Tolerance of Misalignment in Stereoscopic Systems

Elaine W. Jin and Michael E. Miller; Eastman Kodak Company; Rochester, NY/U.S.A.

Abstract

Previous literature demonstrates that alignment errors between two imaging channels in a stereoscopic system can induce discomfort and eye fatigue. While tolerances for these errors have been proposed, they often vary substantially between publications. The effect of three spatial alignment errors (vertical misalignment, magnification difference, and rotation) between the two imaging channels was investigated using two experimental protocols: a short experimental trial (8 s) and a longer experimental trial (18 min). A unified tolerance specification is proposed to link the three types of errors. In addition, the results suggested that the acceptable tolerance for misalignment might be affected by viewing time.

Introduction

Alignment is recognized as an important specification for stereoscopic displays [1]. Misalignment between corresponding objects in a stereoscopic image pair can interrupt image fusion, produce visual fatigue, and produce other undesirable side effects, such as distorted depth perception. Misalignment between the two imaging channels can be produced by vertical shift, magnification difference, and rotation between the two imaging channels. Various authors [2] have investigated the effect of these misalignments. Unfortunately, the results vary substantially between publications. A 2004 publication by Kooi and Toet [3] attempted to clarify the literature by studying the tolerance for various sources of misalignment at strictly specified experimental conditions. However, the authors did not identify a unified metric, their results are not conclusive for all stereoscopic display parameters, and their results have not been independently verified. Further, the authors employed experimental sessions of short duration and, therefore, their results might not be indicative of human performance when using the stereoscopic display in applications that require extended use.

The identification of a unified metric is particularly important. Most stereoscopic systems exhibit some degree of misalignment as the result of differences in vertical and rotational alignment, as well as magnification differences. Further, these errors might be introduced during image capture, display, or intermediate steps, such as image digitization. While it is important to understand the maximum allowable misalignment for each type of misalignment in the absence of other types of misalignment, the system specification must consider the fact that two or more sources of misalignment might be present within a single system.

It is convenient to evaluate the effects of misalignment using short exposure times. Unfortunately, it is possible that effects of system misalignment might be cumulative or latent in nature, such that the effects of these imaging system deficiencies are expressed only after long-term exposure. In this paper, we explore the effect of typical misalignments using a traditional, short-duration experimental protocol. Based on the results of this study, we propose one possible unified metric for system misalignment. We then attempt to verify the results of the short duration study through a study of longer duration.

Method

Participants

Ten volunteers participated in each study. Although each group of ten participants was independently selected, some individuals served as participants in multiple experiments. All participants had normal visual acuity, normal color vision, and normal stereoacuity.

Apparatus

A pupil imaging system for presenting a stereoscopic pair of virtual images at optical infinity was employed [4]. This system was developed to provide a high quality stereoscopic image with no cross talk. The display presented images with a field-of-view of 45° in the horizontal direction and 36° in the vertical direction. This display had 1280 horizontal by 1024 vertical addressable pixels, providing a pixel resolution of 2.1 arc min.

The system was aligned to provide the best attainable physical alignment. Because of imperfections in the optical components, however, there were residual errors in the system. To compensate for these misalignment errors, a software alignment was applied to pre-process the left and right input images so that the two perceived images were well aligned. This software utilized a display displacement map acquired by capturing a set of test targets using a calibrated twin digital video camera set, and a warping algorithm to warp the input images based on the displacement map. The alignment accuracy after algorithmic correction was within 0.5 pixels for each displayed pixel.

Image-Processing Path

Beginning with an image-processing path providing a nearly perfectly aligned image, additional image-processing steps were introduced to simulate misalignment between the left and right images. Each misalignment was introduced into either the left or right eye image to allow the misaligned image to be provided to one eye while a base image was displayed to the other eye. The experimenter originally generated images having extreme misalignment errors of each type and pre-selected conditions that were not expected to cause the participants excessive discomfort.

