Analyzing Effects of Geometric Rendering Parameters on Size and Distance Estimation in On-Axis Stereographics

Gerd Bruder*, Andreas Pusch†, Frank Steinicke‡
Departments of Computer Science and Human-Computer-Media
University of Würzburg

Abstract

Accurate perception of size and distance in a three-dimensional virtual environment is important for many applications. However, several experiments have revealed that spatial perception of virtual environments often deviates from the real world, even when the virtual scene is modeled as an accurate replica of a familiar physical environment. While previous research has elucidated various factors that can facilitate perceptual shifts, the effects of geometric rendering parameters on spatial cues are still not well understood.

In this paper, we model and evaluate effects of spatial transformations caused by variations of the geometric field of view and the interpupillary distance in on-axis stereographic display environments. We evaluated different predictions in a psychophysical experiment in which subjects were asked to judge distance and size properties of virtual objects placed in a realistic virtual scene. Our results suggest that variations in the geometric field of view have a strong influence on distance judgments, whereas variations in the geometric interpupillary distance mainly affect size judgments.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality

Keywords: on-axis stereographics, distance perception, size perception, cue conflicts, head-mounted display

1 Introduction

Modern virtual reality (VR) display technologies enable users to experience a three-dimensional virtual environment (VE) from an egocentric perspective. Such immersive viewing experiences have an enormous potential for a variety of application domains, in which accurate spatial perception during design, exploration or review of virtual models and scenes is required. Head-mounted displays (HMDs) and immersive projection technologies are often used in order to provide a user with near-natural feedback of virtual content, as if the user is present in the virtual scene. In particular, modern realtime rendering systems can create compelling immersive experiences offering most of the spatial visual cues we can find in real-world views, including perspective, interposition, lighting and shadows [Thompson et al. 2011]. However, distance and size perception are often biased in such environments, causing users to over- or underestimate spatial relations in virtual scenes to a much higher magnitude than can be observed in similar scenes in the real world [Loomis and Knapp 2003; Thompson et al. 2004; Interrante et al. 2007].

Obviously, issues with visual rendering have been suggested as a potential source for biased spatial perception. In order for a virtual scene to be displayed stereoscopically on a binocular HMD, the computer graphics system must determine which part of the scene should be displayed where on the two screens. In 3D computer graphics, planar geometric projections are typically applied, which make use of a straightforward mapping of graphical entities within a 3D 'view' region, i. e., the *view frustum*, to a 2D image plane. During the rendering process, objects inside the view frustum are projected onto the 2D image plane; objects outside the view frustum are omitted. The exact shape of each view frustum in on-axis stereographic displays (as used in many HMDs) is a symmetric truncated rectangular pyramid. The opening angle at the top of the pyramid, often denoted as *geometric field of view* (GFOV) [McGreevy et al. 1985], should match the *display's field of view* (DFOV) for the imagery to be projected in a geometrically correct way [Steinicke et al. 2011a]. Another important characteristic of the human visual system is the *interpupillary distance* (IPD), which describes the horizontal separation of the eyes that ranges from 5.77cm to 6.96cm (median: 6.32cm) in adult males (according to Woodson [Woodson 1981]). Since the viewpoints of both eyes are horizontally separated, each eye receives a slightly different retinal image. The brain interprets the binocular inputs and fuses the images, resulting in the impression of a solid space and the perception of depth [Cutting 1997]. Typically, the user's IPD is measured and then applied to the *geometric interpupillary distance* (GIPD) used for stereoscopic rendering, assuming that the HMD's display units are correctly adjusted in front of the user's eyes.

Both geometric rendering parameters, GFOV and GIPD, have to be defined in all on-axis 3D stereoscopic visualization systems. At the same time, they are particularly prone to calibration errors and therefore bear a high risk of accidentally skewing a user's perception in immersive VEs. Common sources for such errors are naïvely applying manufacturer specifications (e. g., the FOV of built-in displays in head-mounted devices [Kuhl et al. 2009; Kellner et al. 2012]) without verification of the physical display characteristics, or by using population means to approximate a user's IPD. Although slight errors in such rendering parameters are quite common in VR environments, it is still not entirely clear as to what effects these discrepancies have on distance and size cues. Moreover, it has been found that when users have direct control over a rendering parameter, they often try to use it to compensate for given perceptual biases that may have been introduced by miscalibration of other parameters [Steinicke et al. 2011b]. It remains a challenging question how rendering parameters are related regarding particular cues, and if they could be used to address perceptual biases.

The motivation of this work is to compare mathematical models for selected cues that are dominated by the two rendering parameters GFOV and GIPD in terms of their mutual influence on size and distance perception in realistic VEs [Thompson et al. 2004; Interrante et al. 2007; Willemsen et al. 2009]. We describe the effects

---

*e-mail: gerd.bruder@uni-wuerzburg.de
†e-mail: andreas.pusch@uni-wuerzburg.de
‡e-mail: frank.steinicke@uni-wuerzburg.de

both rendering parameters have on the theoretical models, and then compare the predictions to relative size and distance judgments collected in a psychophysical experiment.

The paper is structured as follows: Section 2 provides background information about spatial perception and the cues that are primarily affected by the rendering parameters GFOV and GIPD. In Section 3, we describe how the considered cue models are influenced by changes in both stereoscopic rendering parameters. In Section 4, we describe a psychophysical experiment in which we evaluate how subject responses correlate with the predictions of these models when varying the rendering parameters, and discuss the results. Section 5 concludes the paper.

2 Background

In this section, we summarize background information about selected mathematical models for size and distance cues in the scope of the two stereoscopic rendering parameters GFOV and GIPD.

