Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors

David Dunn, Student Member, IEEE, Cary Tippets, Kent Torell, Petr Kellnhofer, Kaan Aksit, Piotr Didyk, Karol Myszkowski, David Luebke, Fellow, IEEE, and Henry Fuchs, Life Fellow, IEEE

Abstract— Accommodative depth cues, a wide field of view, and ever-higher resolutions all present major hardware design challenges for near-eye displays. Optimizing a design to overcome one of these challenges typically leads to a trade-off in the others. We tackle this problem by introducing an all-in-one solution – a new wide field of view, gaze-tracked near-eye display for augmented reality applications. The key component of our solution is the use of a single see-through, varifocal deformable membrane mirror for each eye reflecting a display. They are controlled by airtight cavities and change the effective focal power to present a virtual image at a target depth plane which is determined by the gaze tracker. The benefits of using the membranes include wide field of view (100° diagonal) and fast depth switching (from 20 cm to infinity within 300 ms). Our subjective experiment verifies the prototype and demonstrates its potential benefits for near-eye see-through displays.

Index Terms— Augmented reality, displays, focus accommodation, perception, user study

1 INTRODUCTION

Augmented Reality (AR) [7] overlays computer-generated visuals onto the real world in real time. Near-Eye Displays (NEDs) for AR applications have recently been proposed for widespread public use, such as Meta1, and Microsoft Hololens2. Some of the fundamental limitations [23] of existing NEDs for AR are limited field of view (FOV), low angular resolution, and fixed accommodative state.

Computational methodologies such as light fields [14, 24] can provide accommodative cues while enabling wide FOV. However, light field displays are known to be computationally intensive and limited in angular resolution. Always-in-focus methodologies [1, 28] can imitate accommodative cues in computational means, while providing large FOV with a small form factor, but are limited in angular resolution. Varifocal techniques [27, 34] provide high angular resolution and accommodative cues, but none of these systems have achieved a wide FOV up until now. Recent studies show evidence that supporting accommodative cues through a varifocal mechanism improves visual comfort [16] and user performance [34] while being computationally simpler than volumetric displays. Researchers have also proposed several classical optical designs [2, 19, 32, 38] to address only FOV-related issues without addressing accommodating cues related issues. As demonstrated by Benko et al. [3], combining a NED with projections

1https://www.metavision.com/
2http://www.microsoft.com/microsoft-hololens/en-us
promises larger FOV with no accommodative cues, but it introduces new practical challenges.

In this paper, we tackle the problem of providing wide FOV and accommodative cues together in the context of see-through and varifocal systems. By bringing the idea of hyperbolic half-silvered mirrors [19] and deformable membrane mirrors [30,31,35] together for NEDs in AR applications, we propose a new hybrid hardware design for NEDs that uses see-through deformable membrane mirrors. We present a complete prototype that promises to address Vergence-Accommodation Conflict (VAC) [11] caused by lack of accommodative cues. We validate the performance of our accommodation control in a subjective experiment.

### 1.1 Contributions

**Single Element Optics:** Our design employs a single optical element as the varifocal relay optics, simplifying the design of see-through varifocal optical systems for NEDs in AR applications. We present a ray tracing model for exploring the design space of our proposal.

**Wide Field Of View:** With respect to other varifocal optical components, our optical element is unique due to its large aperture size, leading to wide FOV NED solutions for AR applications. We present different design scenarios leading to wide FOV, accurate defocus blur, and demonstrate a wide FOV prototype.

**Vergence-Accommodation Conflict:** We verify our gaze tracked prototype through a subjective test. Our findings indicate the ability to address Vergence-Accommodation Conflict (VAC) in a gaze-driven way.

**Complete Prototype:** As a proof of concept, we demonstrate a binocular varifocal NED prototype with gaze tracking capability, created by modifying off-the-shelf items, and in-house custom built deformable see-through mirrors. We provide details of our implementation.

Unlike for other methodologies, the computational requirements of image generation for a varifocal system are almost the same as today’s conventional NEDs. Thus, we believe a varifocal system is very likely to be a design choice in next generation NEDs. We hope our easy-to-follow manufacturing and implementation processes provide a reproducible methodology for researchers and manufacturers.

## 2 Related Work

Enabling accommodative cues is known to cause major changes in a NED’s optical design. We revise the designs that have enabled accommodative cues, investigate their characteristics, and provide a comparison of these solutions in Table 1.

