Sampling artifacts in perspective and stereo displays

J. D. Pfautz

Massachusetts Institute of Technology, Research Lab for Electronics
Cambridge, Massachusetts 02139 USA

ABSTRACT

The addition of stereo cues to perspective displays is generally expected to improve the perception of depth. However, the display's pixel array samples both perspective and stereo depth cues, introducing inaccuracies and inconsistencies into the representation of an object's depth. The position, size and disparity of an object will be inaccurately presented and size and disparity will be inconsistently presented across depth. These inconsistencies can cause the left and right edges of an object to appear at different stereo depths. This paper describes how these inconsistencies result in conflicts between stereo and perspective depth information.

A relative depth judgment task was used to explore these conflicts. Subjects viewed two objects and reported which appeared closer. Three conflicts resulting from inconsistencies caused by sampling were examined:

- Perspective size and location vs. stereo disparity
- Perspective size vs. perspective location and stereo disparity
- Left and right edge disparity vs. perspective size and location

In the first two cases, subjects achieved near-perfect accuracy when perspective and disparity cues were complementary. When size and disparity were inconsistent and thus in conflict, stereo dominated perspective. Inconsistency between the disparities of the horizontal edges of an object confused the subjects, even when complementary perspective and stereo information was provided. Since stereo was the dominant cue and was ambiguous across the object, this led to significantly reduced accuracy. Edge inconsistencies also led to more complaints about visual fatigue and discomfort.

Keywords: depth perception, stereoscopy, linear perspective, sampling, sampling artifacts

1. INTRODUCTION

For some tasks, providing stereo depth information in addition to other depth cues can significantly improve performance. There has been much discussion of the utility of stereo depth information in the literature[1 2]. Clearly, the value of stereo is a function of the task to be performed with the stereo display. This paper presents another way of evaluating the usefulness of stereo when spatial resolution is limited and the task involves relative depth judgements.

If stereo information is presented using a flat display (rather than any number of other stereo display technologies), then the display's pixel array samples the stereo information. Sampling results in artifacts that hinder the correct perception of binocular disparity information in 3D computer-generated imagery (CGI). Typically, antialiasing methods are used to ameliorate these artifacts, but antialiasing can represent an unwanted computational cost, especially in systems where an interactive frame rate must be maintained.

Stereo depth cues are seldom presented without linear perspective depth cues. Although different types of projection may be used (e.g., orthographic), perspective projection is the most common in 3D CGI. Sampled scenes presented with perspective projection will have artifacts in the projected size and location of objects, resulting in inaccuracies and inconsistencies in the presentation of depth[3]. Using stereo and perspective projection in a system with limited spatial resolution results in additional inaccuracies and inconsistencies. Inconsistencies in stereo and perspective depth may not occur simultaneously, resulting in conflicts in the depth information.

In this paper, we identify and discuss artifacts in spatially sampled stereo and perspective depth cues. First, we describe assumptions about the display system and the viewing geometry. Then, we present the artifacts that occur in perspective depth information and stereo depth information. We describe the effects of sampling on the combination of these two depth cues. A set of four experiments is presented to examine the interactions between the artifacts in perspective depth cues and the artifacts in stereo depth cues when they represent complementary and conflicting depth information. The implications for stereo display design and task performance are also discussed.

Other author information: Email: jpfautz@mit.edu; Telephone: (617) 253-2517; Fax: (617) 258-7003.

2. BACKGROUND

This paper discusses artifacts that would occur in a time- or space- multiplexing stereo system based on a flat screen, such as shutter glasses systems[4], many stereo HMDs, and some autostereo displays[5]. We also simplify, for purposes of illustration, by assuming pixels are square, the display surface is flat, and pixels behave uniformly across the display surface.

Some previous work has addressed sampling in stereo imagery. Hodges and Davis describe the creation of "stereo voxels" due to the sampling of binocular disparity information on a flat-screen stereo display[6]. Stereo voxels determine what area in space is represented by a particular disparity at a particular location on the screen. They show how the sampling of the on-screen disparity results in sampling of the 3D viewing volume.

