



Published in final edited form as:

IEEE Trans Vis Comput Graph. 2013 April ; 19(4): 691–700. doi:10.1109/TVCG.2013.36.

Perceptual Calibration for Immersive Display Environments

Kevin Ponto [Member, IEEE],

Department of Computer Sciences, University of Wisconsin, Madison. kpono@cs.wisc.edu.

Michael Gleicher [Member, IEEE],

Department of Computer Sciences, University of Wisconsin, Madison. gleicher@cs.wisc.edu.

Robert G. Radwin [Senior Member, IEEE], and

Department of Biomedical Engineering, University of Wisconsin, Madison.

radwin@engr.wisc.edu.

Hyun Joon Shin

Division of Digital Media, Ajou University. joony@ajou.ac.kr.

Abstract

The perception of objects, depth, and distance has been repeatedly shown to be divergent between virtual and physical environments. We hypothesize that many of these discrepancies stem from incorrect geometric viewing parameters, specifically that physical measurements of eye position are insufficiently precise to provide proper viewing parameters. In this paper, we introduce a perceptual calibration procedure derived from geometric models. While most research has used geometric models to predict perceptual errors, we instead use these models inversely to determine perceptually correct viewing parameters. We study the advantages of these new psychophysically determined viewing parameters compared to the commonly used measured viewing parameters in an experiment with 20 subjects. The perceptually calibrated viewing parameters for the subjects generally produced new virtual eye positions that were wider and deeper than standard practices would estimate. Our study shows that perceptually calibrated viewing parameters can significantly improve depth acuity, distance estimation, and the perception of shape.

Keywords

Virtual reality; calibration; perception; distance estimation; shape perception; depth compression; stereo vision displays

1 Introduction

The viewing model used in Virtual Reality (VR) provides a head-tracked, stereo display to the participant. Conceptually, the virtual cameras are placed where the participant's eyes are, providing a view that reproduces the "natural" viewing conditions. The parameters of this model, illustrated in Figure 1, include the *center of projection offset (PO)* that measures the displacement between the position of the head tracker to the center of the virtual eyes, and a binocular disparity (*BD*) that measures the distance between the center of projection of each camera. In principle, this model should provide a viewing experience that matches the natural world, allowing the participants to use visual cues to correctly perceive depth, shape, and the motion of objects.

In practice, the viewing experience in virtual reality is a poor approximation to the real world. Well documented artifacts include depth compression, a lack of depth acuity, poor perception of shape details, and "swimming" effects where stationary objects appear to move as the participant changes their viewpoint. The literature provides much speculation

about the sources of these artifacts, but little concrete evidence to their causes or suggestions on how to best address them [28]. Our goal is to improve the VR viewing experience by mitigating these artifacts. In this paper, we consider how to make such improvements by generating a novel method to determine viewing parameters.

Conventional wisdom suggests that the plasticity of the perceptual system can accommodate errors in the viewing parameters. Therefore, standard practice in Virtual Reality typically uses the participant's measured inter-pupillary distance (IPD) as the binocular disparity (BD), and uses fixed values for the tracker offset. Indeed, many VR systems simplify these parameters further: using nominal standard values for IPD (and therefore BD), and/or ignoring the center-of-projection offset (PO). Empirically, these simplifications have not been shown to be problematic: the limited experiments in the literature fail to show sensitivity in these parameters.

Our premise is that having proper viewing parameters is important to the VR viewing experience. If we can determine the viewing parameters (PO , BD) properly, we can reduce many of the troublesome viewing artifacts in virtual reality. We reconcile this premise with the prior literature by noting that the limited prior attempts to improve the viewing experience through determination of viewing parameters simply did not provide adequate quality in the estimation of the parameters. Therefore in this paper we consider both parts of the viewing parameter problem. First, we introduce a perceptual calibration procedure that can accurately determine participant-specific viewing parameters. Second, we introduce a series of measurements to quantify virtual reality viewing artifacts. These measurements are used in a study to determine the effect of the perceptually calibrated viewing parameters.

One of our key observations is that the viewing parameters are specific to the participant and internal. Even if we were to locate the optical centers of the participant's eyes, they would be somewhere inside their head and could not be measured from external observation. In actuality, we do not believe that there is a simple optical center for the eye: the pinhole camera (or thin-lens) approximation is simplistic, and the actual viewing geometry is only part of the interpretation by the visual system. Therefore it is important to determine the parameters *psychophysically*. That is, we need to make subject-specific measurements to determine the parameters based on what they perceive.

This paper has three key components. First, we provide a perceptual calibration procedure that can accurately determine the subject-specific viewing parameters by asking them to perform a series of adjustments. Second, we describe a series of measurement techniques that allow us to empirically quantify several of the problematic artifacts in VR. Third, we describe an experiment that uses the artifact quantification measurements to show the improvement of using our perceptual calibration procedure over the standard practice of using physical measurements to approximate viewing parameters.

Our contributions include:

- A novel method to psychophysically calibrate an immersive display environment that has been empirically demonstrated to reduce viewing artifacts;
- A novel set of methods to quantify the viewing artifacts in a virtual environment;
- Evidence of the importance of accurately determining viewing parameters for the VR experience, providing an explanation for some of the causes of these artifacts.

2 Background

Much work has focused on general human vision, perception, and psychology [6, 30, 34]. Our work is primarily focused on virtual environments and is motivated from several different areas described below.

2.1 Geometric Models for Virtual Environments

While geometric models of viewing are known, very little research has gone into understanding how these models affect perception in virtual environments. Woods et al. used geometric models to describe the distortions with stereoscopic displays [37]. Pollock et al. studied the effect of viewing an object from the wrong location, meant to better understand the perception of non-tracked users in a virtual environment [23]. They found that the errors were in fact smaller than what the models would have suggested.

Banks et al. showed that these models could predict perceptions in virtual environments and users were unable to compensate for incorrect viewing parameters [1]. Didyk et al. created a perceptual model for disparity in order to minimize disparity while still giving a perception of depth [3]. Held and Banks studied the effects when these models fail, such as when the rays and retinal images do not intersect [7].

We use these geometric models in a very different way from the previously mentioned work. Instead of using these models to determine a perceived point in space, we instead enforce a perceived point in space and inversely compute the correct viewing parameters. For example, while Pollock et al. used geometric models to determine the degradation the virtual experience for non-tracked individuals, we instead attempt to improve the experience for the individual who is tracked.

2.2 Calibration of Virtual Reality Systems

Commonly, Virtual Reality systems attempt to calibrate BD , while less interest has been shown in attempting to calibrate PO . While some research has shown that incorrect BD can be adapted to [35] and does not significantly change a user's experience in a virtual environment [2], other research has shown this is not the case. As an example, Utsumi et al. were able to demonstrate that the perception of distance did not show adaptation when a participant was given an incorrect BD [31]. Furthermore, it has been shown that for tasks in which the user must switch back and forth between the virtual and physical environment, perceptual mismatches can lead to detrimental performance and leave the participant in a disoriented state [33].

Other research has analyzed more calibration mechanisms outside of BD . For instance, Nemire and Ellis used an Open-Loop pointing system to attempt to calibrate an HMD system [20]. Others have used optical targets to calibrate gaze tracking systems [21]. The majority of work for the calibration of Virtual systems has mostly been studied in the field of Augmented Reality [5]. Tang et al. provide evaluations of different techniques for calibrating augmented reality systems [29].

Recently, Kellner et al. used a point and click method for a geometric determination of viewing parameters for a see-through augmented reality device [10]. While Kellner et al. were able to show a great deal of precision in determining the position of the eyes, these viewing parameters were only able to aid in distance estimation tasks, while depth underestimation was still prominent.

