Alternating attention in continuous stereoscopic depth

Steven Poulakos[1,2]*, Gerhard Roethlin[1], Adrian Schwaninger[3], Aljoscha Smolic[1], Markus Gross[1,2]
[1] Disney Research Zurich, [2] ETH Zurich, [3] University of Applied Sciences Northwestern Switzerland

Abstract

The decoupling of eye vergence and accommodation (V/A) has been found to negatively impact depth interpretation, visual comfort and fatigue. In this paper, we explore a hypothesis that placement of visual cues within a scene can assist a viewer in the process of maintaining the V/A decoupling. This effect is demonstrated through the use of a continuous depth plane that connects spatially distinct scene elements. Our experimental design enables us to make the following three contributions: (1) We show that a continuous depth element can improve the time it takes to transition visual attention in depth. (2) We observe that the subjective assessment of fatigue emerges before we detect a quantitative decline in performance. (3) We aim to motivate that stereoscopic 3D content creators may learn scene composition, framing and montage from visual psychophysics.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Display Algorithms I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—3D/stereo scene analysis;

Keywords: stereoscopic 3D, visual attention, depth continuity

1 Introduction

Viewing stereoscopic 3D is inherently an unnatural experience. It has been shown that the decoupling of eye vergence and accommodation experienced during stereoscopic image viewing can lead to depth misinterpretation, discomfort and fatigue [Hoffman et al. 2008]. We explore the impact stereoscopic image viewing can have on the ability to change visual attention in depth. Our hypothesis is that scene composition can help compensate for the eye-vergence and accommodation conflict and facilitate the viewing of stereoscopic content.

To explore this question we attempt to mimic a ubiquitous cinematic scene setting: the basic dialog shot (aka. two-shot) in which viewer attention transitions between two actors [Mascelli 1965]. Figure 1 provides several visualizations of a two-shot (see insets A-D) with stereoscopic scene depth represented on the left. Figure 1-(A) represents an over-the-shoulder shot without depth continuity. The remaining three insets (B-D) provide depth continuity through changes to the camera placement and scene composition. We introduce a variable that corresponds to the difference in scene composition between a mid-level shot (inset A) and a down-shot that provides continuity between two actors (insets B-D). In these examples, continuity is provided by a table, wall or floor element that spans the depth range between the two actors. We hypothesize that these continuous visual depth cues visually link the two

---
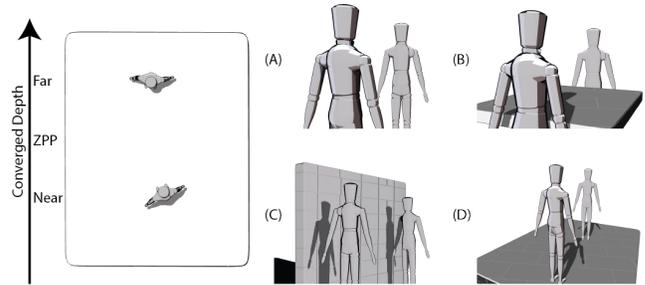
Figure 1: *Visualization of different forms of the two-shot. Actor positions are fixed in depth. (A) Represents an over-the-shoulder shot without depth continuity. Insets (B-D) represent potential ways to provide depth continuity.*

actors and provide an intermediate element that the viewer can use to smoothly saccade from one actor to another.

Our experiment also explores an additional question: can fatigue be measured using a quantitative technique, such as measuring the time required to change visual attention from one actor to another. One would presume that as fatigue grows, the viewer will tire and the speed with which they can change vergence or accommodation will decline, thus providing an indirect measure of fatigue. We administered questionnaires in order to compare self-assessed measure of fatigue against our quantitative approach.

Contributions. Our work makes the following contributions:

- Demonstrate that a continuous depth element can reduce the time required to change visual attention in a stereoscopic scene.

- Compare performance and subjective, questionnaire-based measures of visual fatigue.

- Motivate that stereoscopic 3D content creators may learn scene composition, framing and montage from visual psychophysics.

Paper Organization. Related work is presented in Section 2 followed by the experimental design in Section 3. Results, including our experimental setup and a user evaluation, are presented in Section 4. Limitations and future work are presented in Section 5 before concluding in Section 6.

2 Related Work

Binocular spatial perception is among the most demanding and energy consuming visual tasks viewers perform [Parker 2007]. In natural image viewing, eye vergence, accommodation and pupil diameter work together to form a clear image. The interrelated change is known as a near-triad response [Howard 2002]. The primary benefit of the coupling is an improved visual performance reducing the amount of time to transition visual attention.

Stereoscopic image viewing disrupts natural viewing behavior due

---

*e-mail:steven.poulakos@disneyresearch.com

to the decoupling of eye vergence and accommodation. Many researchers have stated the visual conflict between vergence and accommodation influences visual comfort, depth interpretation or fatigue [Emoto et al. 2005; IJsselsteijn et al. 2006; Lambooij et al. 2009; Patterson 2007; Ukai and Howarth 2008; Yano et al. 2004]. Some have experimentally observed an effect of discomfort or fatigue by comparing stereoscopic image viewing to 2D image viewing [Emoto et al. 2005; Kuze and Ukai 2008; Yano et al. 2002]. However, those findings do not prove the discomfort is caused specifically by the vergence accommodation conflict. Kooi and Toet [2004], for example, explored a variety of additional perceptual distortions caused by stereoscopic viewing that can cause visual discomfort. More recently, Du et al. [2013] explored the influence of stereoscopic motion on comfort.

