Perceptual Tolerance to Stereoscopic 3D Image Distortion

ROBERT S. ALLISON and LAURIE M. WILCOX, York University

An intriguing aspect of picture perception is the viewer's tolerance to variation in viewing position, perspective, and display size. These factors are also present in stereoscopic media, where there are additional parameters associated with the camera arrangement (e.g., separation, orientation). The predicted amount of depth from disparity can be obtained trigonometrically; however, perceived depth in complex scenes often differs from geometric predictions based on binocular disparity alone. To evaluate the extent and the cause of deviations from geometric predictions of depth from disparity in naturalistic scenes, we recorded stereoscopic footage of an indoor scene with a range of camera separations (camera interaxial (IA) ranged from 3 to 95 mm) and displayed them on a range of screen sizes. In a series of experiments participants estimated 3D distances in the scene relative to a reference scene, compared depth between shots with different parameters, or reproduced the depth between pairs of objects in the scene using reaching or blind walking. The effects of IA and screen size were consistently and markedly smaller than predicted from the binocular viewing geometry, suggesting that observers are able to compensate for the predicted distortions. We conclude that the presence of multiple realistic monocular depth cues drives normalization of perceived depth from binocular disparity. It is not clear to what extent these differences are due to cognitive as opposed to perceptual factors. However, it is notable that these normalization processes are not task specific; they are evident in both perception- and action-oriented tasks.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Ergonomics*; H.1.2 [**Models and Principles**]: User/Machine Systems—*Human Factors*; J.5. [**Computer Applications**]: Arts and Humanities—*Fine Arts*

General Terms: Human Factors

Additional Key Words and Phrases: Stereoscopic 3D, scaling, film, interaxial, convergence

ACM Reference Format:

Robert S. Allison and Laurie M. Wilcox. 2015. Perceptual tolerance to stereoscopic 3D image distortion. ACM Trans. Appl. Percept. 12, 3, Article 10 (July 2015), 20 pages.
DOI: http://dx.doi.org/10.1145/2770875

1. INTRODUCTION

Appreciation of the subjects portrayed in photographs, paintings, and 2D film relies critically on our ability to tolerate inconsistencies in viewing geometry. Although our appreciation of 2D film seems almost effortless, it is a remarkable feat given that there is only one point in space—the center of

---

This work is supported by grants to the 3DFlic project from the Ontario Media Development Corporation and the Ontario Centres of Excellence. Thanks to Ali Kazimi for the S3D content and to Karim Benzeroual for conducting the experiments that collected the data reported here.

Authors' addresses: R. S. Allison, Department of Electrical Engineering and Computer Science, York University, 4700 Keele Street, Toronto ON M3J 1P3; email: allison@cse.yorku.ca; L. M. Wilcox, Department of Psychology, York University, 4700 Keele Street, Toronto ON M3J 1P3; email: lwilcox@yorku.ca.

DOI: http://dx.doi.org/10.1145/2770875

projection—from which the geometric layout of the 2D scene is correct. At all other viewing distances, viewing angles, and display sizes, there are substantial inconsistencies between the geometry of the scene and projection on the retina. Recent research has shown that viewers seem to use knowledge, or perception of the orientation of the image plane, to partially compensate for distortions produced by off-center viewing of 2D pictures [Kelly et al. 2013; Sedgwick 1993; Vishwanath et al. 2005; Vangorp et al. 2013]. However, additional information must be used to achieve tolerance to other perceptual distortions, such as those of size; theatergoers routinely accept enlarged human forms and faces in close-up shots without effort or discomfort in conventional 2D movies.

All of the distortions common to 2D pictures are present in stereoscopic 3D (S3D) media. However, S3D content is influenced by additional parameters that may cause further image distortion [Banks et al. 2009; Benzeroual et al. 2011a; Du et al. 2013; Wartell et al. 1999; Woods et al. 1993]. Some of these are associated with the camera arrangement (e.g., separation, orientation), and still others are related to the viewing context (e.g., viewing distance, screen size).

In general, the predicted position in 3D space of a disparate image on the screen can be obtained trigonometrically for a given viewing position. Similarly, geometry determines the amount of on-screen disparity from the scene layout and camera rig configuration [Benzeroual et al. 2011a; Spottiswoode and Spottiswoode 1953; Woods et al. 1993]. However, perceived depth in complex 3D scenes often differs from these geometric predictions. There is some evidence that the compensation process that works in part for 2D picture perception is not as effective for off-center viewing in S3D [Banks et al. 2009]. Kelly et al. [2013] recently reported that displacement of the viewer from the geometrically correct center of projection creates perceptual distortions for 2D content, but these are exaggerated in S3D. As outlined in the following, we evaluate the impact of display and capture parameters on perceived depth in S3D footage.

1.1 Display Parameters

The current renaissance of S3D filmmaking, S3D gaming, and S3D for education and scientific visualization has made a wide variety of content available for consumers of 3D media. At the same time, options for displaying this content have multiplied. Content originally created for theaters may be viewed on a large cinema screen, a 3D television, a 3D laptop or tablet, or even a 3D smartphone display. Consequently, S3D content must either be targeted or adapted for a wide variety of display sizes. Further, in most instances, viewing distance for these displays is a compromise between angular pixel size and field of view, thus distance from the screen co-varies with screen size. Therefore, creators of stereoscopic media face the complex challenge of preserving the 3D artistic context intended for the theater while avoiding distortions and artifacts at multiple scales.

Many image properties change when content is scaled from a large screen to a small screen. For example, an advantage of large format film is often claimed to be the resulting sense of immersion [Allison et al. 2013]; however, the same impact is unlikely from a small display even when viewed from a close distance. Further, small displays, especially handhelds, are more likely to be used in uncontrolled environments, and issues such as window violations may be more problematic. Given the close relationship between display size and viewing distance, for the most part, effects of display size are more properly considered effects of display distance. In the studies reported here, we maintain viewing angle by varying viewing distance with display size and compute the anticipated distortions in perceived depth given these parameters.

When 3D media is scaled from a large to a small display, geometry predicts that roundness, proportionality, and spatial layout should be affected. When an object moves closer to an observer, the size of the image on the retina increases in inverse proportion to its distance. The amount of disparity for a given depth interval also increases as an object moves from far to near. However, rather than being

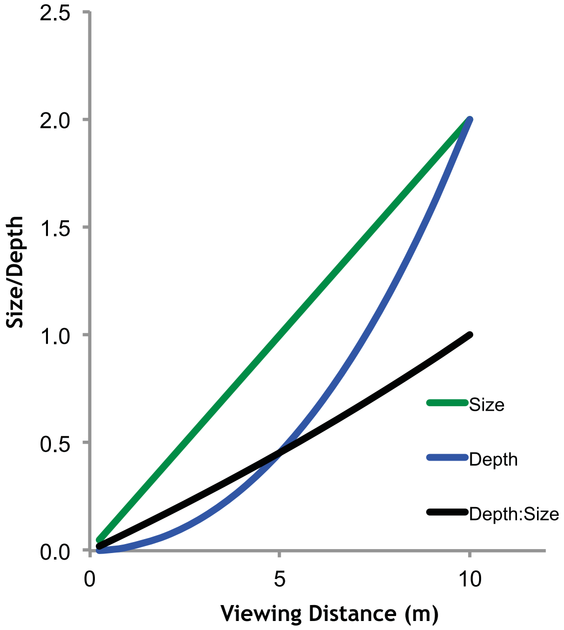


Fig. 1. Expected depth, size, and ratio of depth to size of a stereo image as a function of distance based on a depth:size ratio of 1.0 at 10m. The figure assumes that the display is scaled to maintain the same angular size at all distances (approximately what happens when the same content is viewed on different displays at different distances).

inversely proportional to distance, disparity for a fixed depth interval varies in inverse proportion to distance *squared*. Thus, if we compare a stereoscopic display at a given distance with a display of the same angular size at half the distance (i.e., a display half the size of the more distant display), then we expect the linear size of objects on the small display to be half that of the large display—but we expect depths to be quartered. Therefore, assuming that screen angular size is maintained as a screen is both scaled down and brought nearer, then we expect that size and depth should both decrease. However, the depth should decrease more quickly than the size, and this should perceptually “flatten” the image by decreasing the ratio of depth to width (Figure 1).