While Kellner et al. were able to improve greatly on the precision of geometric calibration techniques compared to previous work, Tang et al. suggested that errors in human

performance tend to make these types of methods very error prone [29]. Additionally, Tang et al. suggest that calibration procedures for two eyes should be designed to ensure that there is no bias towards either of the eyes. We use these recommendations in the design of our perceptual calibration procedure that utilizes perceptual alignment described in Section 3. This method removes human psychomotor errors and as shown by Singh et al., humans can perform alignment tasks with very little error (less than one millimeter) in the physical environment [27].

2.3 Perceptual Judgment in Virtual Reality

Artifacts occur in Virtual Reality when the participant's experience does not match their expectations. We describe three well-known artifacts and methods for quantifying them.

2.3.1 Depth Acuity—The ability of subjects to perceive objects in virtual spaces has been shown to produce discrepancies between the physical and virtual worlds. Mon-Williams and Tresilian determined depth estimations using reaching tasks [19]. Rolland et al used a forced choice experience to determine depth acuity [26]. Liu et al. monitored the movements of subjects locating nearby targets, comparing the time to completion for virtual and physical tasks [17]. Singh et al tested the depth acuity for near field distances of 0.34 to 0.5 meters [27] comparing both physical and virtual environments. Singh et al. found that while users were very good at perceptual matching techniques in physical environments, but tended to over-estimate in the virtual case.

Our measure of depth acuity, described in Section 4.5.1, uses object alignment, similar in concept to [4] and [27]. We built our testing setup to account for the recommendations of Kruijff et al. [14]. As opposed to using a single point in space, we used a disparity plane to determine the perceived position from the user. Additionally we measured another VR artifact, we call “swimming”. While swimming is often associated with tracker lag, we define swimming more broadly, as occurring whenever a static object appears to change position as the participant's viewpoint changes. We quantitatively measure this as the change in distance of a static virtual object changes as the participant moves in the virtual space. To our knowledge, this artifact, while prominent, has not been quantified inside of a virtual environment.

2.3.2 Distance Estimation—Distance estimation has been well studied in virtual environments [9, 10, 15, 28, 36, 38]. Many studies have found that depth compression occurs in virtual environments, resulting in objects appearing at a closer distance [15, 36]. The principal factors responsible for distance compression have remained relatively unknown [28].

Unfortunately, many of the previous methods to monitor distance estimation were unfeasible given the limited space of our immersive display environment (described in Section 4). While Klein et al. attempted to use triangulated blind walking as a means to test walking in small spaces, their results were very mixed [12]. Based on our constraints we designed a novel method of measuring distance estimation that does not rely on locomotion, described in Section 4.5.2.

2.3.3 Perception of Shape—The ability of subjects to perceive the shape of objects in virtual spaces has been shown to produce discrepancies between the physical and virtual worlds. Luo et al. studied size consistency in virtual environments [18] by having subjects attempt to match the shape of a virtual Coke bottle using a physical bottle as reference. Kenyon et al. also used Coke bottles to determine how well subjects could determine size consistency inside of a CAVE environment finding that the surrounding environment

contributed a significant effect [11]. Leroy et al. found that increasing immersion, such as adding stereo and head tracking improved the ability of subjects to perceive shape [16].

Our experiment uses some components of these shape matching experiments in the sense that the subject is able to interactively change the virtual shape of an object using a wand device. However our test does not use a reference object, instead the information needed to determine the shape of the object is determined inside of the virtual environment. For our experiment, described in Section 4.5.3, the participant is given a size in one axis of an object and must reshape the object so that it is the same size in all dimensions. This allows the testing of many different sized objects rapidly and removes any issues of familiarity the subject may have about the shape of an object.

3 Perceptual Calibration

In this section, we discuss our perceptual calibration method. The goal of our calibration technique is to find a proper PO and BD such that we account for an individual's perception of the virtual space. As these values are internal to the participant, they can not be physically measured. Instead we approach this problem using perceptual geometric models.

Constructing perceptual geometric models for immersive display environments is relatively simple for a single point in space as shown in Figure 2. First the intersection of a ray is constructed between each virtual camera and the point in space. Next the intersection of each ray and the projection screen is found. Finally, two new rays are constructed from these intersection points to the correct location of the eyes (shown in orange in Figure 2). Where these two rays intersect determines the perceived location of the point.

In this model, two factors are not inherently known: the position of the perceived point and the correct viewing parameters. As our motivation is to determine the correct viewing parameters, we can force the perceived point to a desired position. As we utilize both virtual and physical objects, our methods are designed specifically for CAVE systems. Future work will aim to provide methods for other types of VR systems.

By aligning the virtual object to a known physical location, we can now determine the position of the perceived point in space. Unfortunately, simply knowing the position of this point in space does not provide enough information to fully determine the viewing parameters. As shown in Figure 3, the correct position of the eyes could lie anywhere along the rays intersecting the point. We call this the BDZ Triangle which we will use for calibration purposes described below.

In the remainder of this section we provide series of steps used to determine the viewing parameters PO and BD based on perceptual experiments. First PO_x and PO_y are calibrated in the environment. Second, a BDZ triangle is generated to determine the relationship of BD and PO_z for the participant. This triangle is an important intermediate step to determining the correct viewing parameters. Finally, using this relationship the calibrated BD and PO_z values are determined based on the reduction of viewing artifacts.

3.1 Calibrating PO_x and PO_y

The first step in our calibration procedure is to determine the viewing parameters such that the height (PO_x) and left-right shift of the eyes (PO_y) are perceptually aligned. As point and click calibration procedures have been shown to be error prone [29], we instead derived a new calibration procedure that utilizes two aspects of the CAVE system:

- While physically discontinuous, the CAVE environment is meant to be perceptually continuous (i.e. items should not appear different when projected between different screens).
- As there is no disparity for virtual objects positioned to lie on the physical display screen, their depth will appear aligned independently of how well the system is calibrated.

Using these aspects we constructed a virtual scenario with two infinitely long posts, one vertically aligned and one horizontally aligned. The posts are positioned so that they were at the front CAVE wall, making all calibration adjustments affect the surrounding walls only. When the participant looks at the horizontal post, any incorrect calibration in PO_y causes the post to appear to bend up or down at the boundary of the CAVE walls. The participant can interactively change the PO_y parameter until the post appears straight. The participant is given a straight edge for comparison, as shown in Figure 4.

In this case, incorrect calibrations in PO_z may also affect this perception of the vertical bend, although most of the perceptual errors generated by PO_z make the post appear to slant towards the participant. To compensate for this, we return to this calibration after PO_z is determined. PO_x is calibrated similarly to PO_y using the horizontal bend of the vertical post.

3.2 Constructing the BDZ Triangle

After temporally fixing PO_x and PO_y , we can then construct the BDZ triangle based on the participant's perception. As mentioned earlier, a BDZ triangle describes the relation between BD and PO_z in which a point is perceived at the same position for all values. Understanding this relationship enables a means to determine perceptually correct viewing parameters. To construct a BDZ triangle correct for a single point in space, we empirically determine a set of PO_z values corresponding to a set of sampled BD values that provide correct perception of the point.

We used a physical alignment object to enforce the position of the perceived point in the model. To construct the BDZ Triangle, we first position a physical box 0.51m wide \times 0.65m tall \times 0.1m deep at a fixed position of 0.26 meters away from the front CAVE screen and position the the participant 1.18 meters away from the front CAVE screen, directly behind the object. We draw a virtual vertical plank that contains a wood texture that provides enough features to make ocular fusion easier for the participant. This plank is virtually positioned at a location that matches the alignment object.

While the plank is virtually positioned such that it is always aligned with the physical object, the viewing parameters dictated whether or not this was perceptually the case. Through the determination of what viewing parameters allow the virtual plank to appear at the correct position, the BDZ Triangle described above can be empirically derived.

For each trial, a BD is procedurally assigned and the participant is able to interactively change PO_z . For the participant, this created the effect that the plank was moving closer and farther from the participant, even though the virtual position of the plank was unchanged. Once the participant had completed the alignment objective, their corresponding BD and PO_z information was fed into a linear regression model and a new trial was given.