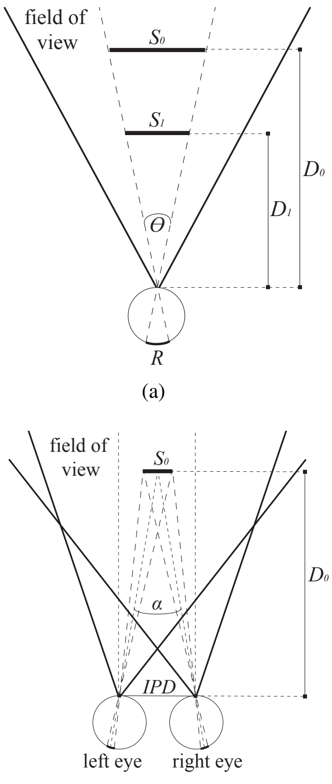
2.1 Egocentric Perspective

Human visual perception of distance, size and spatial relationships relies on prior knowledge (i. e., our lifelong experience in the physical world), individual preferences and neurophysical properties, as well as on given visual stimuli. Among the most important of these are perspective cues [Goldstein 2009], including object retinal size scaled by distance, object distance as a function of relative position to the ground plane and sky (both of which extend toward a visible horizon), convergence of parallel lines in vanishing points, and many more [Thompson et al. 2011].

Emmert's Law [Lou 2007] provides a simple approximation of the inherent assumptions of the perceptual system regarding size, distance, and retinal size. Emmert observed that afterimages, although having a constant retinal size, appeared to be larger, if the viewed background was farther away. In other words, for the same retinal size, an object's perceived size will depend on its perceived distance. This phenomenon is illustrated in Figure 1(a). In order to judge an object's size, the visual system needs to evaluate the object's distance from the viewpoint, or vice versa. The perceived size $\tilde{S_0} \in \mathbb{R}^+$ of an object with linear size S_0 is proportional to the product of its perceived distance $\tilde{D_0} \in \mathbb{R}^+$ (D_0 is linear distance) and angular size $\Theta \in \mathbb{R}^+$, and can be expressed in simple form as

$$\tilde{S_0} \sim \tilde{D_0} \cdot \tan\left(\frac{\Theta}{2}\right). \tag{1}$$

The phenomenon of perceived-size constancy [Gilinsky 1951] denotes the tendency of an object to maintain the same perceived size, even if its retinal size changes, i. e., if the object moves farther away or approaches the observer. In such cases, e. g., for objects with familiar size, the perceived distance of an object can be described using the above relation. This size-distance-invariance hypothesis can describe many properties of size and distance perception, including cases of apparent misperception. However, Murray et al [Murray et al. 2006] have recently replaced the retinal size in the relation by the perceived retinal size. This updated hypothesis states that the ratio of perceived linear size to perceived distance is not necessarily a simple function of the visual angle, but that the visual angle can be subject to perceptual biases as well. These findings underline that the controversial size-distance invariance hypothesis and some other aspects of visual perception are still not well understood [Mon-Williams and Tresilian 1999].



(a)

(b)

Figure 1: *Illustration of (a) visual size-distance ambiguity with object size S_i being a linear function of the distance D_i in case of a constant retinal size R with corresponding angular size Θ, and (b) object size and distance as a function of eye convergence.*

2.2 Stereopsis

While the visual system faces an inherent ambiguity of size and distance in case of monocular vision (given unfamiliar objects or objects of uncertain size), the binocular configuration of the eyes provides humans with absolute stereoscopic cues. As mentioned in Section 1, since the eyes are horizontally separated by an interpupillary distance, the brain receives two slightly different images of a scene. The perceptual system can then make use of these cues, in particular, of binocular disparity and convergence. The former refers to differences in the retinal image locations when light from an object is projected into the left and right eye of an observer. Solving the disparity correspondence problem, the brain relates retinal image contents from the two eyes to one another, allowing the perceptual system to compute the distance of seen objects using simple triangulation (see Figure 1(b)). Further, when focusing on an object, the eyes need to rotate toward that object to bring it to the fovea of each retina. Turning the eyes inward when fixating on a closer object leads to larger convergence angles. The convergence state of the eyes, changed by extrinsic muscle exertion, thus provides an absolute cue about the distance of an object from an observer. In a simplified setting, the distance $D_0 \in \mathbb{R}^+$ of an object can be computed from the convergence angle using the following equation:

$$D_0 = \frac{\mathrm{IPD}}{2 \cdot \tan(\frac{\alpha}{2})}, \tag{2}$$

with the user's IPD and the convergence angle $\alpha \in \mathbb{R}^+$. Figure 1(b) shows how size and distance ambiguities can be resolved with binocular viewing.

Stereoacuity is naturally limited [Steinman et al. 2000]. Researchers assume a conservative threshold of about 10 seconds of arc (approx. 0.003 degrees) [Palmisano et al. 2010] for stereoacuity. Referring to Equation 2, the maximum distance at which stereopsis may produce usable data would be approx. 1.24km. Beyond this distance, the human visual system cannot sufficiently differentiate binocular information. However, in VR environments the angular resolution of a pixel on a display surface may act as an artificial cut-off to the capabilities of human natural vision. That is, although the retina may be able to respond to much smaller visual stimuli, the size of the smallest digital element of a visual display may prevent this capacity from being exploited. To which degree stereopsis effectively contributes to size or distance judgments in immersive VEs depends on various factors, including user characteristics, virtual scene, display technology, and many more. Cue reliabilities and the resulting weights during cue integration ultimately determine the impact of stereopsis [Ernst 2006]. The effects of stereoscopic display and IPD on distance judgments in realistic scenes have been found less important than predicted by Equation 2 [Willemsen et al. 2008; Creem-Regehr et al. 2005].

