Evaluation of Field of View Calibration Techniques for Head-mounted Displays and Effects on Distance Estimation

Carolin Walter*, Gerd Bruder†, Benjamin Bolte*, Frank Steinicke†

*Department of Computer Science   †Department of Psychology II
University of Münster
{carolin.walter|b.bolte}@uni-muenster.de

†Immersive Media Group
Department of Computer Science
University of Würzburg
{gerd.bruder|frank.steinicke}@uni-wuerzburg.de

Abstract: Users in immersive virtual environments (VEs) with head-mounted displays (HMDs) often perceive compressed egocentric distances compared to the real world. Although various factors have been identified that affect egocentric distance perception, the main factors for this effect still remain unknown. Recent experiments suggest that miscalibration of the field of view (FOV) have a strong effect on distance perception. Unfortunately, it is not trivial to correctly set the FOV for a given HMD in such a way that the scene is rendered without mini- or magnification. In this paper we test two calibration techniques based on visual or visual-proprioceptive estimation tasks to determine the FOV of an immersive HMD and analyze the effect of the resulting FOVs on distance estimation in two experiments: (i) blind walking for long distances and (ii) blind grasping for arm-reach distances. We found an impact of the FOV on distance judgments, but calibrating the FOVs was not sufficient to compensate for distance underestimation effects.

Keywords: Head-mounted displays, field of view calibration, distance estimation

1 Introduction

Head-mounted displays (HMDs) coupled with tracking systems for measuring a user’s position and orientation in a virtual reality (VR) setup can provide users with a realistic spatial impression of computer-generated immersive virtual environments (IVEs). Mapping tracked movements of a HMD to motions of a virtual camera for rendering three-dimensional content allows users to explore IVEs from an egocentric perspective. Such freedom of exploration and natural movements have great potential as an enabling technology in many domains, such as architecture or engineering. For such domains it is essential to provide HMD users with an accurate spatial impression of virtual content. However, researchers often observe a discrepancy between spatial judgments in IVEs compared to the real world.
For example, experiments revealed that subjects often judge distances to be compressed in IVEs [IRLA07, LK03, TWG+04]. Over the last decade, a large body of literature has focused on identifying contributing factors for such misperception of three-dimensional content in HMD environments. Some factors were identified focusing on the appearance of virtual content, e.g., the importance of realistic rendering [PRI+09], and the impact of familiar cues on distance estimation [IRLA07].

Recently, experiments by Kuhl et al. [KTCR09] and Steinicke et al. [SBK+09] show that accurate calibration of the visual rendering process to the display hardware can have an essential impact on a user’s distance estimation. Researchers analyzed the impact of different parameters of the visual rendering process on distance estimation, such as the interpupillary distance [WGTCR08] or optical distortions [KTCR09]. A main contributing factor for distance estimation was found in the field of view (FOV) used to specify perspective view frustums in HMD environments [KTCR09, KBB+12]. In this context, the display’s field of view (DFOV) refers to the horizontal and vertical angles subtended by the display units located in front of the user’s eyes. The term geometric field of view (GFOV) describes the horizontal and vertical opening angles of the perspective view frustum, which is used to render the three-dimensional content on the HMD [SBK+09].

Under optimal viewing conditions, the VE is rendered with a GFOV that matches the DFOV, i.e., the optical rays between the user’s eyes and the VE are not refracted by an increased (GFOV > DFOV) or decreased (GFOV < DFOV) opening angle of the virtual frustum. However, in order to apply the correct GFOV to the rendering system it is essential to determine the DFOV of the HMD, which may deviate from the values specified by the manufacturer. Unfortunately, measuring the DFOV of immersive HMDs is not trivial, since non-see-through HMDs do not provide direct visual comparisons between the real world and overlaid virtual objects as with see-through HMDs [AB94].

In this paper we test two calibration techniques to determine the DFOV of an immersive HMD and analyze the effects of the resulting DFOVs on distance estimation. Section 2 provides background information on geometric rendering parameters. Section 3 describes the two considered calibration techniques. In Section 4 we describe the conducted distance estimation experiments. Section 5 provides a general discussion of the results. Section 6 concludes the paper and gives an overview of future work.

