HAL
open science

Virtual Distance Estimation in a CAVE

William E. Marsh, Jean-Rémy Chardonnet, Frédéric Merienne

To cite this version:

William E. Marsh, Jean-Rémy Chardonnet, Frédéric Merienne. Virtual Distance Estimation in a CAVE. Spatial Cognition 9th, Sep 2014, Germany. pp.354-369. hal-01070539

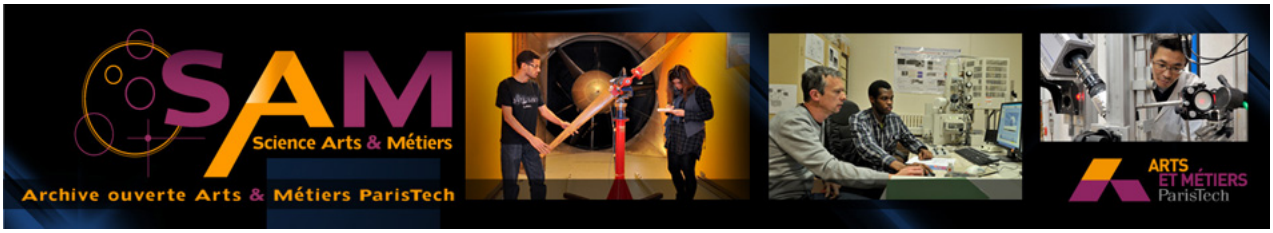
HAL Id: hal-01070539

https://hal.science/hal-01070539v1

Submitted on 1 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers ParisTech researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://sam.ensam.eu
Handle ID: .http://hdl.handle.net/10985/8656

To cite this version :

William E. MARSH, Jean-Rémy CHARDONNET, Frédéric MERIENNE - Virtual Distance Estimation in a CAVE - In: Spatial Cognition 9th, Germany, 2014-09-19 - Spatial Cognition IX, Lecture Notes in Artificial Intelligence - 2014

Any correspondence concerning this service should be sent to the repository Administrator : archiveouverte@ensam.eu

Virtual Distance Estimation in a CAVE

William E. Marsh, Jean-Rémy Chardonnet, and Frédéric Merienne

Arts et Métiers ParisTech – CNRS Le2i, Institut Image,
Chalon-sur-Saône, France
william@wemarsh.com, {jean-remy.chardonnet,frederic.merienne}@ensam.eu

Abstract. Past studies have shown consistent underestimation of distances in virtual reality, though the exact causes remain unclear. Many virtual distance cues have been investigated, but past work has failed to account for the possible addition of cues from the physical environment. We describe two studies that assess users' performance and strategies when judging horizontal and vertical distances in a CAVE. Results indicate that users attempt to leverage cues from the physical environment when available and, if allowed, use a locomotion interface to move the virtual viewpoint to facilitate this.

Keywords: Virtual reality, Distance perception, Spatial cognition, Scale, Architecture, CAVE.

1 Introduction

When interacting with a virtual architectural model, decision makers must accurately perceive the spatial layout of a virtual environment (VE) and make decisions, often regarding a building that has not yet been built. However, the current state of virtual reality (VR) technology presents challenges. One such challenge is that perceived spaces are routinely compressed in VR.

There are two broad categories of distance judgements: egocentric and exocentric. In egocentric tasks, distances are judged from the observer to an object. In exocentric tasks, distances are judged between two objects in the environment, neither of which is collocated with the observer. Past literature has sometimes shown different biases depending on whether a judgement is egocentric or exocentric. When movement is allowed, the distinction between egocentric and exocentric distances is less clear because a person could move in such a way that an exocentric judgement becomes egocentric, or vice-versa.

Many authors have investigated biases related to physical or virtual distance spans on, or parallel to, the ground plane. However, though they are particularly relevant to architectural decisions, there is a notable lack of studies involving vertical distances, and this dearth is even more profound in VR.

1.1 Methods to Assess Perceived Distances

In the following sections, we describe past findings regarding distance perception, and then we discuss our own studies investigating distance perception in a CAVE. This work hinges on identifying appropriate ways to measure perceived distances.

A common way of assessing perceived distances is blind walking, in which participants are asked to walk without vision to a previously viewed landmark [e.g. 11]. However, motor recalibration effects have been reported [14] by which participants learn to improve their estimates through training. These may reflect high-level cognition, as opposed to overall calibration. The effects can be avoided through triangulated blind walking or by withholding performance feedback [16].

Perceptual matching is another means of testing the accuracy of perceived distances. In this paradigm, participants are asked to reproduce a distance span, often by directing the experimenter to place a target at an equivalent distance [e.g. 9]. Due to the difficulty in blind walking a remembered vertical distance, perceptual matching was chosen for the studies described in this paper.

