Richard H. Y. So*
W. S. Wong
Romeo Yip
Department of Industrial Engineering and Logistics Management
The Hong Kong University of Science and Technology
Clearwater Bay, Kowloon
Hong Kong SAR

Andrew K. C. Lam
Patrick Ting
School of Optometry
The Hong Kong Polytechnic University
Kowloon, Hung Hom
Hong Kong SAR

Benefits of Matching Accommodative Demands to Vergence Demands in a Binocular Head-Mounted Display: A Study on Stereo Fusion Times

Abstract

Current head-mounted displays (HMDs) provide only a fixed lens focus. Viewers have to decouple their accommodation and vergence responses when viewing stereoscopic images presented on an HMD. This study investigates the time taken to fuse a pair of stereoscopic images displayed on an HMD when the accommodative demand is matched to the vergence demand. Four testing conditions exhausting the factorial combinations of accommodative demands (2.5 D and 0.5 D) and vergence demands (2.5 MA and 0.5 MA) were investigated. The results indicate that viewers take a significantly shorter amount of time to fuse a pair of stereoscopic images (i.e., fusion time) when the accommodative demand and the stereoscopic depth cues match. Further analysis suggests that an unnatural demand for the eyes to verge toward stereoscopic images whose stereo depth is farther than the accommodative demand is associated with significantly longer fusion time. This study evaluates the potential benefits of using a dynamically adjustable lens focus in future designs of HMDs.

1 Introduction

1.1 Background

Head-mounted displays (HMDs) can present stereoscopic images and are used in virtual reality (VR) training applications (e.g., medical surgery; Ilie et al., 2004) and personal entertainment (e.g., movies and computer games). In a typical HMD, images are projected at a virtual distance in front of the viewer as determined by the lens focus (Cakmakci & Rolland, 2006). This virtual distance sets the accommodative demand for the viewer. Typical commercially available HMDs have a fixed lens focus and, therefore, a fixed accommodative demand (Eichenlaub, 2005). For example, HMDs produced by Integrated Microdisplays Limited[1] have an accommodative demand of 2 m and the V8 HMD from Virtual Research Systems Inc.[2] has an accommodative demand of 0.9 m.

*Correspondence to rhyso@ust.hk.
[1]www.hkimd.com
[2]www.virtualresearch.com

Presence, *Vol. 20, No. 6, December 2011*, 545–558

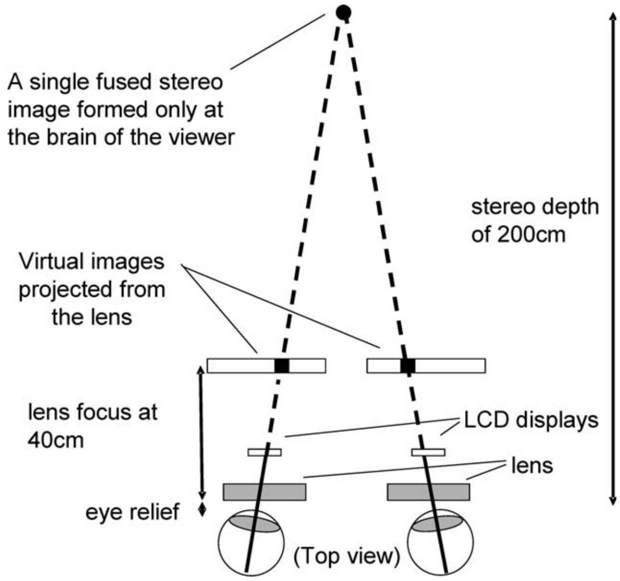

Figure 1. *A ray tracing diagram illustrating a situation where a viewer is viewing a pair of stereo images at 0.5 MA (i.e., a stereo depth of 200 cm) presented on a binocular LCD display with an accommodative demand of 2.5 D (i.e., a lens focus of 40 cm).*

1.1.1 De-coupling of Accommodation and Vergence when Viewing HMDs. When viewing a stereoscopic image presented on a binocular HMD and the stereo depth of the image is less than infinity, both eyes will turn inward (converge; see Figure 1). As the stereo depth decreases, the amount of convergence will increase.

Biologically, convergence is associated with accommodation. This is the subject of many studies concerning near triad (e.g., Ciuffreda & Kenyon, 1983; Krishnan & Stark, 1977; Stark, 1983). As humans watch an object move toward them, both convergence and accommodation increase so that sharp images of the object are formed successively on the retinas and their corresponding positions enable the brain to see a single object with stereoscopic view. However, when viewing stereoscopic images with varying stereo depths presented on an HMD with a fixed accommodative demand, the coupling between the accommodation and vergence of the eyes can no longer be maintained. Figure 1 illustrates a situation in which a viewer is watching a pair of stereo images when the stereo depth is farther than the accommodative demand. To avoid double images, the eyes would verge toward the target by moving either toward each other (converging) or away from each other (diverging). It has been reported that a convergence response is twice as fast as a corresponding divergence response (Hung, Zhu, & Ciuffreda, 1997). This suggests that fusion requiring additional divergence eye movements might take longer.

Rushton and Riddell (1999) hypothesized that the decoupling of accommodation and vergence when using HMDs puts an unnatural demand on the visual system of the viewers and may cause eye stress in the long run. Similar problems have been observed and reported in the use of stereoscopic cameras (Lipton, 1982). Lipton reported that it was necessary to avoid divergence in the viewers' eyes as it would cause eye fatigue. Unlike stereoscopic movies in which image depths are predetermined, 3D computer games expose viewers to image depths that change in real time, depending on user interactions. In other words, while stereo images with serious mismatch between demands in accommodation and vergence were skillfully avoided in filmmaking, they are inevitable in stereoscopic computer games.

1.1.2 Potential Problems Associated with the Decoupling of Vergence and Accommodation when Using HMDs. The decoupling of vergence and accommodation with HMDs has been the subject of many discussion papers (Patterson, Winterbottom, & Pierce, 2006; Roscoe, 1987, 1988; Wann & Mon-Williams, 1997; Wann, Rushton, & Mon-Williams, 1997; Stanney, Mourant, & Kennedy, 1998). However, these discussion papers did not report any empirical data. Shibata et al. (2004) did measure the accommodation and vergence responses in binocular display users. However, the effects of matching the accommodative demand to stereo depth on performance measures such as stereoscopic fusion times were not reported.