<table>
<thead>
<tr>
<th>Focus mechanism</th>
<th>See-through</th>
<th>FOV</th>
<th>Angular resolution</th>
<th>Optics</th>
<th>Form factor</th>
<th>Computational demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-eye light field displays [24]</td>
<td>light fields</td>
<td>yes</td>
<td>small</td>
<td>low</td>
<td>simple</td>
<td>thin</td>
</tr>
<tr>
<td>Light field stereoscope [14]</td>
<td>light fields</td>
<td>no</td>
<td>large</td>
<td>low</td>
<td>simple</td>
<td>moderate</td>
</tr>
<tr>
<td>Pinlight displays [28]</td>
<td>always-in-focus</td>
<td>yes</td>
<td>large</td>
<td>low</td>
<td>simple</td>
<td>thin</td>
</tr>
<tr>
<td>Pinhole displays [1]</td>
<td>always-in-focus</td>
<td>no</td>
<td>large</td>
<td>low</td>
<td>simple</td>
<td>thin</td>
</tr>
<tr>
<td>Holographic optical elements [18]</td>
<td>holographic</td>
<td>yes</td>
<td>N/A</td>
<td>N/A</td>
<td>complex</td>
<td>N/A</td>
</tr>
<tr>
<td>Multi-focal plane displays [12]</td>
<td>multi-plane</td>
<td>yes</td>
<td>small</td>
<td>high</td>
<td>complex</td>
<td>bulky</td>
</tr>
<tr>
<td>Focus tunable light engine [27]</td>
<td>varifocal</td>
<td>yes</td>
<td>small</td>
<td>high</td>
<td>moderate</td>
<td>N/A</td>
</tr>
<tr>
<td>Focus tunable lenses [34]</td>
<td>varifocal</td>
<td>no</td>
<td>small</td>
<td>moderate</td>
<td>moderate</td>
<td>moderate</td>
</tr>
<tr>
<td>This work</td>
<td>varifocal</td>
<td>yes</td>
<td>large</td>
<td>moderate</td>
<td>simple</td>
<td>bulky</td>
</tr>
</tbody>
</table>

Integral Imaging, first proposed by Lippmann [26], deals with the capture and the reproduction of light fields which with enough angular resolution can provide correct accommodative cues to a viewer. Hua and Javidi [13] demonstrate a NED for AR applications that combines recent advancements of free-form relay optics with a computational integral imaging methodology, achieving 15° of diagonal FOV with a maximum image resolution of 640 × 360 px, leading to 10 – 20 cpd. Although rendering of images is instant, the free-form optics in their design use 3 different 10th order polynomial lenses made of Polymethyl methacrylate (PMMA), which requires an access to precision machinery for replication of the work.

Lanman and Luebke [24] introduce a Near-Eye Light Field Display (NELD) that uses microlenses as the relay optics, showing a prototype with a screen of 146 × 78 px and a FOV of 29.2° × 16.0°, leading to a resolution of 2 – 3 cpd. More recently, Huang et al. [14] developed NELDs for virtual reality (VR) applications further, demonstrating a light field stereocope with a diagonal FOV of 110°, an accommodation range of 5.26 to 0.81 diopeters, and a maximum image resolution of 640 × 800 px (3 – 4 cpd). The prototype from Huang et al. employs two Liquid Crystal Displays (LCDs) and a pair of classical magnifiers. The introduced technique also promises a continuous depth information with a computational overhead that demands usage of high-end GPUs, and presents online images at a typical rate of 20 – 35 fps. Always-in-focus mechanisms also offer sharp imagery across different focal planes. The work of Akşit et al. [1] uses a pinhole mask in front of a display as a NED for VR applications, and demonstrates full color images at a diagonal FOV of 83° with a resolution of 460 × 260 px (2 – 3 cpd). The “Pinlights” always-in-focus AR display, by using a see-through sparse backlight mechanism from Maimone et al. [28], introduces a single color prototype with a diagonal FOV of 110°, and a resolution of 2 – 3 cpd. The work of Maimone et al. can also provide full-color imagery with 12 Hz refresh rate. Both of these implementations suffer the primary disadvantage of poor angular resolution.

Researchers have shown a growing interest in the use of Holographic Optical Elements (HOEs) in NED designs [18]. Holography promises a good angular resolution with a thin form factor, but to our knowledge, no implementation of HOEs inside a complete NED has yet been demonstrated. We believe high FOV will be the major practical challenge in holographic NED research.