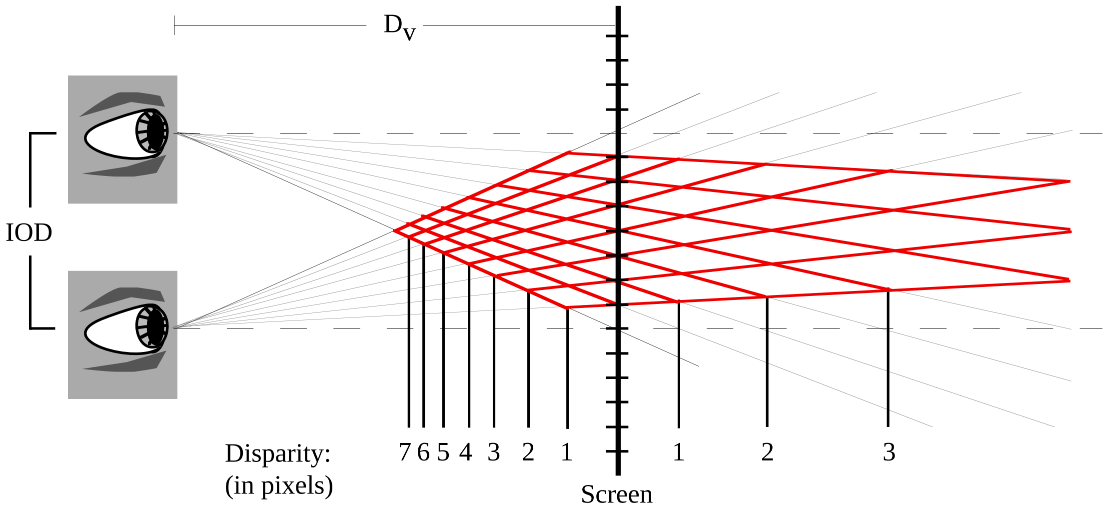

Figure 1: Plan view of spatially sampling a stereoscopic viewing volume. Adapted from Davis and Hodges[7].

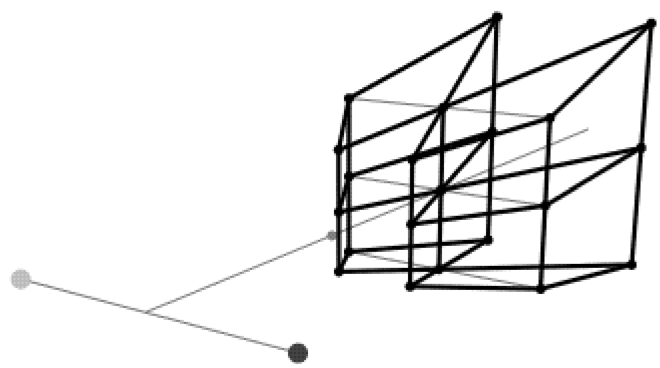

Figure 2: Spatially sampling (two pixels by two pixels) a 3D stereoscopic viewing volume.

Ledley & Frye[7] refer to the depth represented by a particular sampled disparity as a "depth plane". These sources, however, look specifically at stereo information; in this paper, we focus on sampling in both stereo and perspective depth cues.

Stereo information is usually presented in conjunction with linear perspective information[8]. Displaying stereo without perspective cues (or any geometric projection) causes distortions due to the viewing geometry. For example, an object presented with only stereo information appears smaller when in front of the screen, and larger when behind because of the geometry of the stereo viewing volume[9]. In this paper, we assume that both stereo and perspective cues are used to present depth in CGI.

2.1. Sampling Artifacts in Perspective Depth Cues

When perspective depth information is sampled, two types of errors occur: *inaccuracies* and *inconsistencies*. The 2D projected location of a 3D point will be rounded to the nearest pixel, introducing some inaccuracy in the depth the point represents. However, inaccuracy in position is often less of a problem than inconsistency in size[3]. The size of an object will be determined by the rounded projection of its horizontal and vertical endpoints. The endpoints may fall such that the object grows or shrinks by a pixel from the correctly rounded size, as in Figure 3. This means that two objects at the same depth may appear to be different sizes. Worse yet, an object that is at a greater distance may appear larger than a closer object.