After 5 trials, the error of the linear regression model was analyzed. If the R^2 value was shown to be greater than 0.9, the model was considered to contain an acceptable error and the participant was moved to the next calibration step. If the R^2 value was shown to be less than 0.9, the participant was given four extra trials to attempt to improve the model. If after nine trials the R^2 value was shown to be greater than 0.5, the participant was moved to the

next step, but the model was recorded as not well fitted and therefore not well trusted. If after nine trials the R^2 value was shown to be less than 0.5, the trust that the linear regression model would be able to provide correct viewing parameters was considered too low for the subsequent calibration. In this case, the test was reset and the participant was allowed to try again.

3.3 Calibrating PO_z and BD

With the relationship between PO_z and BD now known for a given participant, the perceptually correct viewing parameters could be determined. As these viewing parameters were meant to reduce perceptual artifacts, we chose to use a test which utilized the negative aspects of incorrect calibration. As defined above, swimming artifacts are created when a static virtual object appears to move as a participant changes their viewing position. Figure 5 shows how these artifacts also occur for rotations of the head. We note that these swimming artifacts are perceptually heightened for rotations of the head as a participant has a strong expectation for the optical flow of a static object.

We presented participants with two rows of virtual pillars in order to give the participants many distinct points in space in which they monitor the swimming artifact. The participant was instructed to rotate their head from left and right and to notice the perceived motion of the pillars. While continuing to rotate their head, the participant was able to interactively modify their viewing parameters sampling their derived BDZ Triangle. When the pillars were perceived to no longer be swimming, the viewing parameters were stored. With these newly determined PO_z and BD values, the participant was again presented with the posts described in Section 3.1 so that the PO_x and PO_y in order to fine tune these parameters. This process of determining PO_z and BD was repeated three times in order to ensure consistency of the calibrated viewing parameters.

4 Experiment

We hypothesize that with perceptually calibrated viewing parameters, subjects in immersive display environments will have improved depth acuity, distance estimation, and an improved perception of shape. To test this hypothesis, we created an experiment in which we could measure each of these common viewing artifacts.

4.1 Equipment

The study was administered in a 2.93m×2.93m×2.93m six-sided CAVE consisting of four rear-projection display walls, one solid acrylic rear-projection floor, and one rear-projection ceiling. Two 3D projectors (Titan model 1080p 3D, Digital Projection, Inc. Kennesaw, GA, USA) with maximum brightness of 4500 lumens per projector, with total 1920 × 1920 pixels combined, projected images inside each surface of the CAVE. The displayed graphics were generated by a set of four workstations (2 × Quad-Core Intel Xeon). Audio was generated by a 5.1 surround sound audio system.

The data acquisition system consisted of an ultrasonic tracker set (VETracker Processor model IS-900, InterSense, Inc. Billerica, MA, USA) including a hand-held wand (MicroTrax model 100-91000-EWWD, InterSense, Inc. Billerica, MA, USA) and head trackers (MicroTrax model 100-91300-AWHT, InterSense, Inc. Billerica, MA, USA). Shutter glasses (CrystalEyes 4 model 100103-04, RealD, Beverly Hills, CA, USA) were worn to create stereoscopic images. The head trackers were mounted on the top rim of the shutter glasses.

4.2 Subjects

Subjects were recruited with flyers and announcements and were not given monetary compensation. To ensure that subjects could perform the calibration procedure, a TNO Stereopsis Test was administered [32]. The subject's height, height of the subject's eyes, and interpupillary distance were all measured and recorded. Subject's age, gender, dominant hand, and eye correction were also recorded. Subjects were also asked to report any previous experiences with virtual reality and 3D movies, issues with motion sickness and exposure to video games.

The study consisted of 20 subjects, 10 female and 10 male between the ages of 19 and 65, with an average age of 29. Of the 20 subjects, 10 of them wore glasses, five wore contacts, four had 20/20 vision and one had previously had Lasik eye surgery. On top of the 20 subjects reported in the paper, two other subjects were recruited, but failed to meet our screening criteria for depth perception.

4.3 Conditions

The experiment consisted of three different conditions. The first condition, labeled the **measured** configuration, used viewing parameters in which BD was set to the measured IPD value and the PO_z was set to zero. As the tracker rested very close to the subject's forehead, this zero point represented a point that was very close to the front of the subjects eyes. The second condition, labeled the **calibrated** configuration, used BD and PO_z values derived from the perceptual calibration procedure described in Section 3. The final condition, labeled as the **inverse** configuration, was used to bracket the measured and calibrated configurations. For this configuration the offset from both BD and PO_z from the calibrated to measured configuration is found and is doubled to set the viewing parameters (i.e. $inverse_{BD} = 2 \times calibrated_{BD} - measured_{BD}$)

4.4 Hypothesis

From these conditions, our hypothesis is the psychophysically calibrated configuration will show smaller perceptual errors than the measured or inverse configurations. Furthermore, for directional error we expect that the calibrated configuration will be bracketed between the measured and inverse configurations.

4.5 Procedure

The study was a repeated measures randomized block design with a block size of three trials. In this way, each condition was shown for each group of three trials, but the order in which the conditions were shown were randomized. This design was selected to detect and mitigate any effects of learning during the experiment. All subjects were presented with the tests in the same order. The subject was first presented with the first iteration of the depth acuity test, then the distance estimation test, then the first iteration of the shape perception test, then the second iteration of the depth acuity test before finishing with the second iteration of the shape perception test. The specifics of the tests are described below.

4.5.1 Depth Acuity—This experiment tested our hypothesis that the Calibrated configuration would improve depth acuity. To determine this we had each subject position a virtual board against a physical alignment object from four different positions as shown in Figure 6. The subjects used the joystick to translate the virtual board backwards and forwards until they felt it was correctly positioned. While the goal of aligning a virtual object with a physical one was identical to the goal in the calibration step (Section 3.2), the method to achieve this was quite different. In this case, the object moved as opposed to

changing the viewing parameters. Most subjects reported that this movement was much more natural.

The first iteration of this test aligned the physical alignment box (described in Section 3.2) 0.26 meters away from the front wall of the CAVE. The subject was positioned at a distance of 0.88 meter from the alignment object (Location 1). The subject positioned the board for nine trials, three times for each method. The subject was then moved back 0.8 meters (Location 2) and the process was repeated for nine trials, three trials for each configuration.

In the second iteration of this test, the subject and the alignment box were moved to new locations as shown in Figure 6. The box was moved to a position of 0.72 meters away from the CAVE screen and the subject was moved to a closer distance of only 0.62 meters away from the alignment box (Location 3). After nine trials, the subject was moved back 0.8 meters (Location 4) and for nine more trials were given, three trials for each configuration.

On top of determining the subjects ability to correctly position objects, this test was also designed to test for swimming, defined in Section 2.3.1. To quantify this result across all subjects, we first found the median value for their three trails and then found the change in the median for each configuration per location per subject. We then found the average difference between corresponding alignment positions from the two different locations.

4.5.2 Distance Estimation—This experiment was used to test our hypothesis that the calibrated configuration would improve the estimation of distance. Unfortunately, the size of the CAVE prevented many measurement devices from being used.

Based on our constraints we developed a new method of testing distance estimation, using the idea of measuring the incline of the perceived flat (i.e. level) surface. The concept, shown in Figure 7, is that a flat surface will appear to incline upwards if depth compression is occurring as the surface moves away from the viewer. Depth expansion, on the other hand, would cause the flat surface to appear to inclined downwards. In this way, only when the system is correctly configured will a flat surface appear flat.