3 Cue Conflicts in On-Axis Stereographics

In this section, we discuss the relationship between field of view and interpupillary distance in *on-axis* stereoscopic rendering configurations, which are used in many HMD settings. In contrast, projection screens use *off-axis* stereoscopic rendering, which accounts for the situation that only one display surface exists to present the left and right eye views. In the simplest case of an on-axis binocular display design, the display surfaces are oriented orthogonally to the parallel optical axes for both views, which intersect with the display surfaces in their center. If we disregard optical distortions (e. g., pincushion distortion), the binocular configurations of such displays can be illustrated as shown in Figure 2. We assume a coordinate system in which the camera is represented by a position, look direction vector, up vector, and strafe vector. As illustrated in Figure 2, the frustum of the camera model for HMDs is delimited by the near and far clipping planes, as well as by the size of the virtual image, which appears at the display optics' accommodation distance from the eyes. The size of the display surface defines the horizontal and vertical geometric fields of view[1]. The camera frustums for the left and right eye are separated by the geometric interpupillary distance along the strafe vector.

3.1 Introducing GFOV and GIPD Conflicts

When displaying computer generated images on a physical display, we have to distinguish between the virtual rendering setup and the physical display setup. In order to provide a view to a virtual scene on a head-referenced display surface that matches what a user would see in a corresponding real-world scene, the computer graphics rendering system has to be calibrated to the physical display characteristics. In particular, the GFOV in the rendering environment has to be set to the visual angle covered by the display units in front of the user's eyes. The interpupillary distance of the user has to be applied to the binocular camera model as shown in Figure 2. In case of any discrepancy, various size and distance cues are affected, and can cause perceptual shifts.

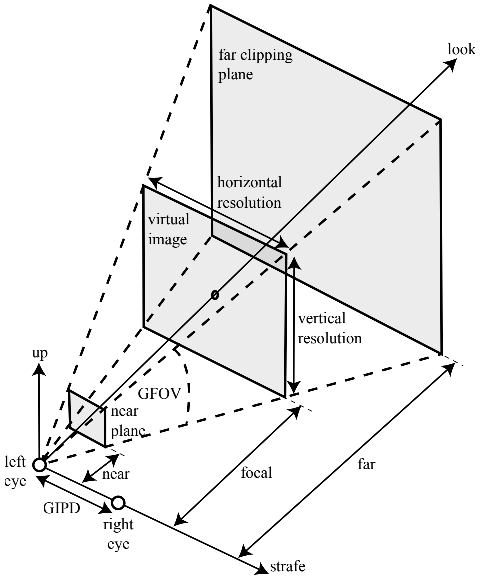


Figure 2: *Idealized binocular camera model in three-dimensional computer graphics for head-mounted displays [Robinett and Holloway 1994]. For better legibility, only the view frustum for the left eye camera object is shown. The right eye camera frustum results from a translation by GIPD along the strafe vector.*

Field of View Gains

The DFOV refers to the visual angles subtended by a display unit in front of the user's eyes (cf. Section 1), whereas the GFOV refers to the visual angles of the view frustum in the computer graphics rendering model illustrated in Figure 2. If GFOV matches DFOV, the virtual space is mapped onto the physical display in such a way that users see a natural perspective. Mapping differences can be described via GFOV gains $g_F \in \mathbb{R}^+$ as

$$\mathrm{GFOV} = g_F \cdot \mathrm{DFOV}. \tag{3}$$

If the GFOV differs from the DFOV with $g_F \neq 1$, the retinal size of displayed objects is scaled. For $g_F > 1$, the virtual image is rendered over a larger visual field, but compressed onto a smaller physical display surface. For $g_F < 1$, a smaller visual field is rendered, and up-scaled onto a larger physical display surface.

Interpupillary Distance Gains

Figure 2 illustrates the displacement of the eye points in strafe direction by GIPD. Assuming the physical display characteristics are correctly applied to the left and right eye camera frustums, differences between users only occur in the GIPD between the centers of projection of the two cameras. If anthropometric population means are applied as GIPD, but deviate from a user's individual IPD, this either results in an increased or decreased GIPD. This relative difference can be described via GIPD gains $g_I \in \mathbb{R}^+$ as

$$\mathrm{GIPD} = g_I \cdot \mathrm{IPD}. \tag{4}$$

For $g_I = 1$, GIPD and the user's IPD match, thus providing a natural perspective. Changes to GIPD with $g_I \neq 1$ lead to systematic changes of convergence angles (see Section 2.2), but have limited effects on retinal size. Increasing the GIPD over the user's IPD (with $g_I > 1$) results in up-scaled convergence angles, i. e., virtual objects should appear closer, whereas decreasing the GIPD (with $g_I < 1$) results in down-scaled convergence angles. For $g_I = 0$ the left and right images are exactly the same.

---

[1] If not stated otherwise, we refer to the GFOV as the vertical geometric field of view, since the horizontal field of view can be computed using the fixed aspect ratio of a display surface.

3.2 Effects of GFOV and GIPD Conflicts

In the following, we describe effects on distance and size that occur in theory when geometric field of view and interpupillary distance of the rendering cameras do not comply with the physical configuration. Although virtual renderings often deviate from the physical display properties, it is difficult to estimate how discrepancies may influence the perceived size or distance of objects in binocular displays. In order to compute effects introduced by GFOV and GIPD gains on size and distance cues of stereoscopically displayed objects, we developed an OpenGL simulation for on-axis stereoscopic rendering. In our test environment, we use the following procedure:

1. We render a scene for both eyes using a virtual camera configuration (with GFOV and GIPD gains applied), i. e., the virtual scene is projected onto the image surfaces using planar perspective projections (see Figure 3, images in right column).

2. We compute the transformation that is introduced by displaying the planar rendered images for the left and right eye on the physical display surfaces; the FOV and IPD in the physical configuration can differ from the virtual configuration (see Figure 3, images in left column).

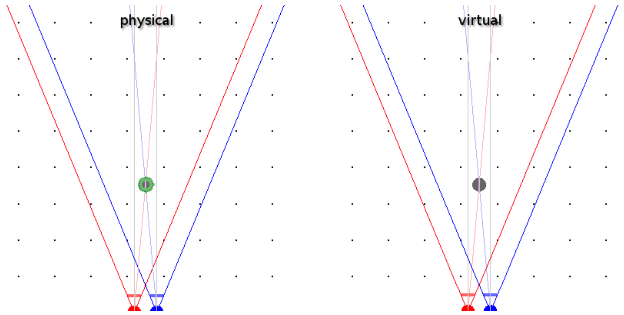
3a. By testing ray intersections from the eye positions to the positions of objects that are displayed on the left and right physical displays in front of the user's eyes, we compute the object's size and distance from stereopsis as described in Section 2.2 (see Figure 3, gray objects).

