The Effects of Matching Lens Focus with Stereoscopic Depth Cues when Viewing Images Presented on a Head-Mounted Display

W.S. Wong[1], Romeo Yip[1], R.H.Y. So[1] H.C. Huang[2], Andrew Lam[3] and Patrick Ting[3]

[1]Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology, Hong Kong SAR, PRC.

[2]Department of Electronics Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, PRC.

[3]School of Optometry, The Hong Kong Polytechnic University, Hong Kong SAR, PRC.

[1]rhyso@ust.hk

Keywords: Stereoscopic image, lens focus, head-mounted displays.

Abstract. Most head-mounted displays (HMDs) only provide fixed lens focus. Viewers are force to decouple accommodation and vergence when viewing stereoscopic images. A study has been conducted to investigate the benefit of matching lens focus with stereoscopic depth cues on the time taken to form a single stereoscopic image. Ten university students participated four conditions that were exhaustive combinations of lens focus (40cm and 200cm) and stereoscopic depth cues (40cm and 200cm). Results indicated that viewers took significantly shorter time to form a single stereoscopic image when the lens focus matched with the stereoscopic depth cues (p<0.05). Further analyses of the data suggest that unnatural demand for the eyes to diverge (i.e. when the lens focus was smaller than the stereoscopic depth cues) is associated with significantly longer periods of double images (p<0.05). The potential benefits of applying dynamically adjustable lens focus to the future design of HMD are discussed.

Introduction

Head-mounted displays (HMDs) can present stereoscopic images to a viewer. With the help of head position tracking system, the orientations of the stereo images can be made appropriate to the orientation of the viewer's head (So and Griffin, 2000). The objective is to immerse a viewer in a computer-generated virtual environment in a natural way (Kiryu and So, 2007). Unfortunately, most HMDs have fixed lens focuses (Cakmakci and Rolland, 2006; Elchenlaub, 2006). When a HMD with a constant lens focus is used to present stereo images with varying virtual depth, viewers are forced to decouple accommodation and vergence responses. This posts an unnatural demand on the viewers' eyes (Rushton and Riddle, 1995) and are major short-comings of the HMD technology (Roscoe, 1987, 1988; Wann *et al.,* 1997a,b).

Shibata *et al.* (2004) reported a HMD with dynamically adjustable lens focus. However, the HMD was costly. A review of literature indicates that there has been no laboratory experiment reporting the effects of mismatch between lens focus and stereo depth cues on visual task performance with HMDs. To encourage HMD manufacturers to invest in producing HMDs with dynamically adjustable lens focus, its benefits needed to be verified.

Experiment

Objectives and hypotheses. The experiment aim to examine the effects of matching lens focus and image depth on the time taken for viewers to form a clear single stereo image (referred to as 'single image formation time'). The possible interaction effects with binocular visual parameters of the viewers will also be studied.

The study has seven hypotheses: (i) when lens focus is 40cm, the single image formation time of viewing images with stereo cues appropriate to a depth of 40cm will be significantly shorter than viewing images of stereo cues appropriate to a depth of 200cm (H1), (ii) when lens focus is 200cm, the single image formation time of viewing images with stereo cues appropriate to a depth of 200cm will be significantly shorter than viewing images with stereo cues appropriate to a depth of 40cm (H2), (iii) with images of stereo cues appropriate to a depth of 40cm, the single image formation time when viewing these images through a lens focus of 40cm will be significantly shorter than that of viewing through a lens focus of 200cm (H3), (iv) with images of stereo cues appropriate to a depth of 200cm, the single image formation time when viewing these images through a lens focus of 200cm will be significantly shorter than that of viewing through a lens focus of 40cm (H4), (v) individual fusional reserve will be negatively correlated with single image formation times (H5), (vi) individual levels of lateral phoria will be correlated with single image formation times (H6), and (vii) the effects of matching lens focus and stereo cues on single image formation times will not reduce with repetitions (H7).

