

Depth Perception in Virtual Environment: The Effects of Immersive System and Freedom of Movement

Adrian K.T. Ng^(✉), Leith K.Y. Chan, and Henry Y.K. Lau

Department of Industrial and Manufacturing Systems Engineering,
The University of Hong Kong, Hong Kong, China
{adriang, lkychan, hyklau}@hku.hk

Abstract. Concerns over the use of virtual reality (VR) systems in experimental psychological research exist. It is found that human egocentric depth perception in a virtual environment (VE) has significant errors compared to real physical environment. It is hypothesized that due to the presence of a human body as a size reference in a mixed reality CAVE-like system, the accuracy of depth estimation will improve. The second hypothesis proposes that when a participant is allowed to move around the VE, motion parallax will supplement the depth perception ability. Results showed that the features of an immersive system did not aid the estimation. Around 40 % underestimation of actual distance was observed above 15 m. By using a 3D-military jet instead of 2D-wall as the judgment object, a significant improvement in the accuracy is found. Pictorial cues were hence, suggested as the improvement basis for next part of the study.

Keywords: Depth perception · Distance estimation · imseCAVE · Perception · Virtual environment · Virtual reality

1 Introduction

The basic elements of virtual reality (VR) systems as defined by Burdea and Coiffet, “Immersion-Interaction-Imagination” (I^3) [1], are well developed as the technology of computer graphics and 3D displays advances. For example, at the Department of Industrial and Manufacturing Systems Engineering at the University of Hong Kong, an immersive, interactive VR system was developed called the imseCAVE based on the concept of the Cave Automatic Virtual Environment (CAVETM) [2]. The system with its automatic trackers and 3D image rendering can provide physical immersion projection to users. As for mental immersion, the success of such immersion depends on the design of the contents and the involvement of the users [3]. With the hardware and software readied, interaction and imagination in a virtual environment (VE) are endless upon creation of applications. MagicPad and MagicPen, build on top of imseCAVE, is a pen and paper like tangible user interface that allows users to interact and create things in VE, for example, drawing 2D and 3D objects, and playing physics game [4].

With the basic elements of immersive systems developed, many researches have started to look at practical applications of VR. In Psychology, VR tools have been used more frequently in psychological experiments and clinical applications for its ability to

present stimulus in environment with greater control over the variables [5]. For example, psychologists used VE to test the likelihood for people to help under different situations [6]. Variables of virtual bystanders and the virtual person that need helps can be experimentally manipulated easily. The bystander effect, a sophisticated high-level social behavior, is observed. VE is in fact, a useful interface compromising between experimental control and ecological validity. In Cognitive Science, VE technology could also be viewed as a unique asset, which provides the possibility of creating and presenting dynamic objects and environments for precise measurement of human cognition and interaction [7].

1.1 Depth Perception Error and Its Importance

Despite the usefulness of VR, some raised concerns over the use of VR systems as the tool in experimental psychological research. Increased attention had been brought to the aspect of human spatial perception. It is found that egocentric depth estimation (i.e. subjective distance perception to an object from the human subject) in VE have lower accuracy compare to real physical environment [8, 9].

Such error is worth investigating as depth perception is an important visual component for perceiving objects, navigating, reaching, and performing size judgments, to name but a few [10]. For experimental researchers, it is important to ensure that the VR systems could be used as an experimental tool in all aspects without any other adjustment. A useful experimental environment should provide an accurate measurement while containing minimum cues. In the above case, a VE should provide viewer an accurate depth perception, the ability to perceive object in three dimensions and the distance of the object; while containing minimum visual depth cues — anchors that assists perception [11].

If such error between VE and real environment differs significantly, researches may not be confident in using VE for many types of cognitive experiments, in particular those that require many spatial awareness, such as way-finding and visual perception. Researches may then investigate on the effects related to their particular cognitive aspect and whether there are any alleviation or improvement possible related to those kinds of cognition.

For Cognitive Scientists, such error is also worth investigating as such error maybe due to a different pathway of cognitive processes in human brain being simulated by VE, different from normal environment. The findings may help improve our understanding on cognition.

From Industrial Engineering (IE) perspective, simulation and visualization are the major uses of VR systems. For example, a container terminal quay crane simulation system in the imseCAVE allows users to simulate loading and unloading of containers [2]. It requires users' depth estimation to complete the tasks. Although previous literatures have not suggested any user-experience problem in simulation and visualization in VE reported attributes to the perception error, the user experience may still be able to improve if such error is improved or eliminated. Users will not know it is a problem until the error is adjusted and that they can compare the difference. To improve in the

system perspective, the design of VE content and rendering matrix of VR software could be adjusted to be more effective and accurate [12].