Theoretically, when a viewer is presented with a pair of virtual stereo images located at a depth inappropriate to the accommodative demand, two things could happen. The viewer could verge to fuse the images, in which case the images will appear out of focus (Bruce, Atchison, & Bhoola, 1995). Or, the viewer could accommodate to see a clear image and verge at a depth appropriate to the accommodation but inappropriate to the virtual images. This will result in seeing clear but double images (Ciuffreda & Kenyon, 1983). The former is related to convergence-accommodation (CA) responses and the latter is related to accommodative-convergence (AC) responses. More discussion on this can be found in Sections 2 and 3 on hypotheses and results.

The study reported by Hoffman, Girshick, Akeley, and Banks (2008) is the one that is closest to ours. They conducted an experiment to examine the effects of matching accommodative demand to stereo depth on the time taken to determine the orientation of a cyclopean stimulus in a random-dot stereogram. Their results indicated that the amount of time required to identify a stereoscopic stimulus was reduced when the lens focus matched the stereo depth. Unfortunately, only three participants were tested, and no statistical result was reported.

This study aims to explore the effects of the mismatch between accommodation and vergence demands on the time taken to fuse a pair of stereoscopic images (i.e., the fusion time). The objective is to collect statistically-verified evidence to justify the application of automated lens adjustment to HMDs. Partial and preliminary results were presented at the First International Symposium on Visually Induced Motion Sickness, Fatigue, and Photosensitive Epileptic Seizures (Wong et al., 2007). The fusion time was chosen as the dependent variable in this study because it was relatively easy to measure. Also, it was a ratio-scaled measurement that could be used to quantify how quickly a viewer was able to see a clear stereoscopic image. If the fusion time is too long, viewers may suffer from eye strain and headaches (e.g., Bando, Iijima, & Yano, 2012; Karasuda & McMains, 2005; Lambooij, IJsselsteijn, Fortuin, & Heynderickx, 2009; Pickwell, 1984; Ware, Gobrecht, & Paton, 1998). The inability to form a single clear stereoscopic image had also been associated with inappropriate depth perception (e.g., Pickwell; Siderov & Harwerth, 1993). With the rapid development of commercial stereoscopic 3D televisions (e.g., Samsung, 2010; Toshiba, 2011), the current study can provide data to indicate how mismatches between the accommodative demand and the vergence demand could affect the fusion of stereoscopic images.

1.1.3 Engineering Solutions to Reduce the Decoupling Between Accommodative and Vergence Demands when Using HMDs. The most direct solution to the mismatch between accommodation and vergence is to eliminate the mismatch. In other words, HMDs with a computer-controlled, dynamically adjustable accommodative demand should be developed. Rolland, Krueger, and Goon (2000) proposed the use of a multi-focal plane liquid crystal display (LCD) to adjust the accommodative demand of an HMD in real time. Hendriks, Kuiper, Van As, Renders, and Tukket (2005) proposed the use of a liquid-based lens to develop an HMD with adjustable accommodative demand. Akeley, Watt, Girshick, and Banks (2004) and Shevlin (2005) reported a prototype fixed-viewpoint stereoscopic volumetric display with multiple focuses. Using this prototype display, the target identification time was reduced when the accommodative demand was matched to the stereo depth demands. However, only three viewers were tested (Hoffman et al., 2008). In summary, engineering solutions for HMDs with adjustable accommodative demands exist but their benefits have not been statistically verified. This study intends to quantify the benefits of dynamically adjusting the accommodative demand of an HMD. The resulting data can help HMD manufacturers decide whether they should invest in developing techniques to adjust accommodative demands.

1.2 Objective

The objective of the present work is to investigate the effects of matching the accommodative demand to the vergence demand on the time taken to fuse a pair of stereoscopic images.

2 Methods

2.1 Variables and Conditions

The main dependent variable was the time taken to fuse a pair of stereoscopic images (i.e., the fusion time). Visual parameters related to vergence eye movements were also measured. The two independent variables were the accommodative demand and the vergence demand.

Table 1. *Fusion Times (25th, 50th, and 75th percentiles) of the Two Matched and the Two Unmatched Conditions Between Accommodation and Vergence Demands**

		Accommodative demand (focus) (Diopters, D)		
		0.5D		2.5D
Vergence (Meter-Angle, MA)	0.5MA	matched 1s - 1.8s - 3s (25th - 50th - 75th %-tiles)	H1d	unmatched 1.4s - 2.8s - 5s (25th - 50th - 75th %-tiles)
		H1b		H1a
	2.5MA	unmatched 1.3s - 2s - 4s (25th - 50th - 75th %-tiles)	H1b	matched 1.2s - 1.6s - 3s (25th - 50th - 75th %-tiles)

*NOTE. The four matched versus unmatched demands hypotheses H1a, H1b, H1c, and H1d are also illustrated.

Two accommodative demands were used: 2.5 diopters (D) and 0.5 D which correspond to setting the lens focus to 40 cm and 200 cm, respectively. Two vergence demands were used: 2.5 meter-angles (MA) and 0.5 MA which correspond to stereo images with stereo depths of 40 cm and 200 cm, respectively. These were paired to create four different conditions. Two of them were matched conditions: 2.5 D and 2.5 MA and 0.5 D and 0.5 MA. The other two were mismatched conditions: 2.5 D and 0.5 MA and 0.5 D and 2.5 MA (see Table 1).

The two accommodative demands of 0.5 D and 2.5 D were chosen to test the effects of matching with near and far accommodation. The accommodative demand of 0.5 D is adopted by a micro-display manufacturer (iMD Micro-display, Ltd.) as the fixed accommodative demand for its display products.

2.2 Hypotheses

Since the matching effects were the two-way interaction effects between the accommodative demand and stereo depth, the effects of matching were inevitably confounded by either the effects of accommodative demand or the effects of stereo depth (i.e., vergence demands). The hypotheses were therefore formulated in such a way that either of the two types of confounding effects were controlled for. It was hypothesized that the fusion time would be significantly shorter when the accommodative demands and the vergence demands were equal. Specifically, the matched versus unmatched demands hypotheses (H1a to H1d) would apply to each of the four settings: (a) a fixed accommodative demand of 2.5 D (H1a); (b) a fixed accommodative demand of 0.5 D (H1b); (c) a fixed vergence demand of about 2.5 MA (H1c); and (d) a fixed vergence demand of about 0.5 MA (H1d). These four hypotheses are consistent with past discussion papers (Roscoe, 1987, 1988; Wann & Mon-Williams, 1997, Wann et al., 1997). The relationships between the four hypotheses (H1a to H1d) and the four experimental conditions are illustrated in Table 1.

