Towards Varifocal Augmented Reality Displays using Deformable Beamsplitter Membranes

David Dunn*†, Praneeth Chakravarthula*, Qian Dong*, Kaan Akşit†, and Henry Fuchs*

***University of North Carolina - Chapel Hill, †Nvidia Research**

Abstract

Growing evidence in recent literature suggests gaze contingent varifocal Near Eye Displays (NEDs) are mitigating visual discomfort caused by the vergence-accommodation conflict (VAC). Such displays promise improved task performance in Virtual Reality (VR) and Augmented Reality (AR) applications and demand less compute and power than light field and holographic display alternatives. In the context of this paper, we further extend the evaluation of our gaze contingent wide field of view varifocal AR NED layout [1] by evaluating optical characteristics of resolution, brightness, and eye-box. Our most recent prototype dramatically reduces form-factor, while improving maximum depth switching time to under 200 ms.

1. Introduction

Generation of realistic visual stimuli plays an important role towards enabling prolonged usage of Augmented Reality (AR) Near-Eye Displays (NEDs). Of various challenges to be addressed in an AR NED, a major challenge in the way of achieving natural looking scenes, and a key cause of discomfort is vergence-accommodation conflict (VAC) [2]. There is increasing recent evidence that suggests gaze contingent AR NED designs enhance visual comfort and task performance [3], [4]. This class of systems which dynamically adjusts the focus of the synthetic imagery based on the user's focal state is widely known as varifocal NEDs. The dramatically less compute and power demands of varifocal NEDs offer a major advantage compared to other accommodation supporting NEDs such as light field [5], [6] and holographic NEDs [7], [8].

We tackle the challenge of VAC through our novel varifocal NED design [1], which brings the idea of hyperbolic half-silvered mirrors and deformable membrane mirrors together in a NED. In this paper, we introduce our latest prototype with faster deformable membranes and improved control mechanisms and further evaluate our design's optical performance. Our latest prototype improves the near eye form factor of our prototype dramatically, replaces the bulky air compressor with much smaller loud speakers, replaces two power hungry cameras for controlling membrane with low-power pressure sensors, and it does not require pressure regulators.

2. Related Work

Our design is an accommodation supporting AR NED design [1] aimed at providing comfortable synthetic visual stimuli with the least amount of compute and power demand. Thus, we review a multitude of accommodation supporting AR NED architectures from recent literature [9], [10].

Early on, Akeley et al. [11] demonstrate the benefits of fixed-viewpoint volumetric desktop displays using multiple display planes and generate near-correct focus cues without tracking eye position. Recently such displays were revisited with improved scene decomposition, and gaze-contingent varifocal multi plane capabilities [12], [13]. However, these displays have large power and computational demands with a complex hardware that doesn't lead to a wearable form factor. The work of Hu et al. [14] demonstrates a time-multiplexed multi-plane display in the form of a wearable near eye display. However, such a display layout offers good resolution, but only with a small field of view ($30^\circ \times 40^\circ$).

Lanman and Luebke [15] introduce a Near-Eye Light Field Display (NELD) that uses microlenses as the relay optics, showing a prototype with a resolution of 146×78 px and a FOV of $29.2^\circ \times 16.0^\circ$, leading to a resolution of 2–3 cycles per degree (cpd). More recently, Huang et al. [16] improved NELDs for VR applications, demonstrating a light field stereoscope with a diagonal FOV of 110°, an accommodation range of 5.26 to 0.81 diopters, and a maximum image resolution of 640×800 px (3–4 cpd). The work of Akşit et al. [17] uses a pinhole mask in front of a display as a NED for VR applications, and demonstrates images at a diagonal FOV of 83° with a resolution of 460×260 px (2–3 cpd). By using a see-through sparse backlight mechanism, the work of Maimone et al. [18] introduces a single color prototype with a diagonal FOV of 110° and a resolution of 2–3 cpd. All of the above-mentioned light field approaches offer limited resolutions.

Researchers have shown a growing interest in the use of Holographic Optical Elements (HOEs) as a replacement of bulky optics. HOEs have recently been part of retinal NEDs [8], [19], enabling almost glasses-like thin form factor, and a field of view as large as 80°. For such displays, a small eyebox, large compute demand, and theoretically limited resolutions (at most 8–12 cpd) are remaining issues to be resolved.