Hoffman et al. [2008] were the first to convincingly demonstrate an effect of the vergence-accommodation conflict on visual discomfort, depth interpretation and fatigue. This was achieved through the development of a volumetric stereoscopic display, which enables the control of both vergence and (approximate) accommodation cues [Akeley et al. 2004]. Accommodation was interpolated between several fixed states. This experimental system makes it possible to isolate and control the degree of vergence and accommodation conflict. The research was continued by Shibata et al. [2011] who demonstrated the influence of viewing distance and disparity sign (e.g. in front or behind the screen) on discomfort and fatigue.

There are several approaches to compensate for the vergence-accommodation conflict. One approach involves presenting microstereopsis, which is a minimally required disparity [Siegel and Nagata 2000; Didyk et al. 2011; Didyk et al. 2012]. Another approach is to apply linear and nonlinear disparity remapping operators to recompose the scene depth and better utilize the limited depth budget [Lang et al. 2010; Kim et al. 2011]. These concepts have also been applied to remapping multiview autostereoscopic displays to the comfort zone [Masia et al. 2013; Chapiro et al. 2014]. Others have sought to ease attention transitions by aligning the depth position of visually salient scene elements between cuts [Koppal et al. 2011]. Our aim is to demonstrate that it is possible to maintain a large depth volume and utilize other visual cues to improve the time to change visual attention.

3 Methods

We use a restricted cinematic domain, similar to a dialog (aka. two-shot) between two spatially separated actors, to motivate visual attention transitions. We do so because the two-shot is one of the most widely used cinematic shots. We measure the response time necessary to change attention between the two scene elements. To ensure visual attention at the appropriate depth, we used random dot stereogram targets [Julesz 1960] to represent the two actors. An example is provided in Figure 3. We asked the viewer to determine if the center portion of the target emerges or recedes from the target background. This is a task that requires binocular fusion to discriminate between these two conditions.

Subjects Twelve adult volunteers (8 male, 4 female, ages 21-37), with normal vision (corrected and uncorrected) participated in the study. An evaluation of the subjects spatial perception was assessed before the test was administered requiring consecutively correct evaluation of the RDS test target at increasing disparities until the subject demonstrated their ability to correctly perceive these targets at the disparities used in the test.
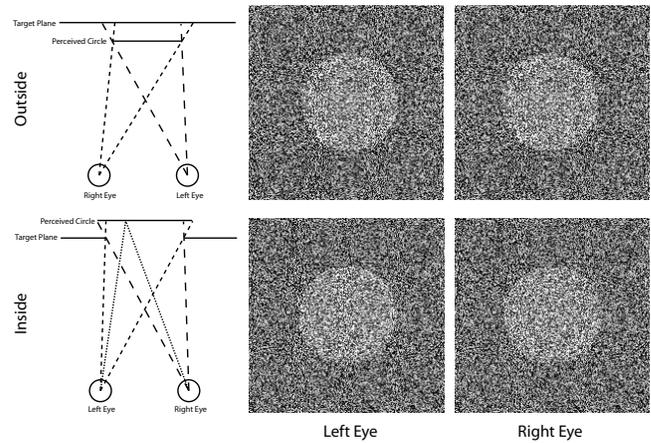
Figure 2: *Modified random dot stereogram (RDS) target stimuli. The circle appeared either in front of the square (Outside) or within the square (Inside). A modified RDS was used to make it easier to fuse the circle.*

Stimuli The stimulus consists of two modified random dot stereograms (RDS) presented at 3 possible stereoscopic depths. Because some participants had trouble viewing binary random dot stereograms, a modified RDS was used. This phenomena was noted by Julesz [1960]. Edge detection and matching is a critical component of disparity tuned cortical vision to produce stereoscopic fusion. In order to keep our potential subject pool as large as possible, we encoded the stereogram with 8 shades of grey (see Figure 2). The darker seven shades were used to encode the dots on the target plane and the lighter seven were used to encode the shape portion of the RDS. Because this shape can be perceived monoscopically, subjects do not identify the encoded shape, but rather interpret the depth location of the shape relative to the target plane (e.g. emerging or receding). The shape was encoded with a 12 pixel disparity and the entire stereogram is 193 pixels square, corresponding to an approximate angular width of 5.6 degrees.

The targets were placed on a floor plane, which is textured with a checkerboard pattern for the continuous depth trials (50%) and not visible (no texture) for the others (see Figure 3). When visible, the continuous depth plane (CDP) extended from the nearest target location to the farthest target. Targets were mounted on either the left or right side of the plane at three depth locations. The depth locations were chosen to test three disparities (-1.72, 0, and +1.69 degrees), corresponding to on-screen disparities of approximately -59, 0 and 58 pixels. The relative disparity for the far target was set so as not to exceed average inter-pupillary distance and was reduced to avoid divergent viewing. The near target was set to provide disparity symmetry across the zero parallax plane.