When viewing stereoscopic images presented on an HMD with a fixed accommodative demand, the eyes should verge toward the images according to the accommodative vergence response (e.g., Semmlow & Venkiteswaran, 1976; Ciuffreda & Kenyon, 1983). Under an unmatched condition, double images may be formed due to different demands for accommodation and vergence. In these cases, the eyes will need to utilize their fusional vergence ability to change their vergence. These changes in vergence movements are made at certain costs: (i) the visual system will need to process the fusional cues and exercise fusional vergence. This would increase the time taken to fuse a pair of stereoscopic images. It was not even clear if every individual viewer has the ability to perform these actions. Another cost is that (ii) the repetition of this process might generate fatigue in the visual system.

Since the fusional vergence ability could vary among individuals, two further hypotheses were tested. Viewers with larger fusional reserves (the maximum amount of vergence movement needed to maintain fusion) were hypothesized to form a single stereoscopic image more quickly ($H2_{\text{fusional reserves}}$) because people with greater fusional reserves were expected to have a greater ability to verge toward the presented pair of stereoscopic images. On the other hand, viewers with larger lateral phoria (i.e., the amount of latent squint) were hypothesized to form a single stereoscopic image more slowly ($H2_{\text{lateral phoria}}$) because people with larger lateral phoria were expected to need most of their vergence capacity to maintain binocular fixation. Consequently, little vergence capacity is left to facilitate the additional vergence movements required to form a single stereoscopic image.

Furthermore, as none of the studies has reported the effects of matching accommodative demand to stereo depth, it was not clear whether viewers can learn to fuse the stereo images more quickly with repeated runs. It was hypothesized that with repeated runs, viewers would learn to fuse the stereo images more quickly (the learning effects hypothesis: H3).

When viewers were presented with mismatched accommodative demand and stereoscopic depth cues, they might have used AC responses to verge toward the images which resulted in blurriness. This concern is revisited in Section 3.

2.3 Design of Experiment

As explained in the previous section, the experiment was conducted under four conditions exhausting the different combinations of two accommodative demands and two vergence demands. The experiment used a within-subjects design and had two repeated sessions conducted at least 3 days apart. Three repeated blocks of the four conditions were tested during each session. Within each condition, participants had to repeat the task of forming a single stereoscopic image five times. The times taken to fuse the pair of stereoscopic images in the five dependent repeated tasks were averaged to form a better mean estimation of the fusion times. The order in which the four conditions were presented was randomized and participants were allowed to rest for 5 min in complete darkness between each condition. Each participant performed 120 tasks ($120 = 2$ sessions $\times$ 3 blocks $\times$ 4 conditions $\times$ 5 dependent repeats) in all, giving 24 averaged data points. Ten healthy Chinese volunteers took part in the experiment (six males and four females). Their mean age was 23 years and the experiment was approved by the Human Subject Experiment Committee of the Hong Kong University of Science and Technology. Each participant received HK$50 per hour (about US$7 per hour) as compensation for their time and travel costs incurred.

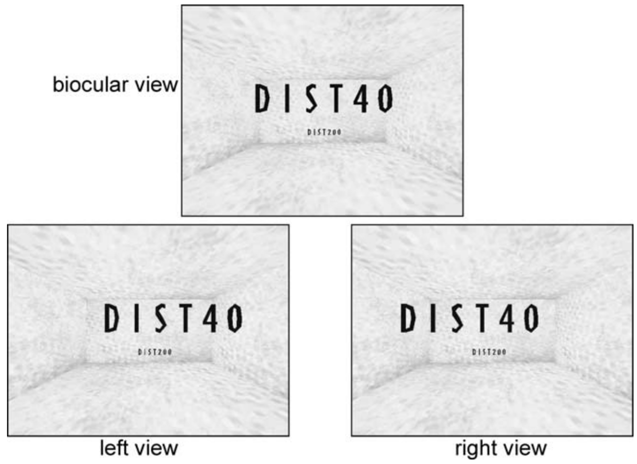

Figure 2. *The upper image is a biocular view of the visual stimuli (a virtual room with two 3D text objects): DIST40 is a stereo object placed at 2.5 MA (i.e., a virtual depth of 40 cm) and DIST200 is a stereo object placed at 0.5 MA (i.e., a depth of 200 cm). The bottom right image is the image for the right eye and the bottom left image is the image for the left eye.*

2.3.1 Stimuli and Apparatus. A virtual environment (VE) was constructed using WorldToolKit (WTK) running on a Silicon Graphics Infinite-Reality II Workstation. The VE consisted of a room (400 cm in width $\times$ 800 cm in height $\times$ 2000 cm in depth) with two virtual 3D text objects (DIST40 and DIST200) placed at virtual depths of 40 cm and 200 cm (i.e., 2.5 MA and 0.5 MA), respectively (see Figure 2). The VE was programmed to be viewed in either the biocular mode, in which the left and right views were identical, or the binocular mode with the inter-pupillary distance (IPD) customized to fit each participant. This VE was viewed through an HMD custom-built by Integrated Microdisplays Limited, Hong Kong. Its resolution was 688 (horizontal) $\times$ 480 (vertical). The diagonal field-of-view (FOV) was 30° for each eye. A processor-controlled accommodative demand adjustment system was custom-built for this study (see Figure 3). This system used stepper motors to adjust the lens arrangement to achieve different accommodative demands (Melzer & Moffit, 1997). It should be pointed out that, while the location of the microdisplay was adjusted to the required location of the virtual images using the processor-controlled unit, the optics

