# Accommodation-invariant Computational Near-eye Displays

ROBERT KONRAD, Stanford University
NITISH PADMANABAN, Stanford University
KEENAN MOLNER, Stanford University
EMILY A. COOPER, Dartmouth College
GORDON WETZSTEIN, Stanford University

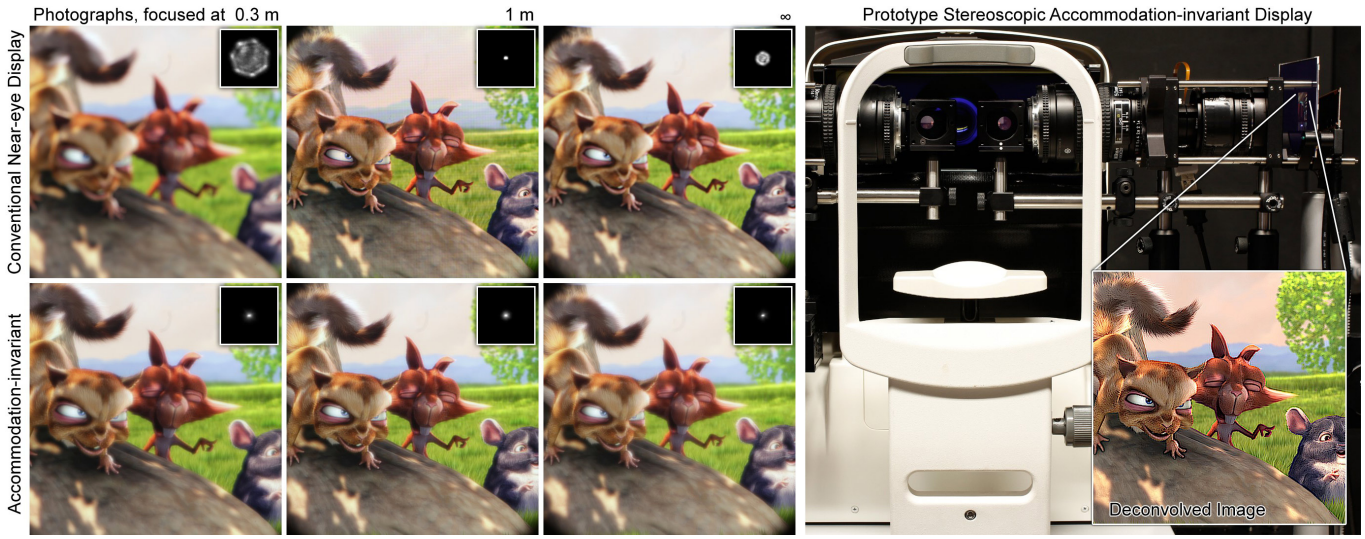


Fig. 1. Conventional near-eye displays present a user with images that are perceived to be in focus at only one distance (top left, 1 m). If the eye accommodates to a different distance, the image is blurred (top left, 0.3 m and ∞). A point spread function (PSF) illustrates the blur introduced for a single point of light at each distance (insets). The fact that conventional near-eye displays have a single sharp focus distance can be problematic, because it produces focus cues that are inconsistent with a natural 3D environment. We propose a computational display system that uses PSF engineering to create a visual stimulus that does not change with the eye's accommodation distance (bottom left). This accommodation-invariant display mode tailors depth-invariant PSFs to near-eye display applications, allowing the eye to accommodate to arbitrary distances without changes in image sharpness. To assess the proposed display mode, we build a benchtop prototype near-eye display that allows for stereoscopic image presentation (right). An autorefractor is integrated into the prototype to validate the accommodation-invariant display principle with human subjects.

Although emerging virtual and augmented reality (VR/AR) systems can produce highly immersive experiences, they can also cause visual discomfort, eyestrain, and nausea. One of the sources of these symptoms is a mismatch between vergence and focus cues. In current VR/AR near-eye displays, a stereoscopic image pair drives the vergence state of the human visual system to arbitrary distances, but the accommodation, or focus, state of the eyes is optically driven towards a fixed distance. In this work, we introduce a new display technology, dubbed accommodation-invariant (AI) near-eye displays, to improve the consistency of depth cues in near-eye displays. Rather than producing correct focus cues, AI displays are optically engineered to produce visual stimuli that are invariant to the accommodation state of the eye. The accommodation system can then be driven by stereoscopic cues, and the mismatch between vergence and accommodation state of the eyes is significantly reduced. We validate the principle of operation of AI displays using a prototype display that allows for the accommodation state of users to be measured while they view visual stimuli using multiple different display modes.



CCS Concepts: • **Computing methodologies** → **Perception**; **Virtual reality**; *Mixed / augmented reality*; Image processing;

Additional Key Words and Phrases: vergence–accommodation conflict, computational displays

**ACM Reference format:**
Robert Konrad, Nitish Padmanaban, Keenan Molner, Emily A. Cooper, and Gordon Wetzstein. 2017. Accommodation-invariant Computational Near-eye Displays. *ACM Trans. Graph.* 36, 4, Article 88 (July 2017), 12 pages.
https://doi.org/10.1145/3072959.3073594

## 1 INTRODUCTION AND MOTIVATION

Emerging virtual and augmented reality (VR/AR) systems offer unprecedented user experiences. Applications of these systems include entertainment, education, collaborative work, training, telesurgery, and basic vision research. In all of these applications, a near-eye display is the primary interface between the user and the digital

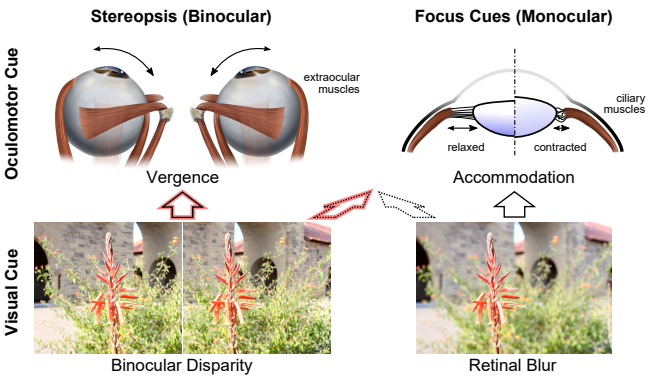


Fig. 2. Overview of relevant depth cues. Vergence and accommodation are oculomotor cues whereas binocular disparity and retinal blur are visual cues. In normal viewing conditions, disparity drives vergence and blur drives accommodation. However, these cues are cross-coupled, so there are conditions under which blur-driven vergence or disparity-driven accommodation occur. Accommodation-invariant displays use display point spread function engineering to facilitate disparity-driven accommodation. This is illustrated by the red arrows.

world. However, no commercially-available near-eye display supports natural focus cues. *Focus cues* refer to both the pattern of blur cast on the retina and the accommodative response of the eyes (see Fig. 2). In a natural 3D environment, as the eyes look around, the accommodative system adjusts so that the point being fixated is in focus. Objects closer or farther than the current accommodative distance are blurred. These cues are important for depth perception [Cutting and Vishton 1995; Hoffman et al. 2008] and a lack of them usually results in conflicting cues from the vergence and accommodation systems. Symptoms of this vergence–accommodation conflict (VAC) include double vision (diplopia), reduced visual clarity, visual discomfort, and fatigue [Kooi and Toet 2004; Lambooij et al. 2009; Shibata et al. 2011].

The challenge of providing natural focus cues in VR/AR is a difficult one. Substantial engineering efforts have been invested into developing displays that can generate focus cues similar to the natural environment. Generally, these approaches can be divided into several categories: dynamic focus, volumetric or multi-plane, light field, and holographic displays (see Sec. 2 for details). The development of each of these technologies poses different challenges, which at present have prevented them from being adopted in practice. For example, dynamic focus displays require eye tracking, multi-plane volumetric displays require extremely high display refresh rates, and light field displays currently offer a limited image resolution. We propose a new computational optical approach that does not attempt to render natural focus cues, but that creates visual stimuli that have the potential to mitigate symptoms of the vergence–accommodation conflict. By circumventing the goal of *natural* focus cues, accommodation-invariant (AI) displays open up a new set of tools for solving focus-related problems in VR/AR, including the often overlooked issue of user refractive errors such as near- and far-sightedness.

Conceptually, the idea of accommodation invariance can be illustrated by imagining that a user views a display through pinholes—the depth of focus becomes effectively infinite and the eyes see a sharp image no matter where they accommodate. Such a Maxwellian-view display [Westheimer 1966] would severely reduced light throughput and prevent the user from seeing an image at all when moving their pupil by more than half the pupil diameter (i.e., the eyebox corresponds to the size of the pupil). AI displays provide a large eyebox and uncompromised light throughput; we design and implement strategies to maximize perceived image resolution. We then ask whether it is possible to drive human accommodation with disparity instead of blur, using engineered point spread functions in a near-eye display system.