We believe this alternative test for distance estimation has many advantageous features. For one, the test does not require any locomotion from the subject, thus enabling depth comparisons in small virtual environments. Secondly, there is a great deal of literature on the perception of slant for objects such as hills [24, 25]. Proffitt et al. provide a good overview on the subject [25]. Likewise, the components of the perception of inclination have been studied by [8, 13]. Finally, research has shown that angular declination is a valid perceptual tool for subjects to determine distances in physical environments [22]. Additionally, research by Ooi et al. were able to show that shifting the eyes with prisms could alter the perception of distance, mirroring the results seen for distance estimation in virtual environments.

The experiment created a scenario in which the subject looked out on five horizontal beams of size 4.75m wide \times 0.46m tall \times 0.52m deep (Figure 8). The beams were positioned such that each was one meter behind the previous with a random perturbation of half a meter (± 0.25 m). The beams were also randomly perturbed by \pm one meter left and right to prevent the subject from using perspective cues. The front beam always stayed in the same position with its top edge at a height of 0.72m from the floor. The subject iteratively is able to change the position of the beam indirectly by changing the angle of inclination which in turn set the beams height based on the product of their distance and tangent of the inclination angle. As an aid, the subjects were also given a physical box matching the height of the front beam. The test consisted of one iteration of nine trials, three trials for each viewing configuration.

Additionally we note that our initial iteration of the distance estimation test had subjects change the tilt of a planar surface until it appeared to be level. Unfortunately, the optical flow from the rotation added additional cues for depth, thus creating the situation of cue conflict. Therefore we created the previously mentioned experiment described in Section 4.5.2 that used five horizontal beams. Even though the vertical position of these beams was determined by the inclination angle, the optical flow corresponding to a rotation was thusly removed.

4.5.3 Shape Perception—This experiment was used to test our hypothesis that the calibrated configuration would improve the perception of shape. To accomplish this, we showed the subject a floating block of wood in the middle of the CAVE environment (Figure 9). The depth of the block of wood (z direction) was fixed, while the subject was then able to resize the width and height of the board using the joystick. The goal for the subject was to resize the block of wood such that it was a cube (i.e. the same size in all dimensions). The depth, and therefore subsequent size, was randomly set for blocks between 0.15 meters and 0.6 meters in each dimension.

The error for each trial was determined by looking at the L2 Norm of the vector from the desired cubic corner and the provided result (Equation 1).

$$\text{Error} = \sqrt{(l_x - l_z)^2 + (l_y - l_z)^2}$$

For the first iteration, the subject was put at a fixed location at a distance of 1.1 meters away from the floating block. For the second iteration of the test, the subject was allowed to move freely around the environment. Each iteration consisted of nine trials, three trials for each viewing configuration.

5 Results

We will first present the results from the calibration procedure before presenting the results of each experiment.

5.1 Calibration

Subjects first completed the perceptual calibration procedure (Section 3) before undertaking the viewing artifact tests. We found that subjects were able to complete the calibration procedure on average in 7 minutes and 18 seconds with an average of 14 trials for subject to generate a linear regression model. While the difficulty of the calibration varied greatly between subjects, all 20 of the subjects were able to complete the calibration procedure.

5.1.1 Grouping by Trust—While all 20 subjects were able to complete the calibration procedure, some subject's calibration data had large amounts of variance. This led to the possibility that the calibrated parameters were not fully optimized. We therefore break our subjects into three separate groups based on the *trust* of the calibration using the R^2 generated from our linear regression model described in Section 3.2.

The first group consisted of those subjects in which we were only moderately trusting their calibration. We defined this group as any subject who had produced an $R^2 > 0.5$, which included all 20 subjects. The second group consisted of those subjects in which we were highly trusting about their calibration, consisting of the set of subjects who had produced an $R^2 > 0.9$. This group included 11 of the original 20 subjects. Of these subjects, three wore contacts, five wore glasses, and three reported 20/20 vision. The final group consisted of

subject in which we were exceptionally trusting about their calibration, consisting of the subjects who had produced an $R^2 > 0.95$. This group included five of the original 20 subjects. Of these subjects, three wore contacts, one wore glasses, and one reported 20/20 vision

Figure 10 shows the distribution of the viewing parameters labeled for these groupings. For the entire subject pool, we found an average PO_z of 96mm a BD of 83mm, on average 24mm wider than their measured IPD. For the subjects who were in the highly trusted group we found an average PO_z of 105mm and an average BD of 76mm, on average 16 mm wider than their measured IPD. For the subjects who had were in the exceptionally trusted group we found an average PO_z of 86mm and an average BD of 75mm, on average 15mm wider than the measured IPD. The adjustment of PO_x and PO_y after PO_z was shown to be very small, each on average less than a millimeter.

5.2 Depth Acuity

We test our hypothesis that subjects will be able to more accurately position objects when correctly calibrated. Figure 11 shows the average positioning errors for all locations. We found these errors to be significantly related to the configuration ($F(2,176) = 101.1$, $p < 0.001$). No significant learning effects were observed (i.e. errors not significantly smaller for later blocks and trials). However, as a significant trend was found between the trust of the calibration and performance, we present the results of each group based on trust separately.

For the moderately trusted group, we found that on average subjects overestimated the depth by 27mm for the calibrated configuration, over-estimated the depth by 51mm for the inverse configuration, and underestimated the depth by 21mm for the measured configuration. TukeyHSD analysis showed the difference in error between calibrated and measured configuration was highly significant ($p < 0.001$) and the difference in error between the calibrated and inverse was also highly significant ($p < 0.001$).

For the highly trusted group, we found that on average subjects overestimated the depth by 13mm for the calibrated configuration, over estimated the depth by 38mm for the inverse configuration, and underestimated the depth by 23mm for the measured configuration. TukeyHSD analysis showed the difference in error between calibrated and measured configuration was significant ($p < 0.01$) and the difference in error between the calibrated and inverse was highly significant ($p < 0.001$).

For the exceptionally trusted group, we found that on average subjects overestimated the depth by 7mm for the calibrated configuration, over estimated the depth by 36mm for the inverse configuration, and underestimated the depth by 33mm for the measured configuration. TukeyHSD analysis showed the difference in error between calibrated and measured configuration was highly significant ($p < 0.001$) and the difference in error between the calibrated and inverse highly significant ($p < 0.001$).

5.2.1 Swimming—We test our hypothesis that static virtual objects will appear to move less between different viewing position when the viewing parameters are correctly calibrated. In our test we generated to pairs in which we can determine this swimming measurement.

As the distance moved from Location 1 to 2 and from Location 3 to 4 were identical, we combine the results together to find an average difference in position (Figure 12). We find that the position change was significantly related to the configuration ($F(2,117)=12.69$, $p < 0.001$). TukeyHSD analysis showed the difference between calibrated and measured

configuration was significant ($p < 0.01$) while the difference in error between the calibrated and inverse was not significant ($p > 0.10$). No significance was found between the trust of the calibration and performance and no learning effects were observed.

5.3 Distance Estimation

We test our hypothesis that subjects will only be able to correctly position the beams without inclination when they are correctly calibrated. We found the inclination of the positioned beams to be significantly related to the configuration ($F(2,177) = 17.22$, $p < 0.001$). The calibrated configuration significantly provided the most accurate perception of inclination as shown in Figure 13. Given the calibrated configuration, subjects were able to position the beams with an average inclination of 0.3 degrees when they were perceived to be level. On the other hand, the measured configuration resulted in subject's positioning the beams with an average declination of 1.7 degrees. When given the inverse configuration, the beams were positioned with an average inclination of 1.1 degrees. TukeyHSD analysis showed the difference between calibrated and measured configuration was highly significant ($p < 0.001$) and the difference in error between the calibrated and inverse was not significant ($p > 0.10$). No significance was found between the trust of the calibration and performance and no learning effects were observed.

5.4 Perception of Shape

We test our hypothesis that subjects will be able to more accurately understand the shape of an object when correctly calibrated. For the first iteration of the test in which the subject was put at a fixed location, the error, shown in Figure 14, was not significantly related to the configuration ($p > 0.10$).