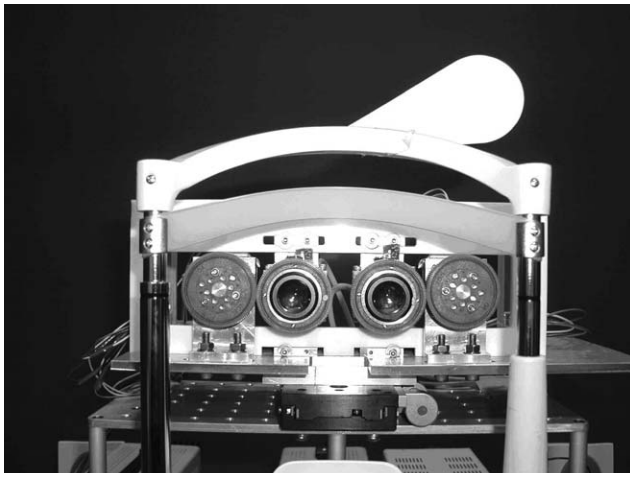

Figure 3. *The custom-built binocular display system used in this study.*

Figure 4. *A double image. The accommodation and vergence responses of the participants were assumed to be at 0.5 D and 0.5 MA. The double image was formed by two DIST40 stereo images both with a stereo depth of 40 cm (i.e., 2.5 MA).*

themselves were designed for the accommodative setting of 0.5 D. Consequently, while the location of the virtual image was being refocused to 2.5 D, the quality of the image might not be equivalent to that for 0.5 D. To ensure that the optics had a high enough spatial resolution to support the experimental tasks, visual acuity tests were conducted using Landolt Cs. It was found that viewers were able to resolve spatial resolutions of at least 2 pixels (0.087° horizontal, and 0.125° vertical) under both 0.5 D and 2.5 D. Since the smallest character size (DIST200 in Figure 2) used in the study was 20 pixels (horizontal) by 10 pixels (vertical), the spatial resolution of the HMD under both accommodation settings was certainly high enough to allow participants to resolve clear images during the experiment.

2.3.2 Procedure and the Expected Status of Accommodation and Vergence During Various Stages of the Experiment. All participants attained at least 20/20 visual acuity using both eyes in the near (36 cm) and distant (6.1 m) viewing conditions (the OPTEC 2000 vision tester, Stereo Optical Co., Inc., USA). None of the participants had color deficiency and all of them attained stereopsis at 40″ of the arc or less. Besides the basic measurements, six visual parameters of the participants were also measured by a registered optometrist. The measured parameters were negative fusional reserve at distance (6 m), negative fusional reserve at near (40 cm), positive fusional reserve at distance (6 m), positive fusional reserve at near (40 cm), lateral phoria at distance (3 m), and lateral phoria at near (1/3 m, i.e., 33 cm).

Before the experiment, participants were briefed on the definitions of biocular images, a single stereoscopic image, and double images. A practice session was given during which the participants learned to indicate the presence of double images by pressing a handheld response button. The detailed procedures of the five dependent repeated runs of each condition, as well as the expected status of accommodation and vergence, are as follows.

Step 1: While the participant was resting in complete darkness, the accommodative demand was preset to an appropriate level. Biocular images of the VE were then presented and the participants were verbally instructed to fixate on either DIST40 or DIST200, depending on the selected condition. Participants were asked whether they could see a sharp single image. They would have accommodated to the predetermined accommodative demand (either 2.5 D or 0.5 D) and their eyes would have verged parallel to each other the moment they saw a sharp single image.

Step 2: After Step 1, a delay period with a randomized duration from 13 s to 30 s was applied. Then, binocular stereoscopic images were presented and participants were instructed to form a single stereoscopic image as fast as they could. Initially, participants would have seen double images (see

Figure 4). They were asked to indicate this by pressing and holding the response button. When a single binocular image was eventually formed, they were asked to release the response button. The time difference between when the binocular images were presented and when the participants indicated they had formed a single stereoscopic image was defined as the fusion time.

Step 3: Following the participants' indication that they had formed a single stereoscopic image, the binocular images were reset back to the biocular images after another delay period of a randomized duration (between 13 s and 30 s). The participants would have seen double images instantly, and they would have had to turn their eyes back to parallel sight to form a single biocular image. Once a single biocular image was formed as indicated by the release of the response button, Steps 2 and 3 were repeated four more times and the participants were then asked to rest for 5 min in complete darkness before performing Step 1 again under a different condition.

3 Results and Discussion

The fusion time data did not follow a normal distribution ($W = 0.6849$, $p < .0001$, Shapiro-Wilk test) even after a Box-Cox transformation ($\lambda = -0.5$) ($W = 0.9542$, $p < .0001$, Shapiro-Wilk test). Consequently, ANOVA analyses were supplemented with nonparametric tests. The median fusion times and the inter-quartile ranges are shown in Table 1. An overall ANOVA was conducted to test for the main effects of accommodative demand and vergence demand. The effects of matched versus unmatched demands appeared to be significant as indicated by the significant levels of the two-way interactions between the effects of accommodative demands and vergence demands, $F(1, 219) = 12.23$, $p < .001$ ANOVA. The pooled difference between the data for the matched condition and that for the unmatched condition was further found to be significantly different in a Wilcoxon signed-rank test ($p < .001$). Further investigation indicated that the effects of matched versus unmatched accommodative and vergence demands were significant in two of the conditions but not in the other two conditions. In the following paragraphs, the significant effects and the nonsignificant effects of matched versus unmatched demands are discussed under separate headings.

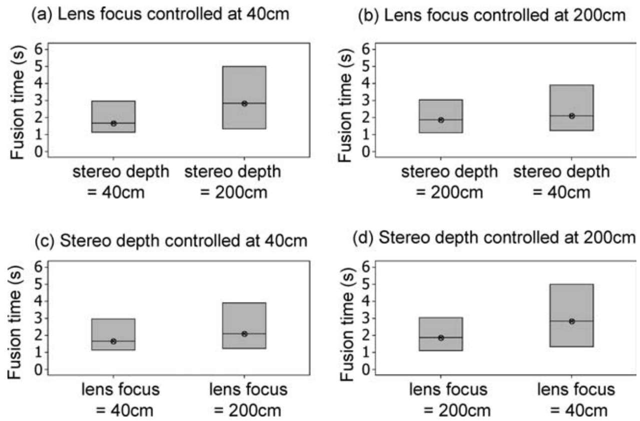

Figure 5. *Pair-wise comparisons of median fusion times between matched and mismatched accommodative demand and convergence demand according to different controlled conditions: (a) an accommodative demand of 2.5 D; (b) an accommodative demand of 0.5 D; (c) a vergence demand of 2.5 MA; and (d) a vergence demand of 0.5 MA. The interquartile ranges are also shown (NB: the experiment only had four conditions and each condition was repeated once in the figure for pair-wise comparisons).*

3.0.1 Significant Effects of Matched Versus Unmatched Accommodative and Vergence Demands and Their Interactions with Gender and Session (Testing of Hypotheses H1a, H1d, and H3).

Figure 5 shows the median and the inter-quartile ranges of the fusion time data collected from the pairs of conditions corresponding to each of the four matched versus unmatched demands hypotheses (H1a, H1b, H1c, and H1d). Please note that the study only had four conditions (see Table 1) and the data from each condition were illustrated twice in Figure 5.