In the virtual architecture domain, it may be important to reproduce distances both egocentrically and exocentrically. Past researchers have often required participants to learn exocentric distance spans and then reproduce them egocentrically. This difference in orientation between learning and recall should prevent participants from using strategies relying on purely visual snapshots, though it may involve higher-level processing. In the studies described here, perceptual matching estimates are made from the same orientation as experienced during learning, but from a different distance, to prevent judgements based solely on visual angle.

1.2 Distance Perception in Real Environments

Most distance perception research has focused on horizontal distances. Real-world distance estimates on the ground are usually very accurate over short ranges, though depth spans are sometimes reportedly compressed relative to frontal spans [10]. In general, absolute distances can be calculated visually with knowledge about the size of familiar objects, the angle between an observer's eye height and an object or, when an object is on the ground, linear perspective. Some additional cues may help when judging distances shorter than approximately 2 m. Binocular disparity provides information, together with the ocular convergence required to target an object and the accommodation required for the lens to focus on an object. Motion parallax and optic flow can also help if the observer is moving.

Vertical distances viewed from the top are often overestimated. Large heights are also overestimated from the bottom, though to a lesser degree. When observing a frontal view of an object, vertical distances are often overestimated by 10–15% relative to horizontal distances [25]. Stefanucci and Proffitt [18] report evidence that these biases may be moderated by an individual's fear of heights.

When constructing the spatial model of a scene, the visual cues above are integrated with proprioceptive, efferent [22], and possibly vestibular information from movements [2]. Some researchers have implicated effort, as well as the capacity and intention to interact [24], in distance perception. For example, wearing a backpack has been shown to lead to increased distance estimates [15], though Durgin et al. [4] point to alternate explanations. If effort is indeed linked to distance perception, it follows that vertical movement requires greater effort

(climbing) than movement along the horizontal plane (walking), and thus vertical distances should be perceived as greater. A related idea, *Evolved Navigation Theory*, is that biases may relate to evolutionary goals of avoiding danger [8]. In either case, there is a link to Gibson's [6] concept of affordances, in which behavioral potential impacts perception.

1.3 Distance Perception in Virtual Environments

Some of the biases described above in Sec. 1.2 also exist in VR, though they are generally greater. Distances are linearly compressed in VEs relative to corresponding estimates in the real world [13] by up to about 50% [12], with reports of rare exceptions [7]. As in the real world, depth spans are often compressed relative to frontal spans [10]. Additionally, Fink et al. [5] showed differences in trajectories between real and virtual movements, indicating greater uncertainty about object positions.

There is a shortage of literature examining vertical distance perception. However, the work referenced above in Sec. 1.2 regarding the horizontal–vertical illusion showed a greater effect in the real world (or VR) than in photos with distances controlled for optical angle, and the magnitude of this illusion increases with physical perceived height on the projection surface [25]. This shows that humans have knowledge of the display surface and that it may affect perception.

Many authors have studied these phenomena, but the causes remain unclear. The problems may be due to some combination of factors involving sensory fidelity, equipment constraints (i.e. weight), and unnatural interfaces. In principle, VE designers should aim to provide the same cues that are available in the real world. However, in practice, this is never quite possible with current technology. For example, graphics fidelity is often reduced in VR. Thompson et al. [20] showed no impact of impoverished graphics on distance estimates, but distance cues including optic flow may still be limited. In VR, accommodation and convergence cues can be in conflict, possibly also impacting virtual distance estimates [3].

Interfaces for large-scale virtual locomotion often do not allow for the types of body-based translation movements, such as walking, that may lead to construction of accurate spatial representations [17]. In fact, Sun et al. [19] showed that providing body-based proprioceptive/efferent information improved path length estimates, even if it was inconsistent with reality as experienced visually. A common class of interface for virtual locomotion is the handheld wand. There are several variations and control schemes, but because they use different muscle groups and actions involve less effort than in the real world, they may lead to non-veridical distance estimates.

Past research has indicated that the intention to interact with an environment and the expected effort associated with that interaction impact distance perception [24]. These studies have primarily focused on horizontal distances, but such theories should predict underestimation of vertical distances in VR because is often easier to travel virtually (by flying) to higher elevations than it is in the real world (by climbing).

Physical characteristics of a VR system may also serve as distance cues. For example, in a CAVE it is usually possible to see corners where the screens meet. In addition, optical accommodation can inform the user about screen position and orientation [21]. It is conceivable that a user, particularly one with a relatively low level of virtual presence [as defined by 23], may use these physical aspects when judging virtual distances. However, most studies on virtual distance estimation have been administered using a head-mounted display (HMD), likely in an attempt to limit these physical cues. For scenarios in which presence is not critical, these cues may be beneficial. Allowing movement may enable users to move into positions to make maximal use of distance cues provided by the physical system. Alternately, cues in the physical environment may introduce additional unwanted biases, so it is important to study users' abilities to use them effectively.