Our primary contributions include

- We introduce accommodation-invariant computational near-eye displays to drive human accommodation in an open loop condition.
- We analyze resolution tradeoffs and introduce a multi-plane AI display mode to optimize image resolution.
- We build a prototype near-eye display using focus tunable optics; we demonstrate a variety of example scenes with the prototype to assess the proposed display modes.
- In a user study, we use an autorefractor to quantify accommodative responses in three different AI display modes, as well as conventional and dynamic focus display modes.
- In a second user study, we take first steps towards assessing visual comfort for AI versus conventional displays.

## 2 RELATED WORK

### 2.1 Computational Displays with Focus Cues

*Two-dimensional dynamic focus displays* present a single image plane to the observer, the focus distance of which can be dynamically adjusted. Two approaches for focus adjustment have been proposed: physically actuating the screen [Padmanaban et al. 2017; Sugihara and Miyasato 1998] or dynamically adjusting the focal length of the lens via focus-tunable optics (programmable liquid lenses) [Johnson et al. 2016; Konrad et al. 2016; Liu et al. 2008; Padmanaban et al. 2017]. Several such systems have been incorporated into the form factor of a near-eye display [Konrad et al. 2016; Liu et al. 2008; Padmanaban et al. 2017]. However, for robust operation, dynamic focus displays require gaze tracking such that the focus distance can be adjusted in real time to match the vergence distance. Gaze or vergence tracking are not supported by commercially-available near-eye displays. AI displays, on the other hand, do not require eye tracking. In addition, although our benchtop prototype uses focus-tunable lenses, our accommodation-invariant optical system can also be implemented with custom optics that do not require dynamic focus adjustments or mechanical actuation.

*Three-dimensional volumetric and multi-plane displays* represent the most common approach to focus-supporting displays. Volumetric displays optically scan out the 3D space of possible light emitting voxels in front of each eye [Schowengerdt and Seibel 2006]. Multi-plane displays approximate this volume using a few virtual planes that are generated by beam splitters [Akeley et al. 2004; Dolgoff 1997] or time-mulitplexed focus-tunable optics [Hu and Hua 2014;

Liu et al. 2008; Llull et al. 2015; Love et al. 2009; Narain et al. 2015; Rolland et al. 2000; von Waldkirch et al. 2004]. Implementations with beam splitters seem impractical for wearable displays because they compromise the device form factor. The biggest challenge with time-multiplexed multi-plane displays is that they require high-speed displays and introduce perceived flicker. Specifically, an $N$-plane display requires a refresh rate of $N\times$ 60–120 Hz. No microdisplay used for commercial near-eye displays today offers suitable refresh rates in color for more than one plane. One of the proposed AI display modes also uses a selective plane approximation to the continuous AI display mode; this method does not require high display refresh rates because each plane shows the same content.

*Four-dimensional light field and holographic displays* aim to synthesize the full 4D light field in front of each eye [Wetzstein et al. 2012]. Conceptually, this approach allows for parallax over the entire eyebox to be accurately reproduced, including monocular occlusions, specular highlights, and other effects that cannot be reproduced by volumetric displays. However, current-generation near-eye light field displays provide limited resolution [Hua and Javidi 2014; Huang et al. 2015; Lanman and Luebke 2013]. Holographic displays may suffer from speckle and have extreme requirements on pixel sizes that are not afforded by near-eye displays also providing a large field of view.

### 2.2 Disparity-driven Accommodation

In natural vision, the accommodative distance of the eyes is thought to be largely driven by retinal blur. Specifically, the eyes act similarly to the autofocus in a camera: the accommodative state is altered until the fixated object appears sharp [Campbell and Westheimer 1960; Fincham 1951; Toates 1972]. However, the accommodative response is also directly coupled to the vergence response, resulting in disparity-driven accommodation that is independent of retinal blur [Fincham and Walton 1957; Schor 1992]. The properties of disparity-driven accommodation (or "vergence accommodation") have been characterized by removing the natural feedback to the accommodative system: placing pinholes in front of the eyes or otherwise altering the visual stimulus so that retinal blur no longer changes noticeably with accommodation [Westheimer 1966]. With exit pupil diameters of 0.5 mm or smaller, the human accommodation system is open looped [Ripps et al. 1962; Ward and Charman 1987]. Under these conditions, it has been shown that the accommodative distance of the eyes will naturally follow the vergence distance [Fincham and Walton 1957; Sweeney et al. 2014; Tsuetaki and Schor 1987].

A near-eye display system that removes the accommodation-dependent change in retinal blur, also known as Maxwellian-view display [Kramida 2015; Westheimer 1966], might allow accommodation to remain coupled to the vergence distance of the eyes, and thus allow for accommodating freely in a scene and mitigating the vergence–accommodation conflict. Unfortunately, pinholes are not useful for practical near-eye display design because they severely reduce light throughput, they can create diffraction-blur of the observed image, and they restrict the eyebox to the diameter of the pupil. The proposed AI display technology uses point spread function engineering and real-time deconvolution to provide high light throughput and a large eyebox for practical accommodation-invariant image display.

### 2.3 Extended Depth of Field

The technique we use to create AI displays is related to extended depth of field (EDOF) cameras. As an alternative to pinhole cameras, EDOF was developed to provide similar depth of field benefits while optimizing light throughput [Dowski and Cathey 1995]. Although Dowski and Cathey's design used cubic phase plates to engineer a depth-invariant point spread function, alternative optical implementations have been proposed, including focal sweeps via sensor or object motion [Häusler 1972; Nagahara et al. 2008] or focus-tunable optics [Miau et al. 2013], multi-focal lenses [Levin et al. 2009], diffusers [Cossairt et al. 2010], chromatic aberrations in the lens [Cossairt and Nayar 2010], and axicons [Zhai et al. 2009].

EDOF displays have also been proposed to extend the focal range of projectors. For example, Grosse et al. [2010] used adaptive coded apertures in combination with image deconvolution to achieve an EDOF effect whereas Iwai et al. [2015] employed focus-tunable optics instead to maximize the light throughput. Von Waldkirch et al. [2005] simulated the depth of field of focus-tunable lens-based retinal projectors with partially and fully coherent light. In general, EDOF cameras differ from EDOF displays in that processing is done after image capture, which allows for larger degrees of freedom and natural image priors to be used for image recovery. The primary limitation of an EDOF display is usually its dynamic range: image contrast may be degraded for pre-processed, projected imagery. Whether applied to cameras or displays, the EDOF principle imposes a fundamental trade-off between the increase in depth of field and image quality. This trade-off applies to our AI displays as well (See Section 3).

AI displays are a new family of computational displays tailored for near-eye display applications. Although the continuous focal sweep created by our display is closely related to Iwai's work, the newly proposed multi-plane AI display mode leverages characteristics of human vision that are unique for near-eye displays. With this paper, we propose a display technology that reaches well beyond what has been discussed in the computational imaging and display communities. Compared to existing volumetric and light field displays, AI displays may provide a practical technology that can be implemented with readily-available components while offering acceptable image resolution, a wide field of view, and a large eyebox.

## 3 MODELING AI DISPLAY SYSTEMS

In this section, we outline the image formation in conventional and accommodation-invariant near-eye displays. We also discuss efficient implementations of the required image deconvolution.

### 3.1 Near-eye Displays with Focus-tunable Lenses

The optical design of most near-eye displays is surprisingly simple. As illustrated in Figure 3, a microdisplay is located behind a magnifying lens. The distance between lens and physical display $d'$ is usually slightly smaller than the focal length of the lens $f$, such that a magnified virtual image is optically created at some larger

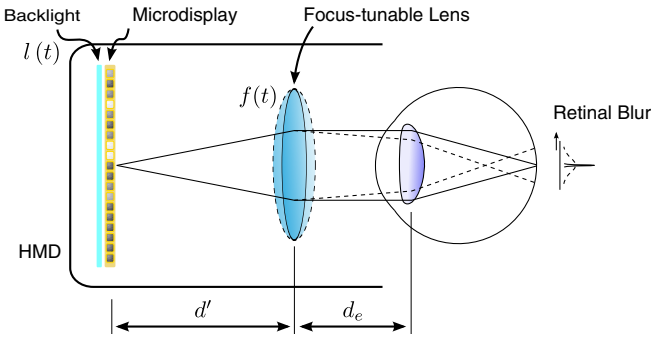


Fig. 3. Near-eye displays place a microdisplay behind a magnifying lens. Using focus-tunable optics, the focal length of the lens can be controlled at a speed that is faster than that of the human accommodation system. This allows for the perceived retinal blur to be controlled, for example to make it accommodation-invariant.

distance $d$ (not shown in figure). Both the magnification $M$ and $d$ can be derived from the Gaussian thin lens formula as

$$\frac{1}{d'} + \frac{1}{d} = \frac{1}{f} \quad \Rightarrow \quad d = \frac{1}{\frac{1}{f} - \frac{1}{d'}}, \quad M = \frac{d}{d'} = \frac{f}{d' - f} \tag{1}$$

This basic image formation model is applicable to most near-eye displays. When focus-tunable lenses are employed, the focal length of the lens $f$ is programmable, so we write the distance to the virtual image as a function of the focal length $d(f)$.