A 50% grey background was used for two purposes. First, a grey background hid some of the crosstalk that is present in our circularly polarized projection system. This crosstalk can be distracting for some viewers, especially when presenting high contrast, black and white images. Second, the use of 50% grey helped to balance the influence of our continuous depth plane. The plane is a black and white checkerboard pattern, which has a local luminance that alternates between black and white and averages globally to 50% grey.
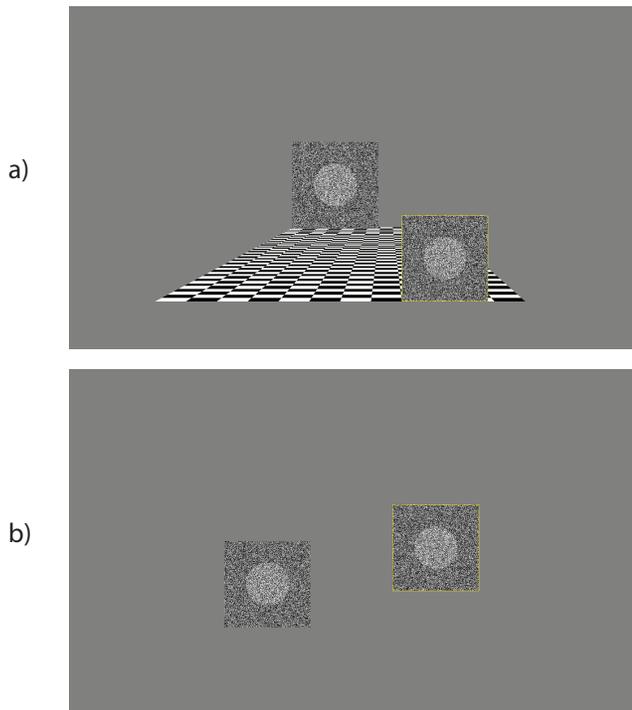
Figure 3: *Example trials in monoscopic view. (a) represents a trial with continuous depth cue. Previous target is located in the far depth and the current target (with yellow border) is in the near depth location. (b) represents a trial without continuous depth cue. Previous target is at near depth and the current target (with yellow border) is at the zero parallax depth location. Note: The continuous depth plane did not exhibit aliasing in actual experiment.*

3.1 Procedure

Participants were tested individually in the same, darkened experiment room. They were seated at a distance of 2 meters from the projected image. The image size was 130 cm x 78 cm (36 degree FOV) and the resolution was 1280x768 pixels. Circular polarizing filters were used to separate the left and right image channels. Participants wore circularly polarized glasses throughout the experiment to view the stereoscopic images and questionnaires. Image brightness on the projectors was reduced to simulate the illumination levels in a dark theater. Approximately 15 cd/m^2 was measured to be passing through the polarizing glasses at a distance of 2 meters. Less light results in a larger pupil size and decreases the depth of focus. User input was provided by a standard computer keyboard, which was placed on the lap of the experiment participant.

Pre-test Participants were first tested to determine if they could perceive the range of depths presented in the study. They were presented near and far targets (alternating one at a time) and asked to respond to the stimulus in the target. Participants were instructed to press specific arrow keys depending on whether they perceived the shape encoded in the RDS to be in front of or behind the RDS target plane. Correct answers resulted in targets that were presented farther from the zero parallax plane. The test ended when the participant demonstrated perception of the depths used in the study. Participants who could not view the required depths were not permitted to continue with the experiment. A short break was given between the pretest and the experiment.

Questionnaire A 23 question survey was given at the beginning, middle and end of the experiment. All questions are summarized in Table 1. The survey first asks a general question about eye fatigue. The next 6 questions originated from a German survey, Kurzfragebogen zur aktuelen Beanspruchen (KAB), which was designed to provide a short scale for assessing stress [Müller and Basler 1993]. The remaining 16 questions are standarized questions from the Simulator Sickness Questionnaire, which was originally designed to assess motion sickness in virtual reality simulators [Kennedy et al. 1993]. Although we lack motion in our experiment, the oculomotor factor of the SSQ is relevant to the viewing of stereoscopic images. Note that the other two factors would be relevant if presenting stereoscopic video. The SSQ enabled us to quickly apply a standard survey that is relevant to stereoscopic eye fatigue. Questions were presented one at a time on the same image screen as the stereoscopic test so that the viewing condition did not change. Using a keyboard, the subject selected the desired response to a question and then confirmed the answer before proceeding.

Experiment Design The experiment is a three factor design. One factor is the *depth change* with 9 levels of change in depth between the two targets (see Figure 4-left and also footnote for depth change codes [1]). The second factor is the presence of the *continuous depth* cue (continuous depth or non-continuous depth), as represented by the checkboard plane. The third factor is the experiment *block* depicted in Figure 5. Blocks are composed of 2 sub-blocks, one for continuous depth trials and one for non-continuous depth trials. The sub-block order is constant per participant and is balanced between participants.

An RDS target is presented in one of 6 locations (3 depths and 2 positions per depth as presented in Figure 4-left). The target to be assessed is outlined with a yellow border. The participant is instructed to decide if the shape encoded on the target is in front of or behind the RDS target plane. Immediately after their response, a new target is presented on the opposite side (left or right) and in one of the 3 depth locations. The new target is outlined with the yellow border as shown in Figure 3. The previous target is still visible, but without the yellow border. Two targets are visible at all times

The entire experiment encompasses 2280 trials, composed of 6 blocks made up of balanced sub-blocks. Each sub-block is composed of 10 cycles. Figure 5 provides an overview of stimuli presentation. A cycle contains every permutation of depth change. Figure 4-right provides an example of a complete cycle. Cycles are used to ensure that the depth change factor is balanced throughout the experiment. Each cycle is composed of 18 trials (9 depth changes and two possible positions: left and right) plus one additional trial (a 19th trial) to transition from one cycle to another. Responses from the additional transition trial are not included in the results because the 19th trial in each cycle would not be balanced. The orientation of the encoded shape (in front or behind the RDS target plane) is randomized. Response time and response are recorded for analysis.