To introduce vertical misalignment, the top 100 pixels were cropped from the top of each image, and the resulting images were overlaid on top of a full screen black image such that one image was offset by one of 0, 5, 10, 15, 20, 25, and 30 pixels vertically, providing vertical misalignment conditions ranging from 0 to 63 arc min. Cropping was applied when introducing vertical shifts and rotations to allow the same image content to be displayed for each experimental condition. To introduce magnification error, the size of one of the images in the image pair was reduced to 100%, 99.5%, 99%, 98%, 97%, 96%, or 95% of their original size and overlaid on a full screen black image. To induce rotational differences, the images were cropped by 50 pixels along each side and a clockwise rotation was applied to the left or right image, with the angles of 0°, 0.4°, 0.8°, 1.2°, 1.6°, 2.0°, or 2.4° before being overlaid on a full screen black image.

Protocol

Two protocols were used in this study. For the short-term study each trial was composed of three consecutive image presentations. The participants first viewed a fixation target as shown in Figure 1. This target encouraged the participants to converge their eyes to optical infinity, a condition that was assured when the two vertical line segments above and below the cross were perceived as being vertically aligned. When the target was fixated, the participant clicked the mouse and was shown the first stereoscopic image pair. After 8 s of presentation the images were replaced with a response screen. The participants indicated the amount of eyestrain experienced when viewing the stereoscopic image pair using the levels "None," "Mild," "Moderate," "Strong," and "Severe." The participants also indicated whether they were able to fuse the images and whether they perceived depth in the image to be "Good" or "Degraded." After providing these responses, the participants pressed a button labeled "Next" to present the next condition.

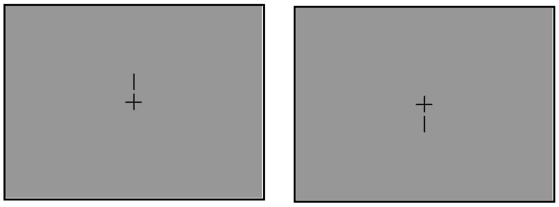
(a) *(b)*

***Figure 1**. Fixation targets for left eye (a) and right eye (b) views.*

Four scenes were used in the short-term study as depicted in Figure 2. Two of the scenes were generated using computer graphics (Chart and Protein) and two were captured scenes (Ocean and Street). The images were generated or captured with parallel cameras and the content in the image was positioned to be comfortable when displayed in this system. It should be noted that the capture system for the scene images was a stereoscopic camera that was designed to minimize optical errors. The images were captured on film and digitized. The digitized images were manipulated to eliminate vertical offsets, magnification, and rotation differences.

During the short-term study, the test conditions were repeated twice. The order of presentation was fully randomized for the scenes, levels, and repetitions. There were 108 trials in total for each manipulation, and it took about 45 min for a participant to complete the experiment. The effects of vertical offset, rotation, and magnification were determined in separate experiments. The experiments were completed in a dark room.

During the long-term study, the participants were asked to view images with different degrees of vertical misalignment. During this study, the participants were asked to complete one 18-min trial within each experimental session, during which the participants viewed a group of images, each image pair having the same vertical misalignment. This larger group of 46 image pairs, which included the 4 images from the short-term experiment and 42 additional captured images, were also hand-edited to remove any initial vertical offsets, magnification differences, or rotations within the captured image pairs.

***Figure 2**. Test images used in the short-term study: Chart, Protein, Ocean, and Street.*

Experimental trials within the long-term study were separated by at least 24 h, and an attempt was made to schedule consecutive sessions on consecutive days. The order of presentation for the level of vertical shift was randomized for all participants. During these experimental sessions, the participants were actively engaged in counting the occurrence of difficult-to-find objects within the scenes. This task required the user to be actively engaged in visual search during the experimental session.