2 Background

Different computational models for transformations required to align a perspective view frustum with the three-dimensional volume in space that is visible to a HMD user through the optics of the display units have been proposed [RH94]. In particular, a computer graphics view frustum can be specified by near and far plane distances in view space, as well as the horizontal and vertical opening angles of a symmetric or asymmetric truncated rectangular pyramid [RH94]. Most HMD manufacturers provide the display’s nominal field of view as angle over the diagonal of the visual field. Assuming a symmetric view volume with known aspect ratio of the virtual image plane, the screen diagonal can be transformed into the
vertical opening angle using the approach described by Steinicke et al. [SBK+09]. If the manufacturers provide information about the accommodation distance to the focal plane of the display units, the vertical opening angle of the display units can be derived using the model described by Robinett and Holloway [RH94].

Usually, the GFOV is specified using the nominal values provided by the manufacturers. If these values are correct, the three-dimensional virtual space is accurately rendered along the user’s optical rays. However, if the nominal values provided by the manufacturers deviate from the head-mounted display’s actual DFOV, this either results in an increased or decreased GFOV relative to the DFOV. This relative difference can be described via geometric field of view gains $g_F \in \mathbb{R}^+$ as $GFOV = g_F \cdot DFOV$. Applying gains $g_F < 1$ can provide a HMD user with a magnifying glass in the virtual world (see Figure 1(a)). Applying gains $g_F > 1$ compresses portions of the virtual world in the user’s view, which results in a minification effect (see Figure 1(b)) [KTCR09, PKB07].

It is often assumed that the human visual system performs optimal in an immersive VE if provided with the exact same cues as in a comparable real-world situation. This assumption was confirmed for familiar size cues in a virtual replica of a physical laboratory room, in which subjects showed the least amount of distance underestimation if the size of the virtual room was exactly the same as in the real world [IRLA07]. In this context, we use two calibration techniques to determine the DFOV of a given HMD, and analyze if applying the determined DFOV to the GFOV reduces the distance underestimation effects.

## 3 Calibrating the Field of View

Different calibration methods can be applied to determine the DFOV of a HMD [MT99]. However, since immersive HMDs do not provide direct visual comparisons between real and virtual views, only few quantitative calibration techniques have been proposed. For instance, Gilson et al. [GFG09] proposed adapting camera-based calibration methods of see-through displays to immersive HMDs, which can provide accurate DFOV values, modulo differences caused by discrepancies of human and camera optics. In the following, we apply two calibration methods that try to reduce such error sources by direct calibration of HMD optics with test subjects.
Figure 2: Calibrations: (a) a subject compares the size of a horizontal stripe in the real world with (b) a virtual stripe using technique C1, and (c) a subject points to the visually perceived angle of (d) a vertical pole displayed on the HMD using technique C2.

3.1 Participants

4 female and 12 male (age 21-55, $\bar{\phi}$: 28.2) subjects participated in both calibrations. Subjects were students or department members of computer science or mathematics. All had normal or corrected to normal vision. 15 of the subjects had experience with HMD setups. We counterbalanced the order of the calibration experiments. The total time per subject including instructions, experiment, breaks, and debriefing was 1 hour.

3.2 Calibration Setup

We performed both calibration experiments in a 10m×7m laboratory room with a Rockwell Collins ProView SR80 HMD (1280 × 1024 @ 60Hz, 80° nominal diagonal FOV) for the visual stimulus presentation. The visual stimuli consisted of a virtual replica of the real laboratory (see Figure 2) and was rendered using the IrrLicht engine and our own software with 60 frames per second on an Intel computer (Core i7 processors, 8GB RAM, nVidia Quadro FX 4800). The subjects received instructions on slides presented on the HMD. In both experiments, subjects faced a wall of the laboratory with their heads fixed at a distance of 6.6m from the wall (see Figure 2).
Figure 3: Pooled results for the left (blue) and right (red) display units for (a) psychophysical calibration process with applied gains on the horizontal axis and pooled estimates for “minified” estimation on the vertical axis, and (b) pointing calibration process with visually presented angles on the horizontal axis and measured pointed angles on the vertical axis (light plots correspond to pooled results from the virtual, and the dark plots from the real world condition).

3.3 Calibration 1 (C1): Psychophysical Calibration

The calibration technique proposed by Steinicke et al. [SBK⁺09] makes use of repeated raising and lowering the HMD to compare visual stimuli from the real with the virtual world, and incorporates a psychophysical two-alternative forced-choice task (2-AFCT) to accurately measure the relation between stimulus intensity and perception reported by a human observer.