1.2 Capture Parameters

Camera parameters have well-known effects on the perception of depth in a shot. For example, zoom lenses with long focal lengths, which are often used to shoot outdoor scenes, compress distances while magnifying far objects [Banks et al. 2014], and the angle of a shot has a large impact on the sense of space produced. In S3D captured using stereoscopic camera rigs, the relative properties of the two cameras, especially their relative location and orientation, are also important factors. Although these latter factors are specific to S3D film, they can also interact with the conventional filmmaking parameters described earlier to enhance or distort perceived depth.

During S3D filmmaking, the camera separation, or interaxial (IA) distance, is determined on a shot-by-shot basis to achieve the desired depth budget (range of binocular parallax), with a particular theater screen size in mind. Typically, large theater screens are targeted, as there is less danger of creating viewer discomfort when content is decreased (rather than increased) in size. The choice of IA is also influenced by the size of the expected viewing environment; standard theaters have a wide range of seating distances, and so the predicted depth in a scene can vary substantially—depending on where the viewer is seated. For instance, a viewer seated in a row near to the screen will experience larger disparities but less depth than another person seated far from the screen.

As noted earlier, disparity varies with the inverse square of viewing distance, so small changes in distance from the scene should have a relatively large impact on perceived depth. The situation is similar in the home viewing environment, but locations relative to the screen are less controlled and somewhat unpredictable. In all cases, when the viewing scenario is inconsistent with the expectations of the media creators, distortions in perceived depth and form may be expected. Further, when the camera separation is too large or small, distortions in perceived object size or puppet theater effects (miniaturization and gigantism, respectively) may be seen. For example, with large IAs and objects presented in front of the screen plane, the disparity signals that the object is near to the viewer. To keep the shape of the object consistent with the scene geometry, the visual system reduces its apparent size, resulting in miniaturization.

The convergence of the camera pair determines the relation of the portrayed space to the screen. If the cameras are rotated with respect to each other so that both are directed toward an object, then that object will be at the center of both images. Thus, the images will have zero image disparity; geometrically, the object should appear to lie at the distance of the screen (referred to as the camera convergence distance). Alternatively, with a parallel camera configuration, in the postproduction phase the images can be translated horizontally in opposite directions (horizontal image translation, or HIT) to bring images of a given object into register, which effectively is the same as converging the cameras on the object.[1] Again, all objects in the scene at this convergence distance map to zero image disparity and should appear at the screen plane. Thus, both means of convergence can be used to shift and center the portrayed scene in depth relative to the screen. For a given viewing arrangement, if we bring the images of objects perceptually nearer, this should theoretically decrease their apparent size and depth as the visual system tries to maintain size and depth constancy. Once again, geometrically, apparent depth should decrease more rapidly than apparent size, and therefore the images should appear flattened when camera convergence is changed from a near to a far object (i.e., when the entire scene is brought closer to the viewer). This article reports six experiments that examine the role of screen size, IA, convergence, and monocular depth cues on perceived size and depth in complex S3D content filmed specifically for these experiments.[2]

2. GENERAL METHODS

2.1 Stimuli

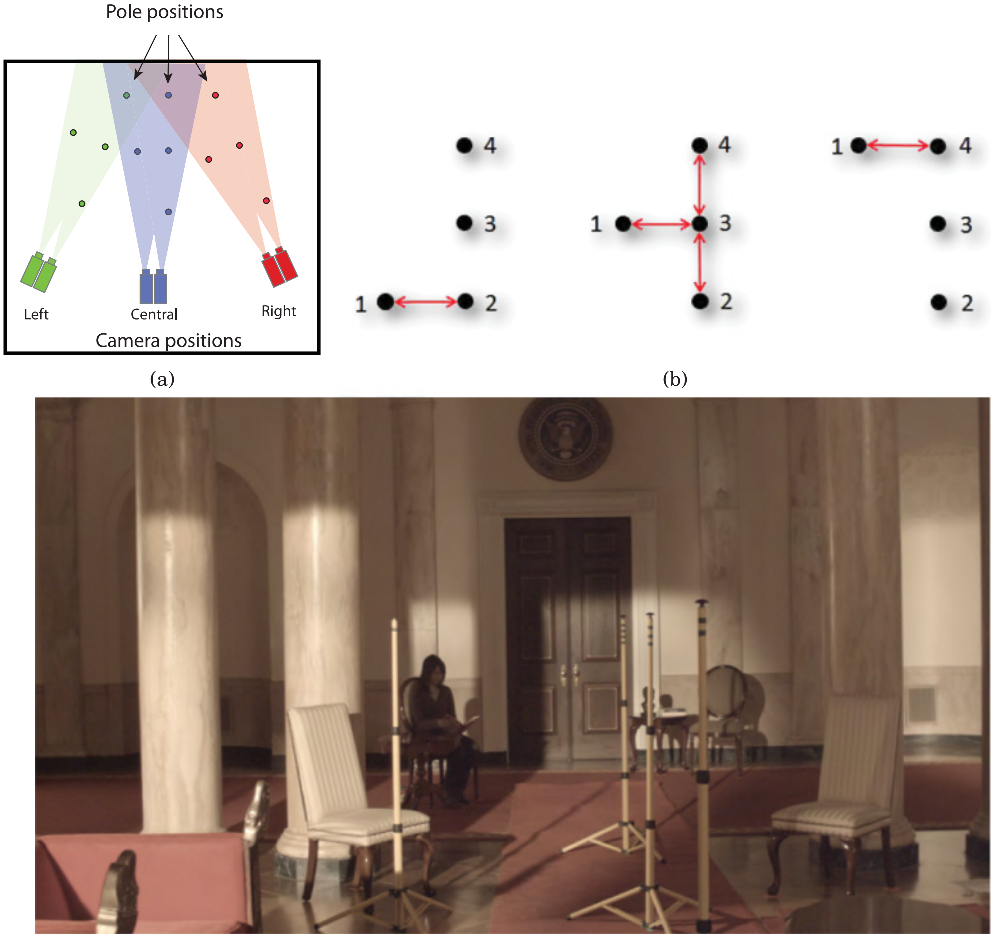
A professional film crew captured stereoscopic movies in an indoor, professional, standing set. The shots portrayed either a woman reading a book in large formal hall (test shots) or a couple playing chess in the same hall (reference shot). The shots were chosen to reflect typical but relatively static cinematic content, as our main concern was perceived depth, not motion. Multiple versions of each shot were captured under a range of camera and stereoscopic rig parameters.

The main test scene was the woman reading, and it was filmed from three different vantage points (Figure 2). The background varied in content (e.g., pictures on the walls), but approximately the same range of distances was visible from the three positions.

---

[1] The relative merits of these two approaches to convergence have been debated vigorously in the applied literature but are beyond the scope of this work. See Allison [2007] for a review and discussion. For a relatively far convergence distance, such as that used in the current work, the differences are actually fairly small.

[2] Experiments 1 and 2 have previously been reported in brief conference proceedings [Benzeroual et al. 2011b] that also discussed the effects of camera focal length.



(a)

(b)

(c)

Fig. 2. Scene used/vantage points. (a) Three positions of the stereo-rig. In each position, the poles were positioned at the same location relative to the cameras. Filming was repeated at each location for all combinations of camera parameters, resulting in three different clips (replications) for each condition. (b) Pole arrangements. Pole 1 was placed to be laterally offset from one of pole 2, 3, or 4. This produced five possible 3D distances to judge: two in depth and three lateral. (c) A still frame from one of the stimulus movies.

Four yellow stands were placed in the scene as targets for stereoscopic judgments (see Figure 2(b)). Because the studio space was large, the poles were in a similar central location in each of the three views. Stands 2 through 4 were arranged in depth along a line perpendicular to the camera baseline with a spacing of 183 cm (poles at 365, 548, and 731 cm from the cameras). The fourth stand, stand 1, was placed at the same distance as one of stand 2, 3, or 4 and laterally separated from the respective stand by 183 cm (i.e., separated in a direction parallel to the camera baseline).