1.4 Architectural Decisions in Virtual Reality

The present work is motivated by the need for *scale-one* distance perception in architectural walk-throughs. The architectural domain has some key requirements that differentiate it from much of the past literature on spatial perception.

First, architectural decisions frequently involve vertical judgements. Most past work has involved only horizontal distances, but the studies in this paper include a vertical condition.

Second, virtual architectural models are often viewed by more than one person at a time. For this reason, walk-throughs often take place on a large screen or in a CAVE, as opposed to an HMD. Most past virtual distance estimation research has been conducted using an HMD, so those results may provide a good starting point, but they fail to predict the interplay between virtual and physical distance cues that are present in a CAVE. The studies reported in this paper were conducted in a four-sided CAVE. Because trials were quick, participants were not expected to achieve high levels of presence, meaning that they might use distance cues from both the VE and the physical system. This paper will not address additional distortions that arise as multiple untracked users move away from the center of projection [1].

2 Study 1: Distance Estimation and the Role of Locomotion in a CAVE

We conducted a study to investigate horizontal and vertical distance estimation in a high-fidelity, fully furnished virtual office environment displayed in a CAVE. This was intended to provide insight into the design space for VR systems used for architectural decisions. Specifically, there were two competing hypotheses in Study 1. First, the ability to locomote may help users judge distances by maneuvering to the optimal vantage point(s). If this is the case, we hoped that the experiment would provide insight into the strategies employed. Alternately, using a handheld wand may lead to greater underestimation, particularly on vertical spans, due to the ease of travel.

2.1 Method

Participants. Twelve participants (4 female, $M = 26.3$ years, $SD = 5.5$ years) were recruited through word of mouth. The only requirements were good (corrected) binocular vision and lack of expertise in the CAVE. Some were familiar with VR, but specific knowledge of the system dimensions was undesirable because it could have biased judgements. There was no compensation for participation.

Apparatus. All distances were learned in a CAVE measuring 3×3 m horizontally and 2.67 m in height, with rear projection on three walls and front projection on the floor. All walls have passive stereoscopic projection. Participants' head positions were tracked with an ARTTRACK2 optical tracking system to control the center of projection. Participants recalled distances in the same room, with their backs facing the CAVE.

Stimuli and Design. The study employed a 2 (interface) × 3 (judgement direction) within-subjects design. Every participant completed two study sessions, counterbalanced, each with a different interface condition (Flystick, none). Three judgement directions (depth, width, vertical) were each experienced three times per session, in random order, such that each participant completed 9 trials per session (18 trials total). Each trial comprised two phases: learning and recall. The learning phase for all trials took place in the same virtual office scene, rendered in Fig. 1a, but each trial in a given session was observed from a different orientation, randomly chosen from a set of nine possible. In the recall phase, exocentric perceptual matching was used to reproduce the learned distance.

The two sessions were on different days to prevent spatial learning in one condition from carrying over to the other. There were two interface conditions: Flystick and none. In the Flystick condition, participants used a Flystick2 wand to translate in the X, Y, and Z directions. The interface did not allow rotation. The hat switch on top of the Flystick2 was used to move along the virtual X–Y plane at a maximum speed of 2.0 m/s, while the trigger button was used to fly at

(a) Virtual office environment. (b) Virtual cone.

Fig. 1. Virtual models used during the learning phase of the experimental trials

a fixed speed of 2.0 m/s. Releasing the trigger button for 2 s allowed gravity to take over, causing the participant to fall to the virtual floor. In the none interface condition, locomotion was not allowed.

In both interface conditions, participants were allowed to move their heads freely, but they were required to keep their feet in the same spot, 192 cm from the front wall of the CAVE (42 cm behind CAVE center), and equidistant from the side walls. This position was chosen in an attempt to maximize the available peripheral information. The virtual floor and the physical floor were co-planar, except when flying, though the virtual floor was larger in extent.

During the learning phase of each trial, distances were presented as a pair of virtual cones, rendered in Fig. 1b, placed at random positions in the virtual world. For depth trials, both virtual cones were placed in front of the participant, the first at a distance of 10–60 cm and the second at a distance of 245–295 cm. The range of possible observed distances was 185–285 cm, with a mean of 235 cm.

For width trials, the virtual cones were also separated by a distance of 185–285 cm, with a mean of 235 cm. The cones were placed at an equal distance to the left and right of the point 342 cm directly in front of the participant's body.

For vertical trials, the cones were also separated by a distance of 185–285 cm, with a mean of 235 cm. The first cone was placed at a distance of between 10 cm and 60 cm from the virtual floor and the second cone was placed at a distance of between 245 cm and 295 cm from the virtual floor. The vertical trials were the only ones in which the cones were not standing upright. In these trials, the tops of the cones pointed parallel to the floor, toward the participant.