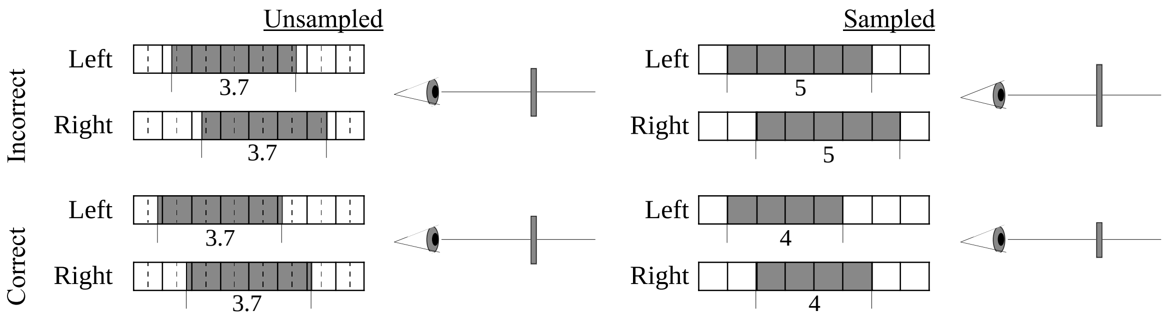

Figure 3: Inconsistencies in projected size due to sampling the endpoints of an object. The top row shows how the size of an object could incorrectly round up, making the object appear larger than intended.

Furthermore, inconsistencies in size do not necessarily occur simultaneously in both dimensions, which results in inconsistencies in the proportions of an object.

The likelihood that inconsistencies occur varies as a function of the size of the object, the distance from the line of sight, and the distance between the object and the viewpoint. A very small object lying on the line of sight experiences few inconsistencies over a range of depth, whereas a large object far from the line of sight experiences many size and proportion inconsistencies.

The perception of these inconsistencies is governed primarily by the size of a pixel. The ability to detect a difference in the projected location of two objects is a function of the separation of the two objects and the pixel size. Similarly, the detectability of a difference in the size or proportions of two objects is governed by the separation of the two objects and the ratio of the size of the object to the size of a pixel. For an in-depth description of these artifacts and how they are perceived, see Pfautz[3].

2.2. Sampling Artifacts in Stereo Depth Cues

Stereo depth information also experiences inconsistencies and inaccuracies due to sampling. Again, we focus on the occurrence of inconsistencies that are more likely to disrupt the accurate perception of depth. When sampling the disparity of a single point, the space between the left and right locations of the point will be sampled much like the size of an object, as described above. This means that the disparity of an object will vary inconsistently in depth. Two objects may be at the same stereo depth but be presented at different stereo depths, as in Figure 4.

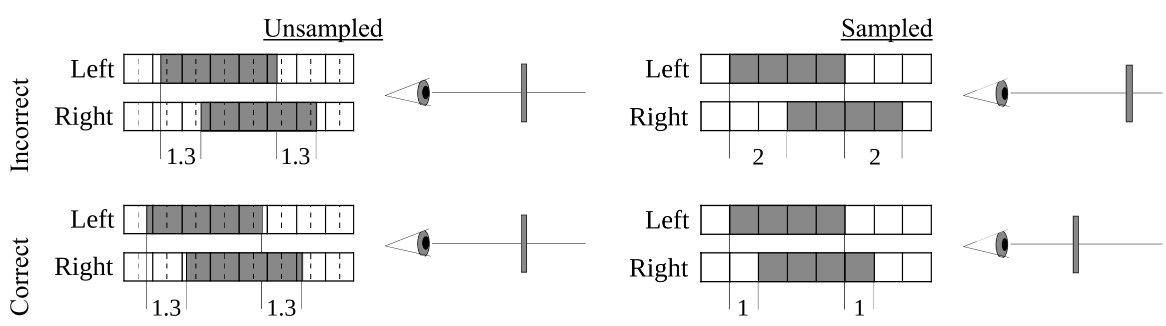

Figure 4: Inconsistencies in stereo depth due to sampling the disparity. The top row shows how the disparity of an object could incorrectly round up, making the object appear further in stereo depth than intended.

The occurrence of this artifact is a function of the distance from the line of sight. Depending where the object is located, the endpoints will round differently.

The detectability of inconsistent stereo depth will depend on whether the system presents steps in stereo depth that are smaller than human stereoacuity (i.e., around 2' of arc[10]). However, many flat-screen based stereo displays do not achieve this goal and steps in stereo depth are perceivable. Thus, the inconsistencies in an object's disparity are likely to be detectable.

2.3. Sampling Artifacts in Stereo and Perspective Displays

When stereo and perspective depth cues are combined, the aforementioned inconsistencies and inaccuracies are present, as well as an additional inconsistency caused when the horizontal size of the object is different in the left and right views. We call this the horizontal edge inconsistency, where one edge has a different disparity than the other, as illustrated in Figure 5.