The scenes were filmed with a pair of synchronized cinema cameras precisely aligned on a stereoscopic mirror rig with carefully matched prime lenses. The camera depth of focus was large, and there was no appreciable defocus blur. All combinations of the following stereo-rig parameters were filmed for each configuration of poles in the test scene:

Interaxial:	3, 25, 50, 75, and 95 mm.
Camera Lenses:	9.5, 12, and 16 mm (35mm equivalents of approximately 37, 48, 63 mm, respectively; 2/3 inch sensor: 9.58 × 5.39 mm).
Point of View:	Center, Left, and Right
Convergence:	At pole 2, 3, or 4. Cameras were parallel for capture, and convergence was achieved through horizontal image translation at the postproduction stage. Specifically, the images in the left and right eye view were shifted laterally in opposite directions to bring the convergence target into register (zero screen disparity). Since the original images were larger (2048 pixels wide) than the displayed images, there was no change in stereoscopic image size after cropping to the screen size. Cropping was performed by selecting pixels centered on the image sensor, and all calculations were based on the effective (cropped) sensor size. This convergence procedure avoids introduction of spurious vertical disparities [Allison 2007].

A shot of the two actors playing chess provided a reference for the distance estimates in the magnitude estimate task. This was filmed with a moderate IA of 50 mm, 12-mm lenses, and convergence on the nearer actor's shoulder. We assigned the apparent distance between the nearest (to camera) shoulder of the man and the nearest shoulder of the woman as a reference distance for subsequent depth and size judgments. Note that this distance was not the same as the interpole distance, and there was no expectation that responses should center on this reference size.

2.2 Apparatus

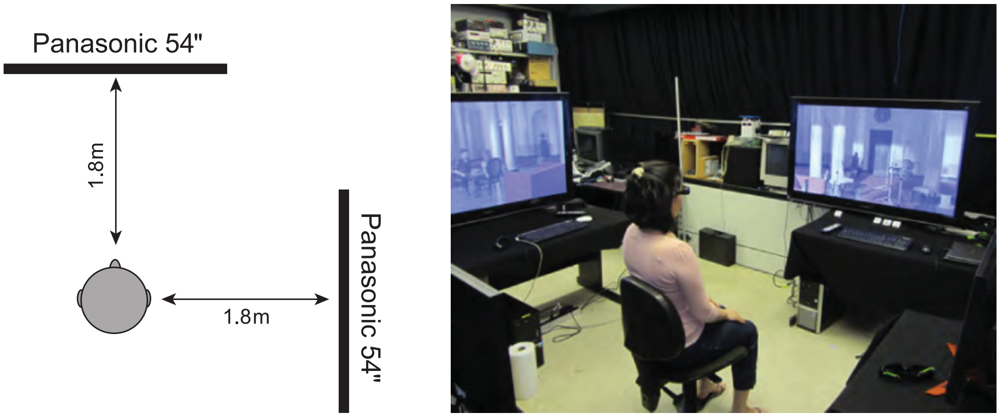
The test scenes were presented to the participant on different screen sizes. Various combinations of the following displays were used to present stimuli:

(1) Panasonic Viera TCP54VT25VT Series 54″ 1080p 3DTV Plasma televisions, 1920 (W) × 1080 (H) pixels, 119.8 (W) × 67.3 (H) cm (Panasonic Corp., Osaka, Japan);
(2) Acer GD235HZ 23.6″ 3D LCD monitors, 1920 (W) × 1080 (H) pixels, 52.1 (W) × 29.3 (H) cm (Acer Incorporated, Taipei, Taiwan) with NVidia 3D Vision glasses; and
(3) Fuji Finepix Real 3D W3 3.5″ autostereoscopic display, 1280 (W) × 920 (H) pixels, 7.1 (W) × 5.3 (H) cm, lenticular panel (Fuji Film Corp., Tokyo, Japan).

The stimuli were viewed through the shutter glasses provided with the displays except for the 3.5″ (W3) display, which was autostereoscopic. The position of the viewer relative to each screen was adjusted so that the horizontal visual angle of the screen was fixed at 37 degrees (180 and 76 cm from the 54″ and 23.6″ displays, respectively), except for the smallest 3.5″ displays.[3] These W3 displays were viewed from 30 cm for a 13-degree horizontal visual angle. Although the horizontal visual angle of this display was smaller, this choice of viewing distance was determined by several considerations, including the normal use of such handheld displays, maintenance of the autostereoscopic image quality, and avoidance of excessive accommodative-vergence conflict, which would be expected if the display were viewed at a closer distance [Shibata et al. 2011].

---

[3] Note that the viewing conditions were not orthostereoscopic, and the subject was not seated at the center of projection. Orthostereoscopic configurations are common in virtual reality but are almost never used in a cinema context. Even if orthostereoscopic presentation is attempted, viewing is accurate for only one seat in the theater.



(a) (b)

Fig. 3. Setup for Experiment 1. A pair of displays was placed at equal distances from the participant at a 90-degree angle from each other. The reference stimulus was shown on one display and the test stimulus on the other.

2.3 Participants

Ten observers participated in Experiments 1 and 3, eight participated in Experiments 2 and 6, six participated in Experiment 5, and nine participated in Experiment 4. Some observers participated in multiple experiments. All observers had good stereoacuity (≤40 arc sec), as tested using the Randot Stereo Test prior to the start of the experiments, and had a normal or corrected-to-normal vision (at least 20/20). Their age ranged from 25 to 35 years.

2.4 Task

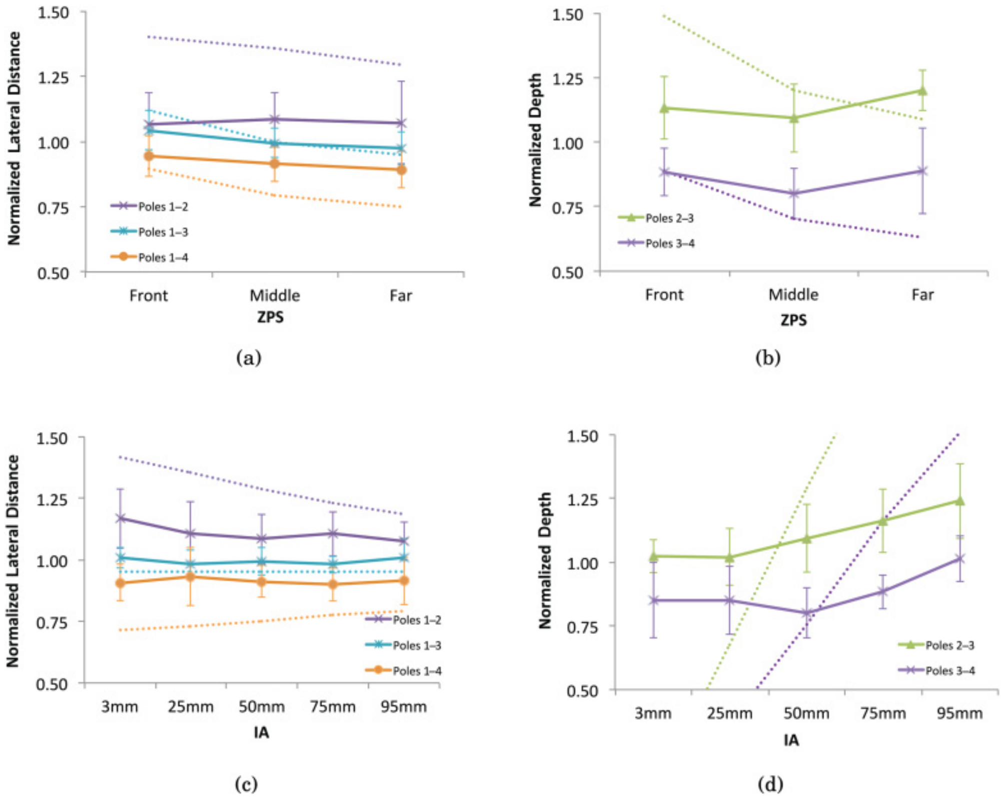
In several of the experiments, participants were asked to estimate 3D distances in the scene using a modified magnitude estimation procedure [Stevens 1962]. In other words, observers were shown the reference scene and told to assign the distance between the shoulders of the chess players a value of "100." They then viewed each of the five different 3D distances (red arrows in Figure 2(b)) and were asked judge the depth based on the reference. Thus, if the test distance looked twice as large, it would be assigned a value of 200, and if it appeared half the size, it would be assigned 50. Tasks specific to individual experiments will be described as they arise.

3. EXPERIMENT 1: CONVERGENCE AND IA

The aim of this experiment was to quantify the effect of both IA and camera convergence or zero-parallax setting (ZPS), the point in the image with zero screen parallax, on perceived depth using a complex scene containing multiple depth cues.