The perceived retinal blur diameter for an observer who is accommodated at some distance $d_a$ is then

$$b(f) = \underbrace{\frac{f_e}{d_a - f_e}}_{M_e} \cdot \zeta \cdot \frac{|d_e + d(f) - d_a|}{d_e + d(f)}, \tag{2}$$

where $\zeta$ is the pupil diameter, $f_e$ is the focal length of the eye, $M_e$ is the magnification of the eye, and $d_e$ is the eye relief (see Fig. 3).

The blur gradient with respect to depth can drive the accommodation state of a viewer with normal vision towards $d(f)$. Note that any software-only approach to changing the rendered image in the display (e.g., gaze-contingent retinal blur) may be able to affect the blur in a perceived image, but not the retinal blur gradient $\partial b / \partial d_a$, which is actually driving accommodation. Only a change in either $f$ or $d'$ affects the blur gradient, which is achieved using focus-tunable optics (varying $f$) or actuated displays (varying $d'$).

Although Equation 2 is a convenient mathematical tool to predict the blur diameter of a focus-tunable near-eye display, in practice one rarely observes a perfectly disk-shaped blur. Optical aberrations, diffraction, and other effects degrade the intensity distribution within the blur circle. Following [Nagahara et al. 2008], this can be modeled by approximating the blur disk by a Gaussian point spread function (PSF)

$$\rho(r, f) = \frac{2}{\pi (c \cdot b(f))^2} \, e^{-\frac{2r^2}{(c \cdot b(f))^2}} \tag{3}$$

where $r = \sqrt{x^2 + y^2}$ is the lateral distance from the blur center and $c$ is a constant.

## 3.2 Accommodation-invariance via Focal Sweep

One convenient way to create a depth-invariant PSF is a focal sweep. These sweeps are easily created with focus-tunable lenses by periodically changing the focal length $f$ of the lens. For near-eye displays, one sweep period would have to be an integer multiple of the display refresh rate (usually 60 Hz). To prevent possible artifacts, the sweeping time should also be faster than the reaction time of the human accommodation system. Since the latter is in the order of hundreds of milliseconds [Heron et al. 2001], this is easily achieved with current-generation tunable lenses.

A focus sweep creates a temporally-varying PSF that the observer perceptually integrates due to the finite "exposure time" $T$ of the visual system. The perceived, integrated PSF $\widetilde{\rho}$ is then given as

$$\widetilde{\rho}(r) = \int_0^T \rho(r, f(t)) \, dt, \tag{4}$$

where $f(t)$ maps time to temporally-varying focal length. Oftentimes, $\frac{1}{f(t)}$, the focal length in dioptric space, is a periodic triangle-like function [Iwai et al. 2015; Miau et al. 2013; Nagahara et al. 2008], ensuring that the blur diameter varies linearly in time.

In practice, the integrated PSF of a depth-invariant near-eye display is calibrated in a pre-processing step and then used to deconvolve each color channel of a target image $i$ individually via inverse filtering as

$$i_c(x, y) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{i(x, y)\}}{\mathcal{F}\{\widetilde{\rho}(x, y)\}} \right\}. \tag{5}$$

Here, $i_c$ is the compensation image that needs to be displayed on the screen such that the user perceives the target image $i$ and $\mathcal{F}\{\cdot\}$ is the discrete Fourier transform. Note that depth-invariant displays are different from depth-invariant cameras in that one does not have to deal with noise, a challenge for all deconvolution algorithms. Therefore, a simple deconvolution technique such as inverse filtering achieves near-optimal results. However, the display has a limited dynamic range, which should theoretically be taken into consideration for the deconvolution problem by integrating the blacklevel and maximum brightness as hard constraints. We show in Section 4 that the difference between inverse filtering and constrained optimization-based deconvolution for the PSFs measured with our prototype are negligible.

*3.2.1 Bounds on Image Resolution.* The closed-form solution of the integral in Equation 4 depends on the specific range of the focal sweep and the sweep function $f(t)$. It is obvious, however, that the integrated PSF has a larger variance than the smallest PSF of a conventional near-eye display. Hence, AI displays impose a fundamental tradeoff between accommodation-invariant range and image resolution. This tradeoff is also observed in all photographed results (e.g., Figs. 1, 8, and supplemental figures). We discuss this tradeoff in more detail in Section 6.

## 3.3 Optimizing Resolution with Multi-plane Invariance

Although focal sweeps create accommodation-invariant PSFs, the inevitable loss of image resolution also degrades the viewing experience compared to the sharpest image produced by a conventional near-eye display. To optimize image resolution while preserving the AI display property, we propose a simple approach that is unique to

near-eye displays implemented with liquid crystal displays (LCDs), liquid crystal on silicon (LCoS), or digital micromirror (DMD) displays. All of these technologies comprise a combination of uniform backlight, most commonly light emitting diodes (LEDs), and the actual spatial light modulator (SLM). While the SLM is often limited in its refresh rate (i.e., LCDs and LCoS usually run up to 240 Hz), the LEDs in the backlight can easily be modulated at rates that are orders of magnitude higher than that. Thus, a clever combination of time-modulated backlight intensity $l(t)$ and displayed image may be a viable approach to optimizing image resolution for accommodation-invariant near-eye displays.

Consider the point spread function model of Equation 4 with the additional option of modulating the backlight $l(t)$

$$\widetilde{\rho}(r) = \int_0^T \rho(r, f(t))\, l(t)\, dt. \tag{6}$$

Assuming that the focus sweep is linear in dioptric space and that a single sweep is completed in time $T$, we can strobe the backlight at $N$ equidistant locations throughout the sweep. The resulting PSF is expressed as

$$\begin{aligned}\widetilde{\rho}(r) &= \int_0^T \rho(r, f(t)) \sum_{k=1}^{N} \delta\left(t - k\frac{T}{N+1}\right) dt \\ &= \sum_{k=1}^{N} \rho\left(r, f\left(k\frac{T}{N+1}\right)\right).\end{aligned} \tag{7}$$

We see that the continuous focal sweep becomes a sum of individual PSFs, each focused at a different distance. A related concept is often used for multi-plane volumetric displays (see Sec. 2), but in that context the SLM is updated together with the backlight such that each depth plane shows a *different* image. Here, we propose to keep the SLM image fixed throughout a sweep but strobe the backlight to create the *same* image at several discrete planes in depth.

Note that human accommodation is a rather imprecise mechanism. The depth of field of the human eye is approximately ±0.3 D, although it varies depending on the properties of the stimulus [Campbell 1957; Marcos et al. 1999]. Multi-plane volumetric displays have been found to drive accommodation naturally with an inter-plane spacing of up to 1 D [MacKenzie et al. 2010]. Thus, our hypothesis for multi-plane AI displays is that a few focus planes may suffice to drive accommodation to the plane closest to the vergence distance. People are unlikely to accommodate in between planes because the retinal blur will drive focus to one of the planes, where the sharpest image is observed. Using this mechanism, a multi-plane implementation of accommodation invariance has the potential to optimize image resolution for the proposed technology. In the following sections, we evaluate point spread functions and human accommodative responses to both continuous and multi-plane implementations.

## 4 IMPLEMENTATION AND ASSESSMENT

### 4.1 Hardware

To evaluate the proposed AI near-eye display modes, we built a benchtop prototype (see Figure 4). The prototype uses two Topfoison TF60010A liquid crystal displays (LCDs), each with a resolution of

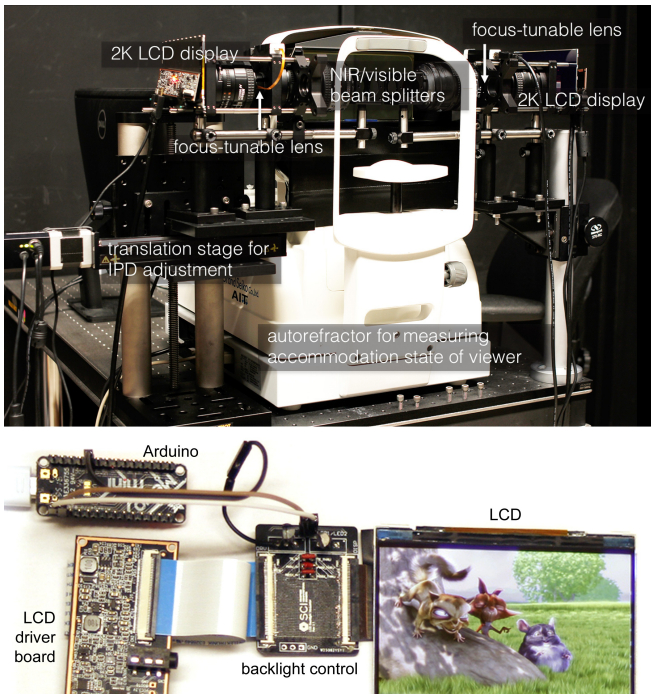


Fig. 4. Photograph of prototype display. Top: a stereoscopic near eye display was table-mounted to include an autorefractor that records the user's accommodative response to a presented visual stimulus. Each arm comprises a high-resolution liquid crystal display (LCD), a series of focusing lenses, a focus-tunable lens, and a NIR/visible beam splitter that allows the optical path to be shared with the autorefractor. The interpupillary distance is adjustable by a translation stage. Bottom: a custom printed circuit board intercepts the signals between driver board of an LCD to synchronize the focus-tunable lens with the strobed backlight via a microcontroller.