3b. From the retinal size of the object on the physical display surfaces in front of the user's eyes, we compute the object size and distance as predicted solely by perceived-size constancy (see Section 2.1), i. e., without considering convergence angles (see Figure 3, green objects).

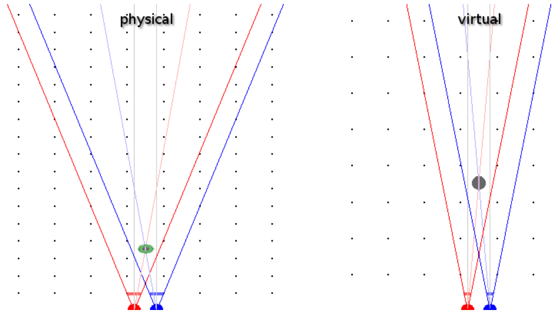
As illustrated in Figure 3, changing GFOV and GIPD from the physical configuration introduces perceptual conflicts between the two cues. Figure 3(a) shows natural viewing with matching rendering and display configurations, (b) shows effects of decreasing the GFOV by a factor of $g_F = 0.5$ when rendering the virtual view, (c) shows effects of increasing the GIPD by a factor of $g_I = 2$ when rendering the virtual view, and (d) shows effects of a GFOV gain of $g_F = 1.77$ and a GIPD gain of $g_I = 2$. The left eye frustums are shown in red, whereas the right eye frustums are shown in blue. The examples show that applying GFOV and GIPD gains both cause changes to the convergence angles of virtual objects, and thus the size and distance cues that can be derived from stereopsis, whereas only GFOV gains have a strong effect on retinal size. Figure 3(d) indicates that if distance cues from stereopsis dominate percepts, it should even be possible to provide a controlled cue conflict stimulus, in which perceptual effects of changing GFOV may be reversed by changing GIPD. However, if retinal size dominates percepts due to perceived-size constancy (see Section 2.1), there should be limited to no effect on distance perception by changing GIPD. In the following, we describe the mathematical effects of GFOV and GIPD gains on the cues, before we describe our psychophysical evaluation in Section 4.

Field of View Gains

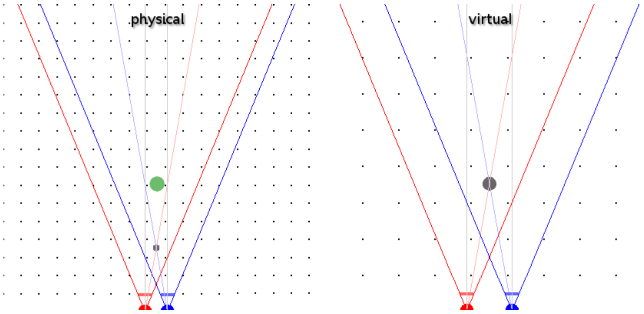
We determined the differences in convergence cues that are introduced by changing GFOV gains. As described in Section 2, increasing convergence angles should result in objects being perceived closer to the observer, as well as smaller, whereas decreasing convergence angles should have reverse effects. As illustrated in Figure 3(b), a gain $g_F = 0.5$ results in the distance to objects being reduced by 50%, whereas the object size appears non-uniformly



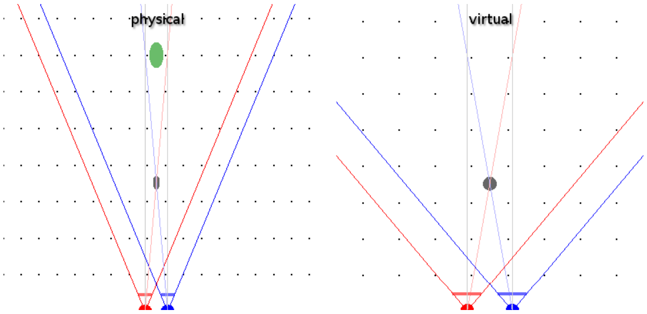
(a) $g_F = 1, g_I = 1$



(b) $g_F = 0.5, g_I = 1$



(c) $g_F = 1, g_I = 2$



(d) $g_F = 1.77, g_I = 2$

Figure 3: *Representations of cue conflict situations: (a) matching size and distance cues with calibrated GFOV and GIPD, (b) effects of reducing the GFOV by 50%, (c) effects of rendering the virtual view with twice the IPD, and (d) effects of gains* $g_F = 1.77$ *and* $g_I = 2$*. We computed the position and size of the gray objects by ray intersection from convergence angles (see Step 3a); the dot pattern illustrates spatial transformations in case of dominance of stereoscopic cues. In contrast, the green objects show the prediction from perceived-size constancy (see Step 3b).*

scaled. The relation between virtual and displayed object distance for convergence cues can be expressed with scaling factor m_f as

$$\hat{D}_0 = D_0 \cdot m_f, \quad m_f := \frac{\tan(g_F \cdot \text{DFOV}/2)}{\tan(\text{DFOV}/2)}, \tag{5}$$

with D_0 being the virtual object distance, and $\hat{D}_0$ the resulting object distance from convergence cues when shown on the physical display. The size of an object results as

$$\hat{S}_s = S_s, \quad \hat{S}_l = S_l \cdot m_f, \tag{6}$$

with S_s the virtual size in strafe direction, S_l the virtual size in look direction (cp. Figures 2 and 3), and $\hat{S}_s$ as well as $\hat{S}_l$ the respective object size dimensions that can be derived from convergence cues when shown on the physical display. In particular, the virtual scene should appear non-uniformly scaled along the look direction in front of the user's eyes.