During the recall phase, participants reproduced distances with exocentric perceptual matching using physical cones nearly identical to those modeled in the VE. For depth matching, the first cone was placed 105 cm from the tips of the participant's feet. The experimenter placed the second cone immediately beyond the first and slowly moved it away from the participant, until directed by the participant to stop.

For width matching, the first cone was placed 150 cm in front of and 182 cm to the right of the participant. The experimenter placed the second cone directly to the left of the first and slowly moved it leftward, along a path perpendicular to the direction the participant was facing, until directed by the participant to stop.

For vertical matching, the first cone was fixed 330 cm above the floor. The experimenter initially placed the second cone 1 m lower than the first. This avoided the need for a ladder, thus eliminating an obvious distance cue in the physical environment that may have led to biases involving anchoring or relative judgements between trials. The experimenter then moved the cone downward until directed by the participant to stop.

Procedure. When participants first arrived, they were told that they would be making a series of basic distance judgements and shown the physical cones that would be used for the perceptual matching tasks. They were given a brief explanation of how the real-world perceptual matching tasks would be used to recall the observed virtual distances.

In each trial, participants were allowed as much time as desired to learn the distances. Once they alerted the experimenter that they were ready, they moved quickly and quietly to the recall area across the room. Participants directed the experimenter to reproduce the learned distance and then a new trial began.

Before beginning the experimental trials, participants completed three practice trials (depth, width, vertical) in the real world. They were told that the tasks were intended for practice, but that the answers would be recorded. For these trials, as for the experimental trials described below, participants were allowed to view the target cones for as long as they wanted and then they were quickly and quietly led across the room to do a perceptual matching task to recall the previously seen distance.

Next, if participants were completing the Flystick session, they were trained to use the Flystick. This included a brief explanation and a practice scenario in which participants were encouraged to move around in a scene different from the experimental scene until they felt comfortable with the interface.

Next, participants completed nine experimental trials. Each of these took place in the same virtual office scenario, but viewed from a different orientation, chosen at random (without replacement) from a set of nine. There were three each of depth, width, and vertical judgements, presented in a random order. In the Flystick session, participants were instructed that they could explore as much as they wanted, but that they must also look at each cone up close.

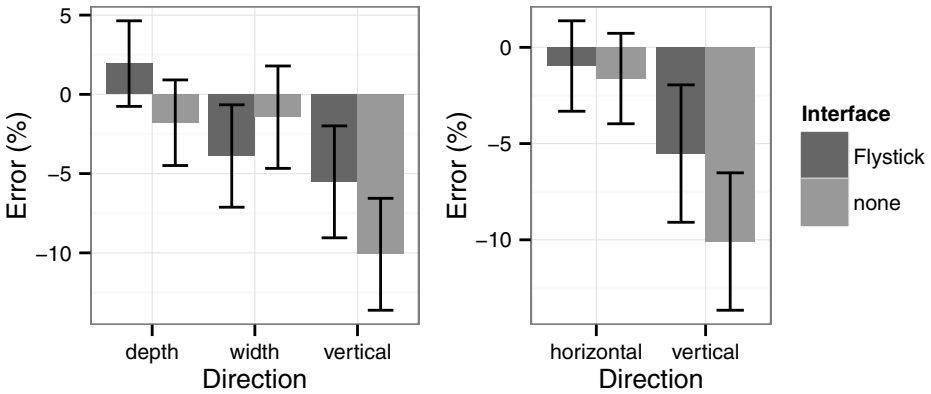
After finishing all trials, participants were asked to complete a post-questionnaire regarding their experiences and problems when estimating the distances. They were also asked to self-report their locomotion performance for the Flystick session. Additionally, after the first session, participants completed a demographic questionnaire with questions about age, sex, height, VR experience, video game experience, and how many hours they exercised per week.

2.2 Results: Distance Estimation Error

The following preparations were made before analysis:

- Because the top cone in the vertical matching task was 330 cm from the floor and because the cones were 14 cm wide, the largest distance that could be recalled was 323 cm. For this reason, all measurements were clipped at a maximum of 323 cm. Likewise, because 323 cm is 88 cm greater than the mean cone distance of 235 cm, very low estimates were clipped at an equal distance from the mean cone distance, 147 cm. This prevented depth and width judgements from being larger than those possible for vertical judgements. This only affected four low and four high judgements.
- The raw distance estimates were converted to percent error estimates, reflecting the percent of under- or over-estimation in a trial.

A two-factor linear mixed-effects model was constructed with percent error as a response, fixed effects for interface and direction combinations (6 means), a random effect for participant, and a by-participant random slope for direction. An



(a) All direction levels separate. (b) Horizontal direction levels pooled.