A varifocal system by Liu et al. [20] uses a tunable lens system combined with a spherical mirror, and demonstrates 28° of diagonal FOV, 800×600 px resolution (10–14 cpd), and an accommodation range of 0–8 diopters switchable within 74 ms. A recent study by Konrad et al. [21] again takes advantage of an electrically tunable lens system as relay optics and demonstrates 36° diagonal FOV. Their solution switches depth from one extreme to another within 15 ms, and provides a maximum image resolution of 488×488 px (5–6 cpd) and an accommodation range of 0–9.5 diopters. Most recently, the work of Akşit et al. [22] proposes HOEs as a part of a AR varifocal NED system, offering improved wearable form factor with 60° field of view, and a resolution

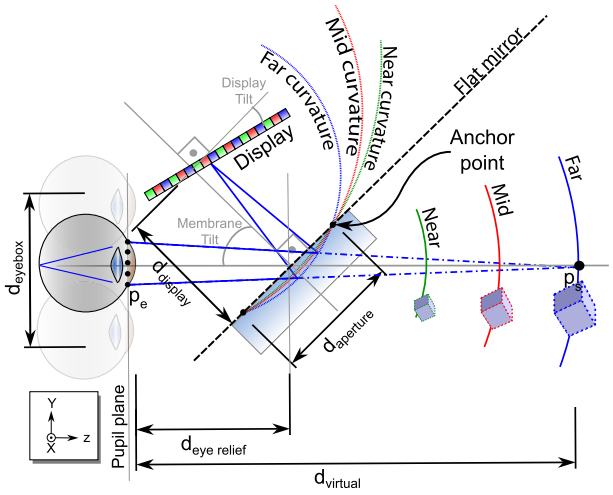

Fig. 1. Optical Layout, which consists of an LCD display panel and deformable beamsplitter at specific angles and distances from the eye.

Table 1. Prototype parameters as related to Figure 1

Parameter	Value
Eye Relief	40 mm
Aperture	57.25 mm x 50 mm
Display Distance	28.629 mm
Membrane Tilt	21°
Screen Tilt	0°

of 18 cpd. All of the above-mentioned varifocal layouts, including our previous proposal, suffer from a large form factor.

3. Hardware Implementation

Our varifocal display proposal relies on the technique of adjusting optical depth of a virtual image by dynamically adjusting the optical power of a semi-reflective membrane to match the gaze of a user. With respect to our first prototype described in [1], we have improved the major drawback of form-factor with a new improved prototype. Our new prototype is based on the same optical configuration depicted in figure 1. Improvements in optical quality and form-factor have lead to the new specifications as listed in Table 1. Overall, the head-mounted portion of our prototype, as shown in figure 2, consumes a much smaller volume (5.5 x 12.5 x 15.2 cm). Additionally, we evaluate optical qualities of our latest prototype in greater detail in section 4.

To provide imagery to both eyes, we use a single Liquid Crystal Display (LCD) panel Topfoison TF60010A-V0 1440x2560 5.98" TFT LCD. The deformable membranes for each eye are manufactured in house using the methodology described in our original proposal [1]. Our most recent implementation does not require air compressors and pressure regulators. Instead, we use a Pyle PLMRW8 8" 400 Watt 4 Ohm Marine Subwoofer to modulate the air pressure in the membrane housing for each eye. A Motorola MPX5010DP pressure sensor provides feedback on the current pressure differential between ambient and our membrane housing, thus our system no longer uses power draining cameras for

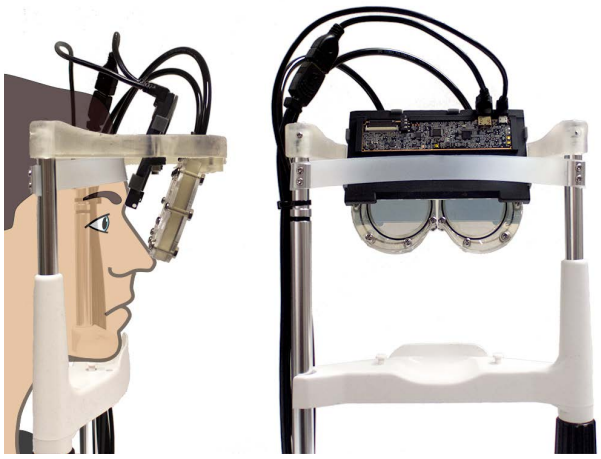

Fig. 2. Side and front view of current prototype with illustrated user to show the form factor.