It can be observed that in all four conditions, the fusion times in the matched cases were consistently shorter than those in the respective unmatched conditions. The results of ANOVA conducted to test the effects of matched demands when accommodative demand was fixed at 2.5 D is shown in Table 2. The

Table 2. *ANOVA Results on the Stereo Image Fusion Times of Participants in Tests Investigating the Effects of Matching the Accommodative and Vergence Demands of Repeated Block, Repeated Session, and Gender, and Their Two-Way Interactions (Data Taken from the Two Conditions of 2.5 MA and 2.5 D, Matched; and 2.5 MA and 0.5 D, Unmatched)*

Source	DOF	ANOVA SS	Mean square	*F* value	Pr > *F*
Matched	1	0.564	0.564	12.91	0.0005
Gender	1	0.259	0.259	5.93	0.0166
Repeated session	1	0.317	0.317	7.26	0.0082
Repeated block	2	0.014	0.007	0.16	0.8494
Matched × session	1	0.002	0.002	0.05	0.8308
Matched × gender	1	0.018	0.018	0.41	0.5254
Matched × block	2	0.034	0.017	0.39	0.6754
Gender × session	1	0.010	0.010	0.22	0.6392
Gender × block	2	0.024	0.012	0.28	0.7600
Block × session	2	0.008	0.004	0.09	0.9104
Model	14	1.251	0.089	2.04	0.0210
Error	105	4.588	0.044		
Corrected total	119	5.839			

main effects of matched demands, gender, and session were all significant but none of the two-way interactions was significant (see Table 2). The significant reductions in fusion times due to matched demands support hypothesis H1a. Similar results were found when the vergence demand was fixed at 0.5 MA (hypothesis H1d: effects of matched demands: $F(1, 105) = 11.02$, $p < .001$; effects of gender: $F(1, 105) = 8.16$, $p < .01$; effects of session: $F(1, 105) = 9.97$, $p < .01$. Again, none of the two-way interactions was significant. The lack of interaction between the effects of matched demands and the effects of gender or session indicates that H1a and H1d are supported by data collected from each session and from either gender. The fusion times collected in the second session were significantly shorter than those collected in the first session ($p < .05$, post-hoc SNK). This supports the learning effects hypothesis (H3).

Since the data were not normally distributed, the results of ANOVAs were verified against nonparametric tests. Wilcoxon signed-rank tests were conducted to test hypotheses H1a and H1d while controlling for the effects of gender and session. Paired *t*-tests were also conducted. The results indicate that matching the accommodative demand to the vergence demand significantly reduced the fusion times collected from either gender and in each session ($p < .01$). This supports hypotheses H1a and H1d and confirms the results of ANOVAs. Results of the paired *t*-tests and the Wilcoxon signed-rank tests were consistent. The confirmation of H1a indicates that when both eyes are parallel and have accommodated to 2.5 D, it takes a significantly shorter time to verge to a pair of stereo images at 2.5 MA than to a pair of stereo images at 0.5 MA. This is interesting because the angular displacement required by the parallel eyes to verge to images with a stereo depth of 2.5 MA was four times that required to verge to images with a stereo depth of 0.5 MA, and yet the median fusion time required was 57% less (median fusion times of images with stereo depths at 2.5 MA and 0.5 MA were 1.6 s and 2.8 s, respectively). In other words, the data suggest that when the vergence demand is appropriate to the accommodative demand, a significantly more efficient vergence response could be achieved. This finding is consistent with that of Semmlow and Venkiteswaran (1976) who reported that a step change in accommodative demand caused accommodative vergence eye movement that shifted the gaze angle away from the target before it was moved back to the target.

The confirmation of the matched versus unmatched demands hypothesis, H1d, suggests that viewers can reduce the time taken to verge to a pair of stereo images at 0.5 MA by allowing their eyes to accommodate to 0.5 D instead of 2.5 D. This demonstrates the benefits of accommodative demand adjustment.

In this study, male participants took significantly longer to fuse the images than did female participants. However, only four female participants were tested and hence no conclusion should be drawn on the effects of gender. Also, the average age of the male participants was 1.2 years older than female participants (male: 23.5 years old; female: 22.3 years old). In this study, the important finding was that the effects of matched demands were significant regardless of the gender of the participants. Inspection of Figure 5 indicates that the interquartile ranges of fusion times with mismatched accommodation-vergence demands (2.5 MA and 0.5 D or 0.5 MA and 2.5 D) are larger than those with matched demands (2.5 MA and 2.5 D or 0.5 MA and 0.5 D) and the ranges are skewed toward longer fusion times. This suggests that some viewers were more affected by the mismatched demands than others. Further analyses of individual effects can be found toward the end of this section.

In Step 1 during the experiment, participants accommodated according to the preset accommodative demand (either 0.5 D or 2.5 D). Accommodative convergence (AC) was triggered from accommodative response (A) according to the AC/A ratio. In Step 2, any misalignment (disparity between accommodative/convergence and the stereoscopic depth cue) was compensated for by the fusional vergence. Accommodation could be triggered according to the convergence accommodation/convergence ratio (the CA/C ratio). However, our CA/C ratios are far smaller than our AC/A ratios (Bruce et al., 1995; Rosenfield, Ciuffreda, & Chen, 1995; Fukushima, Torri, Ukai, Wolffsohn, & Gilmartin, 2009). All our participants were binocularly normal, or at least asymptomatic. The misalignment, even in mismatched conditions, might not have been huge. So the use of fusional vergence was unlikely to have triggered any significant accommodation response to cause blurriness for our participants. Indeed, all of the participants reported that they were able to read focused characters in Step 2 regardless of whether they saw double or single images. In our study, all participants reported that they saw double images for a short period of time before the images were fused. This suggests AC/A was at work. It is possible that minimal CA/C might have been at work which did not cause blurriness and the current results would likely be different if binocularly symptomatic viewers were tested. The accommodative vergence responses are consistent with the findings of Semmlow and Hung (1983), who reported that a near triad response is mainly blur-driven rather than disparity-driven.

3.0.2 Nonsignificant Effects of Matched versus Unmatched Accommodative and Vergence Demands and Their Interactions with Gender and Session (Testing of Hypotheses H1b and H1c).