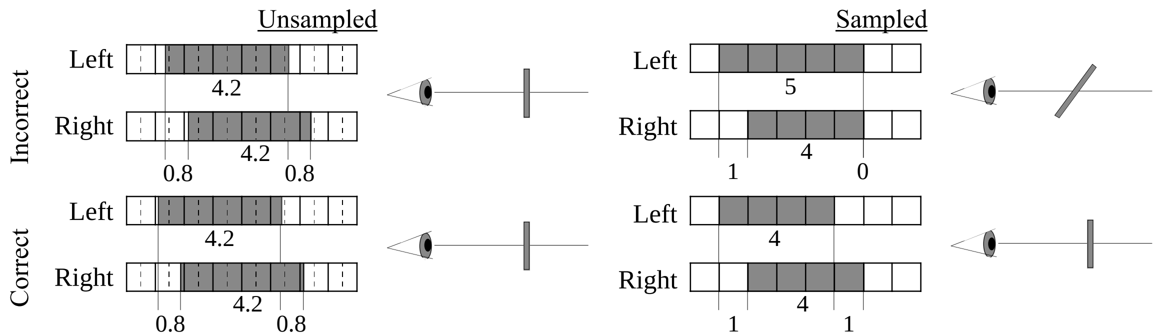

Figure 5: Inconsistencies in disparity and size lead to different disparities at the left and right edges of an object, as shown in the top stereo pair. The top row shows how disparity of one edge could round differently than the other edge, resulting in an object that is slanted in stereo depth.

We conducted an informal experiment to determine how this inconsistency is perceived as the size of the inconsistency increases. Subjects reported that a square's size appeared to grow or shrink when disparity differed by less than 0.38°; for larger disparities, stereo fusion was impossible. That is, subjects saw a change in the horizontal size of the fused square, not a change in its depth. This occurred for both negative and positive differences in disparity and textured and flat-shaded objects. This implies that an object experiencing this inconsistency will not be perceived as slanted in stereo depth, but will be seen as wider or narrower because the perspective cues showing the object as flat dominate the stereo cues.

Importantly, the inconsistencies described above are not likely to occur at the same time. Consequently, conflicting depth cues may be presented:

- Disparity and perspective location cues may conflict with perspective size cues
- Disparity cues may conflict with perspective location and perspective size cues
- Left and right edge disparity cues may conflict with perspective location and perspective size cues

The magnitude of these conflicts will be at most two pixels (i.e., one cue rounds up incorrectly, the other cue rounds down incorrectly). However, one-pixel rounding errors are more common.

3. EXPERIMENTATION

As described above, spatially sampling a scene with perspective and stereo depth information leads to inconsistent presentation of an object's size, disparity and horizontal edge disparities. This section describes a set of experiments designed to assess the detectability of these artifacts. Furthermore, we want to determine which cue dominates when these artifacts lead to conflict in the presentation of depth.

We chose to use a relative depth judgement task that was a variation of the Holway & Boring size-distance experiment[11], with the notable difference that the depths of two objects were compared while they were adjacent. We chose this methodology rather than the Howard-Doleman apparatus[12] used by Nagata[13] and others[14] because we want to evaluate the aforementioned artifacts in static imagery without additional spatio-temporal sampling artifacts found in moving imagery.

3.1. Method

The three inconsistencies, size, disparity and horizontal edge, were treated in four separate experiments run in succession. Size and disparity inconsistencies were tested in one experiment each and horizontal edge inconsistencies in two. We tested the detectability of changes in depth by presenting two objects and asking the subjects to make a judgment about their relative depth.

3.1.1. Size and Disparity Inconsistencies

In these trials, the perspective sub-cues and the disparity were manipulated by one pixel. The one-pixel change in disparity and perspective would result in either complementary or conflicting cues, therefore replicating the size and disparity inconsistencies that occur in typical 3D scenes. To test cue dominance when disparity inconsistencies occur, the perspective location and size were changed separately from the disparity (Experiment 1). For size inconsistencies, the perspective size was changed separately from the disparity and the perspective location (Experiment 2). These conditions are illustrated in Figure 6.

