Computational Focus-Tunable Near-eye Displays

Robert Konrad[1] Nitish Padmanaban[1] Emily Cooper[2] Gordon Wetzstein[1]
[1]Stanford University [2] Dartmouth College www.computationalimaging.org

Abstract

Immersive virtual and augmented reality systems (VR/AR) are entering the consumer market and have the potential to profoundly impact our society. Applications of these systems range from communication, entertainment, education, collaborative work, simulation and training to telesurgery, phobia treatment, and basic vision research. In every immersive experience, the primary interface between the user and the digital world is the near-eye display. Thus, developing near-eye display systems that provide a high-quality user experience is of the utmost importance. Many characteristics of near-eye displays that define the quality of an experience, such as resolution, refresh rate, contrast, and field of view, have been significantly improved in recent years. However, a significant source of visual discomfort prevails: the vergence-accommodation conflict (VAC). This visual conflict results from the fact that vergence cues, but not focus cues, are simulated in near-eye display systems. Indeed, natural focus cues are not supported by any existing near-eye display. Afforded by focus-tunable optics, we explore unprecedented display modes that tackle this issue in multiple ways with the goal of increasing visual comfort and providing more realistic visual experiences.

Keywords: Computational Displays, Virtual Reality, Augmented Reality

1 Motivation and Background

In current VR/AR displays, a stereoscopic image pair is presented to the viewer such that binocular disparity can cues drive the vergence state of the viewer's eyes to arbitrary simulated distances (Fig. 1). However, the accommodation, or focus, state of their eyes is fixed to one specific distance – the optical distance to the magnified microdisplay.

Vergence and accommodation are neurally coupled, because in the physical world these two depth cues work in harmony. VR/AR displays artificially decouple these cues due to their optical image formation. The resulting discrepancy between natural depth cues and those produced by existing VR/AR displays may lead to visual discomfort and fatigue, eyestrain, diplopic vision, headaches, nausea, and compromised image quality [Hoffman et al. 2008].

The benefits of providing correct or nearly correct focus cues not only include increased visual comfort, but also improvements in 3D shape perception, stereoscopic correspondence matching, and discrimination of larger depth intervals. Thus, significant efforts have been made to engineer focus-supporting displays. However, all technologies that can potentially support focus cues suffer from undesirable tradeoffs, such as compromised image resolution, device form factor, brightness, contrast, or other important display characteristics. These tradeoffs pose substantial challenges for high-quality, practical, and wearable displays.

SIGGRAPH '16, July 24-28, 2016, Anaheim, CA,
ISBN: 978-1-4503-4372-5/16/07
DOI: http://dx.doi.org/10.1145/2929464.2929470

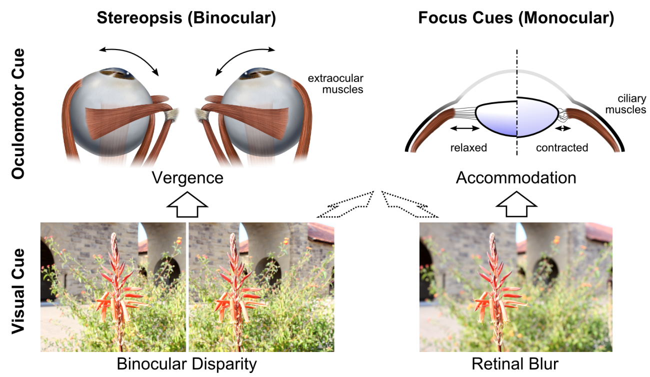

Figure 1: *Overview of relevant depth cues. Vergence and accommodation are oculomotor cues whereas binocular disparity and retinal blur are visual cues. In normal viewing conditions, disparity drives vergence and blur drives accommodation. However, these cues are cross-coupled, so there are conditions under which blur-driven vergence or disparity-driven accommodation occur.*

2 Focus-tunable Near-eye Displays

In this work, we ask whether it is possible to provide natural focus cues and to mitigate visual discomfort using focus-tunable optics, i.e. programmable liquid lenses. For this purpose, we demonstrate a prototype focus-tunable near-eye display system (Fig. 2) that allows us to evaluate several advanced display modes via user studies.

Conventional Near-eye Displays are simple magnifiers that enlarge the image of a microdisplay and create a virtual image at some fixed distance to the viewer.

Adaptive Depth of Field Rendering is a software-only approach that renders the fixated object sharply while blurring other objects according to their relative distance. When combined with eye tracking, this mode is known as gaze-contingent retinal blur [Mauderer et al. 2014]. Due to the fact that the human accommodation system may be driven by the accommodation-dependent blur gradient, this display mode does not reproduce a physically correct stimulus.

Adaptive Focus Display is a software/hardware approach that either changes the focal length of the lenses or the distance between the microdisplay and the lenses [Konrad et al. 2016]. When combined with eye tracking, this mode is known as gaze-contingent focus. In this mode, the magnified virtual image observed by the viewer can be dynamically placed at arbitrary distances, for example at the distance where the viewer is verged (requires vergence tracking) or at the depth corresponding to their gaze direction (requires gaze tracking). No eye tracking is necessary to evaluate this mode when the viewer is asked to fixate on a specific object, for example one that moves.

