Evaluation of Accommodation Responses to Monovision for Virtual Reality

Nitish Padmanaban, Robert Konrad, Gordon Wetzstein
Stanford University
{nit,rkkonrad,gordon.wetzstein}@stanford.edu

Abstract: Emerging virtual and augmented reality (VR/AR) systems can produce highly immersive experiences, but also induce visual discomfort, eyestrain, and nausea for some users. One of the sources of these symptoms is the lack of natural focus cues in all current VR/AR near-eye displays. These displays project stereoscopic image pairs, driving the vergence state of the eyes to arbitrary distances. However, the accommodation, or focus state of the eyes, is optically driven to a fixed distance. In this work, we empirically evalaute monovision: a simple, yet unconventional method for potentially driving the accommodation state of the eyes to two distances by allowing each eye to drive focus to a different distance.

OCIS codes: 330.7322, 330.1400, 330.5020.

1. Introduction

Virtual and augmented reality (VR/AR) systems have promising applications spanning entertainment, education, communication, training, behavioral therapy, and basic vision research. However, the basic optics of these systems have remained largely unchanged since their conception over a century ago. In front of each eye, a small physical display is placed behind a magnifying lens, creating a virtual image at some fixed distance from the viewer, with small differences in the images displayed to the two eyes creating a perception of depth, called stereopsis.

But this simple optical design lacks a critical aspect of 3D vision: changes in stereoscopic depth are also associated with changes in focus. When viewing a near-eye display, users' eyes can change their vergence angle to a range of stereoscopic depths, but to focus on the virtual image, the eyes must accommodate to a single, fixed distance. For users with normal vision, this creates an unnatural condition known as the vergence–accommodation conflict (VAC) [6, 7]. Symptoms associated with this conflict include double vision (diplopia), compromised visual clarity, visual discomfort and fatigue [6, 13]. Furthermore, the presence of accurate focus is a cue that is important for depth perception [2].

Providing consistent stereoscopic and focus cues is a challenging problem and substantial engineering efforts in VR/AR have been invested in developing displays with natural focus cues. Generally, these approaches can be divided into several categories: dynamic focus [4, 5, 9], volumetric or multi-plane [12], light field [3, 8], holographic, and disparity-driven displays (such as monovision).

Monovision Monovision refers to a common treatment for presbyopia, a condition that occurs with age in which people lose the ability to focus their eyes on nearby objects. The goal of current treatments for presbyopia, such as monovision, is to extend the range of distances over which patients can experience clear vision. Monovision accomplishes this goal by focusing one eye to a near fixed distance, and the other to a far fixed distance by placing a different power lens in front of each eye, or through surgery [1].

Monovision shares this goal of extending the range of focus with current VR displays, and as such has been attempted as a method of reducing the VAC [4, 5]. The idea, first introduced by Marran and Schor [10], ideally induces a disparity-driven accommodative response by presenting a different stimuli to each eye. When a user fixates (verges) on a near object, they will ideally focus to the near plane: the more natural viewing condition. When fixating far, they will ideally focus far, relying on the other eye. While previous studies have investigated monovision with subjective user studies, we directly examine users' accommodative responses via a system equipped with an autorefractor.

2. Focus-Tunable Near-Eye Display System

The benchtop prototype (Fig. 1) uses Topfoison TF60010A Liquid Crystal Displays with a resolution of $2,560 \times 1,440$ pixels and a screen diagonal of 5.98 in. The optical system for each eye offers a field of view of 34.48° and comprises three Nikon Nikkor 50-mm f/1.4 camera lenses. In the optical path of each eye is a focus-tunable lens (Optotune EL-10-30-C) that can dynamically place the virtual image at any distance between 0 and 5 D within a 15 ms settling time,