The dependent measures in the long-term study included monitoring changes in visual acuity, stereo visual acuity, dexterity, fusing capability, and subjective ratings that occurred between measures taken before and after the stereoscopic images were viewed. The subjective rating scales were extracted from the simulator sickness questionnaire [5] and required participants to provide ratings on a five-point, continuous rating scale. Participants' ratings included dizziness (eyes open and closed), discomfort, fatigue, headache, eyestrain, difficulty focusing, difficulty concentrating, blurred vision, and confusion.

Results

The participants were able to fuse the vast majority of images. The fusion and depth measurements differed primarily by scene and did not reliably differentiate the experimental conditions. Therefore, the results for the short-term experiments will focus on the eyestrain data. The first step of data analysis was to convert the category scaling data to interval scale data through the application of Torgerson's Law of Categorical Judgment [6]. The scale that was derived was defined such that the boundary between eyestrain levels "None" and "Mild" was assigned a value of zero.

The relationship of the subjective interval scale as a function of the manipulation parameters is shown in Figure 3. The results for vertical shift are shown in the top panel, the results for the magnification difference are shown in the middle panel, and the results for rotation are shown in the bottom panel. Each line in Figure 3 represents the scale values that were obtained for one individual scene.

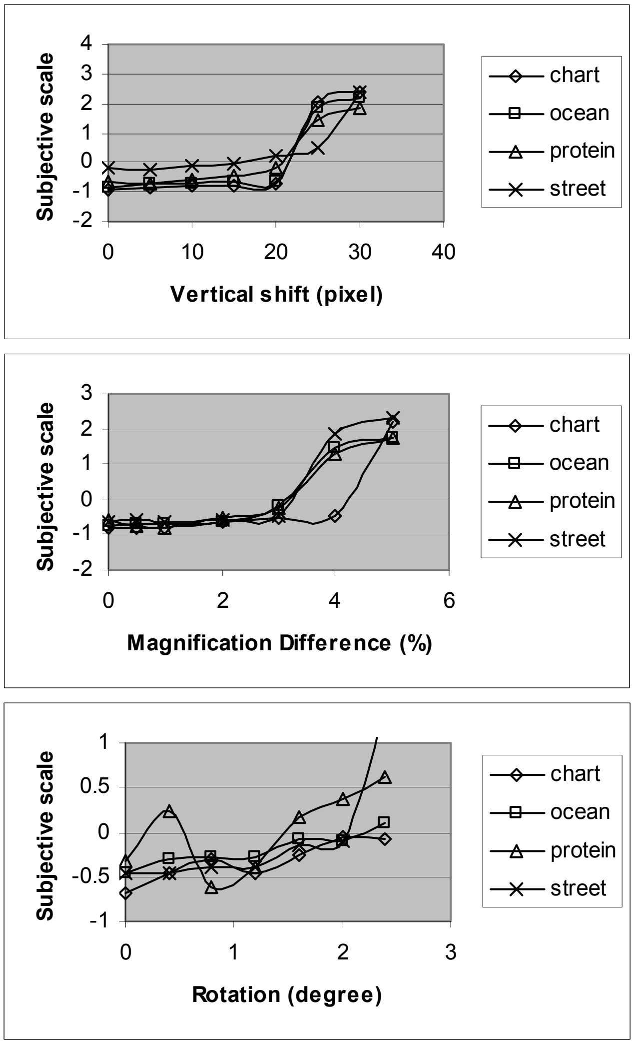

Figure 3. Results of short-term study. Top: vertical shift; Middle: magnification difference; Bottom, rotation.

As shown in Figure 3, the relationship between the degree of vertical (top) and magnification (middle) misalignment with the interval scale of eyestrain is generally monotonic and, while some degree of scene dependence might be observed, a clear psychometric function is observed with a rapid increase in eyestrain reported beyond an experimental level. Unfortunately, the relationship between rotation (bottom) and the interval scale of eyestrain is not as strong and, in fact, the resulting function is not monotonic with increases in rotational misalignment for some of the scenes.