3.1 Methods

Figure 3 shows the setup for Experiment 1. A pair of the 54″ displays were arranged at equal distances in front of and beside the participant. One screen was used as a reference and one as a test. For the test clips, all combinations of five IAs (3, 25, 50, 75, and 95 mm) and three ZPSs (Front/pole 2, Middle/pole 3, and Far/pole 4) were assessed. The three points of view were treated as repeated measures of the IA-ZPS conditions for a total of 45 conditions per participant. For each condition, there were separate clips, with pole 1 aligned with poles 2, 3, and 4, respectively.



(a) (b)

(c) (d)

Fig. 4. Results for Experiment 1. (a, b) Effects of convergence (ZPS) on apparent lateral extent (a) and depth (b) for a fixed IA of 50mm. Geometrically predicted values based on stereoscopic triangulation are shown as dotted lines in matching color. (c, d) Effects of IA camera separation on apparent lateral extent (c) and depth (d) for ZPS fixed on pole 3. Predicted values are shown as dotted lines in matching color. Error bars represent 95% confidence intervals for the mean.

For each condition, participants made five estimates of 3D distance corresponding to lateral distances 1–2, 1–3, and 1–4, as well as in-depth distances 2–3 and 3–4, as shown in Figure 2(b). Test clips were shown on one monitor while the reference clip was displayed on the other. Participants could rotate their seats to look at the reference monitor, as needed, and made magnitude judgments as described in Section 2.

3.2 Results and Discussion

For each observer, in-depth and lateral magnitude estimates were normalized by dividing the raw estimate by the average response over all of the observer's in-depth or lateral estimates, respectively. A repeated-measures analysis of variance (ANOVA) was conducted to analyze patterns in the resultant data.

The predicted normalized 3D distances according to geometrical reconstruction from disparity at the viewer's stereoscopic vantage point are illustrated in Figure 4. Note that the binocular viewing geometry predicts that both the ZPS and the IA will affect the apparent position of the poles in depth relative to the screen plane.

3.2.1 *Predictions—ZPS.* As an object is stereoscopically shifted away from the screen plane, geometry predicts that its apparent size should change proportionally to the change in apparent distance. Therefore, if the ZPS is repositioned to a more distant point in the scene, the poles should appear to be nearer to the participant, and the lateral distances between the poles (i.e., 1–2, 1–3, and 1–4) should appear to decrease. Therefore, we expect a negative relationship between ZPS and lateral distance with the apparent separation equal to the separation in the image when it lies on the screen (Figure 4(a)). Similarly, from geometry alone, we expect to find a negative relationship between depth and convergence, as ZPS on a far object brings the entire scene stereoscopically nearer compared to ZPS on a near object (Figure 4(b)). The perceived lateral separation should also depend on the image size (i.e., lateral separation of the poles in the image), which varies inversely with distance according to perspective. This perspective scaling (coupled with the nonorthostereoscopic viewing that brings everything closer than in the real scene) explains why the predicted lateral distance is smaller for poles that are farther away in the scene (e.g., poles 1–4 vs. poles 1–2).

3.2.2 *Predictions—IA.* Increasing IA also moves objects farther from the screen plane. In this case, geometry predicts that objects behind the screen will appear farther back and objects in front of the screen will appear closer as IA increases. Increasing IA thus results in increases in disparity for a given depth in the scene and hence an expansion of depth in the recreated scene (Figure 4(d)). Once again, the change in apparent lateral separation should be proportional to the change in apparent distance. Thus, apparent lateral separation should increase with increasing IA for far objects and decrease with IA for near objects (Figure 4(c)).

3.2.3 *Results.* In contrast to the ZPS predictions outlined earlier, Figure 4 and the repeated-measures ANOVA indicated that ZPS had no significant effect on either lateral ($F(2{,}341) = 1.397$, $p = 0.249$) or depth ($F(2{,}227) = 0.224$, $p = 0.784$) estimates. The effect of ZPS was also independent of the level of IA (there was no interaction between IA and ZPS).

Similarly, the predicted effect of IA was not observed; Figure 4(c) shows that changes in IA had little influence on judgments of lateral distances, and this was true regardless of ZPS setting (i.e., there was no interaction). Lateral distance estimates decreased slightly on average with increasing IA ($F(1{,}341) = 9.739$, $p = 0.002$), although this is not apparent in the middle convergence data shown in Figure 4(c). As expected, the lateral distance estimates did vary with position of the poles in depth ($F(1{,}341) = 34.040, p < 0.001$), with smaller lateral estimates for poles located farther into the scene, as predicted. However, there was no interaction between pole position and IA; the effect of IA primarily seemed to reduce the separation estimates for the nearer pole pair (poles 1–2) with no sign of an increase in lateral separation for the furthest pair (poles 1–4).

Increasing IA did produce significant increases in apparent depth ($F(2{,}227) = 9.739$, $p = 0.002$), as predicted (Figure 4(d)). However, the effect of IA on depth estimates was much less than predicted from the geometry of stereopsis, as can be seen by inspection of Figure 4(d). Furthermore, although the depth estimates varied with pole position as expected (i.e., poles 2–3 vs. poles 3–4, $F(1{,}227) = 118.701$, $p < 0.001$), the expected interaction between IA and pole position was not significant ($F(2{,}227) = 1.406$, $p = 0.247$).

It is highly likely that the lack of effects of ZPS and the muted effect of IA on both apparent depth and lateral distances reflect, at least in part, the influence of monocular cues to size and depth. Monocular cues to size and depth do not change, as IA and ZPS are varied and so they conflict with the changing stereoscopic cues. Many experiments using relatively simple stimuli, such as lines and random textures, have shown that binocular disparity signals combine with monocular cues to depth to determine overall depth in a stimulus. In cases where these cues conflict, such as if texture signals one depth order and stereopsis another, the result is often a weighted average of the two cues (e.g.,

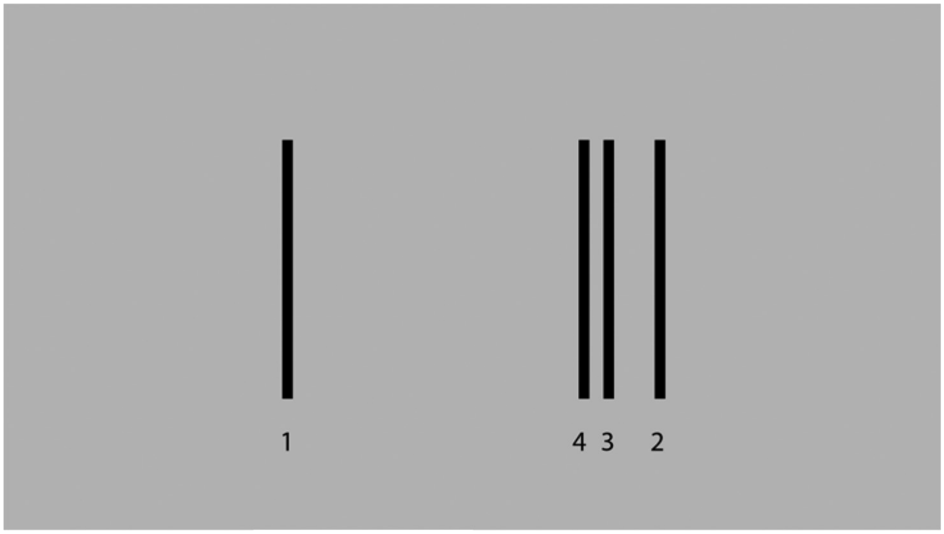


Fig. 5. Stimulus used in Experiment 2. The lines were located in the image position corresponding to the poles in each condition of Experiment 1; otherwise, the display was featureless.

Allison and Howard [2000] and Hillis et al. [2004]). However, the film set used here was a natural architectural scene with multiple rich monocular cues to scale, form, depth, and size.[4] Under these more natural image conditions, it appears that the monocular cues conforming to these conventions outweighed the influence of stereoscopic cues, which suggest unnatural and unusual environment layout. Such "cue vetoing" has been reported in the literature under circumstances where there is substantial conflict between two depth cues [Bülthoff and Mallot 1988] but has not been studied under the complex (multicue) conditions used here.

4. EXPERIMENT 2: SIMPLE LINE STIMULI

To quantify the combined influence of monocular cues in Experiment 1, we replicated that study using stimuli with equivalent disparity but minimal monocular depth cues. If, as outlined earlier, the conflict between multiple monocular cues and the depth signaled by disparity is responsible for the weak effect of IA in the preceding study, then removing that conflict should produce a stronger effect of IA.

