Light-efficient augmented reality display with steerable eyebox

M. KIVANC HEDILI,* BURAK SONER, ERDEM ULUSOY, AND HAKAN UREY
Optical Microsystems Laboratory, Koç University, Rumelifeneri Yolu, 34450 Sarıyer, Istanbul, Turkey
*mhedili@ku.edu.tr

Abstract: We present a novel head-mounted display setup that uses the pinhole imaging principle coupled with a low-latency dynamic pupil follower. A transmissive LCD is illuminated by a single LED backlight. LED illumination is focused onto the viewer’s pupil to form an eyebox smaller than the average human pupil, thereby creating a pinhole display effect where objects at all distances appear in focus. Since nearly all the light is directed to the viewer’s pupil, a single low-power LED for each primary color with 0.42 lumens total output is sufficient to create a bright and full-color display of 360 cd/m² luminance. In order to follow the viewer’s pupil, the eyebox needs to be steerable. We achieved a dynamic eyebox using an array of LEDs that is coupled with a real-time pupil tracker. The entire system is operated at 11 msec motion-to-photon latency, which meets the demanding requirements of the real-time pupil follower system. Experimental results effectively demonstrated our head-mounted pinhole display with 37° FOV and very high light efficiency, equipped with a pupil follower with low motion-to-photon latency.

© 2019 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Displays consume the largest amount of power in smart devices but only <0.1% of the display light actually enters through the viewer’s pupil and nearly all the light is wasted. Similarly, for AR/VR headsets, light efficiency is extremely poor [1]. Another challenge in AR/VR headsets is the requirement for complicated optics and magnifier lenses to relay images from the microdisplay to the retina due to the limited refractive power of the human eye [2–4]. Maintaining large field-of-view and large eyebox using conventional optics is a major challenge [5–7]. As an alternative, microlens array [8,9] and pinhole array based [10] light field displays have been demonstrated. Such solutions provide compactness but they suffer in terms of light efficiency.

To increase the light efficiency, what is known as the Maxwellian view has been used by many groups [11]. In this configuration, since the light source is imaged at the pupil, most of the light is captured by the eye, leading to a more light-efficient display. Many variations of this technique have been demonstrated over the years. Inami et. al [12] showed a stereoscopic AR display using birdbath optics. High brightness potential of the Maxwellian view is demonstrated by Beer et. al. [13]. The main disadvantage of the Maxwellian view is that it provides a small eyebox. If the eye rotates slightly, the focus spot is blocked by the iris and no light is delivered to the retina. To extend the eyebox, Takahashi et. al. focused the light source to the rotation center of the eye at the cost of reduced FOV, solving the small eyebox problem [14]. If one can create a small eyebox that is also steerable, the display system would be extremely light efficient while also providing a large eyebox. In holographic displays the eyebox can be steered by computing an appropriate hologram and multiple eyeboxes can be produced using HOEs but the image quality and FOV is limited [15]. Travis et. al. suggested a waveguide-based display which steers the eyebox; the drawback is that it requires mechanical motion [16].

In Maxwellian view, the light source is focused onto the eye pupil and the display is optically imaged by the eye lens onto the retina. This means the display should be placed...
further than the near-point of the eye or it should be re-imaged to an appropriate distance for the eye to accommodate. This is a challenge for compact optical designs that is required for head-mounted displays. In this work, we developed a method to achieve a large eyebox display by adapting the well-known pinhole camera principle [17].

Unlike the Maxwellian view, the pinhole camera does not require any imaging components to create a sharp image at the retina. An image on a microdisplay that is closer than the near-point becomes resolvable to the eye and, therefore, complicated relay optics are not required. This is illustrated in Fig. 1 where far objects can be focused onto the retina but objects closer than the near-point cannot form a sharp image. A sufficiently small pinhole aperture limits the extent of the rays and creates a sharp image at the back plane. By employing this principle, we can create sharp images at the retina even if the display is closer than the near-point of the eye.

The small eyebox provides focus-free images but it is also a limitation. Assuming the pinhole is initially aligned with the center of the pupil, if the eye rotates more than half the pupil size, the iris blocks all the light, prohibiting image formation on the retina. In the following sections we discuss how to create a pinhole display using single LED illumination and how to steer the eyebox to follow the pupil as it rotates without any moving parts. This is the first report of a low-latency, real-time pupil follower display system.