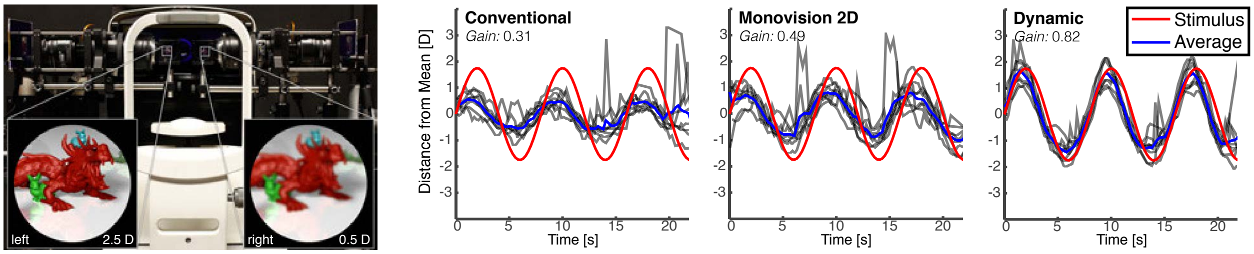

Fig. 1. (left) The benchtop prototype uses focus-tunable lenses to implement the conventional, monovision, and dynamic display modes. An autorefractor is integrated into the optical path to allow measurements of accommodation simultaneously with displayed content. (right) Recorded measurements of the accommodative responses are shown for conventional, monovision (left: 2.5 D, right: 0.5 D), and dynamic, with their respective gains. The stimulus is shown in red, with individual responses in gray and the average across all users in black.

independently per eye. Two additional camera lenses provide a 1:1 optical relay system that increases the eye relief so as to provide sufficient spacing for a near-IR (NIR)/visible beam splitter (Thorlabs BSW20R) that allows measurement of accommodation using a Grand Seiko WAM-550 Autorefractor during active viewing at a rate of 4–5 Hz with an accuracy of $\pm$0.25 D. The system can account for a user's interpupillary distance adjustment via a translation stage.

3. Evaluation

To evaluate the accommodative response to visual stimuli in the monovision display mode, we conducted two user studies comprising 11 volunteers (ages 24–32, 1 female), but 3 people were excluded from the first study due to insufficient data. In each study, we objectively measured user accommodation in response to visual targets using an autorefractor. All had normal or corrected-to-normal vision and normal stereoacuity as assessed with a Randot test. All participants gave informed consent, with procedures approved by the institutional review board of the home institution.

In the first study, we examined the gain of users' accommodative responses while they visually tracked a target oscillating sinusoidally in depth. We compared three different display modes. The first two modes were conventional (both eyes: 0.769 D) and monovision (left eye: 2.5 D, right eye: 0.5 D). The third mode used dynamic focus to provide near-accurate focus cues, and serves as our baseline. The dynamic mode updates the optical distance to the virtual image to match the vergence distance. Several studies have shown [4, 5], via autorefractor measurements [9, 11], that the dynamic display mode is capable of achieving natural accommodative responses.

We predicted users would accommodate most accurately in the dynamic mode, least accurately in the conventional mode, and in between for the monovision mode. For all modes, a 6.2cm Maltese cross oscillated between 0.5 and 4 D (mean 2.25 D, amplitude 1.75 D) at 0.125 Hz. Users were instructed to track the target while autorefractor measurements were recorded. Each user performed this task in each display mode condition once. The order of conditions was randomized per user. The accommodative gain relative to the stimulus was calculated for each condition. The stimulus was presented for 4.5 cycles, and responses were analyzed for 3 cycles after a 0.5 cycle buffer.

Fig. 1 shows individual (black) and average (blue) accommodative responses to the stimulus (red) for each mode. In the conventional mode (left), users exhibited a small average gain of 0.31, despite having a fixed virtual image. Consistent with our prediction, gain in monovision mode (middle) had a higher average of 0.49. In the dynamic mode (right), near-correct focus cues lead to the highest average gain of 0.82. Statistical significance between the gains was assessed using a one-way repeated measures analysis of variance (ANOVA). The ANOVA showed a significant main effect of display mode, $F(2,14) = 36.5$, $p < 0.001$. Follow up t-tests (p-values Bonferroni corrected) indicated that the gains in the dynamic condition were significantly higher than in the conventional ($p < 0.001$) and monovision conditions ($p < 0.05$). Average gain in the monovision condition was also significantly higher than in the conventional condition ($p < 0.05$). These results indicate that disparity-driven accommodation via differently placed virtual images per eye may be possible, although the resulting gain is significantly reduced relative to natural focus cues.

A second study was performed to confirm and extend these results: we recorded accommodative responses to static targets, rendered at different depths, for five different display modes. The three conditions from the first study were retained, and two more monovision conditions were added: one with a 1 D difference between the eyes (left: 1.5 D, right: 0.5 D), and one with a 3 D difference between the eyes (left: 3.5 D, right: 0.5 D). For each trial, users were instructed to fixate on a Maltese cross target that appeared at a fixed distance following a 2 sec dark period. The accommodative response was recorded after at least 3 sec of viewing. The fixation distance was randomized from a set of 9 distances: 0.1 D, and 0.5–4 D in increments of 0.5 D. All distance measurements for each mode were blocked together, with the order of the modes randomized per user.