2560 × 1440 pixels and a screen diagonal of 5.98". The optical system for each eye comprises three Nikon Nikkor 50 mm f/1.4 camera lenses, and a focus-tunable liquid lens. These lenses provide high image quality with few aberrations. The lens closest to the screen is mounted at a distance of 50 mm to create a virtual image at optical infinity. The focus-tunable lens is an Optotune EL-10-30-C with 10 mm diameter and a focal range of 0–5 diopters (D). Without current applied, the focus-tunable lens places the virtual image at 5 D (0.2 m), but with increasing current the curvature of the liquid lens is increased, thereby placing the virtual image at a farther distance from the observer. To create an accommodation-invariant PSF we sweep the lens' focal power in a triangle wave at 60 Hz over the full range. The other two camera lenses provide a 1:1 optical relay system that increases the eye relief to about 4–5 cm. This eye relief also provides space for a near-infrared(NIR)/visible beam splitter (Thorlabs BSW20R) in front of the eyes, which is needed for the autorefractor.

The eyebox provided by this display is 10 mm in diameter, but the integrated PSFs generated for the AI display mode are slightly view-dependent. The *useable* eyebox is therefore restricted to about 5 mm. The resolution provided to each eye is 620 × 620 pixels and

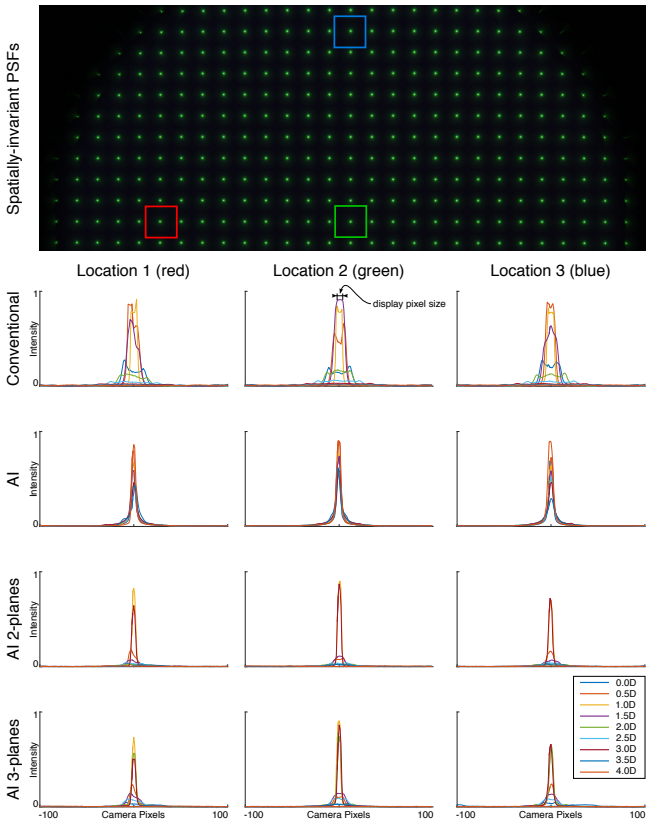


Fig. 5. Captured point spread functions of the green display channel. The plots show one-dimensional slices of captured PSFs at several different locations (top) and depths for conventional, accommodation-invariant (AI) continuous, AI 2-plane, and AI 3-plane display modes. Whereas the conventional PSFs quickly blur out away from the focal plane at 1 D (1 m), the shape of the accommodation-invariant PSFs remains almost constant throughout the entire range. Multi-plane AI PSFs are focused at the respective planes, but not in between.

the monocular field of view is approximately 35° both horizontally and vertically. The mechanical spacing between the lenses, i.e. the interpupillary distance, is adjustable by a translation stage.

A Grand Seiko WAM-5500 autorefractor is integrated into the near-eye display system. The autorefractor uses built-in NIR illumination and a NIR camera to determine the user's accommodative state. The illumination pattern is close to invisible to the user. Accommodation measures are directly transmitted to the computer that controls the visual stimulus. The accuracy of the autorefractor is verified using a Heine Ophthalmoscope Trainer model eye (C-000.33.010).

The LCD backlight is controlled with a custom circuit placed between the display driver board and the LCD panel (Fig. 4 bottom). This circuit allows for 8 MIPI lines, power, and control signals to connect directly between the driver board and the display. The signal to the anode of the LED backlight string is also connected directly from the driver board to the display, but the cathode, coming from the display, is interrupted with a low side NMOS transistor. The source of the transistor is then connected back to the driver board and the gate is connected to an Arduino Uno. The Arduino thus controls the backlight directly and receives an input signal from the Optotune lens driver, which allows for precise synchronization between focal power and backlight illumination.

## 4.2 Assessing Depth Invariance and Spatial Resolution

To confirm invariance of the PSFs created by our prototype as a function of both accommodative distance and lateral displacement on the display, we measured the size and shape of the PSFs in various display modes. Calibration tests were run in four different display modes. First, we tested both the conventional (focal plane at 1 D) and continuous-sweep AI display modes. We also tested two multi-plane AI modes: a 2-plane mode with planes located at 1 and 3 D, and a 3-plane mode with planes located at 1, 2, and 3 D. The data were captured with a Canon Rebel T5 SLR camera and a Canon EF-S 18–55 mm zoom lens set to 35 mm and f/11. In Figure 5, the top panel shows an example captured calibration photograph (a grid of illuminated pixels tiling the display), taken in the accommodation-invariant mode. Additional calibration photographs of both conventional and AI display modes are shown in the Supplemental Information). The line plots below summarize the overall results for different display modes, focus distances, and several pixel locations. Each panel shows results for one display mode (rows) and pixel location (columns). The colored lines show one-dimensional slices through the center of the PSF when the camera was focused a 9 different distances (see legend).

The PSFs of the conventional display (Fig. 5, top row) are narrowest when the camera is focused at the display image distance (1 D, yellow line) and quickly blur out as the camera focal distances moves closer or farther. Their non-uniform shape is due to a combination of non-uniform display pixel shape, non-circular shape of the camera aperture, and imperfections of the optical system. As predicted, the PSFs created by the continuous AI mode are nearly invariant to lateral location on the display and also to accommodation distance (second row). The point spread functions of the multi-plane AI modes are sharpest at the selected planes but they are not constrained in between these planes, and thus blur out (third and fourth rows). Remaining amplitude differences in the depth-invariant PSFs are due to minute imperfections in the waveforms driving the focus-tunable lenses. Note that the plots are shown in camera pixel coordinates—the display pixel size is indicated and provides the highest possible resolution in all cases. All plots are scaled to the same, relative intensity.

To better assess the spatial resolution limits of the AI display we investigated the modulation transfer function (MTF) of the conventional mode (focal plane at 1 D), continuous AI mode, AI 2-plane mode (focal planes at 1 D, 3 D), and AI 3-plane mode (focal planes at 1 D, 2 D, 3 D). Figure 6 shows the MTFs of these modes captured with a camera focused to 1 D, 2 D, and 2.5 D (computed using the slanted edge algorithm based on the ISO 12233 standard). As expected, the continuous AI mode shows a consistent, if reduced, response across the focusing states while the conventional mode is sharp at only the 1 D plane. The AI 2-plane and 3-plane modes provide increased sharpness at discrete planes when compared to the continuous AI mode (seen in the 1 D focusing setting for 2- and 3-plane, and the

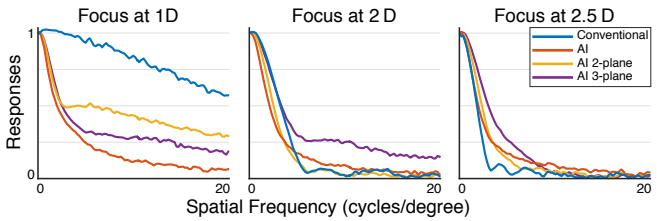


Fig. 6. Modulation transfer function (MTF) measurements. The MTF of our prototype demonstrates that the continuous AI mode has a relatively consistent transfer function across different focus settings. The sharpness can be improved at discrete planes with the AI 2- and 3-plane modes as seen when the camera is focused to 1 D and 2 D (left and center panel). All AI modes outperform the conventional mode as the distance between the focusing plane and conventional focal plane (1 D) increases.