Design, Variables, Stimuli, and Apparatus. This study used a full factorial within-subject design. Dependent repeats were used to obtain better mean estimations of the data. The main dependent variable is the time taken for each participant to form a single stereoscopic image. The experiment has four independent variables: (i) lens focus: 40cm and 200cm; (ii) stereo cues appropriate to image depth of 40cm and 200cm, (iii) repetition block: 3 repeated blocks on the same day, and (iv) two repeated sessions (one session consists of three repeated blocks) separated by at least three days. The choice of 40cm and 200cm is to test the near and far vision. Optical infinity cannot be used because the apparent size of the image will approach zero. Within one block, the viewers watched and formed a single stereoscopic image 20 times exhausting the combinations of two lens focus, two image depths, and five dependent repeats. In other words, each viewer was exposed to four conditions (C1: both lens focus and image depth were 40cm; C2: lens focus was 200cm and image depth was 40cm; C3: lens focus was 40cm and image depth was 200cm, and C4: both lens focus and image depth were 200cm). The order of presentation of the four conditions within each block was randomized and was kept the same across all repeated blocks in both sessions. As stated before, each condition was repeated five times (dependent repeated) and the data collected were averaged to get better mean estimations. The visual stimuli consists of pairs of stereoscopic images showing two groups of three-dimensional characters presented with stereo cues appropriate to virtual depths of 40cm and 200cm. The visual background (also in stereo) was the inside of a room with a depth of 2000cm. Stimuli were generated using World-Tool-Kit^{TM} software and were presented on a home-built HMD with motor-driven adjustable lens focus (see Figure 1). The optical module of this HMD was purchased from iMD Display Ltd. (www.hkimd.com) and had a resolution of 688 pixels (horizontal) by 480 pixels (vertical). The field of view was 30 degrees diagonal. The adjustment range of the lens focus was from 30cm to infinity. The experiment was conducted in an air-conditioned dark room.

Procedure and Participants. Ten university students (six male, four female) took part in the study. Each person viewed and formed single stereoscopic images 120 times (i.e., 120 runs) exhausting the combinations of 4 conditions, 5 dependent repeats, 3 repeated blocks, and 2 repeated sessions. In this study, the HMD was not coupled to a head-tracker because past studies have shown that time delays due to head movement might affected performance (So and Chung *et al.*, 1999; So and Griffin, 1991; 1992; 1995). Vision screening tests were conducted before the experiment using the vision tester OPTEC 2000 (Stereo Optical Co., Inc.). All ten participants attained visual acuity (either normal or corrected) equaled or better than 20/20 in both the near (14 inch, i.e. 36cm) and distance (20 feet, i.e. 6.1 meter) tests. They passed the colour blindness test and attained stereopsis of 40 seconds of arc or less. The purposes of these screening criteria were to ensure that the participants could have sufficient visual ability to perform the task. In additional to these visual parameters, the binocular visual parameters of the ten participants were also measured by a trained optometrist at the School of Optometry, Hong Kong Polytechnic University. Six parameters were measured: negative fusional reserve at distance (6m) and near (40cm), positive fusional reserve at distance (6m) and near (40cm), lateral phoria at distance (3m) and near (1/3m i.e. 33cm).

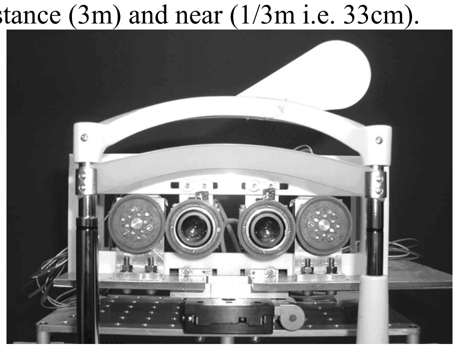

Figure 1. A photograph illustrating the home-built binocular displays with motor-driven adjustable lens focus (adopted from Wong et al., (2008).

During each of the 120 runs, participants were required to fixate and form a single stereoscopic image when viewing the appropriate group of three-dimensional characters as fast as they could. At the beginning of each run, biocular images of the three-dimensional characters were presented first. Then, binocular images were presented and the participants would perceive double images instantly. After a certain short period of time (referred to as the single image formation time), a clear single stereoscopic image should be perceived. After that, biocular images were presented again and the apparatus was ready to show the next run. The experiment was approved by the Human Subject Experimentation Committee at the Hong Kong University of Science and Technology.

Results and Discussion

Effects of Matching Lens Focus with Stereo Depth Cues. The single image formation time data were transformed for normality (Box-Cox transformation, $\lambda = -0.5$). After the transformation, data are not normal. So, non-parametric statistical tests are used where appropriate. Preliminary data are presented in this paper and readers can refer to Wong *et al.* (2008) for more details on the study.

Wilcoxon signed ranks tests have been conducted to test hypotheses H1 to H4. Preliminary results indicate that H1, H3 and H4 are supported (p<0.001) but H2 has been rejected (p>0.09). Since the effects of gender and session have been found to be significant (p<0.05, ANOVAs), the two-way interacting effects among the effects of matching of lens focus and stereo cues, the effects of gender, and the effects of sessions have been investigated. Hypotheses H1 and H4 have been verified to be true for both sessions and genders. H3 is true at both sessions with data from male participants, but false with data from female participants. Hypothesis H2 has been shown to be false at both sessions with female data but true with male data.