As expected, the second iteration in which the subject was allowed to move freely in the environment produced definitively smaller errors for all three configurations as shown in Figure 14. In this iteration we found the error was significantly related to the configuration ($F(2,176) = 3.499$, $p = 0.03$). The average error for the calibrated configuration was 54mm, compared to 70mm for the inverse configuration and 69mm for the measured configuration. TukeyHSD analysis showed the difference in error between calibrated and measured configuration was marginally significant ($p = 0.06$) and the difference in error between the calibrated and inverse was significant ($p = 0.05$). No significance was found between the trust of the calibration and performance and no learning effects were observed.

6 Discussion

We divide our discussion into two separate subsections in which we will discuss the results of our experiments, considerations, and future work.

6.1 Discussion of Results

The results of our experiment correspond with our hypothesis demonstrating that these perceptually calibrated viewing parameters do improve a participant's virtual experience.

Unfortunately, while all subjects were able to complete the calibration procedure, the BDZ Triangle step from Section 3.2 proved to be very difficult for some subjects. The effect of changing PO_z interactively was difficult for some participants as the movement of objects understandably felt very strange and unnatural. For the purposes of time in the experiment, we had allowed subjects to continue the experiment even with a lower confidence that was not ideal.

The demographic information we acquired provided little insight into what may have caused the discrepancy of difficulty for calibration across subjects. Factors such as age, gender, eyewear, measured IPD, height, previous VR experience, and gaming experience were not shown to be significantly correlated with R^2 value. The only piece of demographic information which provided any kind of a marginal correlation was between the R^2 value and the subject's indication of how prone they were to motion sickness ($F(1,18)=3.097$, $p=0.09$). In this case, the subjects who indicated they were prone to motion sickness averaged an R^2 value of 0.79 compared to those who were not who averaged a R^2 of 0.89.

The results of the depth acuity test matched our hypothesis that the calibrated configuration would provide a better depth acuity, although this effect was only true when the confidence of the calibration was taken into account. The trend between confidence and performance on this test was shown to be highly significant for the calibrated and inverse configurations, but not significant for the measured configuration. This in turn means that this exclusion for subjects did not simply represent those who were better than this test.

This also proved to be the only test in which these confidence groups provided a significant difference in the error measurements. We believe this is because while many of the other tests utilized more relative judgements, this was the only test in which absolute judgements were essential, in turn making poorly configured viewing parameters very error prone.

While the swimming test did show significance between the configurations, the improvement for the calibrated configuration was marginal. However, the fact that the calibrated configuration was bracketed around the measured and inverse configuration gives us confidence that with improved calibration and repeated measures this effect would become more pronounced.

The results of the distance estimation test match our hypothesis that the calibrated configuration would improve the estimation of distance. From the results of Ooi et al. [22], we can project that faraway virtual objects in measured configuration will appear closer than they should (depth compression), while faraway virtual objects will appear farther than they should for the inverse configuration (depth expansion). The calibrated configuration would provide a perception in the middle of these two, with objects appearing ever so slightly farther than they should. The fact that the calibrated configuration was not shown to be significantly different from the inverse configuration was likely due to the fact that depth compression and depth expansion are not symmetrical, with depth expansion being a far weaker effect.

The results of the shape perception test match our hypothesis that the calibrated configuration would improve the perception of shape, however we only found significance when the subject was allowed to move around. This could be because when the subject was put in a fixed location, the shape was determined mostly through the judgement of the distances to corners. However, when the user was allowed to move around, the effects of swimming could cause shapes to appear to change when viewed from different positions. In this regard, it seems possible that swimming may play a larger role than precise depth acuity in terms of artifacts from the judgement of shapes in virtual environments.

6.2 Considerations and Future Work

While these perceptually calibrated viewing parameters have been shown to produce an improved virtual experience, the question can be raised of what mechanisms cause these viewing parameters. For many of the subjects, the PO_z value corresponds to a position that is far deeper than one would expect the focal point of eye. It is worth noting that was no

correlation found between this PO_z and other demographic information. Future work will explore the mechanisms that may cause these perceptual viewing parameters.

It would also be of interest to test the difference between the perceptual calibration and the average value found for the highly calibrated group. While our model would suppose that a default offset in IPD would be more appropriate than a default static IPD, the standard deviation between the two was insignificant. In this regard it would be interesting to also test between these two cases.

While the BDZ Triangle corresponds to a single point in space, a different point in space would have a uniquely different BDZ Triangle. In theory, as opposed to using the swimming artifact model to determine BD and PO_z , a series of BDZ Triangles could be constructed for many different points in space. The intersection of these triangles would correspond to the perceptually correct position of the eyes. While initial testing utilized this approach, it was found that small inconsistencies produced large discrepancies in the derived viewing parameters.

While our method is able to account for incorrect positioning of the tracking system, incorrect orientations of the tracking system are currently not accounted for. Initial testing using point and click models found that the rotational inaccuracies were minimal when the tracking device was mounted on top of the stereo glasses. However, in other tracking setups, these errors may be much larger. Future work will aim to create a perceptual experiment to determine these rotational components without point and click methods.

While the current set of methods utilizes the physical discontinuities of the CAVE environment, our calibration method can be easily adapted for tracked stereoscopic displays. By showing linear objects, such as posts, that virtually transition in-and-out of a screen, calibration offsets can still be perceived. Future work will examine alternative methods for these types of devices.

Unfortunately, as many of the methods in this paper use a mix of physical and virtual objects, they are not well suited for Head Mounted Display devices. Additional considerations would need to be taken into account for a similar perceptual calibration technique for these types of devices. However, this calibration technique would be well suited for augmented and mixed reality devices. Furthermore, we believe that perceptual calibration may be necessary for augmented and mixed reality systems to reach their full potential.

While our approach aimed to create a quantitative metric for virtual experiences, it would also be interesting to determine the subjective changes for a participant's experience when using different viewing configurations. For instance, it would be useful to study how perceptual calibration effects the sense of presence a participant has while inside of a virtual environment. Furthermore, studying if perceptual calibration could mitigate negative effects, such as simulator sickness, could prove incredibly useful for the greater VR community at large.

7 Conclusion

We present a perceptual calibration procedure that is derived from geometric models. Our method enables participants to empirically determine perceptually calibrated viewing parameters. Through a 20 subject study we are able to observe these calibrated viewing parameters are able to significantly improve depth acuity, distance estimation and the perception of shape. Future work will aim to create new methods to more easily determine

these perceptually calibrated viewing parameters and will examine how else these viewing parameters affect a participant's virtual experience.

Acknowledgments

We would like to acknowledge the support of the the Living Environments Laboratory, the Wisconsin Institute for Discovery, and the CIBM training program. This project was supported in part by NLM award 5T15LM007359 and NSF awards IIS-0946598, CMMI-0941013, IIS-1162037, and the Industrial Strategic technology development program(No.10041784, Techniques for reproducing the background and lighting environments of filming locations and scene compositing in a virtual studio) funded by the Ministry of Knowledge Economy (MKE, Korea). This paper was also completed with Ajou university research fellowship of 2009. We would specifically like to thank Andrew Morland for his assistance with CAVE activities and Patricia Brennan, Kendra Kreutz, and Ross Tredinnick for their support and assistance as well.