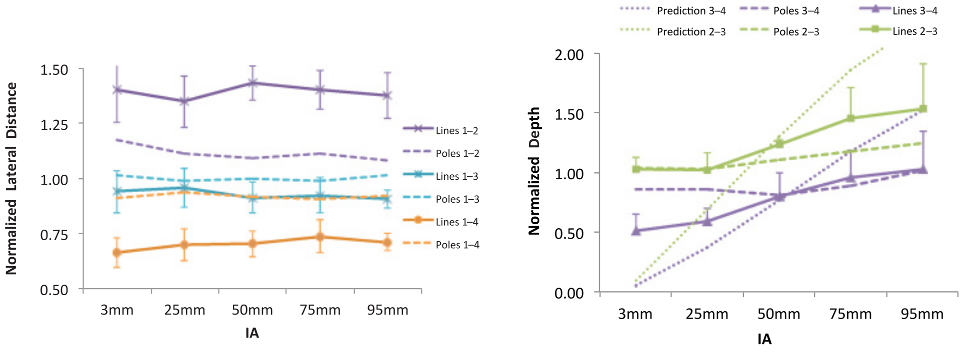
4.1 Methods

Computer-generated stimuli were produced depicting simple black line stimuli (bars) in place of the poles on an otherwise featureless, gray display (Figure 5). Black lines of fixed length and width (566 by 24 pixels) were placed in identical positions (in both the left and right half images) to the poles in the corresponding movie frames in Experiment 1. Thus, the disparities were identical to those in the equivalent IA and ZPS conditions.

Experiment 2 used the magnitude estimation procedure described in Section 2 and used in Experiment 1. Participants were asked to use the same reference stimulus and distance as in Experiment 1 to

---

[4] A variety of monocular cues, including linear perspective, absolute and relative size, occlusion, and height in the field, were available in the monocular images. Others cues, such as accommodation and defocus blur, were not consistent with the portrayed depth (due to the fixed viewing distance and the large depth of field, respectively). We did not attempt to assess the relative effectiveness of the various cues in our stimuli.



(a) (b)

Fig. 6. Results of Experiment 2. Effects of IA camera separation on apparent lateral extent (a) and depth (b) for ZPS fixed on pole 3. Corresponding mean data from Experiment 1 for poles in a real scene are shown as dashed lines in matching color. Predicted values are illustrated in (b) as dotted lines. Error bars represent 95% confidence intervals for the mean.

estimate the depth of the line patterns shown in Figure 5. As in Experiment 1, participants judged the depth of each of the line (pole) combinations in random order. The resulting estimates were averaged for each condition, for each observer.

4.2 Results

Figure 6 shows the reported lateral and in-depth distances as a function of (equivalent) IA. As in Experiment 1, there was no effect of ZPS on either perceived lateral or in-depth distances. Again, we find that the predicted distortions are absent, a result that runs contrary to viewing geometry but is consistent with anecdotal observations that viewers have a poor impression of what lies at the screen plane when viewing stereoscopic content.

As predicted using these simple line patterns, with other depth cues minimized, the IA substantially influenced the estimated depth; the amount of depth for a given disparity was greater for the simple stimuli than for the realistic footage used in Experiment 1. This suggests that the additional monocular cues present in Experiment 1 influenced the perceived estimates of depth and lateral separation, consistent with the scale of the scene. However, we note that even when those additional cues are removed the effect of IA still falls well short of geometric predictions. This shortfall may occur because even in such simplified stimuli, conflicts between stereopsis and monocular depth cues remain. In this case, the monocular cues (accommodation, perspective, size) signal that the lines all lie at the same distance. These conflicts may have reduced the impact of disparity on perceived depth.

5. EXPERIMENT 3: CROSS SCALE

In the preceding experiments, we assessed the impact of IA and ZPS on perceived depth for a given display size. As discussed in Section 1, binocular viewing geometry also predicts that changing the size and distance of the screen should impact viewers' estimates of lateral separation and depth in the displayed content. The relatively weak effects of IA and ZPS shown in Experiment 1 suggest that the visual system is adept at compensating for (or ignoring) disparity-based distortions in the presence of multiple monocular depth cues. Therefore, in the current study, we assessed the influence of display size by asking viewers to make *direct* comparisons of depth in 3D scenes between simultaneously presented displays of different sizes. Any impact of screen size on perceived depth should be most evident under such conditions.

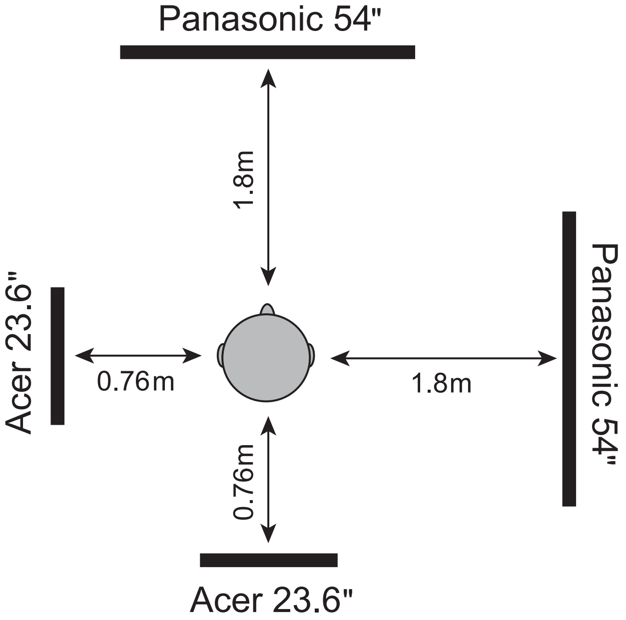


Fig. 7. Display arrangement for Experiment 3. Pairs of each of two different-sized monitors were placed at right angles around the participant. By turning the participant, all combinations of reference and test display size were tested.

5.1 Methods

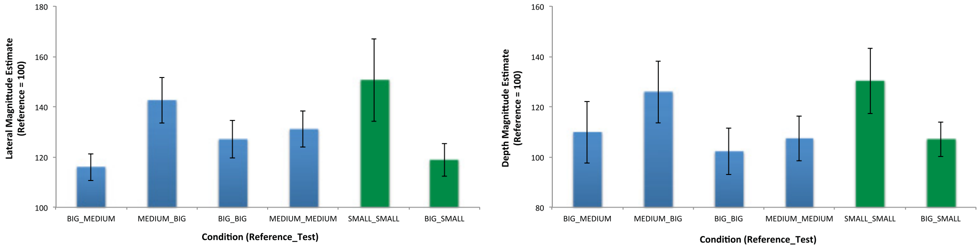
Figure 7 shows the setup for Experiment 3. The participant was placed within an arrangement of four monitors set at 90-degree angles relative to one another. Two were large 54″ television monitors and two were smaller 23.6″ monitors (medium sized) set at the appropriate distance as described in Section 2. Test clips were shown on one monitor while the reference clip was displayed on the other. Participants could freely look at the reference monitor as needed and made magnitude judgments as described in Section 2. By rotating the chair (which was fixed in place) and thus orienting the participant appropriately, we were able to evaluate a variety of combinations of reference and test screen size. All combinations of reference-test display were tested in counterbalanced order in separate blocks of trials (e.g., B-M (big screen for reference and medium screen for test), M-B, M-M, and B-B).

All combinations of three IAs (25, 50, and 75 mm) and two ZPSs (Front/pole 2 and Far/pole 4) were assessed in each block (combination of reference and test screen size). All clips were shot using a 12 mm lens. As in the preceding experiments, the three points of view were treated as repeated measures of the IA-ZPS conditions. For each condition, there were three separate clips presented, with pole 1 aligned with poles 2, 3, and 4, respectively. For each condition, participants made five estimates of 3D distance corresponding to lateral distances 1–2, 1–3, and 1–4, as well as in-depth distances 2–3 and 3–4 in Figure 2(b).

5.2 Results and Discussion

The results were generally consistent with Experiment 1 and showed no effect of ZPS on lateral or in-depth distances. As in Experiment 1, the distance of the poles in the scene had a significant effect on both depth and lateral judgments; however, the effect of IA was not significant.