2 D focus setting for 3-plane), but show a loss in resolution in between the planes (2.5 D focus setting). The AI-display modes trade off sharpness at one focus distance for a more consistent blur across a range of focus distances. This trade off should remove the blur gradients that drive accommodation and allow accommodation to follow vergence.

### 4.3 Software Implementation

All software driving the prototype is implemented in C++. The OpenGL application programming interface is used for 3D rendering and image deconvolution is implemented via inverse filtering in CUDA. For each eye, the deconvolution takes about 5 ms. The total latency for stereo rendering and deconvolution is below 15 ms for the simple scenes used in our examples.

We compare two different deconvolution methods in Figure 7. Inverse filtering (Eq. 5) is the most intuitive approach to deconvolution, but it does not account for constraints imposed by the dynamic range of the physical display. Hence, we compare the results provided by inverse filtering with those generated by the trust region reflective constrained optimization method implemented in the Matlab *lsqlin* function. Although the peak signal-to-noise ratio (PSNR) of the constrained optimization approach is about 5 dB better, the qualitative difference on the prototype (Fig. 7, bottom) is marginal. Faint halos around high-contrast edges are sometimes observed, as seen in the bird's eye and beak. Therefore, we argue that an inverse filter may be appropriate for practical image display and it can be easily implemented on the GPU to provide real-time frame rates.

### 4.4 Results

Figures 1, 7, 8 and S.3–S.7 show results for computer-generated scenes photographed from our prototype display, using the same camera that was used in the calibration measures. The images were captured with an aperture diameter of 3.18 mm, which is comparable to the average human pupil diameter under the given illumination conditions [de Groot and Gebhard 1952].

Figure 1 compares the observed optical blur and corresponding point spread functions (insets) for three different accommodation distances: 0.3 m, 1 m, and optical infinity. As expected, the blur from the conventional display quickly increases away from the

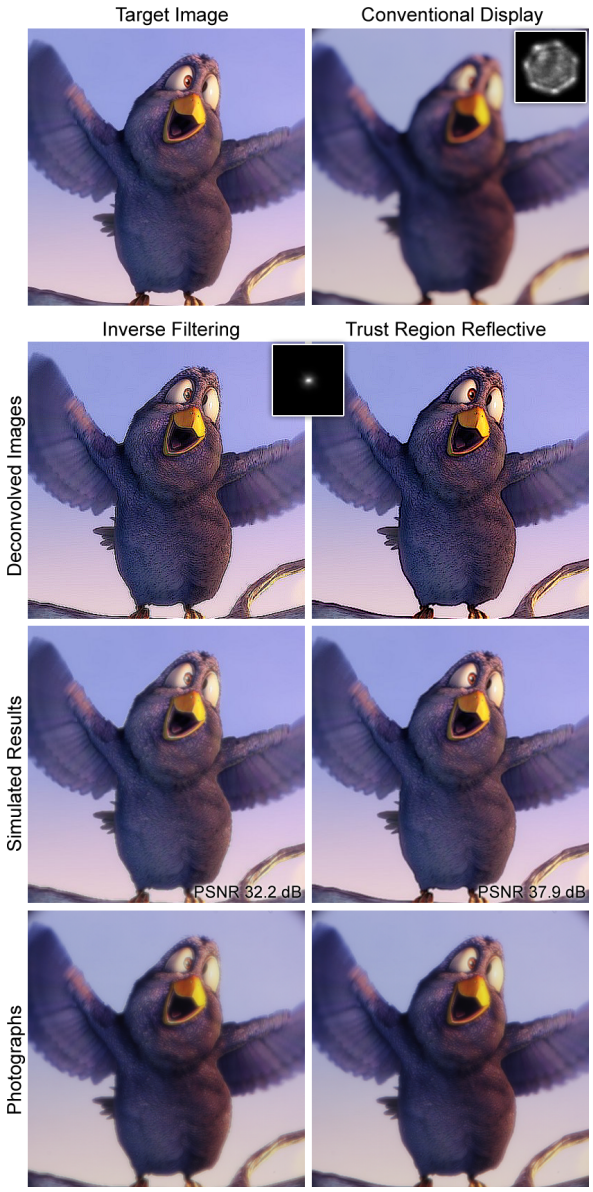


Fig. 7. Comparing deconvolution methods. A target image (top left), creates a sharply-focused image only at a single plane but the perceived blur when accommodated at other distances is severe (top right, focused at 25 cm). Accommodation-invariant displays provide a depth-invariant PSF (center inset) but require the target image to be deconvolved prior to display. We compare two deconvolution methods: inverse filtering and constrained optimization (row 2). The latter provides a baseline for the best possible results, whereas inverse filtering creates near-optimal results in real-time. Photographs of the prototype (row 4) match simulations (row 3).

focal plane at 1 m. The AI display provides a nearly depth-invariant PSF—i.e., it provides a close approximation to the target image with a constant blur gradient. Figure 8 shows results for another scene in the same format (images and PSFs) at six distances ranging from 0–3 D. All display modes are shown, including multi-plane.

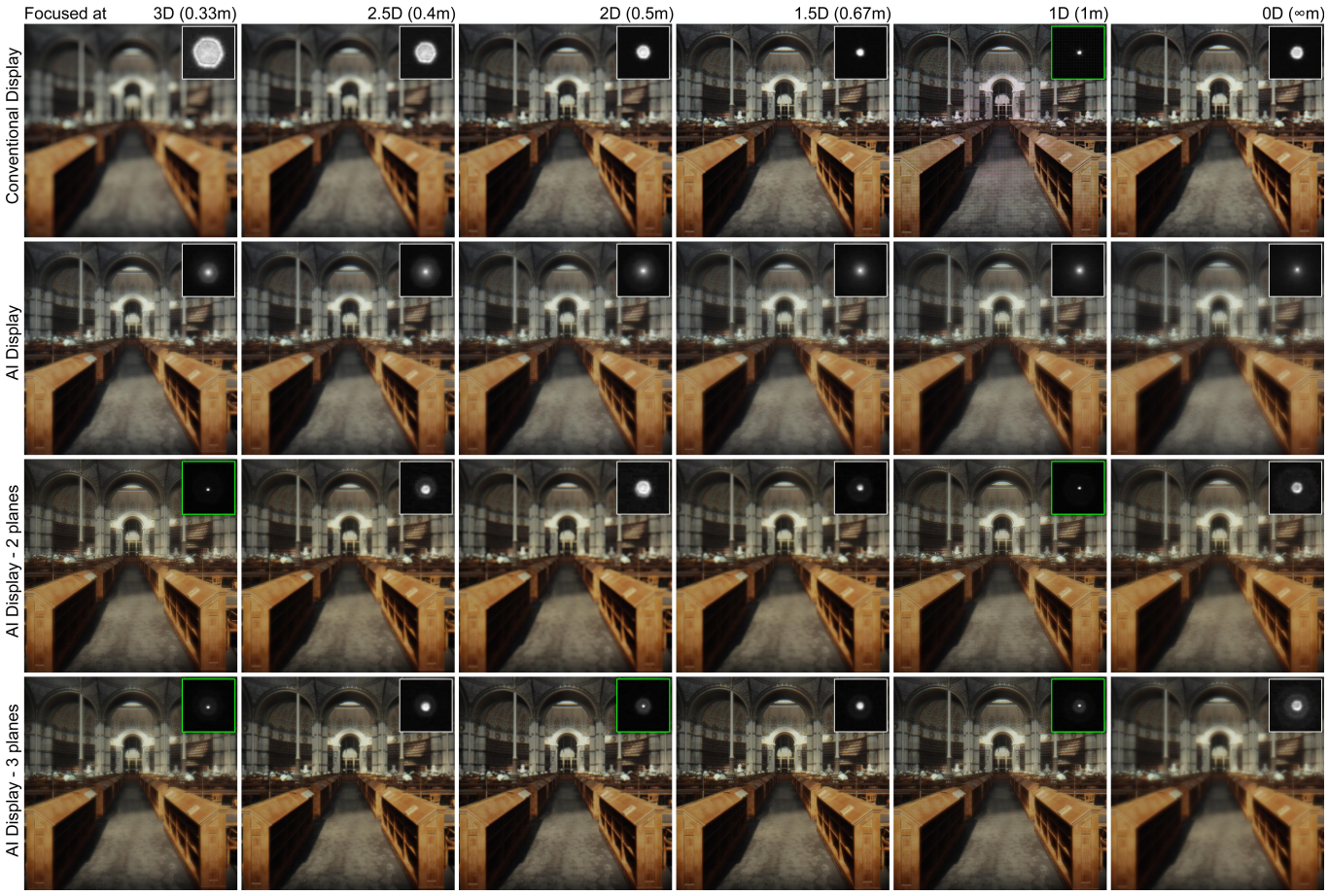


Fig. 8. Photographic results of several display modes for six different focal settings. The conventional mode (top row) provides a sharp image at a single depth plane. Accommodation-invariant displays with a continuous focal sweep equalize the PSF over the entire depth range (second row), but the full image resolution cannot be restored even with deconvolution. Multi-plane AI displays optimize image resolution for a select number of depths, here shown for two (third row) and three (fourth row) planes.