References

- [1]. Banks M, Held R, Girshick A. Perception of 3-d layout in stereo displays. *Information display*. 2009; 25(1):12. [PubMed: 21687822]
- [2]. Best, S. Perceptual and oculomotor implications of interpupillary distance settings on a head-mounted virtual display. *Aerospace and Electronics Conference, 1996. NAECON 1996., Proceedings of the IEEE 1996 National; may 1996; p. 429-434.vol.1*
- [3]. Didyk P, Ritschel T, Eisemann E, Myszkowski K, Seidel H. A perceptual model for disparity. *ACM Trans. Graph.* 2011; 30(4):96.
- [4]. Ellis S, Menges B. Localization of virtual objects in the near visual field. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 1998; 40(3):415–431.
- [5]. Fuhrmann, A.; Schmalstieg, D.; Purgathofer, W. Fast calibration for augmented reality. *Proceedings of the ACM symposium on Virtual reality software and technology, VRST '99; New York, NY, USA. ACM; 1999. p. 166-167.*
- [6]. Gregory, R. *Eye and brain: The psychology of seeing*. Princeton university press; 1997.
- [7]. Held, RT.; Banks, MS. Misperceptions in stereoscopic displays: a vision science perspective. *Proceedings of the 5th symposium on Applied perception in graphics and visualization, APGV '08; New York, NY, USA. ACM; 2008. p. 23-32.*
- [8]. Hillis J, Watt S, Landy M, Banks M. Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*. 2004; 4(12)
- [9]. Jones, JA.; Swan, JE., II; Singh, G.; Kolstad, E.; Ellis, SR. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. *Proceedings of the 5th symposium on Applied perception in graphics and visualization, APGV '08; New York, NY, USA: ACM; 2008. p. 9-14.*
- [10]. Kellner F, Bolte B, Bruder G, Rautenberg U, Steinicke F, Lappe M, Koch R. Geometric calibration of head-mounted displays and its effects on distance estimation. *Visualization and Computer Graphics, IEEE Transactions on*. 2012; 18(4):589–596.
- [11]. Kenyon R, Sandin D, Smith R, Pawlicki R, Defanti T. Size-constancy in the cave. *Presence: Teleoperators and Virtual Environments*. 2007; 16(2):172–187.
- [12]. Klein, E.; Swan, JE.; Schmidt, GS.; Livingston, MA.; Staadt, OG. Measurement protocols for medium-field distance perception in large-screen immersive displays. *Proceedings of the 2009 IEEE Virtual Reality Conference, VR '09; Washington, DC, USA: IEEE Computer Society; 2009. p. 107-113.*
- [13]. Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*. 2003; 43(24):2539–2558. [PubMed: 13129541]
- [14]. Kruijff, E.; Swan, J.; Feiner, S. Perceptual issues in augmented reality revisited. *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on; oct. 2010; p. 3-12.*
- [15]. Lampton, D.; McDonald, D.; Singer, M.; Bliss, J. Distance estimation in virtual environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting; SAGE Publications; 1995. p. 1268-1272.*

- [16]. Leroy, L.; Fuchs, P.; Paljic, A.; Moreau, G. Some experiments about shape perception in stereoscopic displays. Proceedings of SPIE, the International Society for Optical Engineering; Society of Photo-Optical Instrumentation Engineers; 2009.
- [17]. Liu, L.; van Liere, R.; Nieuwenhuizen, C.; Martens, J. Comparing aimed movements in the real world and in virtual reality. Virtual Reality Conference, 2009. VR 2009. IEEE; IEEE; 2009. p. 219-222.
- [18]. Luo X, Kenyon R, Kamper D, Sandin D, DeFanti T. On the determinants of size-constancy in a virtual environment. The International Journal of Virtual Reality. 2009; 8(1):43–51.
- [19]. Mon-Williams M, Tresilian J. Some recent studies on the extraretinal contribution to distance perception. Perception-London. 1999; 28:167–182. [PubMed: 10615458]
- [20]. Nemire, K.; Ellis, S. Calibration and evaluation of virtual environment displays. Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in; oct. 1993; p. 33-40.
- [21]. Ohno, T.; Mukawa, N. A free-head, simple calibration, gaze tracking system that enables gaze-based interaction. Proceedings of the 2004 symposium on Eye tracking research & applications, ETRA '04; New York, NY, USA: ACM; 2004. p. 115-122.
- [22]. Ooi T, Wu B, He Z, et al. Distance determined by the angular declination below the horizon. Nature. 2001; 414(6860):197–199. [PubMed: 11700556]
- [23]. Pollock, B.; Burton, M.; Kelly, J.; Gilbert, S.; Winer, E. The right view from the wrong location: Depth perception in stereoscopic multi-user virtual environments. IEEE Transactions on Visualization and Computer Graphics; 2012. p. 581-588.
- [24]. Proffitt D, Bhalla M, Gossweiler R, Midgett J. Perceiving geographical slant. Psychonomic Bulletin & Review. 1995; 2:409–428. 10.3758/BF03210980.
- [25]. Proffitt D, Creem S, Zosh W. Seeing mountains in mole hills: Geographical-slant perception. Psychological Science. 2001; 12(5):418–423. [PubMed: 11554677]
- [26]. Rolland J, Burbeck C, Gibson W, Ariely D. Towards quantifying depth and size perception in 3d virtual environments. Presence: Teleoperators and Virtual Environments. 1995; 4(1):24–48.
- [27]. Singh, G.; Swan, J.; Jones, J.; Ellis, S. Depth judgment tasks and environments in near-field augmented reality. Virtual Reality Conference (VR), 2011 IEEE; IEEE; 2011. p. 241-242.
- [28]. Steinicke, F.; Bruder, G.; Hinrichs, K.; Lappe, M.; Ries, B.; Interrante, V. Transitional environments enhance distance perception in immersive virtual reality systems. Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization; ACM; 2009. p. 19-26.
- [29]. Tang, A.; Zhou, J.; Owen, C. Evaluation of calibration procedures for optical see-through head-mounted displays. Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR '03; Washington, DC, USA: IEEE Computer Society; 2003. p. 161
- [30]. Thompson, W.; Fleming, R.; Creem-Regehr, S.; Stefanucci, J. Visual perception from a computer graphics perspective. AK Peters Limited; 2011.
- [31]. Utsumi, A.; Milgram, P.; Takemura, H.; Kishino, F. Investigation of errors in perception of stereoscopically presented virtual object locations in real display space. Proceedings of the Human Factors and Ergonomics Society Annual Meeting; SAGE Publications; 1994. p. 250-254.
- [32]. Walraven J, Janzen P. Tno stereopsis test as an aid to the prevention of amblyopia*. Ophthalmic and Physiological Optics. 1993; 13(4):350–356. [PubMed: 8278186]
- [33]. Wann J, Mon-Williams M. What does virtual reality need?: human factors issues in the design of three-dimensional computer environments. International Journal of Human-Computer Studies. 1996; 44(6):829–847.
- [34]. Ware, C. Visual thinking for design. Morgan Kaufmann Pub; 2008.
- [35]. Welch R, Cohen M. Adapting to variable prismatic displacement. Pictorial communication in virtual and real environments. :295–304.
- [36]. Witmer B, Kline P. Judging perceived and traversed distance in virtual environments. Presence. 1998; 7(2):144–167.
- [37]. Woods, A.; Docherty, T.; Koch, R. Image distortions in stereoscopic video systems. Stereoscopic Displays And Applications. 1993.
- [38]. Ziemer, C.; Plumert, J.; Cremer, J.; Kearney, J. Making distance judgments in real and virtual environments: does order make a difference?. Proceedings of the 3rd symposium on Applied

perception in graphics and visualization, APGV '06; New York, NY, USA: ACM; 2006. p. 153-153.

