head. In every case that an outward-facing measure was invalid, an inward-facing measure was available. Additionally, on four out of the 270 trials, only one ending position was recorded in the data file, most likely due to experimenter error. In those cases, we interpreted the single recorded value as if it were a valid outward-facing value as this was the most common occurrence and also the most conservative, since the correction factors, when they were needed, tended to add several centimeters to the distance measured from the inward-facing orientation.

Our first step in analyzing the results was to compute a one-way ANOVA to test for the possibility of a significant effect of hallway type. Although we had tried to construct the three different environments to be as structurally similar as possible while at the same time clearly representing different places, so as to minimize the likelihood of carry-over effects in which distance judgments under later avatar conditions might be affected by an understanding of the environment obtained in earlier-experienced avatar conditions, the possibility remained that some unrecognized characteristics of some of the environments might potentially facilitate or hinder distance judgment accuracy to a different extent than others. For example, Witt et al. [29] report a significant impact of the environmental context beyond a target on the accuracy of peoples' judgments of their distance to the target, in real world scenes. Likewise, rooms with darker colored walls may tend to appear smaller. However, our analysis did not find any significant difference in distance perception accuracy between the different hallway environments used in our experiment {F(2,267) = 0.687, p = 0.504}.

Similarly, but more disappointingly, a one-way ANOVA also found no significant main effect of the avatar condition on distance judgment accuracy {F((2,267) = 0.199, p = 0.819}. These results refute our hypothesis that adding a static or animated virtual human figure to a modeled interior space might evoke or facilitate more accurate egocentric judgments of the interior space. We did, however, find a highly significant main effect of target distance on walked distance {F(4,85) = 21.368, p < 0.001}, providing basic assurance of the robustness of the experimental results and lending weight to the integrity of the findings of non-significance in the tests of hallway and avatar impact. Figure 5 shows plots of the average distance walked, for each distance shown, in each hallway environment and each avatar condition. In addition, we found no significant impact of block order on error rates {F(2,267) = 1.10, p = 0.334}, indicating that participants' performance on the distance judgment task remained generally consistent over time.

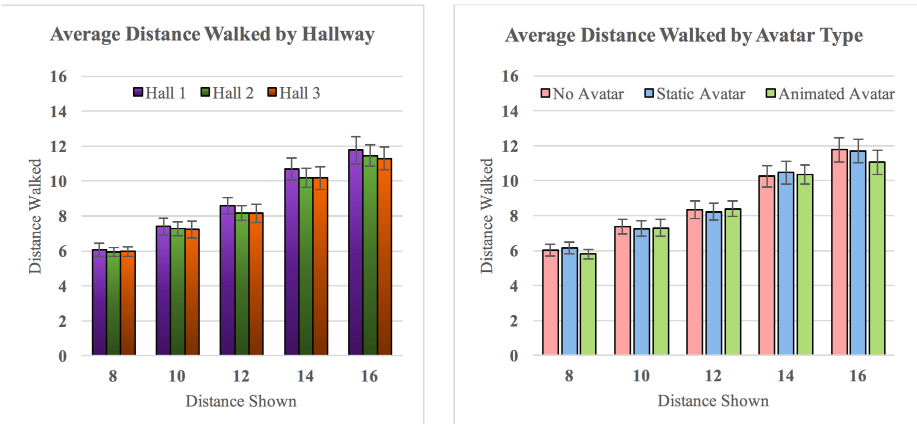

Fig. 5. Left: Average distance walked in each of the three different hallway environments, pooled over all three avatar conditions. Right: Average distance walked in each of the three different avatar conditions, pooled over all three hallway environments. Error bars represent ± 1 standard error.

Results from the SSQ surveys revealed no evidence of cybersickness. We did, however, find significant differences in distance judgment accuracy between participants {F(17, 252) = 43.361, p < 0.001}. Two individuals had exceptionally low average error rates of 0.7% and –1.4%, while one participant had an outlyingly high average error rate of –61.6%, three others had average relative errors in the range of –43.2% to –44.3%, and the remainder were in the range of –11% to –34%. Over all participants and all conditions, the average relative error in distance judgment accuracy was –27.35%.