Fig. 2. Least-squares mean estimation error ± SE in Study 1 as a function of interface and direction

ANOVA using the Satterthwaite approximation for degrees of freedom showed a significant interaction between interface and direction, $F(2, 176.81) = 3.13$, $p = .046$. The least-squares means are plotted in Fig. 2a.

We were also interested in the question of whether distances were underestimated for each direction and interface. These means were compared to zero using planned t tests. This showed that vertical error was significantly less than 0 (underestimation), $t(11.0) = -2.32$, $p = .04$, while depth and width error were not significantly different than 0. Estimates in the none interface condition were less than 0, with marginal significance, $t(12.8) = -1.94$, $p = .07$. All means were compared using t tests with the Bonferroni correction for multiple comparisons, showing no additional significant differences.

When using the Flystick interface, participants were able to position themselves such that depth and width judgements may have been analogous. Additionally, there was no statistical difference between the depth and width judgements in either interface condition. Therefore, another analysis was conducted with the depth and width levels pooled into a single *horizontal* level. A two-factor linear mixed-effects model was constructed with percent error as a response, fixed effects for interface and direction combinations (4 means), a random effect for participant, and a by-participant random slope for direction. An ANOVA using the Satterthwaite approximation for degrees of freedom showed marginally significant main effects of both direction $F(1, 11.0) = 4.48$, $p = .058$, and interface, $F(1, 189.82) = 3.07$, $p = .08$. The least-squares means are plotted in Fig. 2b, showing that vertical error means are less than those for horizontal error. We can also see that the Flystick may have helped participants make better judgements.

2.3 Results: Experimenter Observations and Self-reported Strategies

One question on each post-questionnaire was a Likert scale, regarding perceived performance. The Kruskal-Wallis rank sum test revealed no significant difference in perceived performance between interfaces.

One participant reported problems because virtual objects seemed smaller than physical objects. This may correspond to typical distance underestimation problems in VR.

The experimenter noticed a common strategy when participants used the Flystick interface. Participants often intentionally positioned at least one cone at a physical edge of the CAVE or both cones in the corner between the walls, seemingly as a strategy to help with distance estimation. This was not reported in questionnaire responses. However, participants did comment on difficulty when the distance spans required the use of two physical screens. This problem was reported for both width and vertical judgements. This strategy may reflect attempts to use physical cues when learning the distances.

One participant noted that perceived distances were dependent on point of view, and another reported specifically that objects that were farther away appeared to be a different size when viewed up close. Another participant noted that judgements were easier when using the Flystick interface. This may be because it allows for changing the point of view or because it allows participants to position cones such that they are both on the same physical screen.

One participant reported using the number of imagined footsteps as a metric by which to remember distances. This may be considered a hybrid strategy, as the length of a footstep has meaning in both the virtual and physical environments. This same participant reported trouble on vertical estimates, because they were not easily estimable using this metric, and another stated that she did not know her own virtual height.

2.4 Discussion

Statistically, horizontal estimates were veridical. However, there was underestimation of analogous vertical spans that has not been previously explored in the literature. Note that this is not what would be predicted according to distance estimation literature in the real world, which has shown vertical distances to be overestimated relative to horizontal distances. Because estimates were actually more accurate (though not significantly so) with the Flystick interface, these results fail to support the notion that less-effortful locomotion causes underestimation of distances. However, these results do indicate that the ability to interact improves distance estimates.

While many results were not statistically significant, the patterns shown above and the subjective observations were used to motivate the next study. From the patterns we can see that the Flystick interface may have helped participants make more accurate judgements. In line with experimenter observations, this may be a result of different strategies in which the physical system boundaries are used when a participant is able to maneuver accordingly.

3 Study 2: The Role of CAVE Boundaries in Distance Estimation

We conducted a second study to further investigate the use of information from the physical environment when judging virtual distances. We hypothesized that distance judgements on the plane of the physical floor or walls of the CAVE would be more accurate than those in purely virtual space.

3.1 Method

The design and paradigm used in Study 2 closely resembled those in Study 1.

Participants. Fourteen participants (7 female, $M = 22.6$ years, $SD = 3.9$ years) were recruited through word of mouth. Participant requirements were the same as in Study 1. There was no compensation for participating.

Apparatus. As in Study 1, participants learned distances in the CAVE and recalled them in the same room, from a position with their backs to the CAVE.

Stimuli and Design. The study employed a 3 (physical position) $\times$ 2 (judgement direction) within-subjects design. Two judgement directions (width, vertical) were each experienced six times per session, in alternating order with the starting direction counterbalanced, such that each participant completed 12 trials. As in Study 1, each trial comprised two phases: learning and recall. The learning phase for all trials took place in the same virtual office scene used in Study 1, but each trial with a given judgement direction was observed from a different orientation, randomly chosen from a set of six possible. In the recall phase, exocentric perceptual matching was used to reproduce the learned distance.