On the other hand, perspective changes with GFOV gains also modify the retinal size of a displayed object. As predicted by Emmert's Law, a systematic increase or decrease of retinal size changes the perceived size or distance of seen objects (see Section 2.1). Kuhl et al. [Kuhl et al. 2009] and Steinicke et al. [Steinicke et al. 2011b] observed that changes of GFOVs have an impact on the perceived distance to virtual objects, whereas effects on perceived size have not been reported in the literature. In previous work, researchers mainly studied how distance underestimation with HMDs can be compensated by applying GFOV gains $g_F > 1$. The predicted distance in the case of perceived-size constancy (see Section 2.1) can be described with the following coefficient [Steinicke et al. 2011b]:

$$\hat{D}_0 = D_0 \cdot m_f. \tag{7}$$

The size of an object results as

$$\hat{S}_s = S_s, \quad \hat{S}_l = S_l \cdot m_f, \tag{8}$$

which matches the results from convergence cues. Perceived-size constancy may apply to object features in planes perpendicular to the view direction, but object features may appear stretched or compressed along the view direction.

Interpupillary Distance Gains

With GIPD gains, the distance between the cameras for the left and right eye can be changed leading to altered perspective and convergence cues. As illustrated in Figure 3(c), a gain $g_I = 2$ results in the distance to an object being reduced by 50%, as well as its size being uniformly scaled by 50%. The relation between virtual and displayed object distance for convergence cues can be expressed as

$$\hat{D}_0 = D_0 \cdot g_I^{-1}, \tag{9}$$

with D_0 the virtual target object distance, and $\hat{D}_0$ the resulting object distance from convergence cues when shown on the physical display. The size of an object results as

$$\hat{S} = S \cdot g_I^{-1}, \tag{10}$$

with S the object size in the virtual scene, and $\hat{S}$ the resulting object size from convergence cues when shown on the physical display.

On the other hand, changing the GIPD also changes the Euclidean distance of the cameras to virtual objects, and defines how much the cameras see of the sides of objects (cp. Figure 1(b)). Thus, changing the GIPD also has an effect on the retinal size of virtual objects. However, this depends to a large part on the geometry of the virtual object. In a simple case, for a spherical object, the predicted distance in the case of perceived-size constancy results as

$$\hat{D}_0 = \sqrt{D_0^2 + (g_I^2 - 1) \cdot (\text{IPD}/2)^2} \tag{11}$$

The size of a spherical object in simple form results as

$$\hat{S} = S, \tag{12}$$

with S the virtual size and $\hat{S}$ the resulting object size that can be derived from retinal size in the case of perceived-size constancy.

Observations and Questions

From the above explanations for convergence cues we can see that GFOV gains introduce a non-uniform scaling of the scene in look direction, whereas GIPD gains introduce a uniform scaling of the scene in front of the user's eyes. Indeed, these computational results broadly follow experimental observations in the literature [Kuhl et al. 2009; Steinicke et al. 2011b; Yamanoue et al. 2006]. Moreover, as illustrated in Figure 3(d), one can compute pairs of GFOV and GIPD that–theoretically–exactly compensate for individual effects. However, Figures 3(c) and (d) also show that the perceptual system may need to deal with a cue conflict in the presence of perceived-size constancy (e. g., for objects with familiar size). The main questions that arise are, how such conflicts are resolved, and how much each of the rendering parameters GFOV and GIPD will contribute to size and distance percepts.

4 Perceptual Experiment

In this section, we present an experiment, in which we have investigated the effects of changes to the GFOV and GIPD on relative size and distance judgments.

4.1 Experiment Design

As visual stimulus, we used a virtual hallway scene of 3.8m×2.5m×35m (in width, length, and height), which was rendered with Crytek's CryEngine 3. We used a split screen design of a virtual hallway (see Figure 4), with the left hand side view being rendered with one pair of GFOV and GIPD, and the right hand side view being rendered with another. We did not use the stereoscopic rendering facilities of the CryEngine 3, but added an interface to our own software, with which we handled the generation of the split-screen stereo pair and were able to provide accurate on-axis stereographics. In both virtual scenes we placed a virtual avatar (Caucasian male, 1.85m height) as focus object. We considered the distance of the avatar from the observer as between-subjects variable, and tested three distances: 4m, 6m and 8m.

Hardware Setup

The subjects were equipped with an NVIS nVisor SX60 HMD with a resolution of 1280×1024 pixels per eye at a refresh rate of 60Hz for the visual stimulus presentation. The nVisor HMD uses parallel symmetric on-axis display optics and has a diagonal FOV of 60 degrees, which we used as basis for the GFOV transformations during visual stimuli generation. Views on the real world were blocked with a black lightshield. During the experiment, the subject's head was oriented in view direction along the virtual hallway and presented at the subject's eye height. The experiment did not involve head movements of the subjects in order to provide only static size and distance cues, but no motion cues.

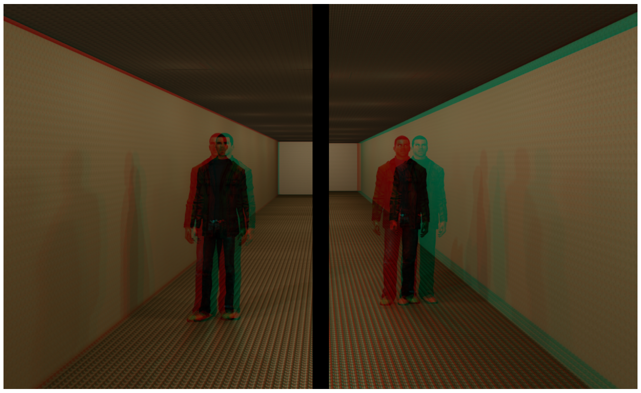


Figure 4: *Illustration of the split-screen visual stimulus used in the experiment (here with red-cyan anaglyphs). Subjects had to compare size and distance of the avatars displayed in the left and right view.*

We used an Intel computer (Core i7 processors, 6GB of RAM and an Nvidia Quadro 4000 graphics card) for rendering, system control, and logging. A standard keyboard served as means for the subjects to enter their perceived size and distance judgments. The subjects received all instructions on slides presented on the HMD. There was no communication between experimenter and subjects during the experiment in order to focus subjects on the task.