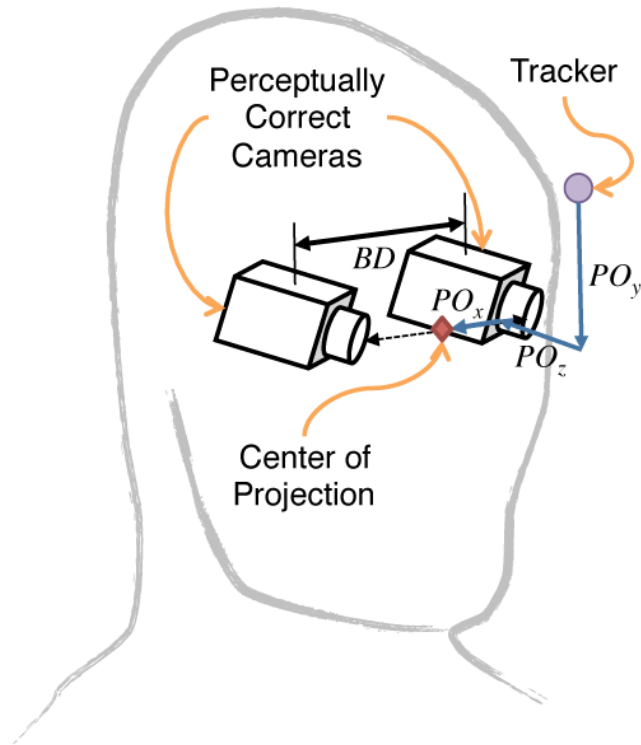


Fig. 1. The position of the eyes is described via two parameters. The binocular disparity (BD) represents the perceptual distance between the eyes. The Center of Projection Offset (PO), represents the point in the middle of the eyes. The coordinated system for these points is defined egocentrically such that the z -direction points along viewing direction.

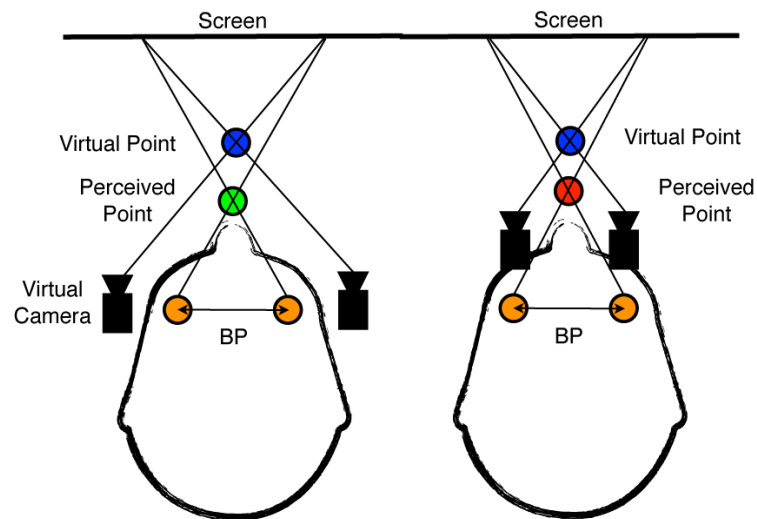


Fig. 2.

A geometric model is demonstrated for the perception for a single virtual point (shown in blue). In the case in which the virtual cameras are positioned with the incorrect Binocular Disparity (BD), the virtual point will be perceived at a different location (shown in green). Cases in which the virtual cameras are positioned with the correct BD but have the incorrect center of projection offset (PO) will also cause the virtual point to be perceived in a different location (shown in red).

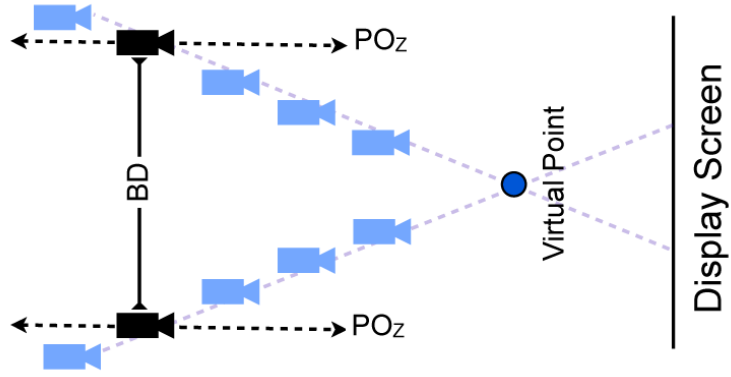


Fig. 3. BDZ triangle: the collection of camera position providing the same perceptive depth.



Fig. 4. Participant calibrating PO_x and PO_y with the help of a straight edge. The participant is able to modify PO_x and PO_y so that the posts are perceptually straight between different CAVE walls.

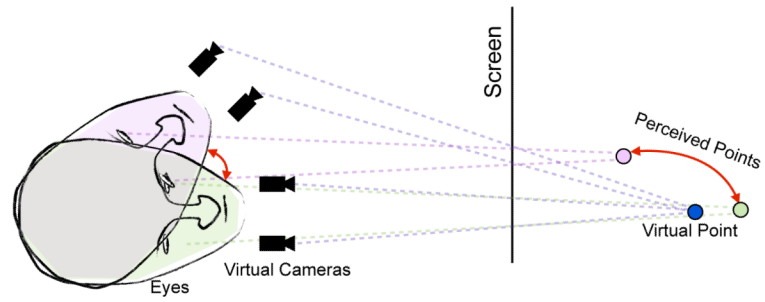


Fig. 5. Swimming artifact: A static virtual object appears to move in space as the viewer moves their head. This artifact is caused when the viewing parameters are not correctly calibrated.

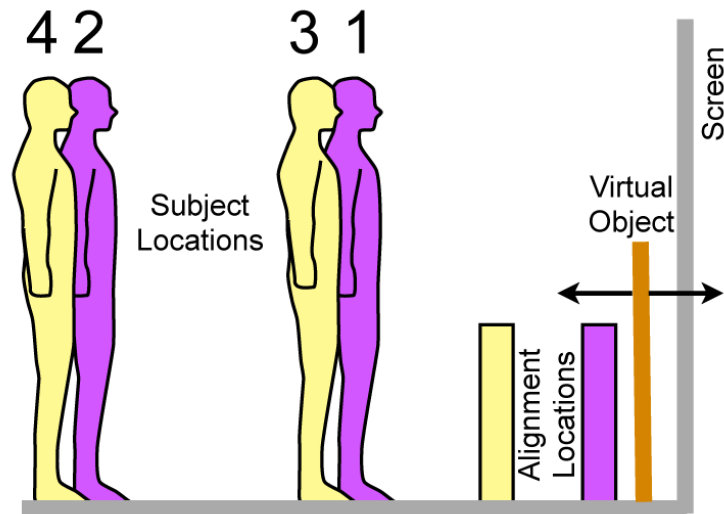


Fig. 6.

Diagram showing the location of the subject and placement of the alignment object for all four locations in the object positioning test. The subject was tasked with moving the virtual object to edge of the purple alignment object for the first two trials and the yellow object for the third and fourth trials. The average error and change between the positions for different locations were measured.

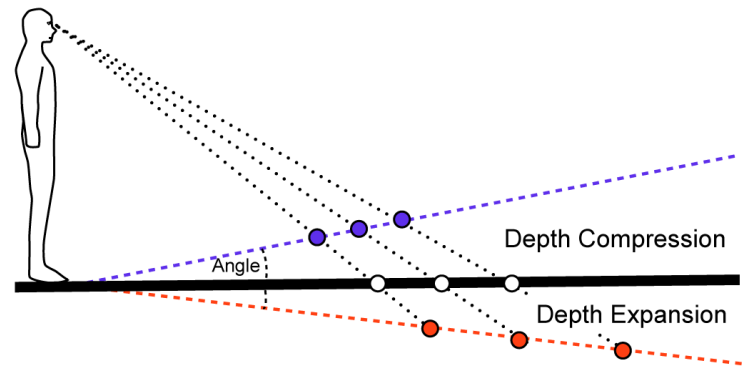


Fig. 7. Modeling distance estimation based on perceived inclination of a large planar floor. If the subject is looking down at the floor and depth compression is occurring, the floor will appear to slant upwards. Conversely, if the depth expansion is occurring the floor will appear to tilt downwards.



Fig. 8. The subject attempts to position the horizontal beams so that they are all level across. The beams move according to an inclination angle. The resulting misperception of level informs depth compression.