Again, we observe that the conventional mode is best-focused at one depth, here at 1 D or 1 m, but quickly blurs out at increasing distances from that plane. The continuous AI mode provides an image sharpness and PSF shape that is approximately constant over the entire accommodation range. However, this invariance comes at the cost of reduced resolution compared to the sharpest plane of the conventional mode. As expected, the multi-plane AI modes provide a significantly increased resolution at the respective focal planes but image quality is degraded between planes. Figures S.3–S.7 show this same scene and four additional scenes photographed at nine focal distances ranging from 0–4 D.

## 5 EVALUATION

Several user studies were conducted to evaluate the AI display modes. All users had normal or corrected-to-normal vision and normal stereoacuity as assessed with a Randot test. All participants gave informed consent, and the study procedures were approved by the institutional review board at the home institution.

### 5.1 Accommodative Responses

To evaluate the human accommodative response to visual stimuli in the various AI display modes, we conducted two user studies. In each study, we objectively measured user accommodation in response to visual targets using the autorefractor. The goal of these studies was to determine whether the AI display modes can stimulate disparity-driven accommodation, allowing users to accommodate to different distances and mitigating the vergence–accommodation conflict. Twelve volunteers participated in both studies, while an additional 4 volunteers participated in only the first study. In total, 16 volunteers participated in the first study (age range 22–32, 1 female), but data from 5 users was discarded due to artifacts in autorefractor recording. Twelve volunteers participated in the second study (age range 23–34, 2 females).

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 and continuous AI. The

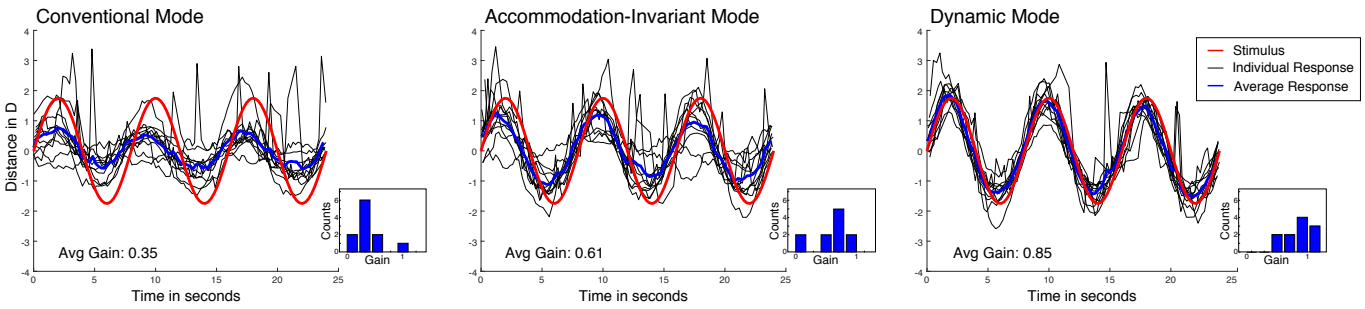


Fig. 9. Accommodative gain in the first user study. Each panel shows the individual (black lines) and average (blue lines) accommodative responses to the oscillating stimulus (red lines) for each display mode (conventional, AI, and dynamic focus). In the conventional mode, the virtual image distance was fixed at 1.3 m. Data are shown for 3 cycles after a 0.5 cycle buffer at the start of each trial. The ordinate indicates the accommodative and stimulus distance with the mean distance subtracted out. This is done to account for individual offsets in each user's accommodative measures. Inset histograms show the distribution of gains for each condition.

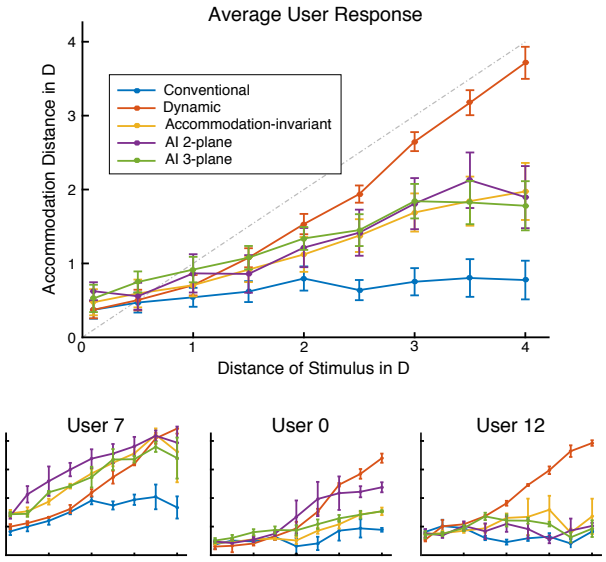


Fig. 10. Accommodative responses in the second user study. The upper panel shows the between-subjects mean accommodative response for each distance in each display mode (see legend). Target distance is on the abscissa and accommodative distance is on the ordinate. Error bars indicate the standard error of the mean. The three lower panels show the results in the same format for three example users.

third display mode used dynamic focus to effectively remove the VAC. For this purpose, the virtual distance of the target was updated dynamically to match the stereoscopic distance, providing accurate and natural focus cues. The dynamic mode is also known as varifocal display and has been demonstrated, via autorefractor measurements, to achieve natural accommodative responses. (See Liu et al. [2008] and Padmanaban et al. [2017] for more details on varifocal near-eye displays.)

We predicted that users would accommodate most accurately to targets in the dynamic focus mode, least accurately in the conventional mode, and that their responses to the AI mode would fall somewhere in between. In all modes, the target was a Maltese cross of size 6.2 cm that 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 with their eyes, and each user performed this task in each display mode once. The order of conditions was randomized per user. The gain of each user's accommodative response for each condition was calculated as the ratio of the amplitude of accommodation at the frequency of the stimulus to the amplitude of the stimulus itself. A gain of one would indicate that the user accommodated to the full range of the stimulus, and temporal lag was not taken into consideration. The stimulus was presented for 4.5 cycles, and responses were analyzed for the 3 cycles after a 0.5 cycle buffer.

The line plots in Figure 9 show the individual (black lines) and average (blue lines) accommodative responses to the stimulus (red lines) for each display mode, and histograms show the distribution of gains. In the conventional mode (left panel), although the virtual image distance was fixed, users still exhibited a small gain, with an average of 0.35. Consistent with our prediction, the AI mode (in which natural focus cues are removed) resulted in substantially increased accommodative gains over the conventional mode (middle panel, average gain of 0.61). In the dynamic mode (right panel), which provides near-accurate blur cues, users exhibited an even higher average gain of 0.85. To examine the statistical significance of these differences in gain, we conducted a one-way repeated measures analysis of variance (ANOVA). The ANOVA showed a significant main effect of display mode, $F(2,20) = 35.579$, $p < 0.001$. Follow up t-tests were performed to compare each pair of display modes, and the p-values were corrected for multiple comparisons using the Bonferroni method. These tests indicated that the gains in the AI condition were significantly higher than in the conventional condition ($p < 0.01$). Gains in the dynamic condition were significantly higher than those in both the conventional and AI conditions ($p < 0.001$ and $p < 0.05$, respectively). These results indicate that disparity driven accommodation via the removal of focus cues in a near-eye display can be achieved, although the resulting accommodative gain is not quite as high as with natural focus cues.

We conducted a second study to confirm and extend these results. In the second study, we compared accommodative responses to static targets, at different depths, in five different display modes. Three display modes were the same as described for the first study. The

two additional modes were the 2-plane AI and 3-plane AI modes described in the previous sections. On each trial in this study, a single Maltese cross target appeared statically in a scene and users were instructed to fixate on the target. The target appeared in a random order at one of 9 different distances: 0.1 D (10 m), 0.5 D (2 m), 1 D (1 m), 1.5 D (0.67 m), 2 D (0.5 m), 2.5 D (0.4 m), 3 D (0.33 m), 3.5 D (0.29 m), 4 D (0.25 m). After a minimum of 3 seconds, the accommodative response was recorded. Once recorded, the users were presented with a blank screen for 2 seconds, after which the target would reappear at a different depth. The order of the modes presented to the user was randomized, and within each mode, the target distances was presented in a random order. Each combination of display mode and target distance was repeated 3 times for each user and the responses were averaged.

The results for this second study are shown in Figure 10. The upper panel shows the mean and standard errors of the accommodative distances across users for the five conditions (see legend). The dashed line shows what the predicted results would be if the users always accommodated exactly to the stimulus distance. The results are as expected, and consistent with the first study. In all of the AI conditions (continuous and multi-plane), users accommodate 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 AI conditions. This is illustrated with data from three individual users in the lower panels of Figure 10. Some users responded to the AI conditions very well, as shown in the bottom left panel, while others seemed to exhibit very little response, as shown in the the bottom right panel. Other users fell somewhere in between (bottom middle panel). This variability may reflect individual differences in the strength of the cross-coupling between vergence and accommodation.