Tabulating the results of the presence questionnaire, we found only moderate agreement with most questions, evidenced by an average Likert score of 4.63, after reversing the scores on the two questions whose scales ran from high to low rather than low to high. We also found no clear correlation between presence rating and distance judgment accuracy ($r = 0.12$), with the slope of the trendline being driven primarily by the single outlying participant. Figure 6 illustrates the relevant data.

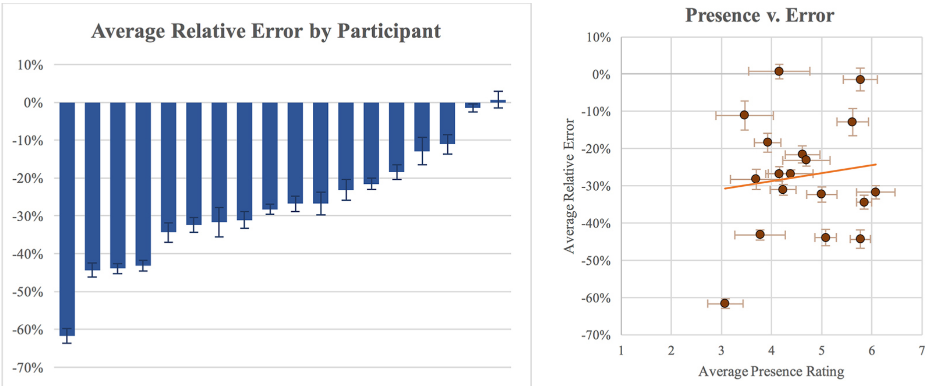

Fig. 6. Left: A plot of the average relative error in the distance judgments made by each participant, pooled over all hallway and avatar conditions and ordered from greatest to lowest average error. Right: A scatter plot, in which each disk represents the averaged data for one participant, showing little correlation between participants' subjective ratings of presence and the accuracy of their distance judgments in the virtual environment. Error bars represent ± 1 standard error in each chart.

In the exit survey, in response to the question: "Did all three hallways seem equally realistic, or did one seem more realistic than the others?", we found a few participants with a preference for hallway 2, and a few who complained about hallway 3 and about the animated avatar, but the majority of participants (10 of 18) reported that the hallways appeared equally realistic (8), or equally unrealistic (2). Specifically, of the four participants who singled out one hallway as being more realistic than the others, three of them chose H2 (with all combinations of avatars) and one chose H1 (seen without an avatar). Of the four participants who singled out one hallway/environment as being less realistic than the others, two mentioned H3, one mentioned H1 (in comparison to H2), and the other two specifically provided complaints about the animated avatar.

Overall, participants were not impressed with the realism of the virtual human. In response to the question: "Does the virtual person appear realistic/human-like or did it feel more un-human"?, the majority (11 out of 18) directly replied that the VH seemed unrealistic or un-human, and six of these, in elaboration, explicitly complained about the movement being unnatural, jerky, glitchy, or "obviously animated". In fact, only three of the 18 participants gave an unqualified response of "realistic" or "human-like" to this question; three more gave qualified responses ("somewhat", or "realistic with the exception of …"), which might be explained by politeness. A final participant remarked that the VH felt human-like only because it resembled the experimenter; they said that they were not sure if they would have thought the same if the virtual human had been a stranger.

When asked “What strategy did you use to arrive at the target square”, 10 out of 18 participants responded with some variation on estimating the number of steps that would be required to arrive at the target. The other 8 of 10 reported using a more holistic strategy, as in “I tried to remember the scene and imagine it when my eyes were closed”, or “I looked at the visual landmarks of the hallway and thought about passing them”.