There were three physical position conditions, as shown in Fig. 3: back, middle, and front. In all three conditions, participants were allowed to move their heads freely, but they were required to keep their feet in the same spot, with an equal distance from each side wall. The virtual floor and the physical floor were coplanar, though the virtual floor was larger in extent. In the middle position, participants stood 187 cm from the front CAVE wall. During the learning phase, distance spans were displayed at the front CAVE wall.

In the back condition, participants stood 50 cm behind the middle position, 237 cm from the front CAVE wall. During the learning phase, distance spans were displayed 50 cm in front of the front CAVE wall, on the physical floor (for width judgements) or right wall (for vertical judgements) of the CAVE.

In the front condition, participants stood 50 cm in front of the middle position, 137 cm from the front CAVE wall. During the learning phase, distance spans were displayed 50 cm past the front CAVE wall.

During the learning phase of each trial, participants viewed a pair of randomly positioned virtual cones, as in Study 1, though with a slightly improved cone model. The distances in Study 2 were slightly shorter than those in Study 1, because it was necessary to fit the cones for the vertical trials on the physical screen, which is 267 cm tall.

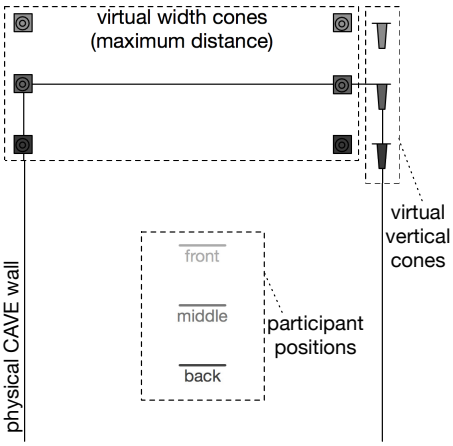


Fig. 3. Top-down diagram of the CAVE, showing the three physical participant positions together with the corresponding virtual cone positions

For width trials, the virtual cones were separated by a distance ranging from 167 cm to 267 cm, with a mean of 217 cm. The first cone was randomly placed at a position between 0 and 50 cm from the plane formed by the left CAVE wall. The second cone was randomly placed between 217 cm and 267 cm from the plane formed by the left CAVE wall.

For vertical trials, the cones were also separated by a distance ranging from 167 cm to 267 cm, with a mean of 217 cm. The first cone was placed at a distance of between 0 cm and 50 cm from the virtual floor and the second cone was placed at a distance of between 217 cm and 267 cm from the virtual floor, both on the plane formed by the right CAVE wall. As in Study 1, the cones did not stand upright in the vertical trials. Instead, the tops of the cones pointed parallel to the floor, toward the participant.

Width matching in the recall phase was identical to that in Study 1, except that the distance specifications were slightly modified. The first cone was placed 150 cm in front of and 168 cm to the right of the participant, a modification intended to expedite distance measurement between trials. The vertical-matching configuration was identical to that in Study 1.

Procedure. The procedure in Study 2 was very similar to that in Study 1, with the following differences.

Before beginning the experimental trials, participants completed only two practice trials (width, vertical) in the real world. These were counterbalanced in the same order as the experimental trials (i.e. if the first experimental trial was to be the width condition, the first practice trial was the width condition).

Participants completed 12 experimental trials. Each of these took place in the same virtual office scenario that was used in Study 1, but viewed from six different orientations chosen at random, such that there were no replacements for trials of a given judgement direction. There were six vertical and six width judgements, interleaved, with the starting direction counterbalanced.

After finishing all trials, participants were asked to complete a post-questionnaire regarding their experiences and problems when estimating the distances and a demographic questionnaire with questions about age, sex, height, VR experience, video game experience, and how many hours they exercised per week. Both questionnaires were slightly modified from Study 1, to more thoroughly investigate problems encountered and strategies employed by the participants.

3.2 Results: Distance Estimation Error

Because the mean actual distance was different, as described above, it was not necessary to clamp raw distance estimates as in Study 1. Two distance estimates were missing, due to experimenter error. As in Study 1, raw distance estimates were converted to percent error before analysis.

A two-factor linear mixed-effects model was constructed with percent error as a response, fixed effects for position and direction combinations (6 means), a random effect for participant, and a by-participant random slope for direction. An ANOVA using the Satterthwaite approximation for degrees of freedom showed main effects of both position, $F(2, 134.22) = 12.23$, $p < .01$, and direction, $F(1, 13.08) = 6.19$, $p = .03$. The least-squares means are plotted in Fig. 4.

Planned comparisons were performed for each of the three pairs of positions in the model. Significant differences were found between the middle and back positions, $t(134.4) = 3.5$, $p < .01$, as well as between the middle and front positions, $t(133.9) = 4.77$, $p < .01$.