Fig. 9.

A participant undertaking the shape test. The participant uses the joystick to reshape the floating block with the goal of making the block a perfect cube. For the first iteration of the test the participant was put at a fixed location. In the second iteration of the test the participant was allowed to move freely inside of the CAVE environment. The tracker has been put on the camera to for photographic purposes.

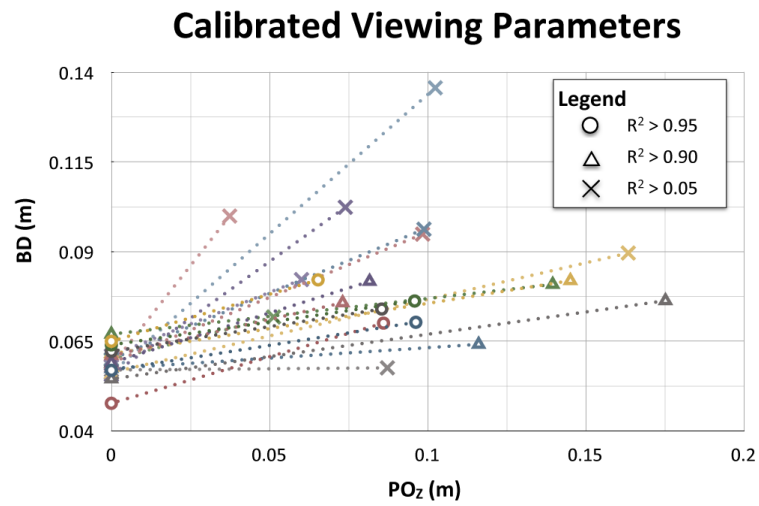


Fig. 10. Comparing the positions of BD and PO_z for the measured configuration and the calibrated configuration. The results are grouped by our trust that the viewing parameters have been correctly calibrated.

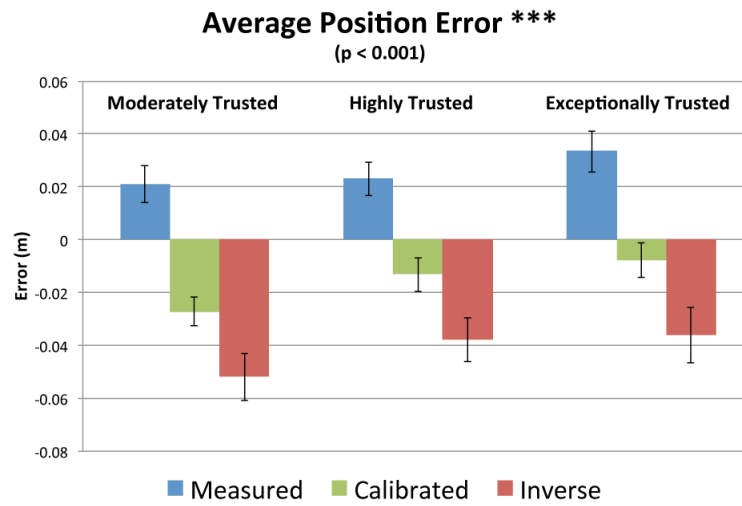


Fig. 11. The average directional error for the object positioning test all locations. Results are grouped by the trust in the viewing parameters have been correctly calibrated. The error bars represent the 95% confidence interval.

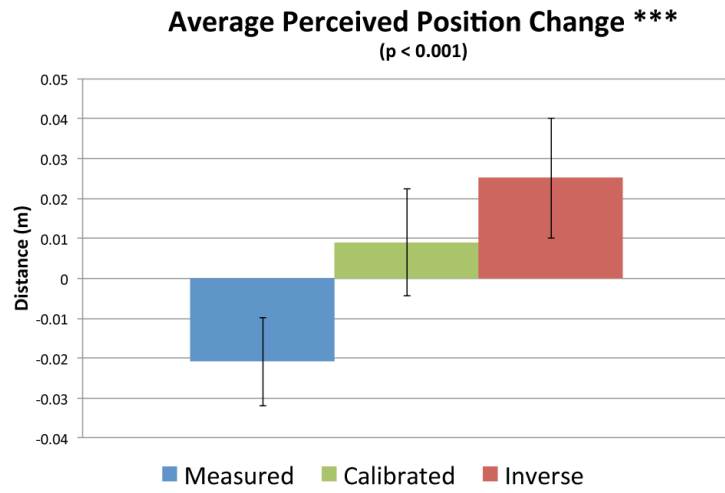


Fig. 12. The average difference in the perceived position of the same object position from Location 1 to Location 2 and Location 3 to Location 4. The error bars represent the 95% confidence interval.

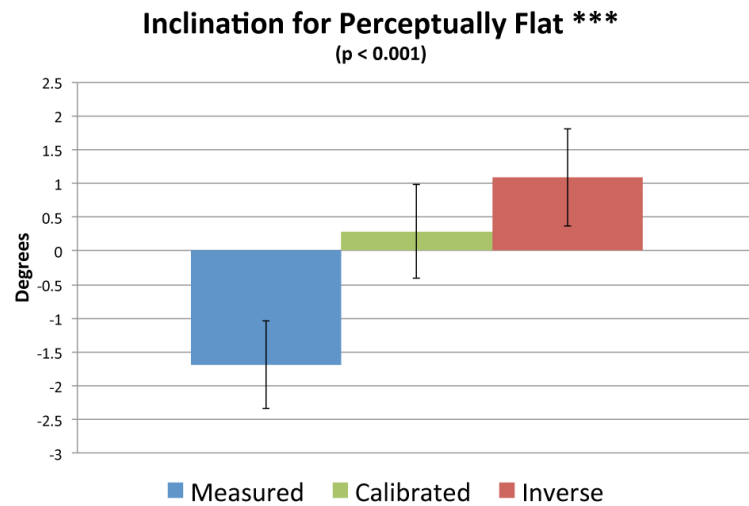


Fig. 13. The average inclination of the beams which the subject perceives to be level. The declination represents of measured configuration represents a compensation for the effects of depth compression, while the inclination angle for the inverse configuration represents depth expansion. The small angle for the calibrated configuration represents a very small of depth expansion for far away objects. The error bars represent the 95% confidence interval.

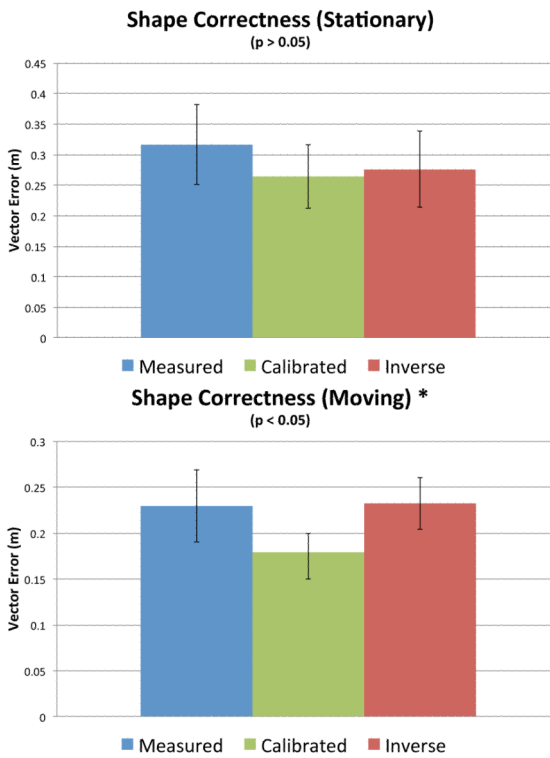


Fig. 14. The error between the correct position of the edge of a box and where the subject positioned the corner. While the test did not show significance when the subject was placed in a static location (top), the calibrated configuration was significantly better when the subject was allowed to move around (bottom). The error bars represent the 95% confidence interval.