### 5.2 User Comfort

Next, we conducted a study to examine whether the stimulation of disparity-driven accommodation in the AI mode improves comfort for users over a conventional display. For this study, we tested 18 users (age range 19–32, 4 female). Each user participated in two sessions, each 20 minutes, separated by a 10 minute break. During each session, users watched one of two videos placed at a stereoscopic distance of 0 D (∞ m), in either the conventional mode or the continuous AI mode. Due to only a subtle observed difference in ratings between natural viewing and viewing with the VAC [Shibata et al. 2011], we maximize the VAC in the conventional mode by setting the focus distance of the display to 3 D (0.33 m). The order of the modes and videos were randomized such that half of the users saw AI first and half saw conventional first. Within these groups, 5 users saw Video 1 in the AI mode first, and 4 users saw Video 1 in the conventional mode first. After completing both sessions, the users were asked to compare the two sessions on the basis of fatigue, eye irritation, headache severity, and overall preference. Each criteria was rated separately on a 5-point scale (session 1 was: much better, better, no difference, worse, much worse). The average responses are shown in Figure 11. While the average response for each question slightly favored the AI mode, a Wilcoxon signed-rank test indicated that these were not significantly different from "no difference" ($ps > 0.5$).

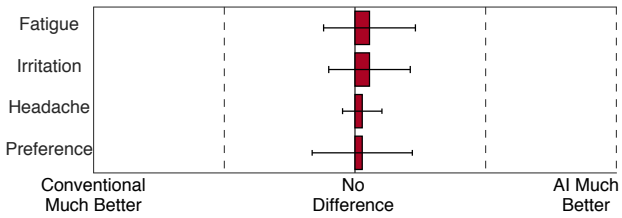


Fig. 11. User ratings comparing two sessions on the basis of fatigue, eye irritation, headache severity, and overall preference. Users viewed videos placed at a stereoscopic distance of infinity in either the conventional (focus distance at 3 D) or AI modes. Conventional and AI modes were not significantly different from each other on any measure. Error bars indicate standard error of the mean across subjects.

Previous work has used active depth judgment tasks and regularly changing depth intervals to induce the symptoms of the VAC, rather than passive viewing (e.g., [Shibata et al. 2011]). It is possible that the passive viewing task with a constant VAC interval (3 D) used in the current study was insufficient to induce substantial discomfort in the conventional condition. Because users were only asked to compare the two sessions, and not to rate the discomfort in each session individually, this possibility cannot be examined with the current dataset. It is interesting to note, however, that users tended to report lighter symptoms in the second session (regardless of which display mode was second), suggesting that they were not accumulating fatigue throughout the study. A follow-up study with longer sessions, more variable stereoscopic content, additional questionnaires, and perhaps a longer break between sessions will be necessary to further examine how AI display modes affect viewing comfort.

## 6 DISCUSSION

In summary, we introduce accommodation-invariant (AI) displays as a new computational optical mode for near-eye displays. Rather than providing natural focus cues to the user, AI displays optically render the perceived retinal blur invariant to the accommodation state of the eye. This approach renders the accommodative system into an open loop condition, which allows stereoscopic depth cues (disparity and vergence) to drive accommodation instead of retinal blur. The proposed display technology is evaluated photographically and its effect on the accommodation of human subjects is validated using refractive measurements. While our comfort study did not indicate that visual discomfort and fatigue were mitigated in our experiments, objective measurements of accommodative responses suggest that disparity-driven accommodation was stimulated. As such, future studies will be required to fully test the effect of the AI display modes on the symptoms associated with the vergence–accommodation conflict, by employing longer-term and more naturalistic viewing paradigms.

### 6.1 Volumetric Multi-plane Displays

Whereas existing multi-plane displays optically scan out a volume with different image content on each plane, multi-plane AI displays aim to present the same content on each plane. We demonstrate that the latter can be easily implemented with displays that have

LED backlights, such as LCDs and LCoS. Using custom optics, AI displays would also support light-emitting displays, such as OLEDs, which are commonly used for VR/AR applications. These display types provide refresh rates of up to 120-240 Hz, so none of them are actually suitable for volumetric multi-plane displays. Overcoming flicker, latency, and other artifacts of volumetric displays represent significant engineering challenges. AI displays on the other hand can be implemented with readily-available microdisplays.

### 6.2 Prescription Eyewear and Multifocal Lenses

A challenging problem for near-eye displays are eyeglasses. Many viewers suffer from near-sightedness (myopia) or far-sightedness (hyperopia or presbyopia) and the small form factor of near-eye displays sometimes makes it inconvenient to wear additional prescription glasses. AI displays provide unique benefits for viewers requiring prescription glasses. The depth-invariant PSF created by these display systems would allow for users with a range of refractive errors to see a clearer image without their eyeglasses than what they normally would see. For example, without having to change anything about the display to tailor it to the user, a myope could accommodate to a near distance and a presbyope could accommodate to a far distance, and both would see the same, relatively sharp, image. Studying this effect in more detail would be particularly interesting in the context of higher-order visual aberrations that are difficult to be corrected with eyeglasses.

Eyewear for presbyopic users is particularly relevant for the proposed technology. Presbyopes, i.e. most people over the age of 45, have a limited accommodation range. This condition is usually treated with either bifocals, monovision, or multifocal contact lenses. The latter option is also available for intraocular lenses after cataract surgery or for LASIK eye surgery. Multifocal lenses combine lenses of different focal powers in a single optical element. This allows a presbyope to see both near and far objects in focus simultaneously. Users sometimes report seeing halos around objects but usually get used to this type of corrective eyewear [Braga-Mele et al. 2014]. With the multi-plane version of AI displays, we have adopted an optical technique similar to multifocal contact lenses to near-eye displays. The ability to perform deconvolution in the display is a major benefit of our technique, which obviously cannot be accomplished in real environments.

### 6.3 Limitations

The primary limitation of AI displays is reduced image resolution compared to the sharpest plane in conventional displays. This effect is due to a fundamental tradeoff between accommodation range and spatial resolution. The larger the accommodation-invariant range of the system, the larger the loss of image resolution. With deconvolution applied to the displayed images and a newly-proposed multi-plane AI mode, we optimize perceived image resolution and introduce the option of tailoring the tradeoff to a particular application or user. It may be possible to engineer point spread functions that are more invertible than those analyzed in this paper, and thus overcome the resolution tradeoff. However, these may not be supported by focus-tunable lenses and require custom optics design.

### 6.4 Future Work

Focus-tunable lenses are useful for a benchtop prototype like ours, because they allow us to dynamically switch between different display modes for the same user. The benchtop setup also makes it convenient to integrate an autorefractor, without which it is not possible to objectively evaluate the proposed display system. However, future implementations of this system would miniaturize the display. To this end, we would like to build custom multifocal lenses that remove the need for strobing the backlight or using focus-tunable lenses. Such custom lenses could be easily integrated into any headset and such a system would also be able to operate with organic light emitting displays, which do not have a backlight at all. Alternative implementations, for example using diffusers [Cossairt et al. 2010] or axicons [Zhai et al. 2009] are also promising directions for future work. Amplitude-coded masks, such as printed transparencies, may also be viable, but the loss of light may degrade the user's experience.

## 7 CONCLUSION

With this work, we hope to stimulate new directions for computational optics design of next-generation near-eye displays. The lack of accommodation is one of the primary sources for visual discomfort in current-generation VR/AR displays. Whereas optically generating accurate focus cues is extremely challenging, it may be possible to overcome the vergence–accommodation conflict with accommodation-invariant displays. AI displays may also provide a robust near-eye display platform for people with refractive errors.

## ACKNOWLEDGMENTS

This project was generously supported by Intel, an Okawa Research Grant, Huawei, and Samsung. R.K. was supported by an NVIDIA Graduate Fellowship. N.P. was supported by a National Science Foundation (NSF) Graduate Research Fellowship. G.W. was supported by a Terman Faculty Fellowship and an NSF Faculty Early Career Development (CAREER) Award.

## REFERENCES

Kurt Akeley, Simon J. Watt, Ahna Reza Girshick, and Martin S. Banks. 2004. A stereo display prototype with multiple focal distances. *ACM Trans. Graph. (SIGGRAPH)* 23, 3 (2004), 804–813.

Rosa Braga-Mele, David Chang, Steven Dewey, Gary Foster, Bonnie An Henderson, Warren Hill, Richard Hoffman, Brian Little, Nick Mamalis, Thomas Oetting, Donald Serafano, Audrey Talley-Rostov, Abhay Vasavada, and Sonia Yoo. 2014. Multifocal intraocular lenses: Relative indications and contraindications for implantation. *J. Cataract Refract. Surg.* 40, 2 (2014), 313–322.

F. W. Campbell. 1957. The depth of field of the human eye. *Opt. Acta (Lond)* 4 (1957), 157–164.

F. W. Campbell and G. Westheimer. 1960. Dynamics of accommodation responses of the human eye. *J. Physiol.* 151 (1960), 285–295.