Effects of repetition. Results of ANOVAs indicate that effects of repeated block are not significant but the effects of session are significant (p<0.001). One reasonable question will be whether viewers will be able to shorten their single image formation time with practice and, hence, eliminate or reduce the benefits of matching lens focus and stereo depth cues. Further examination of the data indicate that as the number of repeated run increase, there exists a trend of shortening the single image formation time and the trend becomes significant when the repetition go into the second session ($p<0.001$, ANOVAs). However, preliminary data also indicate that the rates of reduction of the single image formation times are faster with data collected from the matched conditions than the un-match conditions. In other words, with repetitions, the significant effects of reducing single image formation time by matching lens focus and image depth remain and even get stronger with repetitions. As a result, hypothesis H7 is supported. This finding is important because it suggests that the problems associated with the mismatch between lens focus and stereo depth cues cannot be eliminated by practices.

Effects of Individual visual parameters. The relationship between the individuals' visual parameters and their absolute single image formation time could be learnt through the testing of H5 and H6. Correlation analyses on the preliminary data indicate significant correlations between the single image formation time and negative fusional reserve at both near and far distance ($p<0.001$, Spearman). However there has been no significant correlation for the positive fusional reserves. Therefore H5 is supported only for the negative fusional reserves but false for the positive fusional reserves.

Lateral phoria at near has also been found to significantly correlate with the single image formation time collected in Condition C1 ($p<0.05$) and Condition C4 ($p<0.05$). The positive significant correlation supports the hypothesis H6 made for the matched conditions.

Conclusions and Limitations

For the first time, significant improvement in visual task performance due to matching lens focus and stereo depth cues has been demonstrated. Results also suggest that the improvement remains and becomes stronger with practices.

Future work to examine whether the improvement in single image formation time will lead to reduction of eye fatigue and motion sickness is desirable because about 30% of Chinese population is susceptible to motion sickness (So and Finney *et al.*, 1999) and have been shown to suffer from cybersickness (Lo and So, 2001; So and Lo, 1999).

Acknowledgements

This study was partially funded by the Hong Kong Research Grants Council through the Earmarked Grant HKUST6128/03E and a research grant (A-PA0F) from the Hong Kong Polytechnic University.

References

Cakmakci, O., Rolland J. P. (2006). Head-Worn Displays: A Review. Journal of Display Technology, 2(3), 199-216.

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

Hendriks, B. H. W. , Kuiper S., Van As M. A. J., Renders C. A., and Tukket T. W. (2005). Electrowetting Based Variable Focus Lens for Miniature Systems. Optical Review, 12(3), 255–259.

Kiryu, T. & So, R.H.Y. (2007) Sensation of Presence and Cybersickness in applications of Virtual Reality for Advanced Rehabilitation. Journal of NeuroEngineering and Rehabilitation, 4:34.

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

Melzer, J. E., & Moffitt, K. (1997). Head Mounted Displays: Designing for the User (pp. 55–82). New York: McGraw-Hill.

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.

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.

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.

So, R.H.Y., Chung, K.M. & Goonetilleke, R.S. (1999) Target-directed head movements in a head-coupled virtual environment: predicting the effects of lags using Fitts' law. Human Factors, Vol.41, No.3, 1999, pp. 474-486.

So, R.H.Y., Finney, C.M. & Goonetilleke, R.S. (1999) Motion sickness susceptibility and occurrence in Hong Kong Chinese. Contemporary Ergonomics, 1999, Taylor & Francis, pp.88-92.

So, R.H.Y. & Griffin, M.J. (1991) Effects of time delays on head tracking performance and the benefits of lag compensation by image deflection." Proceedings of Flight Simulation Technologies Conference, New Orleans, Louisiana, 12-14 August, 1991, pp.124-130.

So, R.H.Y. & Griffin, M.J. (1992), Compensating lags in head-coupled displays using head position prediction and image deflection. J. Aircraft. 29(6): 1064-1068, 1992.

So, R.H.Y. & Griffin, M.J. (1995) Head-coupled virtual environment with display lag. Ch5 in: Simulated and virtual realities, (Ed): K.Carr and R. England. Taylor and Francis, London, pp.103-110.

So, R.H.Y. & Griffin, M.J. (2000) Effects of target movement direction cue on head-tracking performance. Ergonomics, Vol.43, No.3, pp. 360-376.

So, R.H.Y. & Lo, W.T. (1999) Cybersickness: An Experimental Study to Isolate the Effects of Rotational Scene Oscillations. Proceedings of IEEE Virtual Reality '99 Conference, March 13-17, 1999, Houston, Texas. Published by IEEE Computer Society, pp.237-241.

So, R.H.Y., Wong, W.S., Yip, R., Lam, A. & Ting, P. (2008) Benefits of matching accommodative demands to vergence requirement in a binocular head-mounted display: a study on stereo fusion times. Submitted to Presence. Awaiting publication.

Tunnacliffe A. H. (1993). Introduction to Visual Optics (pp518-519). The Association of British Dispensing Opticians.

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

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

Webb, N. A. & Griffin M. J. (2002). Optokinetic stimuli: motion sickness, visual acuity, and eye movements. Aviation, space, and environment medicine, 73(4), 351-358.