ANOVAs were also conducted to test the matched versus unmatched demands hypotheses when the accommodative demand was controlled at 0.5 D (H1b) and when the vergence demand was controlled at 2.5 MA (H1c). We were surprised to find that while the main effects of gender and session remained significant ($p < .01$), the effects of matched demands in both situations were not significant (H1b: $F(1, 105) = 1.44$, $p > .2$; H1c: $F(1, 105) = 2.14$, $p > .1$). Again, there was no significant two-way interaction among the effects of matched demands, gender, and session. Consequently, the matched versus unmatched demands hypotheses H1b and H1c are rejected. The rejection of H1b suggests that when viewing stereo images presented on an HMD with a fixed lens focus of 200 cm (i.e., accommodative demand of 0.5 D), there is no significant difference in the time taken to fuse a pair of stereo images regardless of whether their stereo depths are at 2.5 MA (unmatched) or 0.5 MA (matched). This suggests that matching accommodative demands to vergence demands may not have any benefits in some situations.

Further testing of the matched versus unmatched demands hypotheses H1b and H1c using data collected in each of the six repeated blocks was conducted using nonparametric Wilcoxon signed-rank tests (see Table 3). In each test, 10 data points from the matched condition (one from each participant) were compared to 10 data

Table 3. *Summary of* p*-Values Obtained from the Wilcoxon Signed-Rank Tests Conducted to Test the Matched versus Unmatched Effects on Fusion Times for Two Conditions—(a) Accommodative Demand of 0.5 D and (b) Vergence Demand of 2.5 MA**

p-Value of matched effect**	Session 1	Session 2
First repeated block	0.799	0.646
Second repeated block	0.721	0.799
Third repeated block	0.241	0.037
p-Value of matched effect†	Session 1	Session 2
First repeated block	0.139	0.959
Second repeated block	0.139	0.093
Third repeated block	0.013	0.005

*NOTE. The values were collected at each repetition. In each test, 10 data points from the matched condition were compared to 10 data points from the unmatched condition.

**Data from the matched condition (0.5 D and 0.5 MA) were compared against data from the unmatched condition (0.5 D and 2.5 MA). This comparison was for testing the matched versus unmatched demands hypothesis H1b as explained in the main text.

†Data from the matched condition (2.5 D and 2.5 MA) were compared against data from the unmatched condition (0.5 D and 2.5 MA). This comparison was for testing the matched versus unmatched demands hypothesis H1c as explained in the main text.

points from the unmatched condition. With each successive repeat, the p values decreased (see Table 3). Examination of the raw data indicates that with each successive repeat, the time taken to fuse a pair of stereo images decreased more rapidly in the matched condition than in the unmatched condition. Consequently, the benefits of matching the accommodative demand to the vergence demand became increasingly significant with repeated viewings. The results of Wilcoxon tests indicate that data collected after six repeat sessions support both the matched versus unmatched demands hypotheses H1b and H1c ($p < .05$; see Table 3). This finding is important because it suggests that in prolonged tasks where stereoscopic images need to be repeatedly fused, matching accommodative demands to vergence demands is beneficial. The results of the Wilcoxon tests as shown in Table 3 clearly indicate that the effects of repeated blocks and the effects of matching the accommodative to vergence demands have significant interactions. This contradicts the results of the ANOVA analyses which indicated a lack of significant two-way interactions. This discrepancy may be due to the distribution of the data. As the transformed data were not of a normal distribution ($W = 0.9542$, $p < .0001$, Shapiro-Wilk test), the results of the Wilcoxon tests should reflect reality more accurately.

3.0.3 Potential Biases Introduced by the Initial Binocular Image Condition. In this study, viewers were presented with binocular images before the presentation of the stimuli and their resting vergence positions were closer to 0.5 MA (stereo depth of 200 cm) than 2.5 MA (stereo depth of 40 cm). This might have caused a bias in the result. However, further analyses indicated that if this bias had a strong effect, one would expect to see that under matched conditions (i.e., accommodative demand equaled the vergence demand), the fusion times recorded at 2.5 D and 2.5 MA would be larger than those recorded at 0.5 D and 0.5 MA. However, this was not the case ($p > 0.3$, Wilcoxon test). Furthermore, the bias introduced by the parallel resting eye positions would have the same effect on both the matched and unmatched conditions if the vergence demand was controlled either at 0.5 MA (H1d) and 2.5 MA (H1c). When the accommodative demand was controlled at 2.5 D (H1a), the potential bias contradicted the effects of matched demands. Since the effects of matched demands were found to be significant, they could only grow stronger when the bias was removed. It is possible that when the accommodative demand was controlled at 0.5 D (H1b), the bias could have helped to reduce the fusion time when the vergence demand was set to 0.5 MA. Further work is needed to clarify this issue.

3.0.4 Individual Effects (Testing of Hypotheses $H2_{\text{fusional reserves}}$ and $H2_{\text{lateral phoria}}$). It has been reported that humans have a lower ability to fuse stereo images by divergence than by convergence (e.g., Tunnacliffe, 1993) and that a convergence response is faster than a divergence response (Ciuffreda & Kenyon, 1983). Indeed, when the accommodative and vergence demands did not match, participants took significantly longer to fuse a pair of stereo images when the accommodative demand was nearer (as opposed to farther) than the vergence demand ($p < .01$: Wilcoxon signed-rank test). Under an unmatched condition, additional divergence eye movements were needed to fuse stereo images when the accommodative demand was nearer than the stereo depth. For example, when stereo images with depth at 0.5 MA were presented with an accommodative demand of 2.5 D, the participants' eyes first followed their accommodation responses to verge at a stereo depth of 2.5 MA, resulting in double images. To fuse the images, the participants had to diverge their eyes to verge appropriately for images at 0.5 MA. On the other hand, if the accommodative demand were farther than the vergence demand, additional convergence eye movements were required. Since the ability to turn the eyes toward and away from each other can be measured by positive and negative fusional reserves, respectively, fusional reserves may be the critical constraints on the ability to view stereo images presented on an HMD. Spearman correlation analysis between the fusion times and individual visual parameters were conducted. Negative fusional reserve at the near range of 14 to 34 prism diopters was significantly negatively correlated with the fusion times ($p < .05$, Spearman). Figure 6 shows the scatter plot between individual negative fusional reserves at near and the fusion times recorded at the four conditions. Positive fusional reserves (range: 16 to 30 prism diopters at distant; 14 to 38 prism diopters at near), on the other hand, did not significantly correlate with fusion times. Consequently, $H2_{\text{fusional reserves}}$ is supported for negative fusional reserves but not for positive fusional reserves. The lateral phoria at the near range and at the distant range was not significantly correlated with the fusion times. Consequently, hypothesis $H2_{\text{lateral phoria}}$ is rejected.