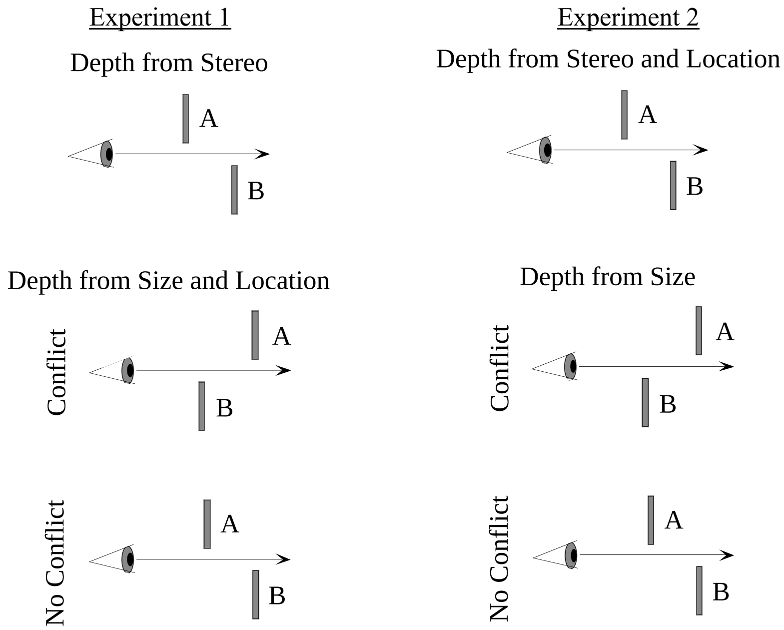

Figure 6: The depths presented with perspective and stereo cues in Experiments 1 and 2.

Subjects were asked, “Which object appears closer?” In this manner, we could determine if a subject is able to see the one-pixel changes in depth and which source of depth information will dominate in conflict situations.

3.1.2. Horizontal Edge Inconsistencies

Informal experimentation on the tolerance for inconsistencies in disparity between an object’s horizontal edges suggests that an afflicted object appears wider or narrower, rather than slanted in stereo depth. If perspective cues are primarily used to resolve depth, then this apparent size change may result in confusion about the distance to the object. In addition, as the width of the object increases, detecting a change in stereo depth will become more difficult since stereo acuity degrades with increased eccentricity[10]. Given these considerations, we designed two experiments, one to assess the detectability of edge consistencies and the other to evaluate the potential effects of cue dominance.

In Experiment 3, subjects viewed two objects with identical perspective depth. A one-pixel edge inconsistency was present on one object on half the trials. Subjects responded to the question, “Are the two objects at the same depth?” In this manner, the detectability of these inconsistencies could be evaluated, regardless of whether they influence perspective or stereo depth.

In Experiment 4, one object was presented with a one-pixel edge inconsistency and therefore an ambiguous stereo depth. The other was matched to either the front or the back edge in stereo depth. Then, perspective information was provided to either assist or hinder the perception of the relative depth of the second object. Subjects were asked, “Which object appears closer?” In this manner, we hoped to determine if the interaction between stereo and perspective cues would be altered when the stereo cue was inconsistent. These conditions are illustrated in Figure 7.

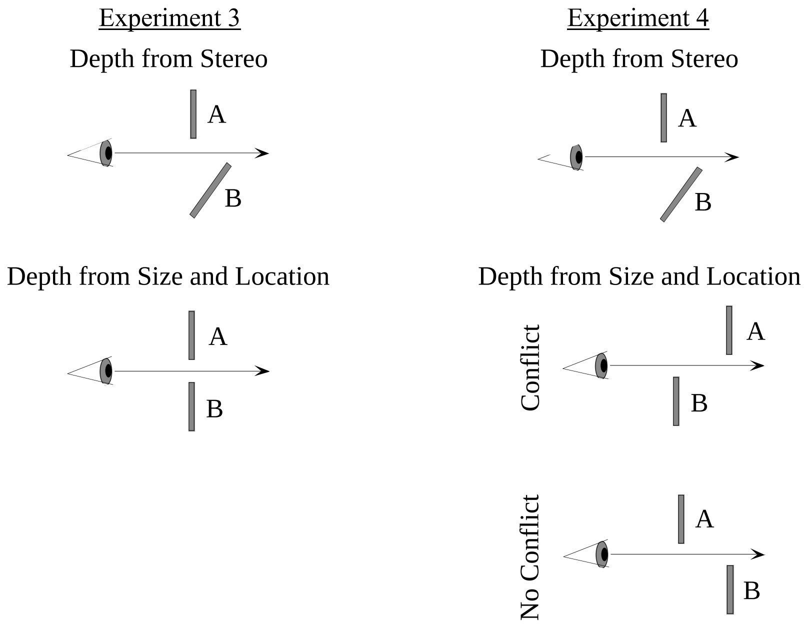

Figure 7: The depths presented with stereo and perspective cues in the two edge inconsistency experiments.

These four experiments were run in succession. Since strategy is an important consideration in cue conflict experiments, we expected a more consistent viewing strategy if the separate experiments were conducted in succession.