The direction of the research is to explore the effects and possible improvements to the accuracy of human egocentric depth perception under CAVE-like immersive system in experimental psychology and IE perspective.

2 Background and Related Works

2.1 Overview

Since there are quite a number of previous researches on depth perception, Lin and Woldegiorgis, and Renner et al. carried out a comprehensive literature review respectively [8, 9]. Their analysis reveals that while egocentric distance estimation in real world is about 94 % accurate, it drops to around 80 % in VE on average, i.e. a 20 % underestimation or compression.

2.2 Comparing Immersive System with Other VR Systems

Although it is shown that the accuracy level in head-mounted displays (HMD) (73 %) and large screens (74 %) are similar [9], the underlying technologies, experiences, mechanism, and psychophysics are different. This paper does not wish to compare them exhaustively but to point out some points related to depth perception. First, a HMD mounts two-lens display in front of the eyes directly. The focus, screen distance and other psychophysical parameters of a HMD are completely different from an immersive system.

Second, as Milgram et al. proposed, the VR technologies fall into a reality-virtuality (RV) continuum, with real environment in one side and virtual environment on the other [13]. A normal HMD displays a complete virtual environment through the two-lens display, while a CAVE-like system, which is an immersive virtual reality system, displays a mixed reality (MR) through the 4-sides stereoscopic displays with the human body in the system. The person in the immersive system sees the virtual world and his body at the same time. This should provide body reference for visual perception of size.

It is true that as error occurs similarly in all hardware systems, it can be ruled out that such error is only limited to a certain system. However, owing to the differences between systems, separated efforts should be made to large screens and immersive systems in depth perception instead of merging it ambitiously with HMD as stereoscopic displays. In the following, immersive system will be focused.

2.3 Distance Perception

Distance estimation tasks require three essential mental processes in the viewer aspect: Perceiving, Analyzing and Reporting [8]. First, a person perceives and analyses to form a perception. A person employs his vision to perceive the distance of the target object from one's location by mental or physical reference point or object. The information is

then processed with other considerations and strategies, for an adjustment to enhance the accuracy. Some previous studies had shown that many factors could affect a person's perception. For example, Naciri et al. found that subjects' performances are affected by the reliance of depth cues. Those that rely on different types of depth cues instead of relying heavily on the apparent size, only one type of depth cue, had more accurate estimation [14]. This is consistent with another study where underestimation was both found in poor and rich cue conditions but less in rich cue condition [15].

The perceptual spaces are generally divided into three egocentric regions [16]. Personal space which is immediately surrounding the observer is defined within 2 m. Action space where human could quickly walk to and interact with, is defined as 2 m to 30 m. Binocular disparity and motion parallax can be used in this space with movement in immersive system or in physical reality [17]. Although motion parallax is an important depth cue in depth perception in physical reality [18], there is no evidence showing that it could aid depth estimation in immersive system.

Vista space is defined as beyond 30 m where binocular disparity and motion perspective are not in use. Depth perception could only be derived from pictorial cues such as relative size and occlusion. The depth estimation error in this space has not been studied in immersive system.

2.4 Measurement Method

In the third phase, reporting, the person needs to decide how far the target is placed as obtained from the previous perception stage and tells the experimenter the estimation. There are a wide variety of protocols to measure a person's depth estimation. In researches related to VE, there are mainly 3 types: verbal estimation, visually imagined action, and perceptual matching [8, 9, 19].

Verbal estimation refers to directly stating the depth estimation in metric units such as meter. Visually imagined action refers to having the subject views the object and imagines walking to the object. Imagination time is recorded with their usual rate of walk for a depth judgment. Perceptual matching refers to having the subject to judge the distances with manipulating or judging an existing object in the screen.

Some suggested that verbal estimation is subjected to bias and noise [20]. However, as previous studies had shown that subjects are able to process depth perception with spatial relation well using perceptual matching [21], the problem of underestimation is not on the VR system itself but the bias and noise. Verbal estimation or other related numerical methods should be used to quantify the bias and noise. Klein et al. compared timed imagined walking, verbal estimation, and triangulated blind walking. They found very similar result in depth judgment in the three methods in both real physical environment and virtual environment [19].