Participants

22 male and 17 female subjects (age $18-44$, mean 23.49, SD 4.93) participated in the experiment. Subjects were students or members of the departments of computer science, psychology or human-computer-media. All had normal or corrected to normal vision; 7 wore glasses and 9 contact lenses during the experiment. We tested visual acuity of all subjects with a simple vision test prior to the experiment based on a classic Snellen chart. We measured the interpupillary distances of our subjects using the mirror method described by Willemsen et al. [Willemsen et al. 2008]. The IPDs of our subjects ranged between $5.0-7.1$cm (mean 6.14, SD 0.48). The eye height of our subjects ranged between $1.51-1.87$m (mean 1.66, SD 0.1). 26 of the subjects had no experience with 3D games, 3 had some, and 10 had much experience. All subjects were naïve to the experimental conditions. 31 of the subjects had much experience with 3D stereoscopic cinema, 7 had some, and 1 had no experience. 4 of the subjects had much experience with HMDs, 3 had some, and 32 had no experience. 3 subjects had participated in experiments involving HMDs before.

We verified all subjects' ability to see stereoscopically prior to the experiment by asking subjects to look at an anaglyphic random-dot stereogram, and report the type of 3D object that was shown. Students obtained class credit for their participation. The total time per subject, including pre-questionnaires, instructions, training trials, experiment, breaks, post-questionnaires, and debriefing, was 1 hour. Subjects wore the HMD for approx. 45 minutes. They were allowed to take breaks at any time; short breaks after every 50 trials were mandatory to rest the eyes for a few moments.

4.2 Methods

We used the method of constant stimuli in a two-alternative forced-choice (2AFC) task [Ferwerda 2008]. In the method of constant stimuli, the applied gains are not related from one trial to the next, but presented randomly and uniformly distributed. We applied GFOV gains $g_F \in \{0.5, 0.75, 1.0, 1.25, 1.5\}$ relative to the DFOV of the HMD, and GIPD gains $g_I \in \{0.0, 1.0, 2.0, 3.0, 4.0\}$ relative to the IPD of the subject. We varied combinations of GFOV and GIPD gains in the left and right views in the split screen design. We tested all combinations of GFOV and GIPD gains against all other combinations for all subjects in random order.

In order to investigate the mutual impact of GFOV and GIPD on size and distance perception, subjects had to answer two 2AFC questions at each trial. They had to choose between one of two possible responses: "Does the left or right avatar appear closer to you?" and "Does the left or right avatar appear smaller?"; responses like "I can't tell." were not allowed. Hence, if subjects cannot detect the signal, they are forced to guess, and will be correct in 50% of the trials. The gain at which the subject favors one response in half of the trials corresponds to the point of subjective equality (PSE), at which the subject judges the size or distance of the avatars that are displayed with different rendering parameters as identical. As the GFOV and GIPD gains decrease or increase from this value, the ability of the subject to detect the differences in distances or sizes increases, resulting in psychometric curves for the discrimination performance. In order to avoid subjects directly comparing the renderings of the virtual scene in subsequent trials, we displayed a blank image for 200ms between the renderings as a short interstimulus interval. Additionally, subjects filled out the Kennedy-Lane simulator sickness questionnaire (SSQ) immediately before and after the experiment, further the Slater-Usoh-Steed (SUS) presence questionnaire, and a demographic questionnaire.

4.3 Results

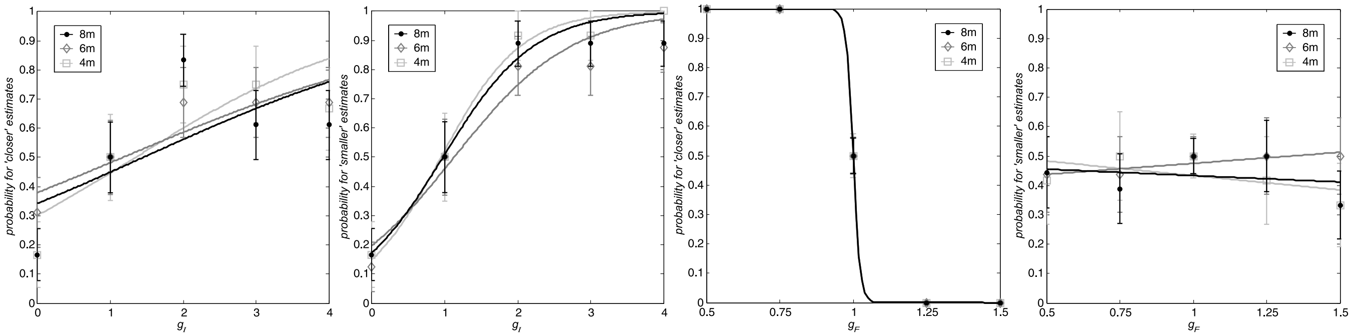
We pooled the data over all subjects for the two 2AFC tasks in the three between-subjects groups. We had to exclude the data of one subject from the 8m target distance group, who showed strong simulator sickness symptoms and made inconsistent judgments over a large part of the experiment. The subjects reported low mean SUS presence scores (mean 2.62, SD 0.97). SSQ mean scores increased from 7.11 (SD 8.60) before the experiment to 28.95 (SD 21.10) after the experiment.