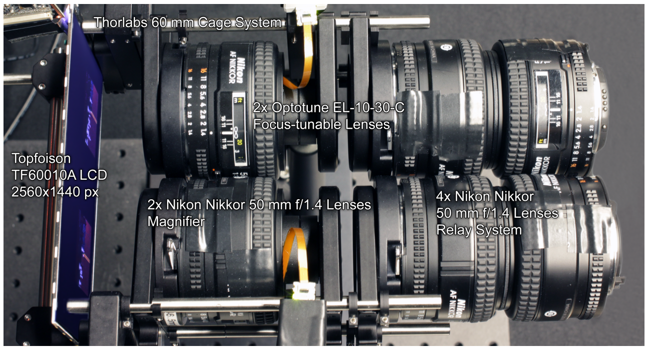

Figure 2: *Photograph of prototype display. This benchtop setup uses a 2K LCD display with relay optics and focus-tunable lenses, allowing the virtual image to be placed as close as 20 cm.*

Monovision refers to a common treatment for presbyopia, a condition that often occurs with age in which people lose the ability to focus their eyes on nearby objects. To improve visual clarity, monovision places two lenses with different prescription values in front of each eye such that one eye dominates for distance vision and the other for near vision. Monvision was recently proposed and evaluated for VR/AR applications [Konrad et al. 2016].

Accommodation-invariant Near-Eye Display is a novel mode not discussed in the literature. Using the principle of a focus sweep, i.e. quickly moving the virtual image throughout the accommodation range of the viewer, a stimulus is created that does not vary w.r.t. the accommodation state of the eye. Technically, this is similar to wearing pinhole contact lenses, but focus-tunable optics provide more light and a wider eye box. This mode "optically disables" the retinal blur cue by rendering it into open loop, and potentially allowing accommodation to be driven by disparity.

3 Discussion

What are the main insights of this study? Both the focus-tunable mode and the monovision mode demonstrate improvements over the conventional display, but both require optical changes to existing VR/AR displays. A software-only solution (i.e. depth of field rendering) proved ineffective. The focus-tunable mode provided the best gain over conventional VR/AR displays. We implemented this display mode with focus-tunable optics, but it could also be implemented by actuating (physically moving) the microdisplay in the VR/AR headset. Please see Konrad et al. [2016] for more details on our studies.

How does this study inform next-generation VR/AR displays? Based on our study, we make several recommendations to near-eye display designers:

- With conventional optical (magnifier) designs, you need to place the accommodation plane somewhere. VR consumer displays often place it at 1.3-1.7 m, but this distance really depends on your application. If you mostly look at virtual objects that are close, place the accommodation plane close. If you mostly look at objects that are far, place it far. If the content could be at any distance, place it at 1-2 m.

- The focus-tunable display mode scored the best in our studies. If you can afford to dynamically change the accommodation plane depending on what the user is looking at – that would be great! We used focus-tunable lenses to achieve this, because they allow us to switch between different display modes easily, which is great for user studies. In practice, you probably want to use actuated microdisplays, because focus-tunable lenses have a small diameter, which results in a small field of view, they have a lot of aberrations, and they are relatively heavy and power hungry. If you build a product, use actuated displays instead to achieve the same effect while maintaining field of view, image quality, etc.

- If you have eye tracking, dynamically adjust the accommodation plane depending on what the user is looking at. We call this mode gaze-contingent focus. With eye tracking in place, this would probably be the engineering solution of choice to achieve visually comfortable VR/AR experiences and plausible focus cues.

- Monovision does not require eye tracking and is much simpler to engineer than gaze-contingent focus – simply use two different lenses for each eye. It has been shown to improve visual clarity and efficiency for some tasks, but this mode requires more long-term studies to find the right applications in VR/AR. An immediate application would be to support presbyopic viewers without having them wear their glasses or contact lenses, which is already a huge benefit for many people. We believe there are other applications for monovision as well, but this needs further investigation. One conclusive statement we can already make is that monovision does not seem to be doing worse than the conventional mode, which is a good starting point for future research.

What is the ultimate VR/AR display? The "ultimate" goal for any VR/AR display is to emit the full 4D light field corresponding to a physical scene into the viewers eye [Wetzstein et al. 2012; Lanman and Luebke 2013; Huang et al. 2015]. The light field models the flow of light rays from the (virtual) scene through the viewers pupil and onto their retina. Light fields are closely related to holograms; people usually refer to holograms as wavefronts created by interference of coherent light whereas the light field is a more general term not associated with one particular technology or with coherence.

References

AKELEY, K., WATT, S., GIRSHICK, A., AND BANKS, M. 2004. A stereo display prototype with multiple focal distances. *ACM Trans. Graph. (SIGGRAPH) 23*, 3.

HOFFMAN, D., GIRSHICK, A., AKELEY, K., AND BANKS, M. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision 8*, 3.

HUANG, F., CHEN, K., AND WETZSTEIN, G. 2015. The light field stereoscope. *ACM Trans. Graph. (SIGGRAPH) 34*, 4.

KONRAD, R., COOPER, E., AND WETZSTEIN, G. 2016. Focus-tunable and Monovision Near-eye Displays. In *ACM SIGCHI*.

LANMAN, D., AND LUEBKE, D. 2013. Near-eye light field displays. *ACM Trans. Graph. (SIGGRAPH Asia) 32*, 6.

MAUDERER, M., CONTE, S., NACENTA, M. A., AND VISHWANATH, D. 2014. Depth perception with gaze-contingent depth of field. *ACM SIGCHI*.

ROLLAND, J., KRUEGER, M., AND GOON, A. 2000. Multifocal planes head-mounted displays. *Applied Optics 39*, 19.

WETZSTEIN, G., LANMAN, D., HIRSCH, M., AND RASKAR, R. 2012. Tensor Displays. *ACM Trans. Graph. (SIGGRAPH) 31*, 4.