3.1.3. Subjects

Eight subjects were recruited from among the students and staff at the University of Cambridge Computer Laboratory. All had either normal vision or vision corrected to normal (self-reported). A simple test was performed to ensure subjects had normal binocular vision. The subjects all had substantial experience viewing and manipulating 3D CGI.

3.1.4. Apparatus and Stimuli

The visual stimuli were produced on a 17-inch CRT using CrystalEyes shutter glasses. The resolution of the image presented to each eye was 1280x491 pixels. A 12cm black frame was constructed to reduce framing effects[15]. The viewing position was located approximately 60cm from the display surface. The field-of-view of the display was 32°x25°. The geometric FOV was chosen to match the real world FOV.

A standard graphics package, OpenGL, was used to generate the images on an SGI Indy workstation. Object sizes and locations were chosen to ensure the appropriate type of inconsistency would be presented. However, since size and separation affect perspective cues, we kept the size and relative location of the objects within the thresholds suggested by previous informal experimentation[3].

Because perspective depth cues are ambiguous, presenting two objects with only perspective information would not provide a convincing sense of depth. Therefore, other cues were presented, provided they did not change the shape or location of the primary objects themselves and thus did not provide conflicting depth information. Following Kim et al.[16], a surrounding box and grid were added. Textures were also added to provide flatness information across the surface of the objects.

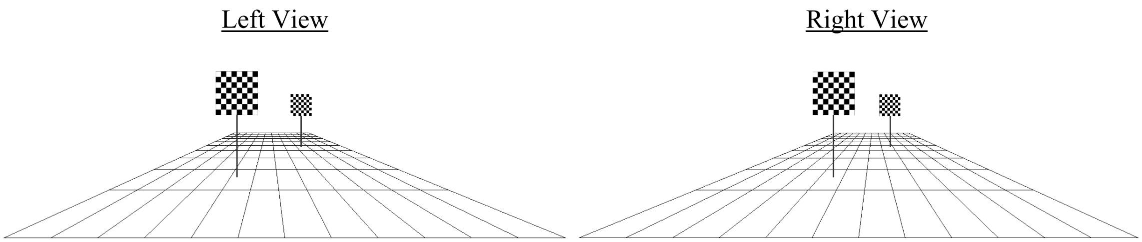

Figure 8: Sample stimulus. Stereo and perspective cues are exaggerated in this figure for increased clarity.

3.1.5. Design

In all four experiments, a two-alternative forced-choice procedure was used for data collection. Similarly, three different depths were used, selected to fall within the range of the binocular disparity threshold (i.e., ±1.5° as in Lipton[8]). The independent variables for each experiment were:

- Experiment 1: Disparity inconsistency, "Which is closer?" (72 trials)
  - Which object moved (left or right)
  - Direction of movement of comparison object (closer or further)
  - Conflicting or complementary cues in the comparison object
- Experiment 2: Size inconsistency, "Which is closer?" (72 trials)
  - Which object moved (left or right)
  - Direction of movement of comparison object (closer or further)
  - Conflicting or complementary cues in the comparison object
- Experiment 3: Horizontal edge disparity inconsistency, "Are they the same depth?" (72 trials)
  - Which object moved (left or right)
  - Direction of movement of comparison object (closer or further)
  - Which edge moved (inside or outside)
  - Type of movement (± one pixel of disparity on selected edge)
- Experiment 4: Horizontal edge disparity inconsistency, "Which is closer?" (96 trials)
  - Which object moved (left or right)
  - Which edge moved (inside or outside)
  - Direction of movement of edge (closer or further)
  - Type of movement (± one pixel of disparity on selected edge)
  - Stereo location of comparison object (same as front or back of slanted object)
  - Conflicting or complementary cues in the comparison object

These cases were presented in random order within each experiment. Because we wanted any development of strategy to be consistent across all subjects, the experiments were always presented in the same order.

For the "Which is closer?" experiments (1, 2, and 4), the dependent variable was the correctness of the response when the cues were complementary and which cue dominated when the cues conflicted. For the "Are the objects the same depth?" experiment (3), the dependent variable was the correctness of the response.

3.1.6. Procedure

Upon arrival, subjects were given simple instructions that carefully avoided suggesting any viewing strategies. They performed a set of sample trials to familiarize themselves with the experimental task. Subjects were seated comfortably and the viewing conditions were checked to ensure consistency. Although their heads were not restrained, subjects were asked to remain as still as possible through a set of trials. The room was darkened by extinguishing overhead lights and blacking out the window.