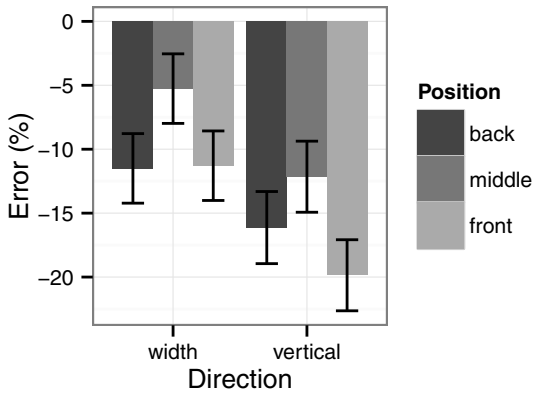


Fig. 4. Least-squares mean estimation error ± SE in Study 2, as a function of position and direction

Table 1. Comparisons between means and 0, with Bonferroni correction for multiple comparisons

Direction	Position	*df*	*t*	*p*	$p < .05$
vertical		13.1	-6.35	$< .01$	*
width		22.1	-3.81	.02	*
	back	17.8	-6.10	$< .01$	*
	middle	17.4	-3.86	.01	*
	front	17.4	-6.91	$< .01$	*
vertical	back	34.7	-5.71	$< .01$	*
width	back	32.9	-4.23	$< .01$	*
vertical	middle	32.9	-4.36	$< .01$	*
width	middle	32.9	-1.93	.75	
vertical	front	32.9	-7.13	$< .01$	*
width	front	32.9	-4.15	$< .01$	*

Post-hoc comparisons were performed on the model, comparing all eleven means (three positions, two directions, six position–direction combinations) to 0, with the Bonferroni correction for multiple comparisons. As shown in Table 1, all were significant except the width–Middle position mean.

3.3 Results: Self-reported Strategies

Questionnaire responses provided insight into strategies employed and problems encountered during the trials. Interestingly, considering the performance results above in Sec. 3.2, no participants made any mention of using cues from the physical environment to help with distance estimation. However, many participants did report using virtual cues, including furniture and windows. Additionally, participants reported using cues that might be thought of as hybrid, in that they can be thought to exist in the physical world and the virtual world depending on the level of presence. Specifically, participants reported judging horizontal distance in terms of the number of steps required to travel between the cones and judging vertical distances relative to their heights.

3.4 Discussion

As in Study 1, there appears to be a difference between horizontal and vertical distances, with vertical distances being underestimated more so than horizontal distances. Note, however, that this study did show some underestimation of horizontal spans, a finding that does not mirror those in Study 1.

Significant differences between positions reflect better performance when cones were positioned along the CAVE corners than when cones were behind the physical screen or co-planar with a single physical wall. This indicates that users were able to make use of information from the physical environment to aid in distance estimation, and this helped equally on the vertical and horizontal judgements.

However, in contrast with our hypothesis, this only happened when cones were placed at the intersection of two CAVE walls, and not when cones were placed on a single wall or the floor.

4 General Discussion and Conclusions

We have presented the results of two studies. Together they represent an investigation into the design space for virtual architecture systems. Past studies on underestimation of virtual distances have often neglected to include vertical spans. Additionally, past research has rarely been conducted in a CAVE, which limits the extent to which findings can be applied to the virtual architecture domain.

Both studies incorporated a slightly different perceptual matching paradigm from that used in previous work. While this should allow for more direct estimates of the distance spans, it may also lead to systematic bias. The pattern of differences within a given judgement direction should still be reliable, but future work should consider possible bias particularly as it may limit the ability to measure differences between width and vertical judgements, for example.

Study 1 results indicate that vertical distances are underestimated in a CAVE. This is seemingly in contrast to findings in the real world, which show an overestimation relative to horizontal spans. Note, however, that the paradigm used here is somewhat different. Past real-world findings involved relative distances, while this study was not designed to force any sort of comparison between horizontal and vertical spans. Observations and performance patterns also indicated that participants attempted to use the physical CAVE boundaries when learning distances. This motivated Study 2.

Study 2 results indicate that users can use distance cues in the physical system to help with virtual distance estimates. Surprisingly, performance improved only when the spans were along the highly-salient CAVE corners, and not simply on a CAVE surface.

These analyses indicate that users know where the physical projection surfaces are, and these studies were interpreted under the assumption that users had a low level of subjective presence in the virtual environment. It is likely, given the above conclusions above, that increased virtual presence will cause users to neglect the physical cues in favor of purely virtual judgements. This should be explored in future work. Additionally, the CAVE used in these studies is a cube shape, with right angles and easily inferred boundaries. With a different configuration, such as a curved screen, it is likely that users will not attempt to use physical cues or that they will use them ineffectively.

Acknowledgements. This project was funded by FUI in the framework of the Callisto project.