The work of Hu and Hua [12] presents a see-through multi-plane NED using Deformable Mirror Membrane Devices (DMMDs) that provide 1 kHz refresh rate. Their prototype provides a 40° diagonal FOV, and an image resolution of 1024 × 768 px, leading to resolvability of 9 – 12 cpd. However, the optical complexity in such approaches has to date challenged their practicality in increasing angular resolution and decreasing form factors.

Closely related to our proposal, a varifocal system by Liu et al. [27] uses a tunable lens system combined with a spherical mirror, and demonstrates 28° of diagonal FOV, 800 × 600 px resolution (10 –
We approached the problem of calculating the required mirror curvature for a given configuration through a three-dimensional (3D) ray tracing model [39]. The objective of our ray tracing model is to find a good mirror curvature that creates the smallest resolvable spot size. The first step of our model defines sample points \( p_e \) inside a given \( d_{\text{eye box}} \). In our ray tracing routine, all sample points from a given \( d_{\text{eye box}} \) collectively represents a forward gazing eye aperture aligned with the optical axis (Z axis). Next, we define a sample point \( p_s \) at a desired depth \( d_{\text{travel}} \). We choose a \( p_s \) aligned with the optical axis again. We define multiple rays from points inside \( p_e \) an eye box traveling to a sample point \( p_s \) in depth. A single one of these rays \( R_0 \) is defined as

\[
R_0 = \begin{cases} 
  p_e = \begin{bmatrix} p_{e_x} \\ p_{e_y} \\ p_{e_z} \end{bmatrix}, \\
  a_{es} = \frac{1}{d_{es}} \begin{bmatrix} p_{e_x} - p_{s_x} \\ p_{e_y} - p_{s_y} \\ p_{e_z} - p_{s_z} \end{bmatrix}, 
\end{cases}
\]

where \( p_e \) indicates a starting point, \( a_{es} \) indicates direction cosines of the ray, and \( d_{es} \) indicates the distance between \( p_e \) and \( p_s \). We trace \( R_0 \) from pupil plane to deformable membrane mirror. Note that Figure 2 shows two anchor points for the deformable membrane mirror. Any given sphere that has such anchor points at both axis (X and Y) by definition is on the line that is perpendicular to the flat mirror surface, and crosses the center of the flat mirror surface. Assuming a configuration as in Figure 2, such a line can be defined as \( z = \frac{-1}{\tan(\beta)}(y + d_{\text{eye relief}}) \), leading to \( M_z = (0, y, z) \). The intersection point between a deformable membrane \( R_0 \) and \( R_n \) can be calculated by finding a ray propagation distance \( d_0 \) that satisfies the sphere equation on the surface of the membrane with a point \( p_{\text{mirror}} = (p_e + d_0 a_{es}) \). Thus, ray propagation distance can be calculated by finding the roots of

\[
\|p_{\text{mirror}} - M_z\| = r,
\]

and choosing the closest root to the plane of the deformable membrane mirror. A surface normal \( R_n \) of the deformable membrane mirror at a point can be calculated as

\[
R_n = \begin{cases} 
  n_{\text{mirror}} = p_{\text{mirror}} - p_{\text{pupil}} - M_z, \\
  a_{\text{mirror}} = \frac{p_{\text{mirror}} - M_z}{\|p_{\text{mirror}} - M_z\|}. 
\end{cases}
\]

Using \( R_n \) and \( R_0 \), we calculate the reflection as a ray \( R_1 \) which can be calculated as

\[
R_1 = R_0 - 2R_n(R_0 \cdot R_n).
\]

To calculate the intersection of \( R_1 \) with a display plane, we need to be able to calculate two things: (1) surface normal of our display plane and (2) ray propagation distance \( d_1 \) from the origin of the ray to a display plane. The surface normal of our display plane \( R_{d1} \) can be calculated as

\[
R_{d1} = \begin{cases} 
  p_{\text{display}} = \begin{bmatrix} 0 \\ \sin(\beta) d_{\text{display}} \\ d_{\text{eye relief}} - \cos(\beta) d_{\text{display}} \end{bmatrix}, \\
  a_{\text{display}} = \begin{bmatrix} 0 \\ \sin(\beta + \alpha) \\ \cos(\beta + \alpha) \end{bmatrix}. 
\end{cases}
\]