2. Extended eyebox using the pupil follower

We propose to achieve pinhole effect in an HMD configuration by using a point light source that is imaged on the viewer’s pupil. Since a very small portion of the eye lens is used, effective aperture is reduced and the depth of field is extended. Due to this large depth of field, everything is in focus regardless of the accommodation distance of the eye. Figure 2(a) illustrates our proposed system. A transparent microdisplay can be placed anywhere on the converging beam path and a sharp image forms on the retina even though there is no magnifier lens between the microdisplay and the eye.

The focus point is the eyebox of this system, thus the system has a very small eyebox. If the eye moves more than half the pupil size in any direction, the focused point is blocked by the iris, which means the image is lost. To extend the eyebox, light sources at different
positions should be turned on according to the pupil position at a given time. To do this we use an array of LEDs that is driven synchronously with a pupil tracker. The pupil tracker algorithm calculates the pupil position using image processing and turns on the appropriate LED as illustrated in Fig. 2.

![Diagram of pupil follower display system](image)

**Fig. 2.** Pupil follower display system operation. (a) The illumination beam is focused at the eye pupil by a focusing lens, which effectively creates pinhole imaging and small eyebox. If the eye rotates slightly no light enters the eye pupil, thus the image is lost. As a solution we propose using multiple light sources that are spaced closer than the minimum pupil size so that all pupil positions can be addressed. (b) We synchronize the array of light sources with a pupil tracker and turn on only the required source so that the display can be seen in an extended eyebox.

Traditional backlit displays diffuse the light on the entire hemisphere but the eye captures only a small portion of it, so most of the light is wasted. In our proposed system, only a single LED is turned on at a given time and the eye captures almost all the light emanating from it, making the system very light efficient. As this architecture is compact and energy efficient, it is suitable for HMDs.

We used a simple pupil tracker algorithm based on the center of mass calculation on dark pupil images under IR illumination [18]. We ran the algorithm on Beaglebone Black and used the programmable real time units (PRU) on board to do the LED switching [19]. Similar pupil following schemes can be built using laser scanning pico-projector based displays but mechanical motion is needed to follow the pupil, which is much slower compared to turning on an LED using real-time computing units. As a result, motion-to-photon latency is minimized in our system by utilizing PRUs.
Figure 2(b) illustrates the timing diagram of the pupil follower system. The pupil tracker algorithm running on Beaglebone is implemented using C language and it finds the pupil position in 10ms after the frame is captured. The tracker program sends this information to PRUs, which turn on the appropriate LED within 1ms. In other words the whole operation takes about 11ms, which means the pupil follower system has the potential to run at 90fps. In our setup the camera speed was 30fps so it is camera limited. Figure 3 shows the pupil tracker camera images for two different pupil positions. The red dot at the pupil center shows the calculated pupil position and a red circle representing the pupil is drawn. On the right the corresponding LED is turned on in real time by the PRU.

Since the pupil position is monitored continuously in our system, the content on the LCD can be shifted based on the gaze direction. The dynamic shift of the content can be used to give binocular disparity and depth cues to the user for 3D perception.

![Figure 3](image)

**3. Design and experimental work**

On the LED array, we use Lumex QuasarBrite-0404 LEDs as the light source because it has three colors (RGB) and small size (1 mm x 1 mm). Due to the physical limitations, LEDs can be placed 2mm apart in each direction. The array consists of 5 rows and 7 columns of RGB LEDs, so there are 105 LEDs at 35 different locations.

The LED array is imaged at the pupil plane with unit magnification; therefore the instantaneous eyebox size is 1mm when the information content in the image is low-pass. Assuming typical pupil size is about 3mm in diameter, the distance between the images of the LEDs should be less than 3mm so that all pupil positions can be addressed without dark transition regions between the LEDs. The optical system that images the LED array is designed using Zemax software. Figure 4(a) shows the Zemax model of the system. Light from the LED array is collimated using a 75mm-focal length lens and then it is reflected from a 45-degree folding mirror. The reflected light is focused by a 75mm-focal length lens, which is relayed to the pupil plane using a beam splitter to make it function as an augmented reality
display. The aperture of the focusing lens determines the field of view (FOV) of the display. In the setup we used a 2-inch diameter lens, which gives about 37° FOV. In any HMD optical design, there is an inherent trade-off between size and FOV. Large diameter lenses are required for large FOV. However, for reduced weight and form factor, Fresnel lenses, thin holographic optical elements (HOE), or any special diffractive component that creates the converging beam can be used to replace the thick standard lenses. Shorter focal length lenses and smaller eye relief can also be selected to miniaturize the optics.