2.5 Improvement Attempts

Some possible alleviative measures are being tested from various aspects. Renner et al. and Ponto et al. proposed that current standard stereo-based physical measurements of eye position are not precise for proper viewing parameters [22, 23]. As an improvement

attempt, users' inter-pupillary distance (IPD) is inputted into a geometric model to predict perceptual errors. Errors can then be inversely calibrated to provide a more accurate image. Improvements were found but such alleviative measures are complex and not optimum with individual differences.

From human experience aspect, manipulations in HMD settings related to learning and familiarization are found to improve the accuracy [21]. The interaction task designed to provide learning and familiarization process significantly reduces users' error to nearly veridical. However, such effects were not found in CAVE-like systems when tested with subject experience and environmental learning [15].

2.6 Aim of the Study

The aim of the study is to first explore the effects of immersive system for the accuracy of human egocentric depth perception under several environments. It is hypothesized that human depth estimation in VE is better when the person is in immersive system than viewing from a single screen stereoscopic display due to the presence of the human body in the MR, and such perception in action space will be even better when participants are allowed to move around since motion parallax should aid the estimation. Human depth estimation error in vista space and the effects of 2D and 3D object will also be investigated. In the later stage of the study, based on the results, measurement method changes and improvement attempts will be proposed and tested.

3 Method

As stated above, this study aims to measure human perception error of egocentric distance estimation in three different environments (single screen, imseCAVE without freedom of movement, imseCAVE with freedom of movement) (Figs. 1 and 2), using two different types of objects (2D- wall and 3D- military jet). Distances in action space (2–30 m) and vista space (> 30 m) were studied.



Fig. 1. Single screen stereoscopic display

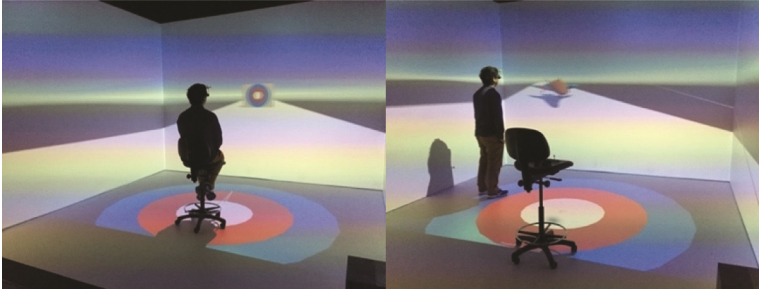


Fig. 2. imseCAVE (a) without freedom of movement (b) with freedom of movement

3.1 Experimental Setup

The experiment is carried in the imseCAVE, a VR system that facilitates the creation of an interactive immersive three-dimensional environment [2, 4]. It is composed of three walls and a floor. The dimension is 4 m deep, by 3 m across and 3 m tall. All projectors use active 3D running at 120 Hz with XGA (1024×768 pixel) resolution. Images are displayed at 120 Hz, alternating between left and right image to create the three-dimensional effect. Stereo image can be perceived with shutter glasses. The optical tracking system is used to track user position with 8 infrared cameras mounted at the ceiling to provide coverage. Markers are mounted on the 3D shuttle glass and handheld controller to enable the tracking systems to measure and calculate the 3D position and orientation of the user and the controller. The images could then be rendered to the viewer realistically.

For the first environment, only the front side is used as stereoscopic display. For the second and third environment, the full system will be used. The visual content is generated using Unity and MiddleVR, where dimension of the objects can be created and displayed in the system accurately.

The type of virtual environment in the study is chosen to be open space, which included a grey floor and a blue sky. The perception of infinity horizon is created by the impression of a horizon with vanishing point induced by the environment; while linear perspective was not available as a secondary depth cue. The texture of the background environment is chosen not to provide any texture gradient information. No metric aid and additional background object were provided.

The experimental condition is designed to be as simple as possible due to two reasons: First, as Armbrüster had tested in unicoloured background by verbal estimation with the type of environment (no space, open space vs. closed space) [11]. There were no significant differences between the environments. Second, the current study tries to represent a practical experimental environment with minimum visual depth cue that attempts to obtain an accurate measure in depth estimation. Hence, such open space with simple cues is selected.

The two different types of objects are drawn with clear and sharp features to be viewed from far away. The two-dimensional object is a wall with circles drawn to resemble a shooting target; while the three-dimensional object is a military jet colored in

orange and blue. Both items are scaled for consistency. Both objects are not objects that could be seen in daily life, such manipulation prevented subjects to guess the object based on their daily experiences by the size.