To study the individual effects of matching accommodative to vergence demands, data collected in the six repeated blocks of the two matched conditions (2.5 D and 2.5 MA; 0.5 D and 0.5 MA) were compared with the data collected from the same individuals in the six repeated blocks of the two unmatched conditions (2.5 D and 0.5 MA; 0.5 D and 2.5 MA). The results of the Wilcoxon signed-rank tests indicate that the amount of time taken to fuse a pair of stereo images in the matched conditions was significantly shorter than that taken in the unmatched conditions for all participants except participants P3, P4, and P7. Surprisingly, P3, P4, and P7 were among those with the smallest negative fusional reserve at distant range (8, 12, 14 prism diopters) and P3 and P4 had the smallest negative fusional reserve at near (14 and 15 prism diopters). This was puzzling because one would expect participants with less negative fusional reserves to have reduced ability to form a single stereo image and, therefore, to benefit more from matching the accommodative demand to stereo depth. Since this study only involved 10 participants, further studies with more participants are needed to confirm the individual effects because large individual differences in stereo depth perception have been reported (Akai, 2007). Further studies to explore the relationships between visual parameters and performance with HMDs are also warranted.

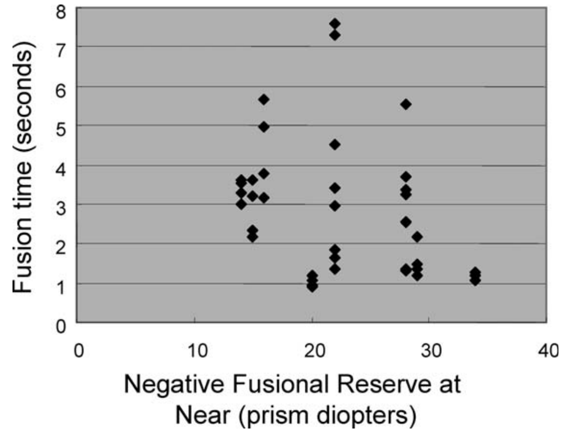

Figure 6. *A scatter plot illustrating the correlations between fusion times (in seconds) and individuals' negative fusion reserves measured at near (in prism diopters). Each data point represents the average of six repeated trials.*

4 Conclusion, Limitations and Future Work

When viewing stereo images presented on an HMD at an accommodative demand of 2.5 D, the median fusion time required to fuse stereo images at the unmatched vergence demand of 0.5 MA (2.8 s) was found to be 75% longer than that required to fuse images at 2.5 MA (1.6 s), even though the former requires less convergence. These differences are significant, $F(1, 105) = 12.91$, $p < .001$: ANOVA and $p < .001$: Wilcoxon signed-rank tests, and the finding holds regardless of the number of repetitions or the gender of the participants.

When viewing stereo images at 0.5 MA, the median fusion time can be shortened by a massive 32% (from 2.8 s to 1.9 s) by changing the accommodative demand of the HMD from 2.5 D to 0.5 D, $F(1, 105) = 11.02$, $p < .001$: ANOVA and $p < .001$: Wilcoxon signed-rank tests. This finding also holds regardless of the number of repetitions or the gender of the participants.

When viewing stereo images at 2.5 MA, the median fusion time can be shortened by changing the accommodative demand of the HMD from 0.5 D to 2.5 D, but the reduction was only significant after six repetitions ($p < .01$: Wilcoxon signed-rank tests).

In general, matching the accommodative demand to the vergence demand can shorten the fusion times when viewing stereo images and the benefits increase with repetitions. Although the fusion times are only a matter of seconds, for viewers who have difficulty in fusing stereoscopic images within a given time, the performance degradation can be substantial. Also, significant increases in fusion time would mean that viewers might be exposed to prolonged periods of double images which are associated with eye strain and headaches (e.g., Karasuda & McMains, 2005; Pickwell, 1984; Ware et al., 1998). Studies on the effects of matching the accommodative demand to the vergence demand on eye fatigue are desirable.

Significant negative correlations between the fusion times and individual negative fusional reserves at near has been found ($p < .05$, Spearman correlations). This suggests that participants with less negative fusional reserves require a longer time to fuse a pair of stereo images—a finding that is consistent with our current understanding (Tunnacliffe, 1993).

Among the 10 participants, seven showed a significant reduction in fusion times due to matching vergence and accommodative demands ($p < .05$, Wilcoxon tests). The three participants who did not show a significant reduction have some of the lowest measured negative fusional reserves. This is a surprising observation because one would expect those with less fusional ability to benefit more from having accommodative demand matched to the stereo depth. Further work in this area is desirable.

The stereo images used in this experiment were static. Future work using images with changing stereo depths would be beneficial. In addition, it would be useful to examine the effects of matching the accommodative demand to stereo depths on eye fatigue during prolonged viewing of stereo images. Visual fatigue among viewers of dynamically moving images has been the subject of concern (Lo & So, 2001; So & Ujike, 2010).

Notwithstanding the limitations stated above, this study provides evidence supporting the benefits of matching the accommodative demand to the vergence demand. In particular, matching the accommodative and vergence demands could lead to a 57% reduction in stereo image fusion time and the improvement from the matched demands increases with repeated viewing. Given the advances in various dynamic accommodation adjustment techniques for HMDs (e.g., the multi-focal LCD lattice of Rolland et al., 2000; the liquid-based lens of Hendriks et al., 2005), manufacturers of HMDs should explore the possible market opportunities for HMDs with an adjustable lens focus.

Acknowledgments

This study was funded by the Hong Kong Research Grants Council through the Competitive Earmarked Research Grant HKUST6128/03E and a research grant (A-PA0F) from the Hong Kong Polytechnic University. The authors would like to thank Prof. H. C. Huang for his advice on optics and display.

References

Akai, C. (2007). Depth perception in real and virtual environments: An exploration of individual differences. Unpublished Masters thesis. School of Interactive Art & Technology, Simon Fraser University, Canada. Retrieved from http://www.siat.sfu.ca/grad/theses/cakai/

Akeley, K., Watt, S. J., Girshick, A. R., & Banks, M. S. (2004). A stereo display prototype with multiple focal distances. *ACM Transactions on Graphics, 23*, 804–813.

Bando, T., Iijima, A., & Yano, S. (2012). Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review. *Displays*. Retrieved from www.sciencedirect.com.

Bruce, A. S., Atchison, D. A., & Bhoola, H. (1995). Accommodation-convergence relationships and age. *Investigative Ophthalmology and Vision Science, 36*(2), 406–413.