Using the surface normal and a vector \( R_2 \) from \( p_{\text{mirror}} \) to \( p_{\text{display}} \), \( d_1 \) can be calculated as

\[
d_1 = \frac{R_{d1} \cdot R_2}{R_{d1} \cdot R_1},
\]

and finally, we can find the intersection point as \( p_{\text{final}} = p_{\text{mirror}} + d_1 a_{\text{mirror}} \). We use the intersection points to calculate the spot size, in

![Fig. 2. A sketch showing our varifocal optical layout with parameters for the single eye case. An image on a display above the user’s eye reflects from our deformable membrane mirror toward the eye. A virtual image can be created at a desired position in space by varying the curvature of our deformable membrane mirror.](image)
which Full Width Half Maximum (FWHM) size of the spot diagonal is calculated using $FWHM = 2.355\sigma$. Using secant method, we optimize the curvature of the deformable mirror membrane by minimizing FWHM size for a given configuration. We choose a new curvature $r_{\text{new}}$ at each iteration as

$$r_{\text{new}} = r_{\text{current}} \left( 1 - \frac{FWHM_{\text{current}} - FWHM_{\text{previous}}}{r_{\text{current}} - r_{\text{previous}}} \right). \quad (8)$$

### 3.2 Design space

Here we explore the design space of our proposal using our ray tracing model to identify the impact of $d_{\text{eye relief}}$, $d_{\text{display}}$, and $d_{\text{aperture}}$. First, we will analyze $d_{\text{aperture}}$, which is defined both in the vertical and the horizontal axis. In our designs, the aperture shapes are chosen as either circular or elliptical. Adult humans have a mean interpupillary distance (IPD) of 63 mm, and their IPDs can range between 50 and 75 mm [8]. Thus, horizontal aperture size is dictated by IPD in the nasal direction. Maximum aperture size at a vertical axis can be of any desired size to cover a larger vertical FOV. Note that user’s eyes can be decentred with respect to the optical axis of a deformable membrane mirror; thus we choose to use $d_{\text{eye box}} = 20\text{ mm}$ to compensate for alignment as in the case of a conventional NED design. Designs with elliptical aperture shapes can introduce perceptible astigmatism in an optical system. Such cases can easily be corrected by placing a single axis lens in between a display and a deformable membrane mirror.

Sample designs in Figure 3 demonstrate our findings on the effects of $d_{\text{eye relief}}$ and $d_{\text{aperture}}$ on FOV. These results suggest that shorter $d_{\text{eye relief}}$ and larger $d_{\text{aperture}}$ promise a larger FOV. We would like to highlight that majority of our designs promise a larger FOV than a typical NED for AR applications. The main limitation of our designs comes from the limited FOV generation towards the brows due to the $\beta$ angle of the membrane mirror causing a more distant reflector in that region. Note that an asymmetrical aperture in different directions (brow, nose, cheek, peripheral), different aperture shapes (square, custom) or offsetting and angling the central axis of the membrane are possible solutions to overcome limited FOV towards the nose and the brow. However, non-elliptical designs require a more complex multi-domain modeling, leading to complex surface deformations largely deviating from regular spherical or aspherical surfaces, while off-axis designs degrade the optical qualities of the reflected image. Increasing the aperture size will also lead to clipping the reflections of the display particularly in the bottom region which reflects the portion of the display that abuts the brow.

We propose a pneumatic system to control the deformations of the membrane mirror. Understanding the required curvature values and maximum displacement for a deformable membrane mirror lets us identify the speed and volume of air movement that dictated the requirements for the pneumatic hardware. We explore the impact of different $d_{\text{eye relief}}$ and $d_{\text{aperture}}$ on curvature, displacement, and resolution by ray tracing to simulate the characteristics of different points in depth aligned with the optical axis. Our ray tracing model suggests that different $d_{\text{eye relief}}$ leads to different $M_r$, and $r$ configurations meaning the deformable membrane displaces different amounts with respect to the flat mirror case. We show the effect of $d_{\text{eye relief}}$ with a sample design in Figure 4. Note that shorter $d_{\text{eye relief}}$ requires less deformation of the deformable membrane mirror, which, as a result, requires more precise pneumatics. On the other hand, larger $d_{\text{eye relief}}$ provides a smaller resolvable pixel size, leading to more resolution, but as noted above decreases the FOV. We conclude the pixel size dictates the required $d_{\text{eye relief}}$ in practical designs. We also evaluate the same sample designs for different $d_{\text{display}}$, as shown in Figure 5. This shows that larger $d_{\text{display}}$ increases resolution while decreasing the required amount of deformation on the membrane, but also increases the overall form factor of the complete system while decreasing FOV.