Figure 5 shows the pooled results of the subjects for the three tested target object distances (8m, 6m, and 4m). In Figures 5(a) and (b) we plotted the effects of varying the g_I in one view, with $g_I = 1$ in the other view, and $g_F = 1$ in both views, on distance and size judgments, respectively. The y-axis shows the probability that subjects judged the avatar displayed with the varied g_I as (a) closer to the observer or (b) smaller. In Figures 5(c) and (d) we plotted the effects of varying the g_F in one view, with $g_F = 1$ in the other view, and $g_I = 1$ in both views, on distance and size judgments, respectively. The y-axis shows the probability that subjects judged the avatar displayed with the varied g_F as (c) closer to the observer or (d) smaller. The solid black functions show the results for the 8m target object distance, the dark gray functions for the 6m distance, and the light gray functions for the 4m distance. The error bars show the standard error. The sigmoid psychometric functions are fitted to the data with $f(x) = \frac{1}{1+e^{ax+b}}$, $a, b \in \mathbb{R}$. The chi-square goodness of fit for the functions is shown in Table 1.

Figure 6 shows the pooled results of the subjects in the 6m target distance condition[2]. We plotted the subjects' responses when seeing one view with $g_F = 1$ and $g_I = 1$, as well as the other view varied with $g_F \in \{0.5, 0.75, 1, 1.25, 1.5\}$ on the x-axis and $g_I \in \{0, 1, 2, 3, 4\}$ on the y-axis. The color gradients show the probability that subjects judged the target object as (a) closer to the observer, or as (b) smaller.

---

[2] Plots for 4m and 8m distances showed the same qualitative distribution.



(a) (b) (c) (d)

Figure 5: *(a+b) Results of $g_F = 1$ in both views, $g_I = 1$ in one view, with $g_I \in \{0, 1, 2, 3, 4\}$ on the x-axis varied in the other view. The y-axis shows the probability that subjects judged the target object displayed with the varied g_I as (a) closer to the observer or (b) smaller. (c+d) Results of $g_I = 1$ in both views, $g_F = 1$ in one view, with $g_F \in \{0.5, 0.75, 1.0, 1.25, 1.5\}$ on the x-axis varied in the other view. The y-axis shows the probability that subjects judged the target object displayed with the varied g_F as (c) closer to the observer or (d) smaller.*

Fig.	8m	6m	4m
5(a)	0.82	0.09	0.37
5(b)	0.18	0.22	0.21
5(c)	1e-14	1e-14	1e-14
5(d)	0.13	0.01	0.10

Table 1: *χ^2 goodness of fit for the psychometric functions for the three target distances plotted in Figure 5.*

4.4 Discussion

Results indicate that the GFOV (within the tested range) had a strong effect on distance judgments. Larger GFOV gains g_F resulted in objects being judged as farther away from the observer (see Figure 5(c)), which is in line with the predictions of the models described in Section 3.2. The GIPD (within the tested range) also showed an effect on distance judgments. Larger GIPD gains g_I resulted in objects being judged as closer to the observer (see Figure 5(a)), although distance discrimination performance appears to be less influenced by the GIPD than by the GFOV (see Figure 6(a)). Figure 6(a) further reveals that the tested GIPD gains had only a slight effect on distance judgments when set in direct relation to the tested range of GFOV gains. These results suggest that distance perception in the tested realistic scene relies less on convergence cues than predicted in Section 3.2 (cp. Figure 3(d)). We observed no consistent effect of the chosen object distances on relative distance judgments.

Results further show an effect of the GIPD on size judgments. Larger GIPD gains g_I resulted in objects being judged as smaller (see Figure 5(b)), which correlates with the predictions of the convergence cue model described in Section 3.2. We observed no consistent effect of GFOV gains on size judgments (cp. Figures 5(d) and 6(b)). We found no consistent effect of the chosen object distances on relative size judgments.

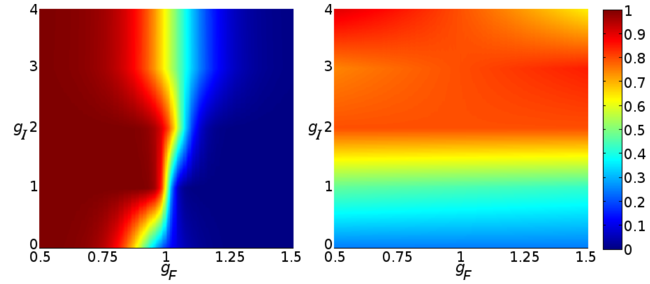
Our results indicate that for cue conflicts introduced with GFOV and GIPD gains in realistic virtual scenes, human distance and size perception differs from the predictions of the models for reduced-cue visual stimuli (i. e., stereopsis and retinal size) described in Section 3.2. In particular, for a typical range of miscalibrated GFOVs, the results indicate a strong effect on relative distance perception, with little effect on relative size perception. In contrast, for a typical range of miscalibrated GIPDs, the results indicate only a slight effect on relative distance perception, but a stronger effect on relative size perception. Figure 6(a) moreover suggests that it may be possible to provide a controlled cue conflict stimulus for distance cues by balancing GFOV and GIPD gains, although merely within a small range.

Limitations

These are interesting results, since they suggest that GFOV and GIPD gains have different effects in realistic scenes than for reduced-cue visual stimuli. However, we have to consider that these quantitative results are likely dependent on a variety of other cues that were specific to the visual stimulus used in our experiment (e. g., retinal size, stereopsis, textures, angle of declination, and height-in-the-picture). Another limitation of the results may be caused by the split-screen design of the visual stimulus. Although the split-screen allowed subjects to directly compare two different renderings with no temporal distortion, this directness could have introduced possible cross-effects due to individual perceptual requirements of alternately viewing the left and right hallways. Finally, the gain ranges we chose for the GFOV and GIPD may have imposed a limitation on our results. It is not unlikely that, for much narrower gain ranges (see Figure 6(a)), we may have found a clearer relationship between GFOV and GIPD for distance perception.