To understand the exact location of the boundary between eyestrain levels of "None" and "Mild," a spline was fit to each of the curves. Table 1 shows the intercept levels for each combination of scene and manipulation. The intercept level for Protein on rotation is in italics because the fitting results are not unique for this combination. To summarize, the vertical shift intercept level is 16 pixels or higher, the magnification intercept level is 3.1% or higher, and the rotation intercept level is 2.1° or higher.

Table 1: Tolerance levels for the short-term study

	Vertical (pixel)	Magnification (%)	Rotation (degree)
Chart	21.6	4.4	2.4
Ocean	21.6	3.1	2.3
Protein	20.8	3.2	*0.1*
Street	15.7	3.3	2.1

Many of the dependent measures for the long-term study were insensitive to the manipulations described in this paper. That is, when exposed to analysis of variance, no statistically reliable trends were observed for the majority of the dependent measures. The change in the subjective rating of eyestrain provided the only exception, providing a monotonically increasing curve as a function of vertical misalignment as shown in Figure 4. This trend was statistically reliable at $p < 0.07$. As shown in Figure 4, no increase in eyestrain occurred when subjective eyestrain ratings that were recorded after the images were viewed were subtracted from the eyestrain ratings that were recorded before the images were viewed for the 0 vertical misalignment condition. However, small but statistically reliable changes in eyestrain were recorded as the vertical misalignment was increased from 0 to 5 pixels (10.5 arc min) and again as misalignment was increased beyond 10 pixels (21 arc min). It is worth noting that the short-term study results demonstrated significant increases in eyestrain only as the vertical offset was increased to more than 16 pixels (33.6 arc min).

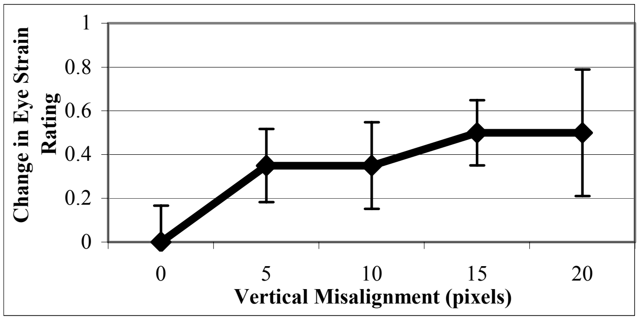

***Figure 4**. Change in eyestrain rating during experimental session as a function of vertical misalignment.*

Discussion and Conclusions

As mentioned in the introduction, it is desirable to provide a single metric that might be used to quantify the acceptability of any stereoscopic system, even when the system suffers from several sources of misalignment. To do this, we hypothesize that because the human visual system is unable to discern horizontal misalignments within the system from true horizontal disparities, the human visual system should be tolerant of horizontal misalignment but is likely to be quite sensitive to vertical misalignment. Implied by this hypothesis is vertical misalignment in the resulting image, regardless of the type of spatial distortion that is present in the imaging system, produces discomfort and, therefore, a value, such as the maximum vertical misalignment in the resulting image, may provide a useful single metric.

To test this hypothesis, it is necessary to determine the vertical misalignment that results from magnification and rotation within our experiment. Vertical misalignment is different for each pixel in the image, but largest at the corners of the image when it is introduced through the magnification or rotation manipulations. The maximum vertical shift at the corner of the image (Δy) due to magnification differences can be calculated using equation (1), where k represents the magnification change in percent. The maximum vertical shift at the corner of the image (Δy) due to rotation can be calculated using equation (2), where d is the diagonal of the display in pixels, α is the rotation angle, and β is the angle formed by the diagonal and the horizontal side of the image. For the test images, $\sin\beta = 924/1180$ and $d = 1498.7$.