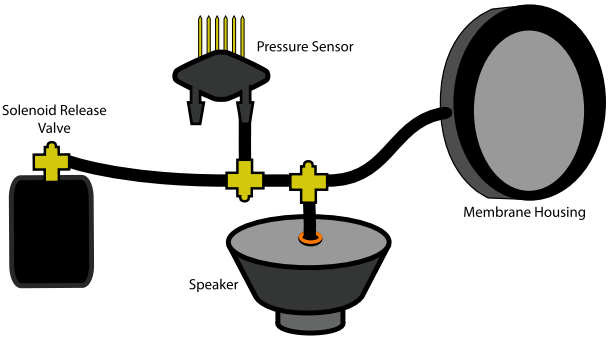

Fig. 3. The vacuum system of our display prototype.

pressure control. A PeterPaul 72B11DGM 12/DC solenoid valve allows for re-pressurizing the system as needed for leak correction, which in our observation typically occurs during continuous operation about thrice an hour. All pressure modules are connected with SMC Pneumatics $\frac{1}{4}$" OD Tubing, one touch fittings, and T-junctions as seen in Figure 3.

We control the vacuum system with an Arduino Teensy 3.6 microcontroller, which uses a software PID controller to hold the membrane at the target depth based on the sensory inputs. A WGCD L298N Dual H Bridge DC Stepper Module with a 12V 5A DC power supply drive the speakers.

4. Analysis

In this section we will evaluate the qualities of our display comparing them to studies of human performance.

Focal Range: Human eye focal range varies by subject and changes with age. The mean focal range of a 10 year old is 13.4 diopters, while at 55 years a mean of 1.3 diopters is reported by [23]. Our display is capable of matching the mean focal range of a 20 year old, or 11 diopters with a focus between 10 cm and optical infinity (represented by 800 cm in all measurements below).

Field of View: The human monocular visual field extends 60° nasally, 60° superiorly, 100° temporally and 70° inferiorly. [24] With a target monocular field of view of 160°

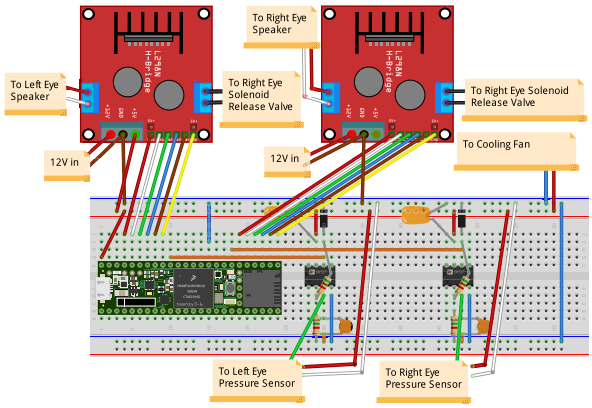

Fig. 4. Circuit design for controlling the speakers.

horizontal and 130° vertical, our display prototype exhibits a 75.5° horizontal and 70.6° vertical field of view.

Eye box: A display must be able to generate an eye box capable of entering the pupil as the eye moves around the visual field with some additional tolerance for imprecise alignment and variations of human anatomy. Most displays target a 20 mm by 20 mm eye box [25] [26].

Our eye box measurement was performed by attaching a camera to a 2 axis linear stage and evaluating the images captured. For an eye relief of 40 mm from the membrane, the eyebox for a 10 cm focal depth is 40 mm horizontal, 20 mm vertical. For all other depths, it is 30 mm horizontal and 20 mm vertical.

Focal Latency: Human accommodation response has several defining characteristics: a latent period of around 300 ms, a main focus adjustment period determined by the distance of focus change, and a settling period where minor corrections are made until a state of micro fluctuations near the target is reached [27].

Our method of measurement used a GoPro Hero 4 camera at 240 frames/second to record the response of the membrane. We indicated the start by displaying a different pattern before sending the new depth signal. In all cases, our display exhibited an initial latent period less than our sampling period of 4.16 ms. Our display also exhibited an initial main focus adjustment period followed a settling period. The main focus adjustment period of our display demonstrated a mean velocity of 55 diopters/second. Mean time for the initial adjustment was 139.5 ms with a maximum of 200 ms. The settling period exhibited several cycles of overshoot, but came to rest in a mean of 201 ms and a maximum of 237.5 ms. Total adjustment times had a mean of 340 ms and a maximum of 438 ms. The long settling times indicate that improvements can be made with a better control algorithm.

Angular Resolution: The consensus of studies explained in [28] show the angle of resolution in the central fovea for humans to be about 0.5 min of arc or 60 cycles per degree (cpd).