5 Conclusion and Future Work

We have investigated the effects of changing geometric field of view and geometric interpupillary distance when displaying a virtual scene on physical displays. We have described selected size and distance cues that can be derived from egocentric perspective and stereopsis, and we have described the effects of manipulating field of view and interpupillary distance on these cues. In a psychophysical experiment, we have evaluated the mutual impact of the two parameters on size and distance perception, and set the results in relation to the models for the cues. In this work we made one step toward understanding what happens, if incorrect GFOVs or GIPDs are applied in on-axis stereoscopic display environments. Future work includes to further disentangle the effects of miscalibration of immersive VR displays on different spatial cues, and further evaluate resulting cue conflicts. In particular, there are many more cues that we did not consider in this work, some of which are changed by GFOV or GIPD gains (e. g., angle of declination [Kuhl et al. 2009] or accommodation blur [Held et al. 2010]).



(a) (b)

Figure 6: *Result plots of* $g_F = 1$ *and* $g_I = 1$ *in one view, with* $g_F \in \{0.5, 0.75, 1, 1.25, 1.5\}$ *and* $g_I \in \{0, 1, 2, 3, 4\}$ *varied in the other view. The colors show the probability that subjects judged the target object as (a) closer to the observer or (b) smaller.*

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG 29160962). We thank the Crytek GmbH for the CryEngine 3, which was used for the visual stimulus generation in this work. We would like to thank Marina Hofmann for her support preparing the experiment.

References

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

CUTTING, J. E. 1997. How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, and Computers 29*, 1, 27–36.

ERNST, M. O. 2006. *Human Body Perception From The Inside Out*. New York, NY: Oxford University Press, ch. A Bayesian View on Multimodal Cue Integration, 105–131.

FERWERDA, J. 2008. SIGGRAPH Core: Psychophysics 101: How to run perception experiments in computer graphics. In *Proceedings of SIGGRAPH*, ACM Press.

GILINSKY, A. S. 1951. Perceived size and distance in visual space. *Psychological Review 58*, 460–482.

GOLDSTEIN, E. B. 2009. *Sensation and Perception*. Cengage Learning EMEA.

HELD, R. T., COOPER, E. A., O'BRIEN, J. F., AND BANKS, M. S. 2010. Using blur to affect perceived distance and size. *ACM Transactions on Graphics (TOG) 29*, 2, 1–16.

INTERRANTE, V., RIES, B., LINDQUIST, J., AND ANDERSON, L. 2007. Elucidating the factors that can facilitate veridical spatial perception in immersive virtual environments. In *Proceedings of Virtual Reality (VR)*, IEEE Press, 11–18.

KELLNER, F., BOLTE, B., BRUDER, G., RAUTENBERG, U., STEINICKE, F., LAPPE, M., AND KOCH, R. 2012. Geometric calibration of head-mounted displays and its effects on distance estimation. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 18*, 4, 589–596.

KUHL, S. A., THOMPSON, W. B., AND CREEM-REGEHR, S. H. 2009. HMD calibration and its effects on distance judgments. *ACM Transactions on Applied Perception (TAP) 6*, 3, 1–19.

LOOMIS, J. M., AND KNAPP, J. M. 2003. *Virtual and Adaptive Environments*. Lawrence Erlbaum Associates, Mahwah, NJ., ch. Visual Perception of Egocentric Distance in Real and Virtual Environments, 21–46.

LOU, L. 2007. Apparent afterimage size, emmerts law, and oculomotor adjustment. *Perception 36*, 1214–1228.

MCGREEVY, M., RATZLAFF, C., AND ELLIS, S. 1985. Virtual space and two-dimensional effects in perspective displays. In *Proceedings of the 21st Annual Conference on Manual Control*.

MON-WILLIAMS, M., AND TRESILIAN, J. R. 1999. The size-distance paradox is a cognitive phenomenon. *Experimental Brain Research 126*, 578–582.

MURRAY, S. O., BOYACI, H., AND KERSTEN, D. 2006. The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience 9*, 3, 429–434.

PALMISANO, S., GILLAM, B., GOVAN, D. G., ALLISON, R. S., AND HARRIS, J. M. 2010. Stereoscopic perception of real depths at large distances. *Journal of Vision 10*, 6, 1–16.

ROBINETT, W., AND HOLLOWAY, R. 1994. The visual display transformation for virtual reality. Tech. rep., University of North Carolina at Chapel Hill.

STEINICKE, F., BRUDER, G., AND KUHL, S. 2011. Realistic perspective projections for virtual objects and environments. *ACM Transactions on Graphics (TOG) 30*, 1–10.

STEINICKE, F., BRUDER, G., LAPPE, M., KUHL, S., WILLEMSEN, P., AND HINRICHS, K. H. 2011. Natural perspective projections for head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 17*, 888–899.

STEINMAN, S., STEINMAN, B., AND GARZIA, R. 2000. *Foundations of Binocular Vision: A Clinical Perspective*, 1st ed. McGraw-Hill Medical.

THOMPSON, W. B., WILLEMSEN, P., GOOCH, A. A., CREEM-REGEHR, S. H., LOOMIS, J. M., AND BEALL, A. C. 2004. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments 13*, 5, 560–571.

THOMPSON, W., FLEMING, R., CREEM-REGEHR, S., AND STEFANUCCI, J. K. 2011. *Visual Perception from a Computer Graphics Perspective*. A. K. Peters / CRC Press.

WILLEMSEN, P., GOOCH, A. A., THOMPSON, W. B., AND CREEM-REGEHR, S. H. 2008. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments 17*, 1, 91–101.

WILLEMSEN, P., COLTON, M. B., CREEM-REGEHR, S., AND THOMPSON, W. B. 2009. The effects of head-mounted display mechanical properties and field-of-view on distance judgments in virtual environments. *ACM Transactions on Applied Perception (TAP) 2*, 6, 1–14.

WOODSON, W. E. 1981. *Human factors design handbook*. McGraw-Hill.

YAMANOUE, H., OKUI, M., AND OKANO, F. 2006. Geometrical analysis of puppet-theater and cardboard effects in stereoscopic HDTV images. *IEEE Transactions on Circuits and Systems for Video Technology 16*, 6, 744–752.