3.2 Experimental Variables

The independent variables are the environment, type of object and distance of the target object. Three different environments were used: single screen stereoscopic display, immersive virtual reality system (CAVE) with and without freedom of movement. The two types of object were used: 2D flat wall and 3D military jet model. The primary dependent variable was the reported verbal depth estimation in meter by the participants. A percentage of error was then calculated by normalizing the result:

$$\% \text{ of Error} = \frac{\text{reported distance} - \text{actual distance}}{\text{actual distance}} \times 100\% \quad (1)$$

As human is found to have around 6 % of error in distance estimation in real world situation [8], an allowance is provided. An error near 0 is veridical, while a value > .1 shows overestimation, and a value < -.1 shows underestimation. The variables are analyzed in the Result session.

3.3 Participants

The research study recruited 40 subjects from the university community of University of Hong Kong; aged between 18 and 24 (with a mean age of 20.9); 20 were male and 20 were female. They volunteered to experience a virtual environment and were not given any payment, credit, or other compensation for their participation. We screened the participants for 20/20 vision of both eyes, either in natural or corrected, and usual normal stereopsis experience in 3D environment. Subjects that did not pass will not be invited to perform the test any further.

3.4 Experimental Task

Before collecting data, participants gave their consent on a form that explained the purpose of the study, and the confidentiality of individual's data sets are ensured. On the form, the subjects were advised that they might experience mild fatigue and discomforts such as motion and cyber-sickness during the procedure. Such fatigue and/or discomforts associated with the environment rendered were kept to a minimum and that should they experience any, they are free to take short breaks or quit at any time.

The visual acuity and stereopsis ability of subjects are checked so as to screen out those who normally do not perceive depth accurately even in physical reality. The subjects are then given a chance to familiar with a rich content VR environment and check if they perceive the rendered 3D content normally. The subjects were then listened to verbal description and demonstration of the task given by the experimenter in the test environment. Subjects were told that they would estimate distances in meters, where the

numbers are random integer. They were also explicitly told that the numbers may not be multiple of 2 or 5. Subjects were informed and shown an object of egocentric distance of 2 m (the distance to wall of the immersive virtual reality system from the seat) as a dimensional reference of the system.

In the first two environments (single screen, imseCAVE without freedom of movement) (Figs. 1 and 2a), subjects are invited to sit in a fixed chair in the middle of the imseCAVE system (2 m from the front screen). The eye height is adjusted to be the same while standing and sitting. The subjects are required to give numerical verbal depth estimation of the object from their seat to the object in front of them in turn. In the third environment (imseCAVE with freedom of movement) (Fig. 2b), subjects are allowed to freely walk around the imseCAVE to view the virtual object and give numerical verbal depth estimation of the object from a designated location to the object. Two types of objects of random sequence were tested as independent variables in each environment. In each trial, the object is displayed in 7 distances ranging from 3 to 60 m. The distances presentation order is randomly permuted given that the same distances do not present twice in a row. After the total of 7 (distance) \times 3 (environment) \times 2 (object) = 42 data point collection, a debriefing was given. Object at some distances were displayed while revealing the true egocentric distances.

4 Results

Every participant made 42 estimations ($7 \times 3 \times 2$). On average, 29.98 ($SD = 11.1$) underestimated ($< .1$), 5.87 ($SD = 3.6$) correctly estimated (between $-.1$ and $.1$), and 6.15 ($SD = 9.2$) overestimated ($> .1$) the distances over all the conditions and distances. Paired t -test results show that there is significant difference between underestimations and correct estimations ($t_{39} = 11.135, p < .001$), and underestimations and overestimations ($t_{39} = 7.503, p < .001$), while correct estimations and overestimations do not differ significantly ($t_{39} = -.206, p < .838$).

The numerical estimations were computed into percentage of error as stated in previous part for comparison. The results of a repeated measure ANOVA with factors of distance (7), environment (3) and object (2), and sex as between-subject factor revealed only a significant main effect for factor of distance ($F[6,228] = 79.801, p < .001$) and type of objects ($F[1,38] = 18.382, p < .001$).