For the open-ended question: “Is there anything else you can tell us to help us improve the virtual environment experience”, 8 of 18 participants had no response or just said “no”. However, three participants complained about the lighting of the virtual model, and two remarked on issues related to distance perception: one said that they felt too short in the virtual environment, and the other complained that the depth seemed “off”, “as though the hallway was not as deep as it was visually”. Also, one participant said that the headset felt heavy. In addition, three participants remarked on factors related to stopping before they reached the target: one reported a fear of walking into things, one said that the automatic appearance of the Vive boundary grid (which they said could be noticed even when their eyes were closed) was distracting, and one participant seemed to suggest that they thought that the grid appeared when they reached the target, saying that “When I got near the point, the display went blue. Don’t do that. Then you kind of know when to stop”.

4 Discussion

The primary result from this work is that we did not find any evidence of a significant improvement in participants’ distance judgment accuracy in the tested virtual hallway models from the addition of a static or animated 3D virtual human, even when the virtual human model was obtained from a 3D photographic scan of a known person. This suggests that people either do not tend to use the sizes of other inhabitants of a shared virtual environment as absolute indicators of the scale of space, or that such information is not useful for resolving the distance underestimation problem in VR.

We also note that most of our participants significantly underestimated distances in this experiment, which is consistent with historical findings, but stands in contrast to some of the more recent results using the Oculus Rift and HTC Vive [12, 3, 30]. We note several limitations of our implementation that may possibly have contributed to these increased errors, as well as several limitations that are unlikely to have had such an effect. Potentially influential factors include: lukewarm participant ratings of presence in the virtual environment, and shortcomings in the visual/experiential realism of the virtual environment model, in the form of notable inconsistencies in the physical plausibility of the lighting appearance and shortcomings of the illumination model used, and visible artifacts in both the static appearance and dynamic motion of the virtual human model. While we did our best to create virtual hallway and virtual human models of the highest possible quality, the bar for photorealism is very high and previous work has suggested that adding even a little unreality to an otherwise plausibly realistic

VR scenario may have significant consequences for distance judgment accuracy [20]. Less likely to be explanatory factors for participants' systematic errors in distance judgments but nonetheless potentially problematic are: the distracting appearance of the boundary grid when participants came close to the limits of the tracking area, the unintentionally slightly uneven distribution of stimulus conditions among participants, and the necessity of using an approximate correction factor to infer several pieces of missing data. The main reasons we doubt these irregularities played a significant role in shaping our findings are that we observed consistent underestimation errors of ~25% at each distance interval tested, including the shortest intervals of 8' and 10' where the boundary grid never appeared, we did not observe even a suggestion of a significant impact of the hallway environment on distance judgment accuracy, the four data points that required correction were a very small fraction of the total number of measurements obtained, and the average correction factor used in those cases was much smaller than the average error.

5 Conclusions and Future Work

While our present results provide scant evidence that the long-studied problem of distance underestimation in HMD-based VR could be solved by populating virtual environments with static or dynamic virtual human agents, we are nevertheless encouraged to continue pursuing efforts to bring virtual architectural models "to life" with autonomous intelligent agents. Just as room size perception may be affected by factors such as ceiling height or furniture placement, judgments of the suitability of the spatial layout of a virtual building model may be significantly informed by the ability to view and experience the interior spaces as they would appear in use, as opposed to empty. For large public buildings, such as schools, libraries, and hospitals, accurately modeling and representing them in their inhabited state could be equally as important and informative to peoples' subjective sense of the suitability of the design as providing an accurate impression of the metric dimensions of the space. Accomplishing this while avoiding the uncanny valley will be a significant future challenge.

Acknowledgments

This work was supported by the National Science Foundation through grants CHS: Small: Transforming the Architectural Design Review Process through Collaborative Embodiment in HMD-based Immersive Virtual Environments (1526693) and REU Site: Computational Methods for Discovery Driven by Big Data (1460620), by the Computing Research Association's Committee on the Status of Women in Research (CRA-W) through its Distributed Research Experiences for Undergraduates (DREU) program, and by the Linda and Ted Johnson Digital Design Consortium Endowment. A previous iteration of this project was begun by Grace Chen and Bezawit Moges in 2014, and continued by Yucai Chang with assistance from John Chen and Peng Liu.