To asses the spatial and angular resolution limits of our display, we evaluate Modulation Transfer Functions (MTF) of our latest prototype at various depth levels. Our measurements are based on the accepted industry standard: ISO

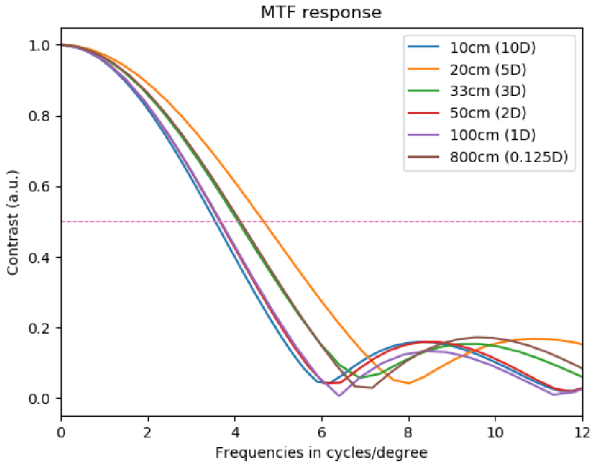

Fig. 5. MTF of our latest prototype measured at various depth levels in accordance with the ISO 122333 slanted-edge method. Our prototype produces 4-6cpd spatial resolutions at various depths.

122333 slanted-edge MTF method [29]. Figure 5 shows the MTF of our prototype at distances 20cm, 33cm, 50cm, 100cm and 800cm, all captured with a Samsung EX2F camera placed 40 mm behind the display. The camera aperture was set to f/3.9 with exposure times of 1/10 second.

First, a high resolution printed checkerboard pattern is used to measure the angles resolved per camera pixel. Then, a specific region of interest from the center of the field of view of a slanted-edge image shown on the display is used to measure the MTF of the display. Low frequencies that add noise to the measurements are filtered by thresholding the edge-spread at 10% and 90% of the measured intensities. This process is repeated for several depths, and it can be seen that the display is capable of consistently producing a spatial resolution of 4 - 6 cpd. The limitation of the spatial resolution of the display primarily comes from the available resolution of the LCD panel used for providing imagery to the eyes. In fact, the individual pixels of the LCD panel are discernible from the reflection off of the deformable beamsplitter membrane. Even a two-fold increase in the resolution of the LCD panel results in a spatial resolution of about 14 cpd, which is the current state-of-the-art for commercially available VR displays. A slightly decreasing trend of MTF is seen with increasing distance of the virtual image, and this behavior is caused by an increase in the magnification of the virtual image.

Luminance: A standard desktop display exhibits a maximum luminance around 250 cd/m^2 and mobile phones which are meant for outdoor generally have a maximum luminance between 450 and 500 cd/m^2. For each focal depth, we measured the luminace of our display prototype using a Photo Research PR-715 SpectraScan spectroradiometer with a MS-55 lens attachment. We set the aperture to 1/2° and using a 1 second exposure obtained several readings. Mean values are reported in Table 2. A decay in the measured values as the focal distance increases is expected because as

Table 2. Luminace values of display prototype given in candela per meter squared for different focal depths.

Depth	Luminance
10 cm	195 cd/m^2
20 cm	135 cd/m^2
30 cm	133.875 cd/m^2
50 cm	131 cd/m^2
100 cm	127.5 cd/m^2
800 cm	115 cd/m^2

our membrane stretches, the distance between silver particles increases causing a reduced amount of reflected light.

5. Conclusion

We have presented a new display prototype using deformable membrane beamsplitters which demonstrates a greatly improved form factor and faster depth switching time. The reported characteristics and performance show that it performs well in many categories including field of view and eye box. While the overall performance is robust, several categories demonstrate room for future work. Our future endeavors will include improving overall response and settling time via an improved PID controller, increasing the angular resolution by utilizing a display panel with a smaller pixel pitch and increased luminance for possible outdoor use.

6. Acknowledgements

The authors wish to thank Jim Mahaney who was invaluable in consulting and assisting with the physical set-up of our display prototype.

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, 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.

References

[1] D. Dunn, C. Tippets, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs, "Wide field of view varifocal near-eye display using see-through deformable membrane mirrors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, pp. 1322–1331, April 2017.

[2] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of vision*, vol. 8, no. 3, pp. 33–33, 2008.

[3] N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein, "Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays," *Proceedings of the National Academy of Sciences*, p. 201617251, 2017.

[4] P. V. Johnson, J. A. Parnell, J. Kim, C. D. Saunter, G. D. Love, and M. S. Banks, "Dynamic lens and monovision 3d displays to improve viewer comfort," *Opt. Express*, vol. 24, pp. 11808–11827, May 2016.

[5] F.-C. Huang, K. Chen, and G. Wetzstein, "The light field stereoscope: Immersive computer graphics via factored near-eye light field displays with focus cues," *ACM Trans. Graph.*, vol. 34, pp. 60:1–60:12, July 2015.