### 4 IMPLEMENTATION

We demonstrate our proposal with an experimental see-through varifocal NED equipped with a gaze tracker as shown in Figure 6. All the hardware components used in our final prototype are presented in a
Fig. 4. A sample design is evaluated for different eye reliefs $d_{\text{eye relief}}$ with a configuration of an aperture size $d_{\text{aperture}} = 50$ mm in horizontal axis, an aperture size $d_{\text{aperture}} = 65$ mm in vertical axis, a mirror tilt $\beta = 45^\circ$, a screen tilt $\alpha = 20^\circ$, an eye box $d_{\text{eye box}} = 20$ mm, and a screen distance $d_{\text{display}} = 60$ mm. For all evaluations, on-axis depth fields as shown in Figure 2 are chosen at different depth levels. A deformable membrane mirror’s curvature is calculated for different depth levels as shown on the left. The maximum amount of displacement required by each depth level is shown in the middle figure. Assuming an eye with an aperture size of 6 mm, resolvable pixel size on a screen inside the given eye box is calculated for different depth levels as shown in the figure on the right. Smaller $d_{\text{eye relief}}$ benefits the design by decreasing required displacement on a membrane, however resolution improves at closer depths with a larger $d_{\text{eye relief}}$.

Fig. 5. A sample design is evaluated for different display distances $d_{\text{display}}$ with a configuration of an aperture size $d_{\text{aperture}} = 50$ mm in horizontal axis, an aperture size $d_{\text{aperture}} = 65$ mm in vertical axis, a mirror tilt $\beta = 45^\circ$, a screen tilt $\alpha = 20^\circ$, an eye box $d_{\text{eye box}} = 20$ mm, and an eye relief $d_{\text{eye relief}} = 50$ mm. For all evaluations, on-axis depth fields as shown in Figure 2 are chosen at different depth levels. A deformable membrane mirror’s curvature is calculated for different depth levels as shown on the left. The maximum amount of displacement required by each depth level is shown in the middle figure. Assuming an eye with an aperture size of 6 mm, resolvable pixel size on a screen inside the given eye box is calculated for different depth levels as in the figure on the right.

4.1 Manufacturing flexible membranes

The task of manufacturing custom flexible membranes is accomplished traditionally through surface micromachining, bulk micromachining, liquid crystals, piezoelectric or electrostrictive actuators as reviewed by Mansell et al. [29]. Pneumatic based systems have also been demonstrated for building tunable microoptics using polydimethysiloxane (PDMS) [42], avoiding the use of high voltages or external fields in operation and precise machining in manufacturing. On the other hand, PDMS has numerous attractive material properties such as outstanding transparency in visible wavelengths, high elasticity, and excellent temperature stability. Inspired by these advantages, we created our own recipe for the task.

We used Sylgard 184 PDMS kit purchased from Dow Corning. Sylgard 184 is a two-part elastomer kit, with PDMS pre-polymer and a cross-linking agent. The prepolymer was mixed with cross-linking agent at a ratio of 10 : 1 and mixed vigorously for 3 minutes. The mixture was then degassed for 15 minutes, to remove bubbles incorporated during mixing. 6" Silicon wafers were purchased from University Wafers. The Wafer was silanized, to ease membrane release, by being placed in a desiccator, with 20 ul of trichloro (1H,1H,2H,2H-perfluorooctyl) silane and evacuated for 30 minutes and left under vacuum for 1 hour. Mixed and degassed PDMS prepolymer is spin cast on the Si wafer for 1 min at 300 RPMs to obtain a PDMS membrane of approximately 240 um. The membrane was then placed in a commercial physical vapor deposition unit (Kurt Lesker PVD 75) and a 20 nm Ag film is sputtered on the membrane. After metalization the film is carefully peeled and stretched taut across the vacuum housing to form the deformable membrane mirror. Fused Deposition Modeling (FDM) based 3D printers that we have tried were not able to accomplish the task of manufacturing...
Transmission and Reflection characteristics of our deformable membrane mirror were captured as in Figure 8 using a J. A. Woollam variable angle spectroscopic ellipsometer. The deformable membrane mirror was aligned and the incident angle was set to 40 degrees to match $\beta$ and $\alpha$ for both the transmission and reflection measurements. Work of Lee et al. [25] highlights that a thickness of an optical combiner plays a crucial role in depth perception, as our membrane mirror has 240 μm thickness, effects described by Lee et al. are expected to be at a negligible level in our implementation.