Oliver Cossairt and Shree K. Nayar. 2010. Spectral focal sweep: Extended depth of field from chromatic aberrations. In *Proc. ICCP*.

Oliver Cossairt, Changyin Zhou, and Shree K. Nayar. 2010. Diffusion coded photography for extended depth of field. *ACM Trans. Graph. (SIGGRAPH)* 29, 4 (2010), 31:1–31:10.

James E. Cutting and Peter M. Vishton. 1995. Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. *W. Epstein and S. Rogers (Eds.), Percept. Space Motion* (1995), 69–117.

S. G. de Groot and J. W. Gebhard. 1952. Pupil size as determined by adapting luminance. *J. Opt. Soc. Am.* 42, 7 (1952), 492–495.

Eugene Dolgoff. 1997. Real-depth imaging: a new 3D imaging technology with inexpensive direct-view (no glasses) video and other applications. *Proc. SPIE 3012* (1997), 282–288.

Edward R. Dowski and W. Thomas Cathey. 1995. Extended depth of field through wave-front coding. *Appl. Opt.* 34, 11 (1995), 1859–66.

E. F. Fincham. 1951. The accommodation reflex and its stimulus. *Br. J. Ophthalmol.* 35 (1951), 381–393.

E. F. Fincham and J. Walton. 1957. The reciprocal actions of accommodation and convergence. *J. Physiol.* 137, 3 (1957), 488–508.

Max Grosse, Gordon Wetzstein, Anselm Grundhöfer, and Oliver Bimber. 2010. Coded aperture projection. *ACM Trans. Graph.* 29, 3 (2010), 22:1–22:12.

G. Häusler. 1972. A method to increase the depth of focus by two step image processing. *Opt. Commun.* 6, 1 (1972), 38–42.

Gordon Heron, W. Neil Charman, and Clifton M. Schor. 2001. Age changes in the interactions between the accommodation and vergence systems. *Optom. Vis. Sci.* 10, 78 (2001), 754–62.

David M. Hoffman, Ahna Reza Girshick, Kurt Akeley, and Martin S. Banks. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J. Vis.* 8, 3 (2008).

Xinda Hu and Hong Hua. 2014. Design and assessment of a depth-fused multi-focal-plane display prototype. *J. Disp. Technol.* 10, 4 (2014), 308–316.

Hong Hua and Bahram Javidi. 2014. A 3D integral imaging optical see-through head-mounted display. *Opt. Express* 22, 11 (2014), 13484–13491.

Fu-Chung Huang, Kevin Chen, and Gordon Wetzstein. 2015. The light field stereoscope: Immersive computer graphics via factored near-eye light field display with focus cues. *ACM Trans. Graph. (SIGGRAPH)* 34, 4 (2015).

Daisuke Iwai, Shoichiro Mihara, and Kosuke Sato. 2015. Extended depth-of-field projector by fast focal sweep projection. In *Proc. VR*.

Paul V. Johnson, Jared A. Q. Parnell, Joohwan Kim, Christopher D. Saunter, Gordon D. Love, and Martin S. Banks. 2016. Dynamic lens and monovision 3D displays to improve viewer comfort. *Opt. Express* 24 (2016), 11808–11827. Issue 11.

Robert Konrad, Emily A. Cooper, and Gordon Wetzstein. 2016. Novel optical configurations for virtual reality: Evaluating user preference and performance with focus-tunable and monovision near-eye displays. In *Proc. SIGCHI*.

Frank L. Kooi and Alexander Toet. 2004. Visual comfort of binocular and 3D displays. *Displays* 25 (2004), 99–108.

Gregory Kramida. 2015. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE TVCG* 22 (2015), 1912–1931. Issue 7.

Marc Lambooij, Marten Fortuin, Ingrid Heynderickx, and Wijnand IJsselsteijn. 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *J. Imaging Sci. Technol.* 53, 3 (2009).

Douglas Lanman and David Luebke. 2013. Near-eye light field displays. *ACM Trans. Graph. (SIGGRAPH Asia)* 32, 6 (2013), 220:1–220:10.

Anat Levin, Samuel W. Hasinoff, Paul Green, Frédo Durand, and William T. Freeman. 2009. 4D frequency analysis of computational cameras for depth of field extension. *ACM Trans. Graph. (SIGGRAPH)* 28, 3 (2009), 97:1–97:14.

Sheng Liu, Dewen Cheng, and Hong Hua. 2008. An optical see-through head mounted display with addressable focal planes. In *Proc. ISMAR*. 33–42.

Patrick Llull, Noah Bedard, Wanmin Wu, Ivana Tosic, Kathrin Berkner, and Nikhil Balram. 2015. Design and optimization of a near-eye multifocal display system for augmented reality. In *OSA Imaging Appl. Opt.*

Gordon D. Love, David M. Hoffman, Philip J. W. Hands, James Gao, Andrew K. Kirby, and Martin S. Banks. 2009. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Opt. Express* 17, 18 (2009), 15716–25.

Kevin J. MacKenzie, David M. Hoffman, and Simon J. Watt. 2010. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *J. Vis.* 10, 8 (2010).

Susana Marcos, Esther Moreno, and Rafael Navarro. 1999. The depth-of-field of the human eye from objective and subjective measurements. *Vision Res.* 39, 12 (1999), 2039–2049.

Daniel Miau, Oliver Cossairt, and Shree K. Nayar. 2013. Focal sweep videography with deformable optics. In *Proc. ICCP*. 1–8.

Hajime Nagahara, Sujit Kuthirummal, Changyin Zhou, and Shree K. Nayar. 2008. Flexible depth of field photography. In *Proc. ECCV*.

Rahul Narain, Rachel A. Albert, Abdullah Bulbul, Gregory J. Ward, Martin S. Banks, and James F. O'Brien. 2015. Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Trans. Graph. (SIGGRAPH)* 34, 4 (2015).

Nitish Padmanaban, Robert Konrad, Tal Stramer, Emily A. Cooper, and Gordon Wetzstein. 2017. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proc. Natl. Acad. Sci. U.S.A.* 114 (2017), 2183–2188. Issue 9.

Harris Ripps, Newton B. Chin, Irwin M. Siegel, and Goodwin M. Breinin. 1962. The effect of pupil size on accommodation, convergence, and the AC/A ratio. *Invest. Ophthalmol.* 1 (1962), 127–135.

Jannick P. Rolland, Myron W. Krueger, and Alexei Goon. 2000. Multifocal planes head-mounted displays. *Appl. Opt.* 39, 19 (2000), 3209–3215.

Clifton M. Schor. 1992. A dynamic model of cross-coupling between accommodation and convergence: simulations of step and frequency responses. *Optom. Vis. Sci.* 69, 4 (1992), 258–269.

Brian T. Schowengerdt and Eric J. Seibel. 2006. True 3-D scanned voxel displays using single or multiple light sources. *J. SID* 14, 2 (2006), 135–143.

Takashi Shibata, Joohwan Kim, David M. Hoffman, and Martin S. Banks. 2011. The zone of comfort: Predicting visual discomfort with stereo displays. *J. Vis.* 11, 8 (2011), 11.

T. Sugihara and T. Miyasato. 1998. 32.4: A lightweight 3-D HMD with accommodative compensation. *SID Digest* 29, 1 (1998), 927–930.

Laura E. Sweeney, Dirk Seidel, Mhairi Day, and Lyle S. Gray. 2014. Quantifying interactions between accommodation and vergence in a binocularly normal population. *Vision Res.* 105 (2014), 121–129.

F. M. Toates. 1972. Accommodation function of the human eye. *Physiol. Rev.* 52 (1972), 828–863.

Tracy K. Tsuetaki and Clifton M. Schor. 1987. Clinical method for measuring adaptation of tonic accommodation and vergence accommodation. *Am. J. Optom. Physiol. Opt.* 64, 6 (1987), 437–449.

Marc von Waldkirch, Paul Lukowicz, and Gerhard Tröster. 2004. Multiple imaging technique for extending depth of focus in retinal displays. *Opt. Express* 12, 25 (2004).

Marc von Waldkirch, Paul Lukowicz, and Gerhard Tröster. 2005. Oscillating fluid lens in coherent retinal projection displays for extending depth of focus. *Optics Communications* 253, 25 (2005).

P. A. Ward and W. N. Charman. 1987. On the use of small artificial pupils to open-loop the accommodation system. *Ophthalmic Physiol. Opt.* 7, 2 (1987), 191–193.

Gerald Westheimer. 1966. The Maxwellian view. *Vision Res.* 6 (1966), 669–682.

G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. 2012. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph. (SIGGRAPH)* 31, 4 (2012), 1–11.

Zhongsheng Zhai, Shanting Ding, QingHua Lv, Xuanze Wang, and Yuning Zhong. 2009. Extended depth of field through an axicon. *J. Mod. Opt.* 56, 11 (2009), 1304–1308.