Figure 8 shows average magnitude estimates for lateral and in-depth judgments in the main experiment (in blue). In this study, the horizontal visual angle of the two screens was equated, and both



(a) (b)

Fig. 8. Results for Experiment 3. The plot shows the average of the lateral (a) or depth (b) magnitude estimates for each of the combinations of reference screen and test screen size. The first four bars (blue) show the data from the main experiment (10 observers), and the last two bars (green) show conditions run in a follow-up study with a subset of 8 of the observers. The graph shows data for the 8 observers who participated in all of the conditions. Error bars show 95% confidence intervals.

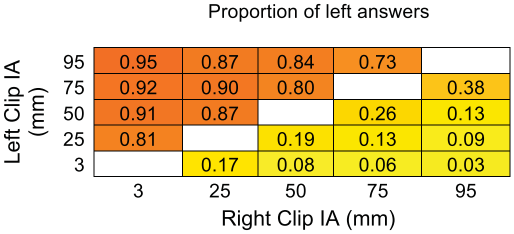
displayed the same content. This means that the angular size and disparity were constant, but the distance was approximately halved. Therefore, the predicted depth would be approximately quartered. Although the size and distance of the reference display relative to test display did significantly affect magnitude judgments ($F(3,167) = 8.329$, $p < 0.004$ and $F(3, 251) = 9.062$, $p < 0.001$, for in-depth and lateral estimates, respectively), the measured effect of display size and distance was again much less than predicted from geometry. When the reference and test display were the same size (M-M or B-B), lateral estimates for a given condition were similar at both display sizes. This was also true for the depth estimates. As expected, when the reference was on the medium display and the test on the large display (M-B), the estimates were on average larger than for the converse condition (B-M). However, the effect was much smaller than the approximately double and quadruple ratios predicted from simple geometry. Furthermore, the effect was smaller for the depth estimates than for the lateral extents, which is contrary to expectations.

Given the relatively small impact of screen size on depth magnitude percepts shown in Figure 8, we conducted a follow-up study to determine if the two displays (54″ vs. 23.6″) assessed a sufficient range of sizes. In this experiment, we tested a subset of eight observers who made the same depth and lateral separation judgments on a small Fujifilm W3 autostereoscopic monitor (3.5″; see Section 2 for details) using either the same sized screen as a reference (S-S) or the large 54″ display as the reference (B-S). Since we have repeatedly found no effect of convergence with these stimuli, in this study the ZPS was maintained on the front pole.

The green bars in Figure 8 show magnitude estimates for the in-depth and lateral separation conditions displayed on the S-S and B-S configurations. These results show that both the lateral and depth judgments are smaller for the small display when the reference is on the big display (relative to when the reference is on the small display; $F(1,125) = 12.949$, $p < 0.001$ and $F(1,83) = 8.524$, $p = 0.005$ for lateral and in-depth estimates, respectively). However, the differences are much less than expected from the 15-fold difference in display size.

6. EXPERIMENT 4: SIDE BY SIDE

In Experiments 1 through 3, participants consistently reported that they perceived a compelling sense of 3D depth and that the strength of this sense of space varied across the shots/conditions. Despite this, participants provided depth and size estimates that were more consistent with pictorial depth cues, or other (2D) image properties, than with expectations from disparity-based reconstructions. To confirm that participants did indeed perceive differences in 3D depth across the scenes, we asked participants



Proportion of left answers

Left Clip IA (mm)	Right Clip IA (mm): 3	25	50	75	95
95	0.95	0.87	0.84	0.73	
75	0.92	0.90	0.80		0.38
50	0.91	0.87		0.26	0.13
25	0.81		0.19	0.13	0.09
3		0.17	0.08	0.06	0.03

Fig. 9. Results for Experiment 4. The proportion of times the left clip was judged as deeper is shown as a function of IA of the left and right clip.

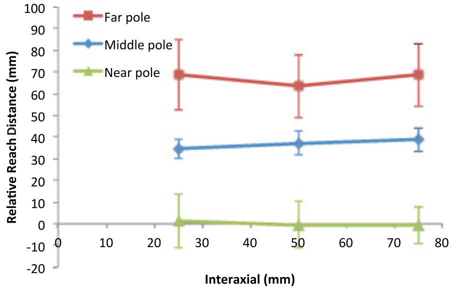


Fig. 10. Results for Experiment 5. Effects of IA camera separation on reached distance relative to the screen (where ZPS places pole 1) averaged across observers. Error bars represent 95% confidence intervals for the mean.

to compare the amount of depth in the scene when shots with different IAs were presented side by side on the same display.

6.1 Methods

A single 54″ monitor was used for this experiment. Pairs of clips were presented at the same time. The clip pairs were displayed side by side in separate windows on the screen (clips were scaled to 942 × 530 pixels with a 13-pixel separation). All clips showed the reading scene described in Section 2, filmed with a 12-mm lens with ZPS on the middle pole. All pairwise permutations of five IAs (3, 25, 50, 75, and 95 mm) were compared, and each permutation was presented 15 times to each participant (five repeats of clips from each of the left, center, and right viewpoints).

On each trial, participants were asked to report which clip (left or right) appeared to have more overall depth. Participants were required to choose even if they could not distinguish the depths.

6.2 Results

Figure 9 shows the proportion of "left clip deeper" response as a function of the left and right clip IA. One subject had a strong bias toward responding predominantly "right deeper" regardless of condition, so his data were excluded (inclusion of this subject did not change the basic pattern of results). The remaining participants reliably made choices indicating that the sensation of depth increased with increasing IA in these clips, particularly for large differences in IA. For smaller differences in IA, especially at the extreme IA of 75 and 95 mm, some participants occasionally reported the reversed depth ordering.

It is clear that participants are sensitive to the disparity differences in the shots used in Experiments 1 through 3. The increased difficulty distinguishing between similar IA at large IA settings (between IA of 75 and 95 mm) may reflect, in part, the upper limits of disparity processing. These IA settings are quite large—much larger than would normally be used to film such a scene in typical film industry practice (disparities of the poles were never divergent). It is well established that as disparity increases beyond an upper limit, depth percepts plateau and eventually decrease [Wilcox and Allison 2009].

7. EXPERIMENT 5: REACHING

The preceding experiments suggest that when presented with naturalistic 3D scenes (with multiple depth cues), observers' estimates of depth and lateral separation are normalized by the monocular information or by contextual information about the nature of the portrayed scene. Thus, subjects make responses consistent with the size and shape of the portrayed scene rather than the geometrically predicted reconstruction of this scene based on the disparities presented. It is possible that this normalization arises from cognitive influences that set the judgments in the context of the "real" scene. If this is the case, we might expect that measures requiring action within or relative to the portrayed stereoscopic space may not show these biases. Previous researchers have demonstrated effects of task on estimations of depth and distance. For example, verbal estimates of distance are often underestimated, particularly in virtual environments and other mediated representations, but participants are much more accurate when walking blindfolded to a remembered target [Loomis et al. 1992]. It is possible that cognitive effects that could have influenced our participants' magnitude estimates may not influence actions made in response to objects in the portrayed world. In our final two experiments, we used active responses of pointing (Experiment 5) and blind walking (Experiment 6), rather than perceptual estimates, to investigate the degree to which such task demands affect our results.

7.1 Methods

Participants viewed the clips on the 8.9 cm (3.5″) screen viewed at 30 cm. The participant's head was stabilized on a chin rest for this experiment. Following each clip, participants used their unseen hand to align a vertical 5-mm brass rod (also unseen) to the perceived position in depth of objects in the scene. The rod could be moved in depth along an optical rail and was tracked to submillimeter precision using a NaturalPoint V120:Trio optical tracker. The pointing apparatus was placed behind black card to block the observers' view of the rod and hand.

Test stimuli were drawn from the original footage captured using the 12-mm lens with convergence on the front pole. Clips for all combinations of three IA (25, 50, and 75 mm), three views (left, center, and right), and three positions for pole 1 (aligned the rod with each of rods 2, 3, and 4) were presented in 27 separate trials. Participants viewed each clip and set the unseen rod to align with pole 1 (thus providing estimates for distance of the near, middle, and far pole positions on separate trials). Trials were presented in a different random order for each participant.