Participants were permitted to take breaks during the experiment. A break was encouraged in the following two ways: First, the participant could take a break after answering all questions in a survey and before beginning the next block of experiment trials. Second, if a break was needed during experiment trials, participants were asked to not respond for at least 10 seconds. Throughout the experiment, the participant remained in the experiment room and under

---

[1]Depth change codes: **F-F** (Far to Far), **N-N** (Near to Near), **Z-Z** (Zero Parallax to Zero Parallax), **F-N** (Far to Near), **F-Z** (Far to Zero Parallax), **N-F** (Near to Far), **N-Z** (Near to Zero Parallax), **Z-F** (Zero Parallax to Far), **Z-N** (Zero Parallax to Near)

Source	Question	Response (Integer or discrete choice)
Ours	How strong is fatigue of your eyes at the moment?	1 ("not at all") - 6 ("very strong")
KAB	At the moment I feel	1 ("tense") - 6 ("relaxed") 1 ("relaxed") - 6 ("queasy") 1 ("worried") - 6 ("untroubled") 1 ("calm") - 6 ("nervous") 1 ("skeptical") - 6 ("trustful") 1 ("comfortable") - 6 ("miserable")
SSQ	General discomfort Fatigue Headache Eyestrain Difficulty focusing Increased salivation Sweating Nausea Difficulty concentrating Fullness of head Blurred vision Dizzy (eyes open) Dizzy (eyes closed) Vertigo Stomach awareness Burping	"none", "slight", "moderate", "severe"

Table 1: *Summary of 23 question survey. The first question is our own. The next six are from the Kurzfragebogen zur aktuelen Beanspruchen (KAB) [Müller and Basler 1993]. The last 16 are from the Simulator Sickness Questionnaire [Kennedy et al. 1993]. The questionnaire was integrated directly in the experiment using the same display and keyboard.*
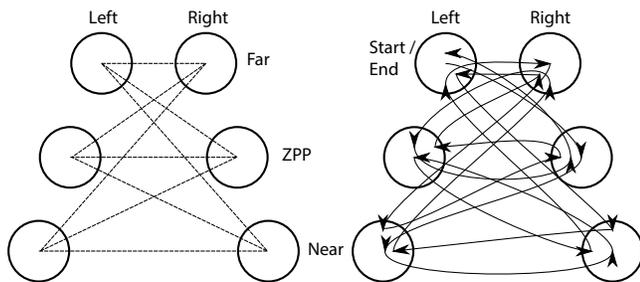
Figure 4: *On left: Possible target locations (circles) and depth change levels (lines). Assuming symmetry, the 18 lines reduce to 9. On right: An example cycle, which is balanced to contain 18 (2x9) balanced depth changes.*

Figure 5: *Six blocks total, each containing two sub-blocks of continuous depth and non-continuous depth trials (balanced order across subjects). Each sub-block contained ten randomly selected, balanced cycles of all depth changes, plus transition trials to the next cycle.*

the same illumination conditions until the experiment finished. The operator of the experiment was present at all times.

4 Results and Discussion

Data Pre-Treatment was required before analysis. To be sure that we excluded all trials where participants took a break, we decided for a conservative response time threshold of 5000 ms. Responses less than 200 ms were also excluded to eliminate accidental responses due to a "double key press". Response time histograms were analyzed to ensure meaningful data was not lost. Our filter thresholds agree with recommendations of Voss et al. [2013].

The collected data enabled us to analyze response time, accuracy and self-assessments from the questionnaires. Analysis was performed with a three factor repeated measure ANOVA, using Greenhouse-Geisser adjusted degrees of freedom when Mauchly's Test of Sphericity was violated. Post-hoc pair- wise comparisons

with Bonferroni corrections were run for multiple comparisons. All three main effects were significant.

Main Effect: Continuous Depth The main effect of Continuous Depth had a significant influence on the response time ($F(1,11) = 10.284$, $MSE = 7.76 * 10^5$, $p < .01$). Participant average response time was 1075 ms with and 1195 ms without the continuous depth plane (CDP), an average performance increase of 10%.

Main Effect: Depth Change The main effect of Depth Change also significantly influenced the response time ($F(1.96,21.58) = 11.404$, $MSE = 1.62 * 10^6$, $p < .001$). A significant interaction was observed between Continuous Depth and Depth Change ($F(2.42,26.61) = 3.822$, $MSE = 9.00 * 10^4$, $p < .05$). This implies that the effect of continuous depth cue depends on the level of depth change. A detailed discussion of the influence of continuous depth on each of the nine possible depth changes proceeds below.

4.1 Depth Change

Response time per depth change are summarized in Figure 6. To facilitate analysis, we classify the depth changes into three different types: lateral, inward, and outward changes.

Lateral Change Trials in which both stimuli are located at the same depth are labeled as lateral change. Three depth changes meet this criteria: F-F (Far-to-Far), N-N (Near-to-Near), and Z-Z (Zero Parallax-to-Zero Parallax). For the Z-Z condition, attention transitions are along the zero parallax plane. We observed an improvement of 2.9% in response time in the presence of the continuous depth plane (CDP), however the effect was not statistically significant ($p = 0.218$) when compared using a Paired Samples T-Test. We hypothesize the small improvement is due to the CDP providing an additional cue to direct attention between the two target stimuli [Egly et al. 1994].

The other two lateral changes exhibited significant improvement in the presence of the CDP. Mean improvement due to the CDP was 11.6% for N-N ($p < .01$) and 7.7% for F-F ($p < .05$) lateral changes. It appears the CDP again provides a visual cue to direct attention. However, we may observe statistical significance because the cue additionally helps the visual system maintain the decoupling of eye-vergence and accommodation during the attention transition. We hypothesize that during the attention transition, the visual system tries to return to a state of natural correspondence between eye-vergence and accommodation.