3.3.1 Material and Methods of C1

The calibration process is based on the method of constant stimuli; the presented stimuli are randomized and uniformly distributed between trials. Visual stimuli presented on the HMD were generated with gains $g_F$, ranging between 0.90 and 1.10 in steps of 0.02, to the vertical opening angle of the camera frustum, computing the horizontal opening angle with respect to the display ratio. We computed the vertical FOV from the nominal 80° diagonal FOV provided by the manufacturers as base for applying relative GFOV gains. We presented the gains in randomized order, each gain occurring 5 times. Figure 2(a) shows a subject in the laboratory, who compares a modeled horizontal stripe of 0.3m × 0.03m size in the virtual scene with a corresponding stripe mounted in 0.7m distance from the subject in the real world by repeatedly raising and lowering the HMD. We separately calibrated the left and right display units of the HMD, i.e., we provided subjects with an eye patch in the real world and blanked out the display of the HMD. This process deviates slightly from the setup used by Steinicke et al. [SBK⁺09], but follows the same general approach.
After comparing the real and virtual visual stimuli, the subject had to answer the question “Do you think the virtual world is minified or magnified?” Subjects were trained in 5 training trials as to what “minified” and “magnified” refer to in this context. At the point of subjective equality (PSE) subjects perceive the real and virtual stimulus as identical and will not be able to distinguish between “minified” and “magnified”, i.e., they will respond “minified” in 50% of the trials on average. As the gain decreases or increases from this value the ability of subjects to detect a difference between both stimuli increases, resulting in a psychometric curve for the discrimination performance.

3.3.2 Results and Discussion of C1

Figure 3(a) shows the mean probability for a subject’s estimation that the virtual stripe was “minified” on the vertical axis, with the range of tested gains on the horizontal axis. The solid lines show the fitted psychometric functions of the form \( f(x) = \frac{1}{1+e^{-ax+b}} \) with \( a, b \in \mathbb{R} \). The blue and red lines correspond to the pooled results for the left and right display unit, respectively. The error bars show the standard errors. The PSE for the left eye is \( g_F = 1.0023 \), and for the right eye \( g_F = 0.9924 \). The results show that the diagonal GFOV judged as correct by the participants is close to the nominal 80° field of view specified by the manufacturers, i.e., 80.184° for the left eye and 79.392° for the right eye.

Comparing the results with those reported by Steinicke et al. [SBK+09], reveal that subjects were not as good at discriminating between virtual and real stripes as in the original experiment [SBK+09]. This may in part be caused by the different setup, i.e., the separation between left and right eyes and using non-expert users to perform the calibration.

3.4 Calibration 2 (C2): Calibration by Pointing

The calibration technique by Ellis and Nemire [EN93, NE93] is based on proprioceptive-visual registration of line of sight angles in a real and virtual condition. Therefore, vertical poles are displayed in the real and virtual world, with the instruction for subjects to point in their respective direction without providing subjects any feedback about their accuracy. The measured differences in visually perceived and proprioceptively responded angles can be interpreted as offset from the visually applied GFOV to the actual DFOV. The real and virtual conditions are used to cancel out the effects of differences in pointing accuracy between subjects, i.e., to reduce systematic errors in pointing direction. We applied the nominal 80° FOV of the HMD provided by the manufacturer to the virtual camera in this experiment, allowing us to compute a relative difference to the actual DFOV.

3.4.1 Material and Methods of C2

In both real and virtual conditions, a subject’s head was fixed in the real world such that the eye level was adjusted closely above a planar surface, providing an unobstructed view over the surface, but blocking the subject’s view below that surface (see Figure 2(c)). A
matching view as in the real world was displayed on the HMD in the virtual condition (see Figure 2(d)). In both conditions, a target (vertical 0.01m$\times$0.06m stripe) was presented at a distance of 0.7m. Between trials, we shifted the target to the left or right, resulting in line-of-sight angles between $-30^\circ$ and $30^\circ$ in steps of 6$^\circ$. The subject’s task was to point with a laser pointer to the targets using their dominant hand. We captured the points to which subjects aimed the laser pointer with a WorldViz PPT X8 tracking system. We tested each of the considered angles 5 times. The task was performed for both eyes separately.

Our calibration method varies from the original experiment [EN93] in that subjects judged the direction using a laser pointer, instead of moving a mechanical tracker. Moreover, we presented subjects with targets in a horizontal rectangular plane for both eyes separately, rather than in a horizontal hemicircular plane centered only on the right eye of subjects.

3.4.2 Results and Discussion of C2

Figure 3(b) shows the mean pointed angles on the vertical axis for the visually presented target angles on the horizontal axis. The solid lines show the fitted linear functions of the form $f(x) = mx + b$ with $m, b \in \mathbb{R}$. The blue and red lines correspond to the pooled results for the left and right display unit, respectively. The light blue and red plots show the results from the virtual condition, whereas the dark blue and red plots show the results from the real world condition. The error bars show the standard errors.