[6] J. Kim, Q. Sun, F.-C. Huang, L.-Y. Wei, D. Luebke, and A. Kaufman, "Perceptual studies for foveated light field displays," *arXiv preprint arXiv:1708.06034*, 2017.

[7] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke, "Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics," 2017.

[8] A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, p. 85, 2017.

[9] H. Hua, "Enabling focus cues in head-mounted displays," *Proceedings of the IEEE*, vol. 105, no. 5, pp. 805–824, 2017.

[10] G. Kramida, "Resolving the vergence-accommodation conflict in head-mounted displays," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 7, pp. 1912–1931, 2016.

[11] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," in *ACM transactions on graphics (TOG)*, vol. 23, pp. 804–813, ACM, 2004.

[12] O. Mercier, Y. Sulai, K. Mackenzie, M. Zannoli, J. Hillis, D. Nowrouzezahrai, and D. Lanman, "Fast gaze-contingent optimal decompositions for multifocal displays," *ACM Trans. Graph.*, vol. 36, pp. 237:1–237:15, Nov. 2017.

[13] R. Narain, R. A. Albert, A. Bulbul, G. J. Ward, M. S. Banks, and J. F. O'Brien, "Optimal presentation of imagery with focus cues on multi-plane displays," *ACM Transactions on Graphics*, vol. 34, pp. 59:1–12, Aug. 2015. To be presented at SIGGRAPH 2015, Los Angeles.

[14] X. Hu and H. Hua, "High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics," *Optics express*, vol. 22, no. 11, pp. 13896–13903, 2014.

[15] D. Lanman and D. Luebke, "Near-eye light field displays," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 220, 2013.

[16] F.-C. Huang, D. Luebke, and G. Wetzstein, "The light field stereoscope," *ACM SIGGRAPH Emerging Technologies*, p. 24, 2015.

[17] K. Akşit, J. Kautz, and D. Luebke, "Slim near-eye display using pinhole aperture arrays," *Applied optics*, vol. 54, no. 11, pp. 3422–3427, 2015.

[18] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," in *ACM SIGGRAPH 2014 Emerging Technologies Booth 203*, ACM, 2014.

[19] C. Jang, K. Bang, S. Moon, J. Kim, S. Lee, and B. Lee, "Retinal 3d: Augmented reality near-eye display via pupil-tracked light field projection on retina," *ACM Trans. Graph.*, vol. 36, pp. 190:1–190:13, Nov. 2017.

[20] S. Liu, D. Cheng, and H. Hua, "An optical see-through head mounted display with addressable focal planes," in *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, pp. 33–42, IEEE, 2008.

[21] G. W. R. Konrad, E.A Cooper, "Novel optical configurations for virtual reality: Evaluating user preference and performance with focus-tunable and monovision near-eye displays.," *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'16)*, 2016.

[22] K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke, "Near-eye varifocal augmented reality display using see-through screens," *ACM Trans. Graph. (SIGGRAPH)*, no. 6, 2017.

[23] A. Duane, "Studies in monocular and binocular accommodation with their clinical applications," *American Journal of Ophthalmology*, vol. 5, no. 11, pp. 865–877, 1922.

[24] P. J. Savino and H. V. Danesh-Meyer, *Color Atlas and Synopsis of Clinical Ophthalmology – Wills Eye Institute – Neuro-Ophthalmology*. Lippincott Williams & Wilkins, 2012.

[25] O. Cakmakci and J. Rolland, "Head-worn displays: a review," *Journal of display technology*, vol. 2, no. 3, pp. 199–216, 2006.

[26] K. Tsurutani, K. Naruse, K. Oshima, S. Uehara, Y. Sato, K. Inoguchi, K. Otsuka, H. Wakemoto, M. Kurashige, O. Sato, *et al.*, "Optical attachment to measure both eye-box/fov characteristics for ar/vr eyewear displays," in *SID Symposium Digest of Technical Papers*, vol. 48, pp. 954–957, Wiley Online Library, 2017.

[27] C. M. Schor and S. R. Bharadwaj, "A pulse-step model of accommodation dynamics in the aging eye," *Vision research*, vol. 45, no. 10, pp. 1237–1254, 2005.

[28] F. W. Weymouth, "Visual sensory units and the minimal angle of resolution," *American journal of ophthalmology*, vol. 46, no. 1, pp. 102–113, 1958.

[29] P. D. Burns, "Slanted-edge mtf for digital camera and scanner analysis," in *Is and Ts Pics Conference*, pp. 135–138, SOCIETY FOR IMAGING SCIENCE & TECHNOLOGY, 2000.