It should also be noted that the lateral distance between stimuli differed for each lateral change condition. The lateral distance was greatest for N-N and smallest for F-F. In both cases, the continuous depth plane improved visual performance to change attention, with a greater improvement occurring when the lateral distance was longer. The CDP appears to both provide a directed attention cue [Egly et al. 1994] and also to help maintain the necessary decoupling between eye-vergence and accommodation [Howard 2002].

Inward Changes Trials in which visual attention changes from a far target to a near target are labeled as inward changes. Three depth changes meet this criteria: Far-to-Zero Parallax (F-Z), Zero Parallax-to-Near (Z-N), and Far-to-Near (F-N). The F-Z condition is interesting in that visual attention is transitioned to the zero-parallax plane, where the eye-vergence and accommodation conflict is minimum. As would be expected, we do not observe a statistically significant improvement ($p = 0.103$). A statistically significant reduction in response time is observed for the other two inward changes. Z-N improved by 14.1% ($p < .05$) and F-N improved by 10.9% ($p < .05$).

Outward Changes Trials in which visual attention changes from a near target to a far target are labeled as outward changes. The remaining three depth changes meet this criteria: Near-to-Zero Parallax (N-Z), Zero Parallax-to-Far (Z-F) and Near-to-Far (N-F). The transition to zero parallax (N-Z) exhibits a similar behavior as F-Z, in which attention transitions to a location of minimum eye-vergence and accommodation conflict. The response time for N-Z was not significantly reduced by the CDP ($p = 0.125$).

The remaining two outward depth changes do show statistically significant improvement. We observe an 14.2% reduction in response time for Z-F ($p < .05$) and 14.4% reduction for N-F ($p < .01$). The CDP reduced the response time by approximately 202 ms for those two conditions. The outward depth changes also show a trend to take longer than all other depth changes.
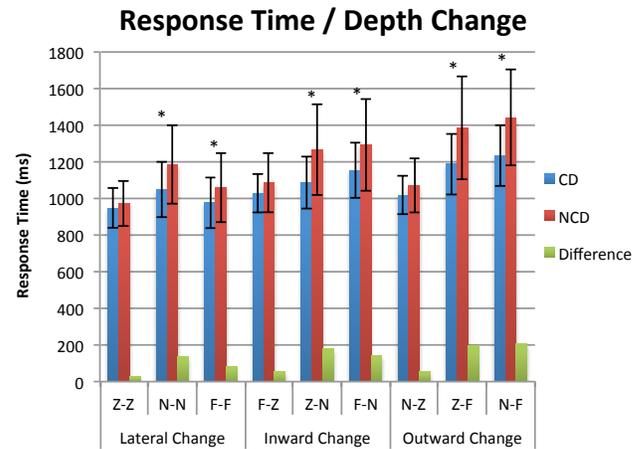
Figure 6: *Comparison of the average response time per depth change. Continuous depth improves the outward depth change most. F: Far, N: Near, Z: Zero Parallax. Asterisks represent significant difference in response time between continuous depth condition ($p < .05$). Error bars represent 95% confidence intervals.*
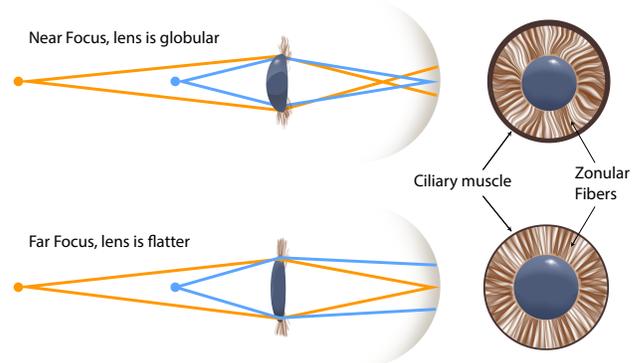
Figure 7: *Top: Ciliary muscle contracts, relaxing zonula fibers. The lens thickens to facilitate accommodation of near objects on retina. Bottom: Ciliary muscle relaxes, placing tension on zonula fibers. The lens stretches and becomes more flat to facilitate accommodation of far objects on retina.*

Observations The inclusion of a continuous depth plane (CDP) linking two targets provided a statistically significant 10% reduction in response time required to change attention. If we exclude depth changes to the zero parallax plane (Z-Z, N-Z and F-Z), we observe a 12.4% reduction in time to change visual attention. A simple change in scene composition can have a significant influence on the viewer's ability to attend to elements within the scene. Viewers are able to change their spatial attention faster when a continuous depth plane is present.

Another observation from analyzing the data is the direction dependence on transitions in depth. Attention transitions from either near or far locations to zero parallax take approximately the same time. However, the other two outward changes tend to take longer than the other two inward changes. This observation appears consistent with a simple biological model of the eye accommodation.

The lens of the eye changes shape to accommodate as represented in Figure 7. This is accomplished by contracting or relaxing the internal ciliary muscle, which adjusts tension on the zonula fibers that

radiate from the lens. Since the ciliary is an annulus muscle, contraction decreases the diameter of the muscle resulting in releasing tension and increasing convexity of the lens. When the ciliary muscle tightens, the eye accommodates to a nearer point. Relaxation of the ciliary muscle increases tension on the zonula fibers resulting in far focus. Since contraction of a muscle is always faster than relaxation, we expect to see an asymmetry in the time to change accommodation. This effect agrees with our data: changes of accommodation from far-to-near are faster than near-to-far.

Since vergence can drive accommodation [Nguyen et al. 2008; Howard 2002], we could reason about the demanding process of changing visual attention in stereoscopic images. When the visual system changes eye-vergence, a natural response causes an initial reaction to adjust accommodation to the new fixation point. However, that adjustment will result in focus deteriorating because all visual information is fixed at the zero parallax plane. The visual system then begins the counter-intuitive response of decoupling eye-vergence and accommodation.