Pairwise comparisons (t -tests) between the zero error and mean percentage error in estimated distances over all conditions show that the error in estimations of 3–7 m were not significant (see Table 1). Yet, for actual distance range above 15 m, the errors are significant. The signs and values indicate an underestimation of 34.6–44.5 % in actual range of 15–56 m (see Fig. 3). For the two types of object, 2D-wall yields $-.351\%$ ($SD = .270$), while 3D-military jet yields $-.214\%$ ($SD = .400$) error on average of estimated distance in all condition. Paired t -test results show that there is significant difference between them ($t_{39} = -4.323, p < .001$).

Table 1. Mean percentage error of estimated distance in all conditions

Actual distance range (meters)	Mean % error of estimated distance	Std. Deviation	t-Value	Sig.
3–4	0.063	0.216	1.836	.074
5–7	-0.034	0.318	-0.680	.500
15–17	-0.346	0.304	-7.201	.005
20–22	-0.393	0.357	-6.954	.005
30–32	-0.439	0.352	-7.886	.005
41–43	-0.384	0.446	-5.452	.005
55–57	-0.445	0.431	-6.533	.005

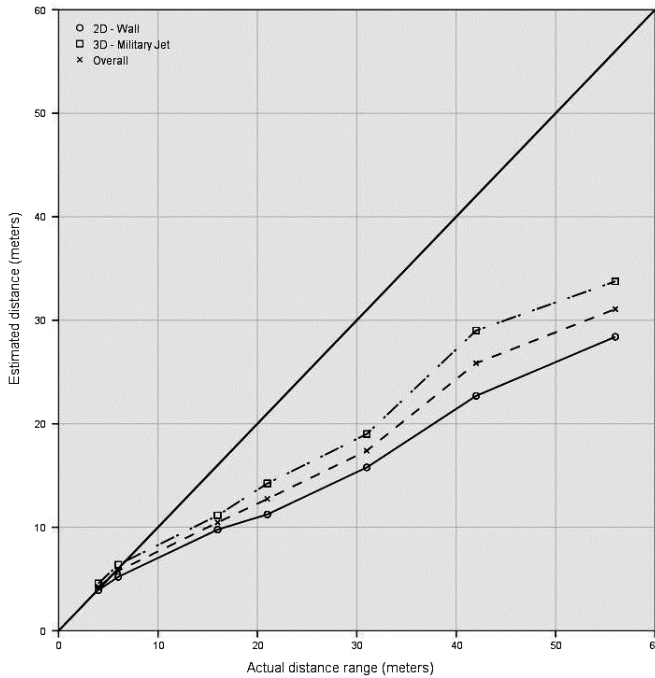


Fig. 3. Distance estimation against actual distance range

5 Discussion

The first part of the study was conducted to investigate the effects of immersive system on the accuracy of human egocentric depth perception by a simple VE that simulates a usual psychological experimental environment. Overall, virtual distances were underestimated by participants.

CAVE-like immersive systems provide human body reference for size and freedom of movement for motion parallax. The hypothesis that these unique features may contribute to the depth perception and improve the accuracy of distance estimation was

tested. The three environments had no differential effect on accuracy of depth perception. The features of immersive systems are not essential elements to depth estimation. Yet, since immersion is an important element in VR (the I³), it should still be preserved when VR is used as an experimental tool to provide a total VR experience to the subjects.

Depth perception estimation in immersive system in such simple virtual environment was found to be inaccurate, including vista space. Researchers should take extra care when using VR systems as a cognitive experimental tool, especially if the tasks involve perception or navigation. One research direction could investigate if our visual system activates a different set of visual pathway in processing VR images.

The object types and distances in vista spaces were added as an exploration as there are no previous attempts. Compared to 2D objects and far away objects, 3D objects and closer objects were found to significantly provide better accuracies of depth estimation. It is speculated that this is due to more pictorial cues those objects provided. The effects of the types of object to the accuracy on depth perception could be investigated in a practical level. Since different kinds of objects were visualized in imseCAVE for virtual prototyping or simulation, it would be practical to investigate if such problem exists in the practical usage. The next step of the study will also, focus on the pictorial cues as an improvement basis. Depth cues should be first isolated individually to evaluate the effect to depth perception. Then, combinations should be tried to produce accurate estimation using the minimum cues. Hopefully, a clear combination of depth cues that are essential to depth perception could be found. Improvements on the content presented, for example, by providing more depth cues of a specific kind, should be a more ideal experimental operation than modifying hardware based on individual differences.