### 4.2 Integration

Our choice of design parameters was mainly constrained by the availability of off-the-shelf components and the costs of custom tooling. Per eye, we use a vacuum source (115 Torr $\sim$ 15 kPa) with a SMC ITV2090-21N2BL5\(^4\) vacuum regulator, a t-junction, and a bleed hole to create a controlled partial vacuum environment inside our vacuum housing. Our vacuum regulators can regulate pressure levels in between $-1.3$ to $-80$ kPa, and each is controlled by a Teensy 3.2 microcontroller\(^5\) (µC). Our combination of µCs and vacuum regulators provides us $\sim$ 60 addressable stable depth planes ranging from 0.2 to 7 diopters according to our empirical experiments. We used a Adafruit Qualia 9.7” LCD\(^6\) with 260 ppi, active region used per eye is 1050 $\times$ 1260 px. Our prototype uses a gaze tracking Pupil-labs camera\(^7\) per eye, running at 120 Hz.

Using an early prototype of the housing, we conducted a deformation test for our deformable mirror as shown in Figure 9. During our deformation tests, we stressed the membrane to deformations that are 10 times larger than the deformations that we have during operation. Large ripples at the edge of the deformable membrane are believed to be caused by a weak attachment to the housing wearing out after 26700 iterations, which we solved in later iterations of the housing with a more secure attachment. Hazing in the images is believed to be caused by a change in surface structure after many iterations. Our deformation test was conducted over a 30 hour time frame. As our membrane underwent strains far greater than during normal operation.

\(^4\)https://www.smcpneumatics.com  
\(^5\)https://www.adafruit.com/product/1652  
\(^6\)https://www.adafruit.com/product/2756  
\(^7\)https://pupil-labs.com/store/
without failing, we can conclude that our deformable membrane mirror and pneumatics control mechanism are suitable for long term usage.

4.3 Software
We developed an in-house software to control our prototype, to conduct subjective experiments, and to render images accordingly. Our software is written in Python programming language taking advantage of GLFW\(^8\) for user interface rendering, OpenCV\(^9\) for image processing tasks, and Pupil-labs library for gaze tracking tasks. Our software runs on an Intel Xeon CPU W5590 @ 3.33 GHz PC with two Nvidia Quadro NVS 420 GPUs and Linux operating system.

Our control methodology for the deformations of the deformable membrane mirror is based on reflection shape detection from an Infrared (IR) Light Emitting Diode (LED) placed above each deformable membrane mirror. An IR camera running at 30 FPS for each deformable membrane mirror is also placed above the deformable membrane mirror as shown in bottom view of Figure 6. Whenever, system is dictated to change the effective focal power, PC electronically controls the vacuum regulator through uCs, and reflection detections from IR cameras act as a feedback mechanism to form a closed loop control mechanism.

For different depth levels, image distortions caused by our deformable membrane mirror are captured by a PointGrey Flea FLEA-HICOL camera\(^10\) with a Fujinon F1 : 1.2 - 2.8 - 8 mm aperture lens. Note that the mentioned camera is for identification of image distortions, and not a permanent part of our system. We characterized image distortions by using the work of Yamazaki et al. [43] and our captures. We applied our findings on the image distortions to our software to present images consistently with the changing focus.

5 Experiments
The goal of our experiment was to verify whether our accommodation support works well, and if users can benefit from it while performing visual acuity task in a monocular viewing scenario. Our hypothesis was that a correct accommodation will allow users to resolve higher spatial details.

5.1 Stimuli
Each stimulus consisted of a pair of white Landolt C shapes shown on a black background (Figure 10). The location of the gaps was either on the top or the bottom side corresponding to the up and the down orientation of the shape. The shapes were separated by 2 visual degrees, and each of them spanned 30 arcmin which imposes the gap size of 6 arcmin, where the normal 20/20 eye can resolve 1 arcmin. Since through our NED calibration its focus state has been precisely setup for each trial, we opted for the larger gap size so that the user response is immediate and effortless, as well as it is not affected by lower display contrast, limited spatial resolution, and possibly imperfect luminance adaptation with respect to the requirements of standard visual acuity test. One shape was presented on one of two physical screens located at 0.25 m (Adafruit Qualia, 9.7”, 2048 × 1536, 23.5 cpd, 60 Hz) and 5.0 m (Sharp Aquos Quattron LC-70LE732U, 70”, 1920 × 1080, 54.3 cpd, 60 Hz) from the viewer. The other Landolt shape was presented on our NED with a focal distance either matching the distance to the physical screen or a modified one to simulate a lack of a correct accommodation cue. The range of considered focal distance offsets was 0.2 to 5 diopters. For the screen located at 0.25 m, we moved the virtual object further from the observer, while for the screen located at 5.0 m, we moved the virtual image closer to the observer.