$$\Delta y = (k/100)*(1024/2) \tag{1}$$

$$\Delta y = d/2*(\sin(\alpha + \beta) - \sin\beta) \tag{2}$$

Using these equations, the maximum vertical displacement given the tolerances for the short-term experiment was calculated and shown in Table 2. With the exception of the rotation condition for the protein scene, the converted maximum vertical shift values are close to the vertical shift result. Overall the vertical shifts for all combinations are in close vicinity to each other. Based upon this result, it appears that the maximum vertical displacement of any two pixels in the resulting stereoscopic image pair might provide a measure that can serve as an initial summary metric. Therefore, when displaying typical pictorial stereoscopic images, the maximum allowable vertical displacement for any two pixels in a stereoscopic image would be on the order of 16 pixels within our display, or 33.6 arc min.

Table 2: Converted tolerance levels for the short-term study

	Vertical (pixel)	Magnification (pixel)	Rotation (pixel)
Chart	21.6	22.3	19.0
Ocean	21.6	16.1	18.3
Protein	20.8	16.4	*0.5*
Street	15.7	16.7	16.5

Overall, the study showed that if individual measures are to be used for each source of misalignment, the short-term tolerance is 33.6 arc min for vertical misalignment, 3.1% for magnification differences, and 2.1° for rotation. These values are relatively consistent with the existing literature [3]. However, when the scene contains multiple sources of misalignment, it is believed that a summary metric must be employed. A total vertical displacement between the stereoscopic representations of any image feature greater than 33.6 arc min should be avoided.

Measurement of the effect of long-term viewing of stereoscopic displays remains a challenge. In this study we probed this effect using a number of performance and subjective measures, most of which did not provide a reliable dependent measure as vertical image misalignment was manipulated. However, differences between "before" and "after" reports of eyestrain increased reliably with increases in vertical misalignment and appeared to increase even when the vertical misalignment was less than 5 pixels (10.5 arc min). Because of the limited observer population and the lack of a strong correlation with other measures, we do not consider this result conclusive. This result does, however, draw into question the use of short-term experimental protocols to determine longer-term effects in stereoscopic displays. Therefore, future research should concentrate on identifying more reliable long-term performance and subjective measures. One could argue that the inability to identify dependent measures, which clearly correlate with changes in display alignment, indicates that alignment differences are inconsequential. However, some participants clearly expressed discomfort when using the misaligned stereoscopic system. Therefore, we believe there is clearly an effect of these parameters on user comfort and likely long-term performance. Once better metrics are defined, it will then be necessary to apply these measures to study long-term effects of stereoscopic displays.

References

[1] J. Merritt and A. Woods, Stereoscopic Display Application Issues, Tutorial for IS&T/SPIE 15th Annual Symposium (2003).

[2] C.E. Rash, Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft (SPIE, Bellingham, WA, 2000) pg. 135.

[3] F.L. Kooi and A. Toet, "Visual Comfort of Binocular and 3D Displays," Displays, 25, 99–108 (2004).

[4] J.M. Cobb, Autostereoscopic Desktop Display: An Evolution of Technology, Proc. SPIE Stereoscopic Display and Applications, pg. 139–149. (2005).

[5] R. Kennedy, N. Lane, K. Berbaum, and M. Lilienthal, Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int'l. J. Aviat. Psychol., 3, pg. 203–220. (1993).

[6] W.S. Torgeson, Theory and Method of Scaling (Robert E. Krieger Publishing Company, Malabar, FL, 1985) pg. 205.

Author Biography

Elaine W. Jin holds a Ph.D. in Optical Engineering from Zhejiang University and a Ph.D. in Psychology from the University of Chicago. She is a human factors and optical engineer in the Display Science & Technology Center of Eastman Kodak Company. Her research interests include vision modeling, color preference, image quality, and optical modeling of display performance. She is a member of IS&T.

Michael E. Miller is a human factors and systems engineer in the Display Science & Technology Center of Eastman Kodak Company. His current research is focused on OLED structures and image processing to improve display quality. He holds B.S. and M.S. degrees in Industrial and Systems Engineering from Ohio University and a Ph.D. in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University.