References

[1] Banks, M.S., Held, R.T., Girshick, A.R.: Perception of 3-D layout in stereo displays. Information Display 25(1), 12–16 (2009)

[2] Chrastil, E.R., Warren, W.H.: Active and passive contributions to spatial learning. Psychonomic Bulletin & Review 19, 1–23 (2012)

[3] Drascic, D., Milgram, P.: Perceptual issues in augmented reality. In: Proceedings of SPIE, vol. 2653, pp. 123–134. SPIE, Bellingham (1996)

[4] Durgin, F.H., Baird, J.A., Greenburg, M., Russell, R., Shaughnessy, K., Waymouth, S.: Who is being deceived? the experimental demands of wearing a backpack. Psychonomic Bulletin & Review 16(5), 964–969 (2009)

[5] Fink, P.W., Foo, P.S., Warren, W.H.: Obstacle avoidance during walking in real and virtual environments. ACM Transactions on Applied Perception 4(1), 2:1–2:18 (2007)

[6] Gibson, J.J.: The Ecological Approach to Visual Perception. Psychology Press (1986)

[7] Interrante, V., Anderson, L., Ries, B.: Distance perception in immersive virtual environments, revisited. In: Proceedings of IEEE Virtual Reality (IEEE VR 2006), pp. 3–11. IEEE Computer Society, Los Alamitos (2006)

[8] Jackson, R.E., Cormack, L.K.: Evolved navigation theory and the environmental vertical illusion. Evolution and Human Behavior 29, 199–304 (2008)

[9] Jackson, R.E., Willey, C.R.: Evolved navigation theory and the plateau illusion. Cognition 128, 119–126 (2013)

[10] Li, Z., Sun, E., Strawser, C.J., Spiegel, A., Klein, B., Durgin, F.H.: On the anisotropy of perceived ground extents and the interpretation of walked distance as a measure of perception. Journal of Experimental Psychology: Human Perception and Performance 39(2), 477–493 (2013)

[11] Loomis, J.M., DaSilva, J.A., Fujita, N., Fukusima, S.S.: Visual space perception and visually directed action. Journal of Experimental Psychology 18(4), 906–921 (1992)

[12] Loomis, J.M., Knapp, J.M.: Visual perception of egocentric distance in real and virtual environments. In: Hettinger, L.J., Haas, M.W. (eds.) Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues, pp. 21–46. CRC Press, Mahwah (2003)

[13] Messing, R., Durgin, F.H.: Distance perception and the visual horizon in head-mounted displays. ACM Transactions on Applied Perception 2(3), 234–250 (2005)

[14] Mohler, B.J., Creem-Regehr, S.H., Thompson, W.B.: The influence of feedback on egocentric distance judgments in real and virtual environments. In: Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization (APGV 2006), pp. 9–14. ACM, New York (2006)

[15] Proffitt, D.R., Stefanucci, J., Banton, T., Epstein, W.: The role of effort in perceiving distance. Psychological Science 14(2), 106–112 (2003)

[16] Richardson, A.R., Waller, D.: The effect of feedback training on distance estimation in virtual environments. Applied Cognitive Psychology 19, 1089–1108 (2005)

[17] Ruddle, R.A., Volkova, E., Bülthoff, H.H.: Walking improves your cognitive map in environments that are large-scale and large in extent. ACM Transactions on Computer-Human Interaction 18(2), 10:1–10:20 (2011)

[18] Stefanucci, J.K., Proffitt, D.R.: The roles of altitude and fear in the perception of height. Journal of Experimental Psychology: Human Perception and Performance 35(2), 424–438 (2009)

[19] Sun, H.J., Campos, J.L., Chan, G.S.W.: Multisensory integration in the estimation of relative path length. Experimental Brain Research 154, 246–254 (2004)
[20] Thompson, W.B., Willemsen, P., Gooch, A.A., Creem-Regehr, S.H., Loomis, J.M., Beall, A.C.: Does the quality of the computer graphics matter when judging distances in visually immersive environments? Presence: Teleoperators and Virtual Environments 13(5), 560–571 (2004)
[21] Todorović, D.: The effect of the observer vantage point on perceived distortions in linear perspective images. Attention, Perception, and Psychophysics 71(1), 183–193 (2009)
[22] Waller, D., Loomis, J.M., Haun, D.B.M.: Body-based senses enhance knowledge of directions in large-scale environments. Psychonomic Bulletin & Review 11(1), 157–163 (2004)
[23] Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments 7(3), 225–240 (1998)
[24] Witt, J.K., Proffitt, D.R., Epstein, W.: Perceiving distance: A role of effort and intent. Perception 33, 577–590 (2004)
[25] Yang, T.L., Dixon, M.W., Proffitt, D.R.: Seeing big things: Overestimation of heights is greater for real objects than for objects in pictures. Perception 28, 445–467 (1999)