Subjects were less accurate at pointing to visually perceived angles in the real world condition and show more variance in responses compared to the results found by Ellis and Nemire [EN93]. Comparing the slopes between the real (right eye $m = 0.89$, left eye $m = 0.9$) and virtual conditions (right eye $m = 1.03$, left eye $m = 1.03$), we computed the GFOVs that have to be applied to the left and right eye to compensate for the discrepancies in pointing performance. The results show that the diagonal GFOV judged as correct by the participants deviates from the nominal 80$^\circ$ FOV specified by the manufacturers: 87.886$^\circ$ for the left eye and 88.724$^\circ$ for the right eye.

4 Effects on Distance Judgments

As discussed in Section 2, applying veridical GFOVs to the rendering process may reduce distance underestimation. In this section we describe the experiments that we conducted to evaluate a range of GFOVs around the values determined in Sections 3.3 and 3.4 for effects on distance judgments.

4.1 Participants

5 female and 6 male (age 21-27, $\varnothing$: 24.7) subjects participated in both distance judgment experiments. Subjects were students or department members of computer science or mathematics. All had normal or corrected to normal vision. All subjects had experience with HMD setups. We randomized the order of the distance judgment experiments. The total time per subject including instructions, experiment, breaks, and debriefing was 1 hour.
Figure 4: Experiment setup: (a) subject performing a blind walking task to the estimated distance of (b) a virtual marker displayed on the HMD, and (c) subject at a fixed position performing a blind grasping task to (d) a virtual marker displayed on the HMD.

4.2 Experiment Setup

For the experiments, we used the setup described in Section 3, but adapted the calibration desk and the virtual replica (see Figure 4) for the blind grasping experiment. In the laboratory, we tracked the position of the HMD with the optical tracking system WorldViz PPT X8, and the orientation of the HMD with an InterSense InertiaCube 3 (see Figure 4(a)).

4.3 Experiment 1 (E1): Blind Walking

In this experiment we evaluated effects of the GFOVs with a blind walking task.

4.3.1 Material and Methods of E1

The blind walking experiment was divided into a baseline phase in the real world and a test phase in the virtual replica. We used a within-subject design. At the beginning of each trial in the virtual condition, subjects were guided to a fixed start position. Once the subject had assumed the start position, the virtual view changed to display the virtual replica of the laboratory. We tested four different diagonal GFOVs: 80°, 83.33°, 86.67° and 90°. For each trial, we displayed a target marker at a distance of 3m, 4m or 5m (each presented 4 times in randomized order), and instructed subjects to walk to the target after the display turned black. Once the subject walked without vision to the estimated target position, a button
Figure 5: Pooled results of (a) blind walking and (b) blind grasping experiment for the real world and virtual conditions with different tested GFOVs. The horizontal axes show the visually presented target distances, and the vertical axes show the judged distances.

Press on a Wii controller indicated the end of the trial. We measured the Euclidean distance from the start position to the tracked position at the end of the trial.

In the baseline phase, we used a physical marker similar to the marker shown in Figure 4(b). Blindfolds prevented subjects from seeing the laboratory in the blind walking phases. We randomized the order of baseline and test phase between subjects.

4.3.2 Results and Discussion of E1

Figure 5(a) shows the pooled results for the blind walking experiment with the target distances on the horizontal axis, and the mean walked distances on the vertical axis. The black plots show the results from the baseline phase in the real world. The blue and green plots show the walked distances for the diagonal GFOVs in the virtual conditions. The gray line corresponds to optimal distance judgments. The error bars show the standard errors.

The results show that subjects estimated distances in the real world on average as 2.81m (3m target distance), 3.83m (4m target distance) and 4.95m (5m target distance), which is quite accurate compared to the results in the virtual conditions. For the applied diagonal GFOV of 80° we found results of 2.21m, 2.91m and 3.73m for the 3m, 4m, respective 5m target distance, corresponding to distance underestimations of 26%, 27% and 25%. For the diagonal GFOV of 83.33° we found results of 2.25m, 2.96m and 3.83m, i.e., underestimations of 25%, 26% and 23%. For the diagonal GFOV of 86.67° we found results of 2.29m, 3.08m and 3.81m, i.e., underestimations of 24%, 23% and 24%. For the diagonal GFOV of 90° we found results of 2.34m, 3.14m and 3.98m, i.e., underestimations of 22%, 22% and 20%. In general, subjects underestimated distances in all virtual GFOV conditions and the underestimation decreased with increasing the GFOV.
4.4 Experiment 2 (E2): Blind Grasping

In this experiment we evaluated effects of the GFOVs with a blind grasping task.