7.2 Results and Discussion

The mean pointing position as a function of IA, averaged across participants, is shown in Figure 10. It is clear that the settings with unseen hand were noisy; even so, there is little evidence of an effect of IA—an observation that was confirmed statistically using ANOVA ($F(2,152) = 0.079$, $p = 0.924$). Thus, reaching measures appear to confirm the attenuated effect of IA on perceived depth (the effect was most strongly attenuated for the smallest displays) reported in the preceding experiments. As expected, the effect of pole position in the scene was significant ($F(2, 152) = 75.360$, $p < 0.001$). As predicted for a small, near display, the estimated distances were small (around 7 cm for the far pole relative to the front pole).

It is possible, albeit unlikely given the results of Experiments 1 through 4, that the limited effect of IA seen using the pointing task in Experiment 5 was due to the small range of predicted perceived depths for this display configuration. To evaluate this possibility, in Experiment 6 we use a blind walking measure to assess the effect of IA on perceived depth for larger displays where the predicted apparent depths were much larger.

8. EXPERIMENT 6: BLIND WALKING

8.1 Methods

Participants viewed movie clips in this study on the large 137 cm (54″) 3D display at 1.8 m. The stimuli consisted of clips for all combinations of five IA (3, 25, 50, 75, and 95 mm), three focal lengths (9.5, 12,

(a) (b)

Fig. 11. Setup for Experiment 6. (a) Participants viewed the display, turned, and (b) walked without vision to displace a pole from the viewing position to a distance corresponding to the perceived depth.

and 16 mm), and three views (center, left, and right). As in Experiment 5, in all cases convergence was on the front pole. For each combination, the clip was presented twice in separate trials for matching either the front pole to middle pole depth or the middle pole to back pole depth (distances 2–3 and 3–4 in Figure 2(b)). These 90 trials were presented to each observer in a different random order.

On each trial, participants were first told whether they were estimating the front or back interpole distance. They then watched the clip and formed an impression of this distance. The participants then had to move a real pole (identical to those in the stimuli) from a starting position to an end position so that the distance between the initial and final position of the pole was matched to the perceived depth in the clip.

To do so, the participants viewed the clip, closed their eyes, turned 90 degrees, and grasped the test pole. They then walked down a long hallway without vision and set the pole at a distance that matched depth interval in the clip (Figure 11). Infrared tracking markers (tracked with submillimeter accuracy) were mounted on the top of the pole, as in Experiment 5.

To obtain the distance traveled, we captured tracking coordinates for three points near the starting position, orthogonal to the direction of walking. These formed a triangular arrangement defining a vertical plane, which we defined as the starting position. By taking the cross product of an arbitrary two vectors between these points, we found the normal to this vertical plane. We then calculated the distances with reference to this plane by projecting the tracker coordinates onto this normal and subtracting the origin of the reference plane.

8.2 Results and Discussion

The mean depth matches, averaged across participants and across both the near and far pair judgments, are shown in Figure 12. Repeated-measures ANOVA found significant main effects of IA ($F(4,683) = 38.68$, $p < 0.001$), focal length ($F(2,683) = 52.512$, $p < 0.001$), and depth interval (front vs. back pair, $F(1,683) = 27.93$, $p < 0.001$) but no significant interactions.

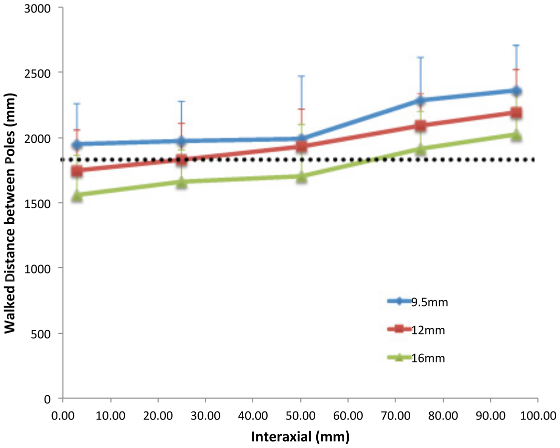


Fig. 12. Results for Experiment 6. Effects of IA camera separation and focal length on the walked distance between the poles averaged across observers. The separation of the poles in the actual scene is shown as a dotted line. Error bars represent 95% confidence intervals for the mean.

Depth reproduced by blind walking clearly increased with IA, although it fell well below that expected from disparity reconstruction (Figure 4). Post hoc *t*-tests (with Holm correction for multiple comparisons) showed significant differences in depth estimates between all levels of IA except between 25 and 50 mm.

On average, as expected, depth percepts for smaller focal length lenses were larger than for long lenses, and the effects of IA can be seen at all focal lengths. Depth was significantly larger for the 9.5-mm lens compared to the 12-mm lens and for the 12-mm lens compared to the 16-mm lens. These patterns were similar for both depth intervals estimated. However, depth reproduction for the interval between the further pair of poles was on average less than for the interval between the nearer pair of poles (difference 14.3cm).

Thus, with a natural, active measure of depth in the scene, presented on a large display, there was a clear effect of IA; however, as in our preceding experiments, this effect was much smaller than anticipated from the geometry of the scene. There was also a clear effect of monocular cues, as demonstrated by the effect of focal length. These findings were consistent with perceptual judgments from earlier experiments.

9. GENERAL DISCUSSION

This series of experiments consistently showed no effect of convergence point on perceived depth, with both full-cue and simple S3D content. This is inconsistent with a simple geometric reconstruction of depth from disparity (e.g., Woods et al. [1993]). It is consistent, however, with the common introspection that it is difficult to determine the point of convergence when viewing a stereoscopic movie. This phenomenon commonly occurs in darkened theaters, where there is limited reference information that can anchor the relative disparity cue. Similarly, the experiments reported here were conducted in dim lighting under which the edges of the displays, when visible in the periphery, would provide a relatively weak frame of reference. It is likely that in a fully lit room, the surrounding context would have provided a better frame of reference and effects of convergence point may have been more apparent. However, in that case, the results would have been less generalizable to S3D film. In fact, the difficulty in localizing the convergence point in isolated stereoscopic media is well appreciated in the film industry. For example, professional monitors for S3D review, edit, and quality control typically have superimposable grids or reference points in the monitor software. When reviewing S3D content, these are used to define the screen plane and to meter disparity in the images.

In addition to this weak reference plane information, there is the fact that overall perceived depth in complex S3D footage like that used here also relies on multiple monocular cues to depth. As discussed later in the context of IA and screen size, the presence of these conflicting monocular cues likely serves to further reduce the impact of convergence in our experiments.

The principal parameter that content creators use to control the degree of stereoscopic depth is the IA. Although this is particularly true in live action stereoscopic photography, a laterally separated pair of cameras is also the most common model for computer-generated S3D content. In our experiments, we found an effect of this critical parameter, but one that was much smaller than predicted from geometric reconstruction of the disparities in the image. Reported size and depth were resilient to variation over a large range of IA (3 to 95 mm) and display size (3.5″ to 55″). Banks et al. [2009] and Vishwanath et al. [2005] have found tolerance for lateral position in 2D but not S3D picture viewing. However, whereas their 2D images contained multiple monocular depth cues, the 3D imagery was relatively simple (a hinged grid) with few additional monocular depth cues. The present findings suggest significant tolerance for scale and disparity. We believe that this reflects the fact that monocular cues play a significant role in setting context for interpretation of parallax cues in complex real-world scenes. Importantly, the environment that we used was a familiar man-made environment with both expectations and cues to overall scale and layout. These includes cues such as familiar size, which might be the key difference with the stimuli used by Banks et al. and Vishwanath et al. Although our findings suggest that monocularly available cues to size and scale can modulate the appearance of stereoscopic depth, it is important to note that our study used a single man-made environment. Results of pilot experiments conducted with a smaller indoor set suggest that these findings generalize to other man-made environments. Similar cues to scale exist in natural scenes, but we cannot rule out the possibility that expectations of the structures and regularities of the carpentered world might play a role in our results.

The discrepancy between "calculated" depth from disparity and apparent depth shown here, and by others [Foley 1980], is mediated by perceptual and cognitive factors promoting maintenance of natural proportion (the "dual nature" of pictures; see Sedgwick [1993]). A potentially strong cognitive factor is that judgments are presumably made relative to the scale of the perceived scene. This likely explains the fact that cross-scale size and depth judgments in Experiment 3 showed effects of display size that were much smaller than predicted by the 2:1 ratio of screen size and distance. If observers provided responses consistent with or influenced by the scale of the portrayed scenes, it is not surprising that these were relatively consistent across scales. This might be related to the phenomena that actors in close-up shots on both the cinema screen and the television screen look "normal" despite large differences in scale.