We hypothesize that the continuous depth plane provides additional eye-vergence cues to assist the visual system in compensating for the decoupling with accommodation. In the case of our stimuli, the visual system may saccade via the continuous depth plane to the new target. Without this additional information, the visual system may invoke larger or more time consuming accommodative changes.

Our hypothesis is further reinforced by observation that all depth changes to the zero parallax plane are not significantly improved with the continuous depth plane. There is no, or minimal, vergence-accommodation conflict at the zero parallax plane. If we assume that the continuous depth plane helps compensate for the vergence-accommodation conflict, we should expect to see no improvement when the conflict does not exist. Our data confirms this hypothesis.

4.2 Measuring Fatigue

The second question posed in our study was whether fatigue observed by performance measures correlates with self-assessed questionnaire data. This requires an analysis of not only response time, but also response accuracy and the change in questionnaire data throughout the experiment.

There remains one main effect that we did not discuss in the previous analysis of the influence of continuous depth on depth change. That main effect is the *Block*. Our experiment task consisted of six blocks: each block encompassed all permutations of depth changes as well as two conditions: with and without a continuous depth plane (CDP), which are isolated in sub-blocks. Each block consisted of 380 depth discrimination trials. In total, each subject evaluated 2,280 trials. In addition to these blocks, we administered three questionnaires at three points during this task: before, at the mid-point (after Block 3), and at the end, immediately after Block 6, as summarized in Figure 5.

Main Effect: Block The main effect of the Block factor has a significant influence on response time (F(1.17,12.82) = 8.146, MSE = $2.53 * 10^6$, $p < .001$). Figure 8 presents a trend in the data for the response rate to decrease from an overall average of 1415 ms in Block 1 to minimum at around Block 5 (1010 ms) and a small increase in Block 6 (1022 ms). Subjects tend to become faster throughout the experiment. We interpret this in two ways: First, there seems to be a learning effect in the first three blocks. Second, the time to achieve fusion tends to a slight increase between Block 5 and Block 6 for some of those depth changes that often require the longest time to complete (e.g. N-F and Z-F).
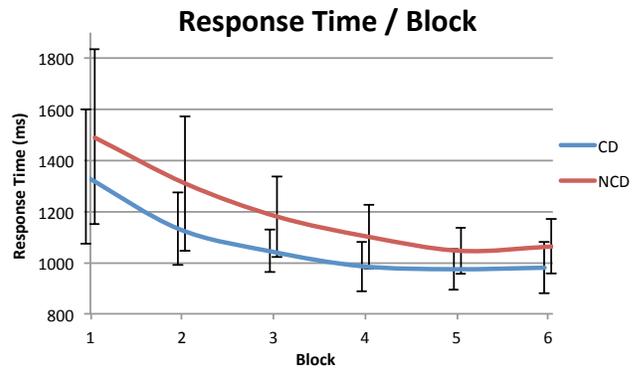
Figure 8: *Response time per block. Comparison of response time across different blocks with and without continuous depth. Error bars represent 95% confidence intervals.*
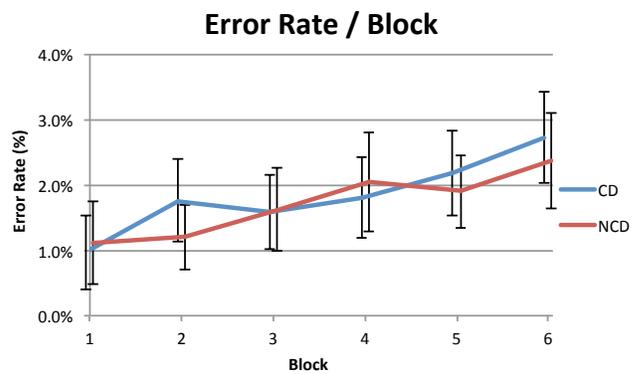
Figure 9: *Error rates per block. Error rate is low for both conditions, increasing in the last block. The error bars represent 95% confidence intervals.*

Analysis of error rate also reveals a significant main effect of Block (F(5,55) = 10.129, MSE = .001, $p < .001$). Figure 9 presents the mean error rates per continuous depth condition over each block. Closer analysis reveals that more errors are made in Block 4 than Block 1 and more errors are made in Block 6 than Blocks 1, 2, and 3 ($p < .05$).

In Block 6, we observed a trend for slower attention transitions for the difficult depth changes and an increase in the error rate. This behavior is expected when subjects experience performance fatigue. Next, we seek to relate these observations with subjective self-assessment data provided by the experiment questionnaire.

Questionnaire The questionnaire (see Table 1) is analyzed in 3 Survey Blocks. Survey Block 1 is in the beginning. The second is at the midpoint, which occurs 15-36 minutes into the experiment. The final block is at the end, which occurs after 30-70 minutes, depending on the participant's response rate throughout the experiment. Note that the substantial difference in elapsed time is indicative of the performance variance observed among test subjects.