Cakmakci, O., & Rolland, J. P. (2006). Head-worn displays: A review. *Journal of Display Technology, 2* (3), 199–216.

Ciuffreda, K. J., & Kenyon, R. V. (1983). Accommodative vergence and accommodation in normals, amblyopes, and strabismics. In C. M. Schor & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinical aspects* (pp. 101–174). Woburn, MA: Butterworth-Heinemann.

Eichenlaub, J. B. (2005). Passive method of eliminating accommodation/convergence disparity in stereoscopic head-mounted displays. *Proceedings of SPIE, 5664*, 517–529.

Fukushima, T., Torri, M., Ukai, K., Wolffsohn, J. S., & Gilmartin, B. (2009). The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. *Journal of Vision, 9* (13), 1–13.

Hendriks, B. H. W., Kuiper, S., Van As, M. A. J., Renders, C. A., & Tukket, T. W. (2005). Electrowetting-based variable-focus lens for miniature systems. *Optical Review, 12* (3), 255–259.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and causes visual fatigue. *Journal of Vision, 8* (3), 1–30.

Hung, G. K., Zhu, H., & Ciuffreda, K. J. (1997). Convergence and divergence exhibit different response characteristics to symmetric stimuli. *Vision Research, 37* (9), 1197–1250.

Ilie, A., Low, K. L., Welch, G., Lastra, A., Fuchs, H., & Cairns, B. (2004). Combining head-mounted and projector-based displays for surgical training. *Presence: Teleoperators and Virtual Environments, 13*, 128–145.

Karasuda, E., & McMains, S. (2005). *Displaying readable object-space text in a head tracked, stereoscopic virtual environment*. CADML Technical Report 05-0722. Computer Aided Design and Manufacturing Laboratory, UC Berkeley, University of California.

Krishnan, V. V., & Stark, L. (1977). A heuristic model for the human vergence eye movement system. *IEEE Transactions on Biomedical Engineering, 24*, 44–49.

Lambooij, M., IJsselsteijn, W., Fortuin, M., & Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology, 53*(3), 030201-10 to 030201-14.

Lipton, L. (1982). *Foundations of the stereoscopic cinema*. New York: Van Nostrand Reinhold.

Lo, W. T., & So, R. H. Y. (2001). Cybersickness in the presence of scene rotational movements along different axes. *Applied Ergonomics, 32* (1), 1–14.

Melzer, J. E., & Moffitt, K. (1997). *Head mounted displays: Designing for the user* (pp. 55–82). New York: McGraw-Hill.

Patterson, R., Winterbottom, M. D., & Pierce, B. J. (2006). Perceptual issues in the use of head-mounted visual displays. *Human Factors, 48* (3), 555–573.

Pickwell, D. (1984). *Binocular vision anomalies: Investigation and treatment*. Woburn, MA: Butterworth-Heinemann.

Rolland, J. P., Krueger, M. W., & Goon, A. (2000). Multifocal planes head-mounted displays. *Applied Optics, 39* (19), 3209–3215.

Roscoe, S. N. (1987). The troubles with HUDs and HMDs. *Human Factors and Ergonomics Society Bulletin, 30* (7), 1–3.

Roscoe, S. N. (1988). The troubles with virtual images revisited. *Human Factors and Ergonomics Society Bulletin, 31* (1), 3–5.

Rosenfield, M., Ciuffreda, K. J., & Chen, H. W. (1995). Effect of age on the interaction between the AC/A and CA/C ratios. *Ophthalmic and Physiological Optics, 15*(5), 451–455.

Rushton, S. K., & Riddell, P. M. (1999). Developing visual systems and exposure to virtual reality and stereo displays: Some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics, 30*, 69–78.

Samsung. (2010). SAMSUNG announces true 3D capability for multiple screen technologies. Retrieved from www.samsung.com

Semmlow, J. L., & Hung, G. K. (1983). The near response: Theories of control. In M. Clifton & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinical aspects* (pp. 175–195). Woburn, MA: Butterworth-Heinemann.

Semmlow, J. L., & Venkiteswaran, N. (1976). Dynamic accommodative vergence in binocular vision. *Vision Research, 16*, 403–411.

Shevlin, F. (2005). A fixed-viewpoint volumetric stereoscopic 3-D display using adaptive optics. *SPIE Conference on Stereoscopic Displays and Virtual Reality Systems,* No. 12, vol. 5664, pp. 22–27.

Shibata, T., Kawai, T., Otsuki, M., Miyake, N., Yoshihara, Y., Iwasaki, T., & Terashima, N. (2004). Development and evaluation of stereoscopic 3-D display with dynamic optical correction. *Proceedings of the 11th International Display Workshops (IDW04),* 1555–1558.

Siderov, J., & Harwerth, R. (1993). Precision of stereoscopic depth perception from double images. *Vision Research, 33* (11), 1553–1560.

So, R. H. Y., & Ujike, H. (2010). Visually induced motion sickness, visual stress and photosensitive epileptic seizures: What do they have in common? Preface to the special issue. *Applied Ergonomics, 41* (4), 491–393.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues on virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments, 7*(4), 327–351.

Stark, L. (1983). Normal and abnormal vergence. In C. M. Schor & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinical aspects* (pp. 3–14). Woburn, MA: Butterworth-Heinemann.

Toshiba. (2011). Toshiba announces a commercial glass-less 3D TV. Retrieved from http://www.businessreviewusa.com/industry-focus/technology/toshiba-announces-glasses-less-3dtv-and-new-tablet

Tunnacliffe, A. H. (1993). *Introduction to visual optics* (pp. 518–519). Fent, UK: The Association of British Dispensing Opticians.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1997). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 35*, 2731–2736.

Wann, J. P., & Mon-Williams, M. (1997). Health issues with virtual reality displays: What we do know and what we don't. *Computer Graphics, 31*, 53–57.

Ware, C., Gobrecht, C., & Paton, M. A. (1998). Dynamic adjustment of stereo display parameters. *IEEE Transactions of Systems, Man, and Cybernetics, Part A: Systems and Humans, 28* (1), 56–65.

Wong, W. S., Yip, R., So, R. H. Y., Huang, H. C., Lam, A., & Ting, P. (2007). The effects of matching lens focus with stereoscopic depth cues when viewing images presented on a head-mounted display. *Proceedings of the First International Symposium on Visually Induced Motion Sickness, Fatigue, and Photosensitive Epileptic Seizures (VIMS2007)*, 151–155.