Figure 4(b) shows the optical setup on the bench. The transmissive microdisplay is positioned right before the focusing lens. The setup is built for the right eye, which looks into the beam splitter. The beam splitter and the mirror fold the whole system towards the ear, which gives the natural shape of eye glasses. Figure 4(d) shows the LED array. LEDs are 1x1mm in size and they are placed 2mm apart in each direction and each LED unit have RGB LEDs in them, as seen in Fig. 4(c). Figure 4(e) shows the image of the LED array at the pupil plane, with all LEDs turned on to show the extended eyebox. The distance between the images of the LEDs is measured as 2mm, which results in an extended eyebox size of 14x10mm.
There are various liquid crystal displays available on the market. Selecting the one with minimum pixel size to achieve maximum resolution seems intuitive in the beginning. To select the optimum pixel size, we ran physical optics simulations based on the angular spectrum approach [20]. We simulated pixel sizes from 60μm to 800μm and calculated the resulting spot size at the retina. As seen in Fig. 5(a), for small pixel sizes we see diffraction spread and for large pixel sizes the spot size converges to values predicted by geometrical optics, which increases linearly with pixel size. According to our simulations 250μm pixel size yields the smallest spot size at the retina, hence the best resolution. The typical spot size criterion is the point where the encircled energy reaches 0.865 of the normalized encircled energy, which corresponds to 1/e² diameter for Gaussian beams. We used full-width at half-maximum intensity (FWHM) as the spot criterion as it defines a smoother view with denser pixel count [21,22]. Figure 5(b) shows the simulated cross-section of the spot at the retina for 250μm pixel size.
Fig. 6. Demonstration of the system outdoors. We can show bright images with a single low power LED as all the light from the LCD is captured by the eye. FOV is roughly circular, 37°.

Fig. 7. Demonstration of the system for an indoor navigation scenario. The full version can be viewed in Visualization 3.
Having calculated the optimum pixel size for our LCD, we used a back-lit LCD module that has the closest pixel size to 250μm. We removed the back light unit along with diffuser films to get the bare LCD. We placed this bare LCD almost touching to the focusing lens to maximize our FOV and demonstrated our pinhole imaging display. Figures 6 and 7 show the experimental results under bright ambient light. The image in Fig. 6 is captured outside in a moderately sunny day. The prototype we built has a measured luminance of 360 cd/m² using an LED with 0.42 lumen total output. Due to the super efficient nature of our display we can show bright images with a single low power LED even in outdoor settings.

Since the effective pupil size is very small, resolution of the display is reduced as seen in Figs. 6 and 7. To quantify this degradation in resolution we simulated and experimentally verified the modulation transfer function (MTF) of the system as seen in Fig. 8. The LCD we used has 200μm pixel pitch and is placed 75mm away from the eye, which means the maximum frequency that can be displayed with this LCD is 3.3 cyc/deg. Using the angular spectrum approach, we simulated a range of spatial frequencies that can be represented by integer number of pixels of our LCD as marked in Fig. 8(a). Fringe contrast gives the MTF value at that frequency. To obtain the continuous MTF curve, third order polynomials are fit.
between the simulated data points. To verify the simulations, we displayed the simulated frequencies on the LCD and captured the images with a camera, as shown in Fig. 8(b). Although some modulation is visible for the highest frequency, which is marked as 1 in Fig. 8(b), we can say that the cutoff frequency is about 2 cyc/deg, which is good enough to display simple symbols that is required in AR applications.

4. Conclusion

We successfully demonstrated a new head-mounted, near-to-eye display architecture based on the well-known pinhole imaging principle. This architecture alleviates the problem of the human eye being unable to directly resolve near-to-eye displays. The proposed system counters the small eyebox problem arising from the pinhole imaging approach by extending the effective eyebox with a pupil follower with only 11msec motion-to-photon latency. The pupil follower also creates a light-efficient display since nearly all the light coming out of the LCD enters the eye, making it suitable for mobile applications. The prototype achieved 37° circular FOV with a luminance of 360 cd/m2 using an LED with only 0.42 lumen output. The MTF cutoff frequency for the display system was measured to be approximately 2 cyc/deg. Good quality experimental results for a real use case were observed on a prototype setup with visually undetectable motion-to-photon latency. Pupil follower concept can be applicable to other HMD architectures such as holographic displays and foveated displays, therefore it will be of interest to a broader audience.