Each stimulus was presented until the subject responded with a mouse click. Subjects were asked to respond to the question "Which object appears closer?" with the corresponding mouse button or to "Are they the same depth?" with the left mouse button for "yes" and the right for "no". A blank screen was then displayed for one second to clear any afterimages. After each experiment, subjects were given a break to help avoid visual fatigue. A complete run took about thirty minutes.

3.1.7. Results – Size and Disparity Inconsistencies

The mean correctness in Experiment 1 (size inconsistency) was 98.4% (SD = 12.5). Similarly, the mean correctness in the Experiment 2 (disparity inconsistency) was 98.4% (SD = 12.5). In these cases, the ability to detect a change in depth when perspective and stereo information were complementary was near perfect.

Strategy played a major effect on whether stereo or perspective information was used. When the disparity was in conflict with the perspective size and location, six subjects used stereo to determine depth on 99.4% (SD = 7.5) of the trials, while one subject used perspective cues on 100% (SD = 0.0) of the trials.

One subject changed strategies in the middle of Experiment 1, switching from using perspective cues to using stereo cues. When the size conflicted with the stereo and location information, seven subjects used stereo to determine the depth on

98.1% (SD = 13.5) of the trials, while the same subject as in Experiment 1 used perspective cues on 83.3% (SD = 37.8) of the trials.

Clearly, the strategy chosen at the beginning of the experiment (presumably during the sample trials) influenced the perception of the objects. However, most subjects used stereo in situations where the cues were conflicting or ambiguous. This suggests that most viewers will see the stereo depth presented when inconsistencies in size or disparity occur.

Two analyses of variance (ANOVA) were performed on the data from Experiments 1 and 2, using correctness as the dependent variable. In both experiments, the displayed depth of the objects had no effect on accuracy or cue use in either the conflicting or the complementary cases.

3.1.8. Results – Horizontal Edge Inconsistencies

On the first of the two edge inconsistency experiments (Experiment 3), subjects were able to determine that the objects were at the same depths on 79.7% (SD = 40.3) of the trials when the objects were unchanged. An ANOVA was performed and this result was statistically significant, $F(1,574) = 37.93$, $p < 0.01$. When one object was slanted, the detectability of the inconsistency when the inside edge differed was significantly better than when the outside edge differed, $F(1,574) = 10.59$, $p < 0.01$. This is consistent with our expectation that the inner edge will dominate the comparison since stereo acuity decreases with increased eccentricity. The depth of the objects had no effect on accuracy.

The second edge inconsistency experiment had no significant results. Subjects were either guessing on most cases or the effects of the various conditions were too subtle to show up in the statistical analysis. While subjects reported visual fatigue and discomfort for all four experiments, they especially mentioned difficulties with the two edge inconsistency experiments.

3.2. Discussion

The four experiments assessed the detectability of changes in stereo and perspective depth cues due to inconsistencies in size, disparity and horizontal edges. These inconsistencies can cause conflicts between stereo and perspective depth cues; therefore, cue dominance was also evaluated for each type of inconsistency.

When size and disparity cues were complementary, subjects achieved near-perfect accuracy. This confirms that combining depth cues improves the perception of depth. When size and disparity were inconsistent, stereo dominated perspective. In both the conflicting and complementary cases, the difference was reliably detected in a typical viewing situation (i.e., pixel width = 1'24" of visual angle). In static images, we expect stereo to be the stronger cue, since perspective cues are particularly aided by disambiguation due to movement.

The inconsistency between the disparities of the horizontal edges of an object confused the subjects. Even when complementary information was present, judgements were still difficult. Again, since stereo was the dominant cue in this situation and was ambiguous across the object, this led to decreased accuracy. The difference in stereo depth was more easily detected when the difference occurred on the inside edges of the objects. This implies that differences that occur too far apart will not be detected. Edge inconsistencies also led to more reports of visual fatigue and discomfort.

4. CONCLUSIONS

The experiments described in this paper have some impact on stereo display design. In systems where antialiasing in software or hardware is impractical and spatial resolution is limited, sampling artifacts occur that affect the accuracy of depth perception. Therefore, the presentation of stereo and perspective depth information in static imagery must account for these artifacts if accuracy is critical for task performance. Furthermore, these artifacts result in conflicts in depth cues which cause additional visual fatigue and viewer discomfort. The value of presenting stereo depth information in a system where these artifacts will occur should be weighed against task requirements; if the task requires fine relative depth judgments in static images and the user cannot tolerate some discomfort, these artifact will severely hinder the effectiveness of the display system. These problems can be ameliorated with typical antialiasing methods, although alternative, more efficient methods have been proposed[3].