4.4.1 Material and Methods of E2

Materials and methods of the blind grasping experiment were similar to the blind walking experiment. Subjects were seated at a desk in the laboratory with a fixed head position as illustrated in Figure 4(c). We tracked the position of an active infrared marker attached to the subject’s pointing finger of the dominant hand with the optical tracking system.

In the virtual condition in each trial we displayed a virtual replica of the desk to subjects, with a marker shown at a randomized distance of either 0.3m, 0.4m or 0.5m (each tested 4 times in randomized order). We tested four different diagonal GFOVs: 80°, 83.33°, 86.67° and 90°, which we applied to the rendering process. After presenting the visual stimulus, we blanked the screen, instructing the subject to touch the respective point of the desk at which they had perceived the marker. After the subject had moved the finger with the infrared marker to the respective position, a button press on a Wii controller indicated the end of the trial. We used the tracked marker position to determine the Euclidean distance on the desk to the target position. After a trial had ended, subjects were instructed to move their hand back to a starting position indicated by a tactile marker on the front side of the desk.

In the baseline phase, we used a physical marker similar to the virtual marker. Blindfolds prevented subjects from seeing the laboratory in the blind grasping phases. We randomized the order of baseline and test phase between subjects.

4.4.2 Results and Discussion of E2

Figure 5(b) shows the pooled results for the blind grasping experiment with the target distances on the horizontal axis, and the mean judged distances on the vertical axis. The black plots show the results from the baseline phase in the real world. The blue and green plots show the judged distances for the diagonal GFOVs in the virtual conditions. The gray line indicates optimal distance judgments. The error bars show the standard errors.

Subjects estimated distances in the real world on average as 0.29m (0.3m target distance), 0.38m (0.4m target distance) and 0.46m (0.5m target distance), which is quite accurate compared to the results in the virtual conditions. For the applied diagonal GFOV of 80° we found results of 0.25m, 0.35m and 0.42m for the 0.3m, 0.4m, respective 0.5m target distance, corresponding to distance underestimations of 17%, 13% and 16%. For the diagonal GFOV of 83.33° and 86.67° we found results of 0.26m, 0.35m and 0.43m, i.e., underestimations of 13%, 13% and 14%. For the diagonal GFOV of 90° we found results of 0.28m, 0.36m and 0.44m, i.e., underestimations of 8%, 10% and 12%. In general, we found similar qualitative results as for blind walking, i.e., subjects underestimated distances for all tested GFOVs, and the underestimation decreased with increasing GFOV. In comparison to blind walking, the subjects showed a generally reduced relative underestimation of arm-reach distances.
5 General Discussion

Both calibration methods described in Section 3 yield to large differences in the calibrated GFOVs. The visual calibration approach by Steinicke et al. [SBK+09] shows no significant difference to the DFOV provided by the manufacturers, whereas the visual-proprioceptive calibration approach by Ellis and Nemire [EN93] shows a large discrepancy, which may indicate a less accurate calibration due to the proprioceptive response task. The calibrations resulted in different values for left and right eyes, which indicate that the eye positions tend to diverge from the optimal exit pupil positions in the HMD, which thus highlights the importance of separate calibration of the eye frustums.

The distance judgment experiments described in Section 4 do not show a local optimum in distance estimation for any of the considered GFOVs, but rather a monotone relation between minification caused by increasing the GFOV, and distance judgments. The monotone progression of the results suggests that a much larger GFOV is needed to compensate for the distance underestimation of the subjects, which cannot be explained by FOV miscalibration.

6 Conclusion and Future Work

In this paper we tested two calibration techniques to determine the FOV of an immersive HMD. The visual calibration technique proposed by Steinicke et al. [SBK+09] resulted in a calibrated GFOV close to the nominal value provided by the HMD manufacturer, whereas the visual-proprioceptive calibration technique proposed by Ellis and Nemire [EN93] resulted in an increased GFOV. Applying the results of the calibration processes to distance judgment tasks, i.e., blind walking for long distances and blind grasping for arm-reach distances, we found the tendency that increasing GFOVs reduces the distance underestimation. However, GFOVs have to be greatly increased for subjects to approximate distance judgments as in the real world. In the future, we will further investigate interrelations between size and perspective cues when applying different geometric rendering parameters.

Acknowledgements

Authors of this work are supported by the German Research Foundation (DFG 29160962).

References