References

1. Anderson, A.T.: On the human figure in architectural representation. Journal of Architectural Education 55(4), 238–246 (2002).
2. Arbitrary Studio. https://gumroad.com/l/uPXbQ. Last accessed Oct 2, 2017.
3. Creem-Regehr, S.H., Stefanucci, J.K., Thompson, W.B., Nash, N. McCardell, M.: Egocentric distance perception in the Oculus Rift (DK2). In Proceedings of the ACM Symposium on Applied Perception, pp 47–50. ACM, New York (2015).
4. Colonesse, F.: Human figure as a cultural mediator in architectural drawings. In: Koç, G., Claes, M.-T., Christiansen, B. (eds.) Cultural Influences on Architecture, IGI Global (2016).
5. Imamoglu, V.: The effect of furniture density on the subjective evaluation of spaciousness and estimation of size of rooms. In Küller, R. (ed.) Architectural Psychology (Proceedings of the Lund Conference), (1973).
6. Interrante, V., Ries, B., Lindquist, J., Kaeding, M., Anderson, L.: Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. Presence: Teleoperators and Virtual Environments 17(2), 176–198 (2008).
7. Jones, J.A., Swan, J.E. II, Bolas, M.: Peripheral stimulation and its effect on perceived spatial scale in virtual environments. IEEE Transactions on Visualization and Computer Graphics 19(4), 701–710 (2013).
8. Jun, E., Stefanucci, J.K., Creem-Regehr, S.H., Geuss, M.N., Thompson, W.B.: Big Foot: using the size of a virtual foot to scale gap width. ACM Transactions on Applied Perception 12(4), Article 16, 12 pages (2015).
9. Jung, E., Takahashi, K., Watanabe, K., de la Rosa, S., Butz, M.V., Bülthoff, H.H. Meilinger, T.: The influence of human body orientation on distance judgments. Frontiers in Psychology: Perception Science 7:217 (2016).
10. Kato, K., Higashiyama, A.: Estimation of height for persons in pictures. Perception & Psychophysics 60(8), 1318–1328 (1998).
11. Kelly, J.W., Hammel, W.W., Siegel, Z.D., Sjolund, L.A.: Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. IEEE Transactions on Visualization and Computer Graphics 20(4), 588–595 (2014).
12. Kelly, J.W., Cherep, L.A., Siegel, Z.D.: Perceived space in the HTC Vive. ACM Transactions on Applied Perception 15(1), Article 2, 16 pages (2017).
13. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lillenthal, M.G.: Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology, 3(3), 203–220 (1993).
14. Langbehn, E., Bruder, G., Steinicke, F.: Scale matters! Analysis of dominant scale estimation in the presence of conflicting cues in multi-scale collaborative virtual environments. In Proceedings of the IEEE Symposium on 3D User Interfaces, pp. 211–220 (2016)
15. Leyrer, M., Linkenauger, S.A., Bülthoff, H.H., Mohler, B.J.: Eye height manipulations: a possible solution to reduce underestimation of egocentric distances in head-mounted displays. ACM Transactions on Applied Perception 12(1), Article 1, 23 pages (2015).
16. Li, B., Zhang, R., Nordman, A., Kuhl, S.A.: The effects of minification and display field of view on distance judgments in real and HMD-based environments. In: Proceedings of the ACM Symposium on Applied Perception, pp. 55–58. ACM, New York (2015).
17. Linkenauger, S.A., Leyrer, M., Bülthoff, H.H., Mohler, B.J.: Welcome to wonderland: the influence of the size and shape of a virtual hand on the perceived size and shape of virtual objects. PLoS ONE 8(7), e68594 (2013).