5.2 Participants
Twelve subjects (2 F, 10 M, 20 to 34 years of age) that had a normal or corrected-to-normal vision, took part in the experiment. To keep participants inside the eyebox of our NED, all participants used a chin and forehead rest.

5.3 Procedure
At each trial, a participant was asked to monocularly fixate at one of the physical screens. To this end, a simple math equation was displayed on the screen using a font of height 15 arcmin, while nothing was displayed on our NED. The user was asked to push one button if the equation was true and another if it was false. This math task was introduced to control the user fixation and give him enough time to correctly accommodate to the distance at which the physical screen was located. Immediately after responding, the stimulus appeared on the reference and the NED at a location central to the equation. The presentation time of the stimulus was set to 300 ms. The time was chosen such that it was just enough to perform the visual acuity task, and it was determined during a pilot experiment. Note that the time is also shorter than the latency before the actual change in the lens shape is triggered, which we discuss in more details in Section 6. Next, the participant was presented with a blank screen and asked to press a button selecting whether the two patterns were of equal or different orientation. Afterwards, the study continued with a next trial. In total, two physical displays and six focus distances for the NED were used in random order which after 20 repetitions gave the total of 240 trials per participant. Each participant took on average 30 minutes to complete the task.

\(^{8}\)http://www.glfw.org/
\(^{9}\)http://opencv.org/
\(^{10}\)https://www.ptgrey.com
The proportion correct as a function of test focal distance of the NED. Two points marked by rectangles are points where the reference and the test distances matched. For such conditions, the performance is expected to be the best. The error bars denote Clopper-Pearson binomial confidence intervals.

5.4 Results
The graph in Figure 11 shows the relation of the NED focal distance and the proportion of correct responses for each of the reference displays. We performed a $\chi^2$-test to analyze differences between different conditions and found a significance influence of the test focal distance on the proportion correct for both 0.2 diopters ($\chi^2 = 82.7$, $df = 5$, $p < 0.001$) and 4.0 diopters ($\chi^2 = 204.7$, $df = 5$, $p < 0.001$) references. A post-hoc analysis with Bonferroni correction and significance level equal to 0.05 revealed that the differences between test pairs were significant for all but the following: 0.2/2.0, 0.2/3.0, 2.0/3.0, 4.0/5.0 for 0.2 diopters reference and 1.0/2.0, 1.0/5.0, 2.0/4.0, 3.0/4.0 for 4.0 diopters reference.

In general, as the test focal distance approached the reference depth, i.e., both stimuli were presented at the the same focal distance, the participants were able to correctly perform the task more often. As our analysis shows, this is a significant drop.

The performance of our display is affected by the system latency. Its velocity for the lens accommodation, but a high variance can be observed in their data. Also, for disaccommodation, the peak velocities are reached. Phillips et al. [33] have observed latencies as short as 200 ms, the probability of their occurrence is very low. They hypothesize that such short latencies can be explained by coincidence or confusion of some subjects who have not carefully followed the experiment protocol. The duration of actual lens accommodation of 500 – 800 ms has been reported [4, 6, 10, 33], which means that the complete accommodation cycle, including the latency, typically requires around 1 second [6].

The velocity of accommodation is a good measure of the lens accommodation dynamics. Bharadwaj et al. [4] observed a smooth increase in velocity to its peak value and then its slightly slower reduction to the steady state. The peak velocity increased with the accommodation magnitude and the highest value of around 10 diopters/second has been reported. Kasthurirangan et al. [17] observed a similar average peak velocity for the lens accommodation, but a high variance can be observed in their data. Also, for disaccommodation, the peak velocities over 20 diopters/second have been measured for the large accommodation magnitudes of 4 – 5 diopters, which are still within the range of focal changes in our NED. The operational velocity of our membrane amounts to 16.6 diopters/second, which might be below the peak velocity for disaccommodation in extreme depth changes. Since our membrane deformation is initiated during the period of eye accommodation latency and its maximum duration is less than 300 ms, we expect that the whole process is completed well before such extreme lens accommodation velocities are reached.