Funding

European Research Council (ERC) under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC advanced grant agreement (340200).

References

1. J. P. Rolland, K. P. Thompson, A. Bauer, H. Urey, and M. Thomas, “See-through head-worn display (HWD) architectures.” Handbook of Visual Display Technology (Springer, 2011), pp. 2929–2961.
2. J. F. Koretz and G. H. Handelman, “How the human eye focuses,” Sci. Am. 259(1), 92–99 (1988).
3. G. Wysecki and W. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, (John Wiley & Sons, 1982).
4. J. F. Koretz and G. H. Handelman, "Modeling age-related accommodative loss in the human eye," Math. Model. 7(5–8), 1003–1014 (1986).
5. J. Yang, W. Liu, W. Lv, D. Zhang, F. He, Z. Wei, and Y. Kang, “Method of achieving a wide field-of-view head-mounted display with small distortion,” Opt. Lett. 38(12), 2035–2037 (2013).
6. W. Song, D. Cheng, Z. Deng, Y. Liu, and Y. Wang, “Design and assessment of a wide FOV and high-resolution optical tiled head-mounted display,” Appl. Opt. 54(28), E15–E22 (2015).
7. G. E. Romanova, A. V. Bakholdin, and V. N. Vasilyev, “Optical schemes of the head-mounted displays,” Proc. SPIE 10374, 103740I (2017).
8. S. Choi, Y. Takushima, and S. W. Min, “Improvement of fill factor in pinhole-type integral imaging display using a retroreflector,” Opt. Express 25(26), 33078–33087 (2017).
9. C. Yao, D. Cheng, T. Yang, and Y. Wang, “Design of an optical see-through light-field near-eye display using a discrete lenslet array,” Opt. Express 26(14), 18292–18301 (2018).
10. K. Akrit, I. Kautz, and D. Luebke, “Slim near-eye display using pinhole aperture arrays,” Appl. Opt. 54(11), 3422–3427 (2015).
11. G. Westheimer, “The maxwellian view,” Vision Res. 6, 669–682 (1966).
12. M. Inami, N. Kawakami, T. Maeda, Y. Yanagida, and S. Tachi, “A stereoscopic display with large field of view using Maxwellian optics,” in Proceedings of The 7th International Conference on Artificial Reality and Tele-Existence (Tachi Lab, 1997), pp. 71–76.
13. R. D. Beer, D. I. MacLeod, and T. P. Miller, “The Extended Maxwellian View (BIGMAX): A high-intensity, high-saturation color display for clinical diagnosis and vision research,” Behav. Res. Methods 37(3), 513–521 (2005).
14. T. Hideya, and S. Hirooka, “Stereoscopic see-through retinal projection head-mounted display,” Proc. SPIE 6803, 68031N (2008).
15. J. H. Park and S. B. Kim, “Optical see-through holographic near-eye-display with eyebbox steering and depth of field control,” Opt. Express 26(21), 27076–27088 (2018).
16. A. R. Travis, L. Chen, A. Georgiou, J. Chu, and J. Kollin, “Wedge guides and pupil steering for mixed reality,” J. Soc. Inf. Disp. 26(9), 526–533 (2018).
17. P. Frank L., and L. S. Pedrotti, Introduction to Optics, 3rd edition (Prentice Hall, 1993).
18. C. H. Morimoto and M. R. Mimica, “Eye gaze tracking techniques for interactive applications,” Comput. Vis. Image Underst. 98(1), 4–24 (2005).
19. G. Coley, Beaglebone Black System Reference Manual (Texas Instruments, 2013).
20. J. W. Goodman, Introduction to Fourier Optics (Roberts and Company Publishers, 2005).
21. A. E. Siegman, “How to (maybe) measure laser beam quality,” in Diode Pumped Solid State Lasers: Applications and Issues (Optical Society of America, 1998) paper MQ1.
22. W. J. Smith, Modern Optical Engineering (Tata McGraw-Hill Education, 1966).