Analysis of the general eye fatigue question as well as the 6 KAB questions resulted in a significant main effect of Survey Block (F(2,22) = 13.221, MSE = 10.111, $p < .001$). The results are as follows: assessment of eye fatigue showed a significant increase between the three questionnaire phases of the experiment ($p < .001$). We observed two other general phenomena. Some assessments in-
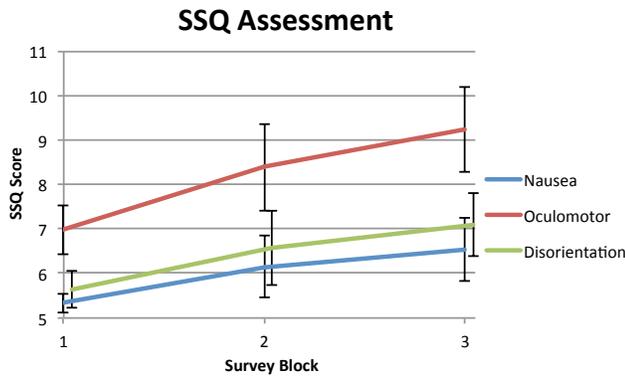
SSQ Assessment

Figure 10: *SSQ Assessment: The Oculomotor factor is most influenced during the experiment. Error bars are 95% conf. intervals.*

creased between the Survey Blocks 1 and 2, but did not significantly change between the Survey Blocks 2 and 3. Those assessments include the subject feeling more queasy ($p < .05$). The following assessments significantly varied between the first and third Survey Block, which we interpret as a more gradual increase: Subject feels more relaxed, feels more miserable, and more nervous ($p < .05$).

The remaining questions are from the Simulator Sickness Questionnaire (SSQ). We first used the three factor analysis defined by Kennedy, et al. [1993] to determine that factors pertaining to oculomotor were most influenced by our test. The results presented in Figure 10 were expected because we did not present moving images that create a conflict between vestibular and visual motion that would influence the other 2 factors: nausea and disorientation.

ANOVA was then conducted on only the 7 questions pertaining to oculomotor factor of the SSQ. The main effect of Survey Block had a significant effect on those 7 questions (F(2,22) = 14.098, MSE = 5.671, $p < .001$). From those questions, General Discomfort gradually increases from Survey Block 1 to 3 ($p < .05$). Eyestrain and Difficulty Concentrating both increased between Survey Blocks 1 and 2 ($p < .052$), but did not change significantly between Survey Blocks 2 and 3. It is likely that these symptoms are perceived as initially worsening before plateauing at a general level of discomfort as stereoscopic viewing continues. Fatigue and Blurry Vision had a tendency to become stronger, but they were not significant. Headache did not increase during the experiment.

The survey data indicates that symptoms such as eye fatigue and general discomfort tend to increase during the stereoscopic viewing activity. Other symptoms may appear when beginning the task, but not worsen through the experiment duration. Since we observe a general improvement in response time during the experiment duration with only a small increase in error rate, we conclude that subjective, self assessment of fatigue emerges before a decline in performance fatigue. We do observe deviation in performance effects in Block 6. However, we need more data or more blocks beyond Block 6 to draw conclusions. This is discussed further in Section 5.

Additional Subject Pool Observations We used a screening procedure to verify that all subjects had normal binocular spatial vision. However, among our sample set (psychology and computer science graduate students) we were surprised by an extremely wide variance among the subjects. Some subjects require up to three times longer than others to achieve fusion. Approximately 30% of our potential subjects were unable to achieve fusion for targets with absolute disparities that were within about 10% of parallel viewing. This variance implies that if stereoscopic 3D is to be successful,

a conservative approach should be taken in order to establish safe boundaries on the dynamic range of depth within a scene.

5   Limitations and Future Work

Conclusions drawn from our research are subject to several limitations. First, we evaluate only one viewing scenario, in which accommodation is always at the same distance to the image screen and three stereoscopic disparities are presented. It would be beneficial to explore other viewing distances and disparities to produce a more general model that is applicable to many common viewing scenarios. Second, we evaluate only one form of continuity between important scene elements. A future direction would be to vary the size and distance of the continuous element. Similarly, there is an open question about the strength of disparity features provided by the continuous element. Our checkerboard pattern provides several strong cues to depth [Cutting and Vishton 1995].

In terms of data analysis, our sample size of twelve subjects is small. However, we compensate for this by observing a large number of trials per subject (2280 trials). Given our current experiment design, it is difficult to draw conclusions about fatigue. The increase in error rate at the end of the experiment could be a symptom of the speed-accuracy trade-off [Wickelgren 1977], however, it could also be a sign that subjects are getting bored during the task. Regarding the questionnaire, responses to several questions show a decrease in comfort and increase in factors related to fatigue. It may be possible that for some questions related to fatigue, for example, subjects may expect fatigue to increase and increment their response score with each instance of the survey block. It would be good to explore other strategies of presenting the questionnaire.

Finally, in terms of response time per block, we observed an improvement (reduction) until Block 6, in which several of the difficult depth change conditions began to take slightly longer. It is not clear if we are about to observe the effects of performance fatigue or that the subjects simply got bored or distracted in doing the task. Our data indicates that a computational model of stereoscopic attention transitions should take into account the sign and magnitude of the disparity change as well as connectivity and perhaps even viewing duration. These are all topics for continued future work.

6   Conclusion

We have shown that changes in scene composition have a significant influence on the viewer's ability to change visual attention among spatially distinct scene elements. Continuous depth improves visual attention transitions to locations where a vergence-accommodation conflict is present, but not to locations where there is no conflict (e.g. the zero parallax plane). We also observe that self-assessment of eye fatigue and general discomfort increase before a decrease in visual performance is observed. For 3D cinema and interactive media to remain as a viable entertainment genre, additional studies of this type may ensure that content will reach the widest possible audience.

The important message is that scene composition, framing and montage can significantly influence visual performance in terms of time to change visual attention. Visually important, or salient, scene elements can be viewed more quickly when connected by visual information that continuously varies in stereoscopic depth. It may not be practical or beneficial to ensure depth continuity in all stereoscopic content. However, our contributions are highly applicable when rapid eye movements are required to achieve very short video scene cuts or rapid visual targeting for stereoscopic game play.