We had considered that such cognitive factors may be reduced with active tasks given the proposed separation of visual processing into "what" (dorsal) versus "how" (ventral) pathways [Goodale 2014]. Although active measures do scale with screen size, we still have to ask why the effect of IA is so weak. Perhaps, even in these tasks, observers act to provide measures that match their perceptual estimates rather than distances directly in the scene. While possible, such an interpretation would be counter to the common usage of these tasks as measures of perception for action in the recent literature.

The fact that the effects of IA and screen size were consistently much smaller than predicted suggests that observers compensate for distortion in the scene. We conclude that the presence of multiple realistic depth cues drives normalization of perceived depth from binocular disparity. Further, these normalization processes are not task specific; they are evident in both perception- and action-oriented tasks. Thus, although content scaling has become an increasingly important issue that can negatively impact the S3D experience, we have considerable tolerance to acquisition and display parameters.

Our results suggest that filmmakers have considerable flexibility in their choice of camera parameters. Conversely, we would argue that the output of "stereoscopic calculators" must be used with caution. Many cinematographers and stereographers rely on devices such as stereoscopic tables or calculators to predict stereoscopic depth and comfort from stereo-rig parameters. Generally, such tools can effectively predict acceptable ranges of binocular parallax, which can be calculated easily and precisely. However, seasoned stereographers know from experience that such tools do not reliably predict the viewer's experience, and so they rely on their own heuristics to interpret the outputs of these tools. If calculators are used indiscriminately, then S3D content as captured may not meet the filmmakers' depth requirements and may need adjustment during an expensive postproduction process.

Taken together, the results of the experiments presented here highlight the critical role of monocular cues to depth, and of scene context, in determining depth perception in S3D content. There is no question that the presence of binocular parallax enhances the sense of space and volume in 3D content; however, this depth information is interpreted in close conjunction with monocular depth information. In most cases, they will augment one another; as shown here, when depth from disparity should be distorted, the monocular depth information serves to help "correct" our perception (e.g., see Glennerster et al. [2006]). Importantly, this compensatory effect seems contingent on the presence of monocular cues (Experiments 1 and 2). The implication is that S3D content captured with relatively sparse scenery and few monocular depth cues will elicit a greater impact of binocular parallax. Consideration of the perceptual consequences of 2D and stereoscopic parameters, as well as their interaction, at the early stages in the creation process will provide obvious advantages to developing a more sophisticated visual approach and to bypass time-consuming and costly attempts at correction in postproduction.

REFERENCES

Robert S. Allison. 2007. Analysis of the influence of vertical disparities arising in toed-in stereoscopic cameras. *Journal of Imaging Science and Technology* 51, 4, 317–327.

Robert S. Allison and Ian P. Howard. 2000. Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research* 40, 14, 1869–1885. DOI:http://dx.doi.org/10.1016/S0042-6989(00)00034-1

Robert S. Allison, Laurie M. Wilcox, and Ali Kazimi. 2013. Perceptual artefacts, suspension of disbelief and realism in stereoscopic 3D film. *Public* 24, 47, 149–160. DOI:http://dx.doi.org/10.1386/public.24.47.14_1

Martin S. Banks, Emily A. Cooper, and Elise A. Piazza. 2014. Camera focal length and the perception of pictures. *Ecological Psychology* 26, 1–2, 30–46. DOI:http://dx.doi.org/10.1080/10407413.2014.877284

Martin S. Banks, Robert T. Held, and Ahna R. Girshick. 2009. Perception of 3-D layout in stereo displays. *Information Display* 25, 1, 12–16.

Karim Benzeroual, Robert S. Allison, and Laurie M. Wilcox. 2011a. Distortions of space in stereoscopic 3D content. In *Proceedings of the SMPTE International Conference on Stereoscopic 3D for Media and Entertainment*. 6.1–6.10.

Karim Benzeroual, Laurie M. Wilcox, and Robert S. Allison. 2011b. On the distinction between perceived and predicted depth in S3D films. In *Proceedings of the 2011 International Conference on 3D Imaging (IC3D'11)*. 59.1–59.8.

Heinrich H. Bülthoff and Hanspeter A. Mallot. 1988. Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A* 5, 10, 1749–1758. DOI:http://dx.doi.org/10.1364/JOSAA.5.001749

Songpei Du, Shimin Hu, and Ralph Martin. 2013. Changing perspective in stereoscopic images. *IEEE Transactions on Visualization and Computer Graphics* 19, 8, 1288–1297. DOI:http://dx.doi.org/10.1109/TVCG.2013.14

John M. Foley. 1980. Binocular distance perception. *Psychological Review* 87, 5, 411–434.

Andrew Glennerster, Lili Tcheang, Stuart J. Gilson, Andrew W. Fitzgibbon, and Andrew J. Parker. 2006. Humans ignore motion and stereo cues in favor of a fictional stable world. *Current Biology* 16, 4, 428–432. DOI:http://dx.doi.org/10.1016/j.cub.2006.01.019

Melvyn A. Goodale. 2014. How (and why) the visual control of action differs from visual perception. *Proceedings of the Royal Society B: Biological Sciences* 281, 1785, 20140337. DOI:http://dx.doi.org/10.1098/rspb.2014.0337

James M. Hillis, Simon J. Watt, Michael S. Landy, and Martin S. Banks. 2004. Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision* 4, 12, 967–992. `DOI:`http://dx.doi.org/10.1167/4.12.1

Jonathan W. Kelly, Melissa Burton, Brice Pollock, Eduardo Rubio, Michael Curtis, Julio De La Cruz, Stephen Gilbert, and Eliot Winer. 2013. Space perception in virtual environments: Displacement from the center of projection causes less distortion than predicted by cue-based models. *ACM Transactions on Applied Perception* 10, 4, 18:1–18:23. `DOI:`http://dx.doi.org/10.1145/2536764.2536765

Jack M. Loomis, Jos A. da Silva, Naofumi Fujita, and Sergio S. Fukusima. 1992. Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance* 18, 4, 906–921. `DOI:`http://dx.doi.org/10.1037/0096-1523.18.4.906

H. A. Sedgwick. 1993. The effects of viewpoint on the virtual space of pictures. In *Pictorial Communication in Virtual and Real Environments*, S. R. Ellis, M. K. Kaiser, and A. J. Grunwald (Eds.). CRC Press, Boca Raton, FL, 460–479.

Takashi Shibata, Joohwan Kim, David M. Hoffman, and Martin S. Banks. 2011. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision* 11, 8, Article No. 11. `DOI:`http://dx.doi.org/10.1167/11.8.11

Raymond Spottiswoode and Nigel Spottiswoode. 1953. *The Theory of Stereoscopic Transmission and Its Application to the Motion Picture*. University of California Press.

Stanley Stevens. 1962. The surprising simplicity of sensory metrics. *American Psychologist* 17, 1, 29–39. `DOI:`http://dx.doi.org/10.1037/h0045795

Peter Vangorp, Christian Richardt, Emily A. Cooper, Gaurav Chaurasia, Martin S. Banks, and George Drettakis. 2013. Perception of perspective distortions in image-based rendering. *ACM Transactions on Graphics* 32, 4, 58:1–58:12. `DOI:`http://dx.doi.org/10.1145/2461912.2461971

Dhanraj Vishwanath, Ahna R. Girshick, and Martin S. Banks. 2005. Why pictures look right when viewed from the wrong place. *Nature Neuroscience* 8, 10, 1401–1410. `DOI:`http://dx.doi.org/10.1038/nn1553

Zachary Wartell, Larry F. Hodges, and William Ribarsky. 1999. Balancing fusion, image depth and distortion in stereoscopic head-tracked displays. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. 351–358.

Laurie M. Wilcox and Robert S. Allison. 2009. Coarse-fine dichotomies in human stereopsis. *Vision Research* 49, 22, 2653–65. `DOI:`http://dx.doi.org/10.1016/j.visres.2009.06.004

Andrew J. Woods, Tom Docherty, and Rolf Koch. 1993. Image distortions in stereoscopic video systems. In *Proceedings of SPIE 1915: Stereoscopic Displays and Applications IV.* 36–48. `DOI:`http://dx.doi.org/10.1117/12.157041

Received February 2015; revised May 2015; accepted May 2015