The total latency of our system remains below the delays of the eye accommodation process and may be sufficiently low for AR applications. This is supported by results of our subjective experiments. We leave more in-depth experiments regarding latency requirements.
being equivalent to optical blur, we have not implemented this solution. Due to the potential of rendered blur not mimicking outside such region can be relaxed. This greatly simplifies the image quality. Using the eye tracking system, we are able to provide a precise focus in the fovea region, while the precision of membrane eccentricity reduces requirements imposed on the membrane design in our display, as relatively high defocus blur can be tolerated outside the central foveal region without causing any perceivable degradation of the image quality. Using the eye tracking system, we are able to provide a precise focus in the fovea region, while the precision of membrane shaping outside such region can be relaxed. This greatly simplifies the high visual quality over a wide FOV.

6.5 Depth of field

Our display is capable of displaying only a single depth at a time, which leads to incorrect views for virtual content at different depths. A simple solution to this would be to apply a defocus kernel approximating the eye’s point spread function to the virtual image according to the depth of the virtual objects. Due to the potential of rendered blur not being equivalent to optical blur, we have not implemented this solution. Future work must evaluate the effectiveness of using rendered blur in place of optical blur.

6.6 Occlusion support

The work of Kiyokawa et al. [20] describes an occlusion cable NED, and introduces an application space that requires occlusion support. Our proposal does not attempt to support occlusion. We leave this challenge as a future work.

6.7 Monocular vs. binocular experiment

We present a binocular display, but verify it only by a monocular experiment. Our monocular experiment has demonstrated that combined real-virtual depth-varying task performance can be improved with focal accommodation. However, binocular experiments would allow us to show a relationship between vergence-accommodation conflict and task performance. We leave binocular experiments for a future work, and are excited by the possible perceptual studies which are now open with this new hardware prototype. In particular, we would like to verify that visual fatigue due to vergence accommodation conflict can be mitigated by our display. Such experiments can potentially also reveal more interesting facts about the vergence-accommodation relationship.

7 Conclusion

In order to provide a high-quality augmented reality experience, it is crucial to design headsets that are capable of reproducing all visual cues across the whole visual field. In this respect, the most challenging is reproduction of accommodation cue as well as providing wide field of view. To address these problems, we propose a novel design of a see-through near-eye display for augmented reality applications. The key to our solution are two membranes with half-mirror properties. Thanks to their deformation properties, their focal power can be adjusted using an airtight chamber to provide accommodation cues matching the observer’s fixation distance determined by an eye tracker. This addresses the problem of visual discomfort which is usually caused by a mismatch between vergence and accommodation. It also improves user task performance as demonstrated in our experiment. Another unique advantage of our membranes is the fact that they enable a significantly larger field of view when compared to other varifocal designs. Despite few limitations of our system, we believe that providing correct focus cues as well as wide field of view are most crucial features of head-mounted displays that try to provide seamless integration of the virtual and the real world. Our screen not only provides basis for new, improved designs, but it can be directly used in perceptual experiments that aim at determining requirements for future systems. We, therefore, argue that our work will significantly facilitate the development of augmented reality technology and contribute to our understanding of how it influences user experience.

Acknowledgments

The authors wish to thank Joohwan Kim, Ward Lopes, Josef Spjut, and Turner Whited for fruitful discussions. Jim Mahaney was invaluable in consulting and assisting with the physical set-up of our display prototype. The authors would also like to thank the participants of our subjective experiments for their participation, and the anonymous reviewers for valuable feedback.

This work was partially supported by National Science Foundation (NSF) Grant IIS-1645463, by NSF Grant A14-1618-001, by a grant from NVIDIA Research, by the Fraunhofer and Max Planck cooperation program within the German pact for research and innovation (PFI), and by the BeingTogether Centre, a collaboration between Nanyang Technological University (NTU) Singapore and University of North Carolina (UNC) at Chapel Hill, supported by UNC and the Singapore National Research Foundation, Prime Minister’s Office, Singapore under its International Research Centres in Singapore Funding Initiative. A portion of this work was performed at the Chapel Hill Analytical and Nanofabrication Laboratory (CHANL), a member of the North Carolina Research Triangle Nanotechnology Network (RTNN), which is supported by NSF Grant ECCS-1542015 as part of the National Nanotechnology Coordinated Infrastructure (NNCI).
REFERENCES