References

AKELEY, K., WATT, S. J., GIRSHICK, A. R., AND BANKS, M. S. 2004. A stereo display prototype with multiple focal distances. *ACM Trans. Graph. 23*, 3 (Aug.), 804–813.

CHAPIRO, A., HEINZLE, S., AYDIN, T., POULAKOS, S., ZWICKER, M., SMOLIC, A., AND GROSS, M. 2014. Optimizing stereo-to-multiview conversion for autostereoscopic displays. *Computer Graphics Forum, proc. of Eurographics 2014 33*, 2.

CUTTING, J. E., AND VISHTON, P. M. 1995. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Handbook of perception and cognition*, vol. 5. Academic Press, ch. Perception, 69–117.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2011. A perceptual model for disparity. *ACM Trans. Graph. 30*, 4 (July), 96:1–96:10.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., SEIDEL, H.-P., AND MATUSIK, W. 2012. A luminance-contrast-aware disparity model and applications. *ACM Trans. Graph. 31*, 6 (Nov.), 184:1–184:10.

DU, S.-P., MASIA, B., HU, S.-M., AND GUTIERREZ, D. 2013. A metric of visual comfort for stereoscopic motion. *ACM Trans. Graph. 32*, 6 (Nov.), 222:1–222:9.

EGLY, R., DRIVER, J., AND RAFAL, R. D. 1994. Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology 123*, 2, 161–177.

EMOTO, M., NIIDA, T., AND OKANO, F. 2005. Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *Journal of Display Technology 1*, 2, 328–340.

HOFFMAN, D. M., GIRSHICK, A. R., AKELEY, K., AND BANKS, M. S. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J. Vis. 8*, 3 (3), 1–30.

HOWARD, I. P. 2002. *Seeing in Depth – Basic Mechanisms*, vol. 1. I Porteous, Thornhill, Ontario, Canada.

IJSSELSTEIJN, W. A., SEUNTIÃNS, P. J. H., AND MEESTERS, L. M. J. 2006. Human factors of 3d displays. In *3D Videocommunication: Algorithms, Concepts and Real-Time Systems in Human Centred Communication*. John Wiley & Sons, Ltd., ch. 12, 219–233.

JULESZ, B. 1960. Binocular depth perception of computer-generated patterns. *Bell System Tech. 39*, 5, 1125–1161.

KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIENTHAL, M. G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology 3*, 203–220.

KIM, C., HORNUNG, A., HEINZLE, S., MATUSIK, W., AND GROSS, M. 2011. Multi-perspective stereoscopy from light fields. *ACM Trans. Graph. 30*, 6 (December), 190:1–190:10.

KOOI, F. L., AND TOET, A. 2004. Visual comfort of binocular and 3d displays. In *Displays*, vol. 25, 99–108.

KOPPAL, S. J., ZITNICK, C. L., COHEN, M. F., KANG, S. B., RESSLER, B., AND COLBURN, A. 2011. A viewer-centric editor for 3d movies. *IEEE Computer Graphics and Applications 31*, 1, 20–35.

KUZE, J., AND UKAI, K. 2008. Subjective evaluation of visual fatigue caused by motion images. *Displays 29*, 159–166.

LAMBOOIJ, M., IJSSELSTEIJN, W., FORTUIN, M., AND HEYNDERICKX, I. 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Tech. 53*, 3, 1–14.

LANG, M., HORNUNG, A., WANG, O., POULAKOS, S., SMOLIC, A., AND GROSS, M. 2010. Nonlinear disparity mapping for stereoscopic 3d. *ACM Trans. Graph. 29*, 4 (July), 75:1–75:10.

MASCELLI, J. V. 1965. *The Five C's of Cinematography*. Silman-James Press, Beverly Hills, CA.

MASIA, B., WETZSTEIN, G., ALIAGA, C., RASKAR, R., AND GUTIERREZ, D. 2013. Display adaptive 3d content remapping. *Computers & Graphics 37*, 8, 983 – 996.

MÜLLER, B., AND BASLER, H. 1993. *Kurzfragebogen zur aktuellen Beanspruchung (KAB)*. Beltz-Test GmbH.

NGUYEN, D., VEDAMURTHY, I., AND SCHOR, C. M. 2008. Cross-coupling between accommodation and convergence is optimized for a broad range of directions and distances of gaze. *Journal of Vision 48*, 893–903.

PARKER, A. J. 2007. Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience 8*, 379–391.

PATTERSON, R. 2007. Human factors of 3-d displays. *Journal of the Society for Information Display 15*, 861–871.

SHIBATA, T., KIM, J., HOFFMAN, D. M., AND BANKS, M. S. 2011. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision 11*, 8, 1–29.

SIEGEL, M. W., AND NAGATA, S. 2000. Just enough reality: Comfortable 3-d viewing via microstereopsis. *IEEE Transactions on Circuits and Systems for Video Technology 10*, 3.

UKAI, K., AND HOWARTH, P. A. 2008. Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays 29*, 2, 106 – 116.

VOSS, A., NAGLER, M., AND LERCHE, V. 2013. Diffusion models in experimental psychology. *Experimental Psychology 60*, 6, 385–402.

WICKELGREN, W. A. 1977. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica 41*, 67–85.

YANO, S., IDE, S., MITSUHASHI, T., AND THWAITES, H. 2002. A study of visual fatigue and visual comfort for 3d hdtv/hdtv images. *Displays 23*, 4, 191 – 201.

YANO, S., EMOTO, M., AND MITSUHASHI, M. 2004. Two factors in visual fatigue caused by stereoscopic hdtv images. *Displays 25*, 141–150.