References

1. Burdea, G., Coiffet, P.: *Virtual Reality Technology*, 2nd edn. Wiley, New Jersey (2003)
2. Chan, L.K.Y., Lau, H.Y.K.: A cost effective virtual reality system for simulating logistics operations. *Int. J. Logist. SCM. Syst.* **6**(1), 71–76 (2012)
3. Sherman, W.R., Craig, A.B.: *Understanding Virtual Reality*. Morgan Kaufmann Publishers, New York (2003)
4. Chan, L.K.Y.: *MagicPad: A Spatial Human-System Interface for Immersive Virtual Environment* (2015)
5. Wilson, C.J., Soranzo, A.: The use of virtual reality in psychology: a case study in visual perception. *Comput. Math. Methods Med.* **2015**, 1–7 (2015)
6. Kozlov, M.D., Johansen, M.K.: Real behavior in virtual environments: psychology experiments in a simple virtual-reality paradigm using video games. *Cyberpsychol. Behav. Soc. Netw.* **13**(6), 711–714 (2010)
7. Rizzo, A.A., Bowerly, T., Buckwalter, J.G., Schultheis, M.T., Matheis, R., Shahabi, C., Sharifzadeh, M.: Virtual environments for the assessment of attention and memory processes: the virtual classroom and office. In: Sharkey, P., Sik Lányi, C., Standen, P.J. (eds.) *4th International Conference on Disability, Virtual Reality and Associated Technology, ICDVRAT 2002*, pp. 3–11. University of Reading, England (2002)
8. Lin, C.J., Woldegiorgis, B.H.: Interaction and visual performance in stereoscopic displays: A review. *J. Soc. Inf Disp.* **23**(7), 319–332 (2015)

9. Renner, R.S., Velichkovsky, B.M., Helmert, J.R.: The perception of egocentric distances in virtual environments- A review. *ACM Comput. Surv.* **46**(2), 23 (2013)
10. Walk, R.D., Gibson, E.J.: A comparative and analytical study of visual depth perception. *Psychol. Monogr.* **75**(15), 1 (1961)
11. Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., Fimm, B.: Depth perception in virtual reality: distance estimations in peri-and extrapersonal space. *Cyberpsychol. Behav.* **11**(1), 9–15 (2008)
12. Saracini, C.: *Spatial cognition in Virtual Environments* (2011)
13. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: a class of displays on the reality-virtuality continuum. In: *Telemanipulator and Telepresence Technologies*. SPIE, vol. 2351, pp. 282–292 (1995)
14. Naceri, A., Chellali, R., Dionnet, F., Toma, S.: Depth perception within virtual environments: comparison between two display technologies. *Int. J. Adv. Intel. Syst.* **3**, 51–64 (2010)
15. Murgia, A., Sharkey, P.M.: Estimation of distances in virtual environments using size constancy. *Int. J. Virtual. Real.* **8**(1), 67–74 (2009)
16. Cutting, J.E.: How the eye measures reality and virtual reality. *Behav. Res. Methods. Instrum. Comput.* **29**(1), 27–36 (1997)
17. Luo, X., Kenyon, R., Kamper, D., Sandin, D., DeFanti, T.: The effects of scene complexity, stereovision, and motion parallax on size constancy in a virtual environment. In: *2007 Virtual Reality Conference, VR 2007*, pp. 59–66. IEEE (2007)
18. Rogers, B., Graham, M.: Motion parallax as an independent cue for depth perception. *Perception* **8**(2), 125–134 (1979)
19. Klein, E., Swan, J.E., Schmidt, G.S., Livingston, M., Stadt, O.G.: Measurement protocols for medium-field distance perception in large-screen immersive displays. In: *2009 IEEE Virtual Reality Conference*, pp. 107–113. IEEE Computer Society (2009)
20. Thompson, W.K., Willemsen, P., Gooch, 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-Teleop. Virt.* **13**(5), 560–571 (2004)
21. Richardson, A.R., Waller, D.: Interaction with an immersive virtual environment corrects users' distance estimates. *Hum. Factors* **49**(3), 507–517 (2007)
22. Renner, R.S., Velichkovsky, B.M., Helmert, J.R., Stelzer, R.H.: Measuring interpupillary distance might not be enough. In: *ACM Symposium on Applied Perception*, pp. 130–130. ACM (2013)
23. Ponto, K., Gleicher, M., Radwin, R.G., Shin, H.J.: Perceptual calibration for immersive display environments. *IEEE Trans. Vis. Comput. Graph.* **19**(4), 691–700 (2013)