18. McManus, E.A., Bodenheimer, B., Streuber, S., de la Rosa, S., Bülthoff, H.H., Mohler, B.J.: The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. In Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization, pp. 37–44. ACM, New York (2011).
19. Mohler, B.J., Creem-Regehr, S.H., Thompson, W.B., Bülthoff, H.H.: The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. Presence: Teleoperators and Virtual Environments 19(3), 230–242 (2010).
20. Phillips L., Interrante, V.: A little unreality in a realistic replica environment degrades distance estimation accuracy. In the Proceedings of IEEE Virtual Reality (Posters), pp. 235-236 (2011).
21. Ragan, E.D., Wilkes, C., Cao, Y., Bowman, D.A.: The effects of virtual character animation on spatial judgments. In the Proceedings of IEEE Virtual Reality: Short Papers and Posters, pp. 141–142 (2012).
22. Renner, R.S., Velichkovsky, B.M., Helmert, J.R.: The perception of egocentric distances in virtual environments - A review. ACM Computing Surveys 46(2), Article 23, 40 pages (2013).
23. Richardson, A.R, Waller, D.: Interaction with an immersive virtual environment corrects users' distance estimates. Human Factors: The Journal of Human Factors and Ergonomic Society 49(3), 507–517 (2007).
24. Ries, B., Interrante, V., Kaeding, M., Anderson L.: The effect of self-embodiment on distance perception in immersive virtual environments. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 167–170. ACM, New York (2008).
25. Ries, B., Kaeding, M., Phillips, L., Interrante, V.: Analyzing the effect of a virtual avatar's geometric and animation fidelity on ego-centric spatial perception in immersive virtual environments. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 59–66. ACM, New York (2009).
26. Steinicke, F., Bruder, G., Hinrichs, K., Lappe, M., Ries, B., Interrante, V.: Transitional environments enhance distance perception in immersive virtual reality systems. In Proceedings of the ACM/SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, pp. 19–26. ACM, New York (2009).
27. Usoh, M., Catena, E., Arman, S., Slater, M.: Using presence questionnaires in reality. Presence: Teleoperators and Virtual Environments 9(5), 497–503 (2000).
28. Witmer, B.G., Singer M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators and Virtual Environments 7(3), 225–240 (1998).
29. Witt, J.K., Stefanucci, J.K., Riener, C.R., Proffitt, D.R.: Seeing beyond the target: environmental context affects distance perception. Perception 36(12), 1752–1768 (2007).
30. Young, M.K., Gaylor, G.B., Andrus, S.M., Bodenheimer, B.: A comparison of two cost-differentiated virtual reality systems for perception and action tasks. In Proceedings of the ACM Symposium on Applied Perception, 83–90. ACM, New York (2014).

Appendix

Presence questions (based on the Université du Québec en Outaouais Cyberpsychology Lab's revision [http://w3.uqo.ca/cyberpsy/docs/qaires/pres/PQ_va.pdf] of the Witmer and Singer [28] Presence Questionnaire and the Slater-Usoh-Steed [27] Presence Questionnaire):

1. In the computer generated world, I had a sense of "being there" (1=not at all; 7 = very much)
2. When you think back to the experience, do you think of the virtual environment more as images that you saw or more as somewhere that you visited? (1=images; 7 = felt like I visited)
3. How aware were you of the real world surrounding while navigating in the virtual world (i.e. sounds, room temperature, other people, etc.)? (1 = Not aware at all; 7 = extremely aware) [reversed for scoring]
4. To what extent were there times during the experience when the virtual environment was the reality for you? (1 = none; 7 = always)
5. How real did the virtual world seem to you? (1 = not real at all; 7 = completely real)
6. How much did your experience in the virtual environment seem consistent with your real world experience? (1 = not consistent; 7 = very consistent)
7. How natural did your interactions with the environment seem? (1 = unnatural; 7 = very natural)
8. How compelling was your sense of objects moving through space? (1 = not at all; 7 = very compelling)
9. How completely were you able to actively survey or search the environment using vision? (1 = not at all; 7 = completely)
10. How compelling was your sense of moving around inside the virtual environment? (1 = not at all; 7 = very compelling)
11. How quickly did you adjust to the virtual environment experience? (1 = not at all; 7 = quickly adjusted)
12. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience? (1 = not reasonably; 7 = very proficient)
13. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? (1 = not at all interfered; 7 = interfered a lot) [reversed for scoring]