ACKNOWLEDGEMENTS

This paper represents work performed as part of a doctoral thesis[2] that was generously supported by Trinity College, the Cambridge University Computer Laboratory, the Cambridge Overseas Trust, and the Overseas Research Student scheme offered by the Committee of Vice-Chancellors and Principals of U.K. Universities. I would also like to express my thanks to my thesis advisor, Peter Robinson, as well as Neil Dodgson and the members of the Rainbow Research Group at the Cambridge University Computer Laboratory.

REFERENCES

1. J. Hsu, Z. Pizlo, C. Babbs, D. Chelberg, & E. Delp, "Design of Studies to Test the Effectiveness of Stereo Imaging Truth or Dare: Is Stereo Viewing Really Better?" in *Proc. SPIE – Stereoscopic Displays and Virtual Reality Systems*, vol. 2177, pp. 211-222, 1994.
2. E.T. Davis & L.F. Hodges, "Human Stereopsis, Fusion, and Stereoscopic Virtual Environments", in *Virtual Environments and Advanced Interface Design*, W. Barfield & T. Furness, eds., pp. 145-174, Oxford University Press, New York, 1995.
3. J.D. Pfautz, *Depth Perception in Computer Graphics*, Doctoral Dissertation, University of Cambridge, UK, 2000.
4. L. Lipton, *The CrystalEyes Handbook*, StereoGraphics Corporation, San Rafael, California, 1991.
5. N. Dodgson, J. Moore, S. Lang, G. Martin, & P. Canepa, "A 50" Time-Multiplexed Autostereoscopic Display", in *Proc. SPIE – Stereoscopic Displays & Applications XI*, vol. 3957, 2000.
6. L.F. Hodges & E.T. Davis, "Geometric Considerations for Stereoscopic Virtual Environments", *Presence*, **2**(1), pp. 34-43, 1993.
7. R. Ledley & R. Frye, "Processing of Stereo Image Pairs: Elimination of Depth Planes Using the 'Cut-Plane' Procedure", in *Proc. SPIE – Stereoscopic Displays and Virtual Reality Systems*, vol. 2177, pp. 66-77, 1994.
8. L. Lipton, "Composition for Electrostereoscopic Displays", in D. McAllister, ed., *Stereo Computer Graphics and Other True 3D Technologies*, pp. 11-25, Princeton University Press, Princeton, New Jersey, 1993.
9. L. Harrison & D. McAllister, "Implementation Issues in Interactive Stereo Systems", in D. McAllister, ed., *Stereo Computer Graphics and Other True 3D Technologies*, pp. 117-151, Princeton University Press, Princeton, New Jersey, 1993.
10. Y. Yeh, "Visual and Perceptual Issues in Stereoscopic Color Displays", in D. McAllister, ed., Stereo *Computer Graphics and Other True 3D Technologies*, pp. 50-70, Princeton University Press, Princeton, New Jersey, 1993

A. Holway & E. Boring, "Determinants of Apparent Visual Size with Distance Variant", *American Journal of Psychology*, **54**, pp. 21-37, 1941.

B. Graham, "Visual Perception", in *Handbook of Experimental Psychology*, S. Stevens, ed., pp. 868-920, John Wiley & Sons, New York, 1951.

11. S. Nagata, "How to Reinforce Perception of Depth in Single Two-Dimensional Pictures", in *Pictorial Communication in Virtual and Real Environments*, S. Ellis, ed., pp. 527-545, Taylor & Francis, London, 1993

A. Utsumi, P. Milgram, H. Takemura, & F. Kishino, "Investigation of Errors in Perception of Stereoscopically Presented Virtual Object Locations in Real Display Space", in *Proc. HFES 38th Annual Meeting*, pp. 250-254, 1994.

12. S. Ohtsuka, Y. Ishigure, Y. Kanatsugu, T. Yoshida, & S. Usui, "Virtual Window: A Technique for Correcting Depth-Perception Distortion in Stereoscopic Displays", in *Proc. SID International Symposium*, pp. 893-896, 1996.
13. W. Kim, S. Ellis, M. Tyler, B. Hannaford, & L. Stark, "Quantitative Evaluation of Perspective and Stereoscopic Displays in Three-Axis Manual Tracking Tasks", *IEEE Trans. on Systems, Man and Cybernetics*, **17**(1), pp. 61-72, 1987.