The results for this study are shown in Fig. 2. The upper panel shows the mean and standard errors of the meas-

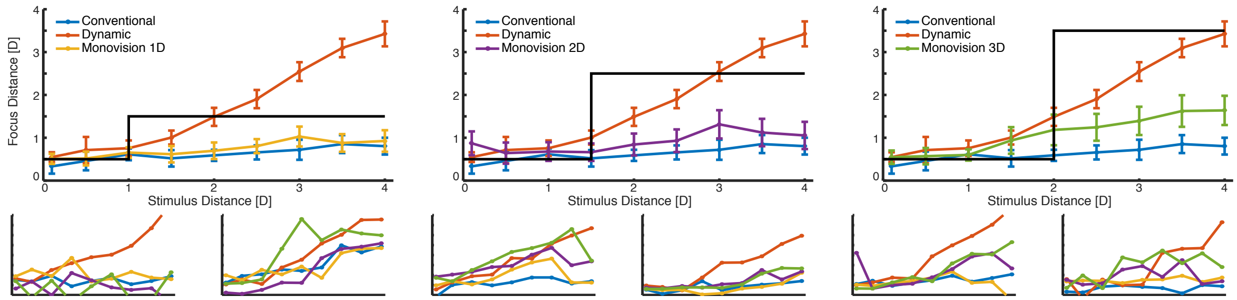

Fig. 2. Accommodative responses in the second user study. The upper panels show the between-subjects mean accommodative response for each distance in each display mode. Error bars indicate the standard error of the mean. The dashed lines correspond to the predicted accommodative response for the monovision mode in each panel. The lower panels show results in the same format for six example users.

ured accommodation across users for each condition. The dashed black line indicates the virtual plane presented to either eye that is nearest the stimulus distance: the theoretical monovision response. Consistent with the first study, in each monovision condition, users accommodated more accurately than in the conventional condition (blue), but less accurately than in the dynamic condition (red). Interestingly, there were large variations in responses between users to the monovision conditions while responses to the dynamic condition remained consistent, illustrated by the lower panels in Fig. 2. The inconsistent response to monovision may reflect the individual differences in the strength of the cross-coupling between the vergence and accommodation, and may indicate its limited use for the general population.

4. Conclusion

Monovision offers VR systems a tantalizingly simple solution to the issue of extending the range of focus. It requires no extra power, no moving parts, and could work equally well for users with or without presbyopia. Previous work has been conflicted on its efficacy, and we aim to resolve this with objective accommodation data. While monovision may be beneficial some users, its effect on accommodation is highly variable and largely inconsistent across users. There seems to be an increase in accommodative gain on average, but the size of this effect falls far short of a near-correct stimulus, such as the dynamic mode. As such, monovision may not be an effective solution for the population at large.

References

1. B. J. Evans. Monovision: a review. *Ophthalmic and Physiological Optics*, 27(5):417–439, 2007.
2. D. Hoffman, A. Girshick, K. Akeley, and M. Banks. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3), 2008.
3. F. Huang, K. Chen, and G. Wetzstein. The light field stereoscope: Immersive computer graphics via factored near-eye light field display with focus cues. *ACM Trans. Graph. (SIGGRAPH)*, 34(4), 2015.
4. P. Johnson, J. Parnell, J. Kim, C. Saunter, G. Love, and M. Banks. Dynamic lens and monovision 3D displays to improve viewer comfort. *Optics Express*, 24:11808–11827, 2016.
5. R. Konrad, E. Cooper, and G. Wetzstein. Novel Optical Configurations for Virtual Reality: Evaluating User Preference and Performance with Focus-tunable and Monovision Near-eye Displays. In *Proc. SIGCHI*, 2016.
6. F. L. Kooi and A. Toet. Visual comfort of binocular and 3d displays. *Displays*, 25:99 – 108, 2004.
7. M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3), 2009.
8. D. Lanman and D. Luebke. Near-eye light field displays. *ACM Trans. Graph. (SIGGRAPH Asia)*, 32(6), 2013.
9. S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *Proc. ISMAR*, pages 33–42, 2008.
10. L. Marran and C. Schor. Multiaccommodative stimuli in vr systems: problems and solutions. *Human Factors*, 39(3):382–388, 1997.
11. N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences*, 2017.
12. B. Schowengerdt and E. Seibel. True 3-d scanned voxel displays using single or multiple light sources. *J. SID*, 14(2):135–143, 2006.
13. T. Shibata, J. Kim, D. Hoffman, and M. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11(8):11, 2011.