applied optics

Slim near-eye display using pinhole aperture arrays

KAAN AKŞIT,* JAN KAUTZ, AND DAVID LUEBKE

NVIDIA Research, 2700 San Tomas Expy, Santa Clara, California 95050, USA
*Corresponding author: kaksit@nvidia.com

Received 4 December 2014; revised 12 March 2015; accepted 12 March 2015; posted 13 March 2015 (Doc. ID 228663); published 9 April 2015

We report a new technique for building a wide-angle, lightweight, thin-form-factor, cost-effective, easy-to-manufacture near-eye head-mounted display (HMD) for virtual reality applications. Our approach adopts an aperture mask containing an array of pinholes and a screen as a source of imagery. We demonstrate proof-of-concept HMD prototypes with a binocular field of view (FOV) of 70° × 45°, or total diagonal FOV of 83°. This FOV should increase with increasing display panel size. The optical angular resolution supported in our prototype can go down to 1.4–2.1 arcmin by adopting a display with 20–30 μm pixel pitch.

OCIS codes: (110.1220) Apertures; (110.1758) Computational imaging; (330.1400) Vision - binocular and stereopsis.

http://dx.doi.org/10.1364/AO.54.003422

1. INTRODUCTION

Design of a near-eye head-mounted display (HMD) includes multiple challenges such as achieving a wide field of view (FOV), high optical angular resolution, slim form factor, light weight, and ease of replication. The use of any reflective, refractive, or diffractive components entails a trade-off in one of these challenges.

Researchers have recently demonstrated various ideas to overcome these challenges. Lanman and Luebke developed an HMD [1] using microlenses in front of a microdisplay for virtual reality (VR) applications, which shares a similar principle with the existing integral imaging displays/cameras. Their HMD also addressed the problem of the accommodation-vergence conflict through a lightfield approach. Maimone and Fuchs demonstrated a unique see-through HMD [2], which uses a stack of LCD panels. The image in this case is formed computationally through modulating each panel with a set of optimized patterns. Most recently, Maimone *et al.* introduced another see-through HMD approach [3], which adopts a sparse array of point light sources with a single spatial light modulator (SLM). This approach achieves a wide binocular FOV of 110° in their prototype.

A recent trend in HMD design is the use of mobile device displays from phones and tablets, which provides very high density, typically around ~300–600 pixels per inch (ppi), and large display dimensions up to around 6 in. diagonal. Some HMD manufacturers take a classical magnifier approach in their design to provide a practical VR experience with mobile devices, often using the same screen for both eyes. Although this design provides simplicity, it comes with a bulky form factor due to the large focal length of the used lens(es), and it also introduces optical image distortion problems caused by the refractive elements.

In this paper, we propose a novel method to build a lightweight, slim near-eye HMD for VR. Our approach pairs a conventional display with an aperture array, or an array of pinhole apertures.

Pinholes have been used for various purposes in the display domain. Huang *et al.* demonstrated a visual aberration correcting display [4], which works with pinhole parallax barriers and microlenses. Akşit *et al.* used pinholes with wavelength selective filters in front of the eye to address the accommodation-vergence conflict in stereoscopic displays [5]. Both cases have a large viewing distance. Sprague *et al.* embed a single pinhole inside a contact lens to avoid bulky imaging optics in HMD design [6]. Song *et al.* propose a lightfield HMD [7] using free-form optics combined with a pinhole aperture array. However, to the best of our knowledge, ours is the first system to use pinhole aperture arrays as a stand-alone image relay layer in an HMD configuration.

We believe our technique is a basis for a framework to build simple HMDs with a small number of components. Our proposal's main contributions to the domain of HMD research are itemized as follows:

- We provide a wide FOV, and high optical resolution HMD design, almost matching the optical capabilities of a human eye.
- We provide the simplest optical design to build HMDs with a small number of optical components.
- We provide a thin-form-factor, cost-effective, lightweight solution, unlike many other designs in industry and research.

• We provide an easily replicable design for researchers interested in lightfield and VR research.

2. PROPOSAL

Our proposal is sketched in Fig. 1, in which a pinhole aperture array is placed at a distance to the cornea of an eye, with a display screen at a fixed distance beyond the pinhole aperture array. When designing such a HMD, we must first choose the distance of the screen from the eye(s). Given the generally accepted 1 arcmin angular visual acuity threshold of a human eye, we propose to place the screen at the tip of the nose, i.e., at 45–50 mm distance from the eye. At this distance, the smallest resolvable feature corresponds to 10–15 μm—a pitch size the display industry should be able to provide in the near future as a standard.

A human eye cannot focus at a distance as small as 45–50 mm. A pinhole can help in this case, as it will bound the cone of rays that enter the eye's pupil as shown in Fig. 1. In other words, each screen pixel will send an almost angularly bounded beam of light to the eye. The bundle of rays from a single pinhole will form an image on the retina, which are called retinal elemental images. In similar configurations, a FOV of a single pinhole is ∼8°–12°, in which the eye pupil diameter plays a dominant role over the FOV. By increasing the number of pinholes in our system, the FOV can be widened. We predict a full FOV with a large enough screen and with enough pinholes.

3. PINHOLE APERTURE ARRAY DESIGN

The mathematical representation of our proposed system helped us to find the correct aperture designs for our prototypes. There are two different variables in our system, which are the cornea to pinhole aperture array distance d_{ae}, and the pinhole aperture array to screen distance d_{ai}. The rest of the variables are either a direct result of the change in these variables or constant values.

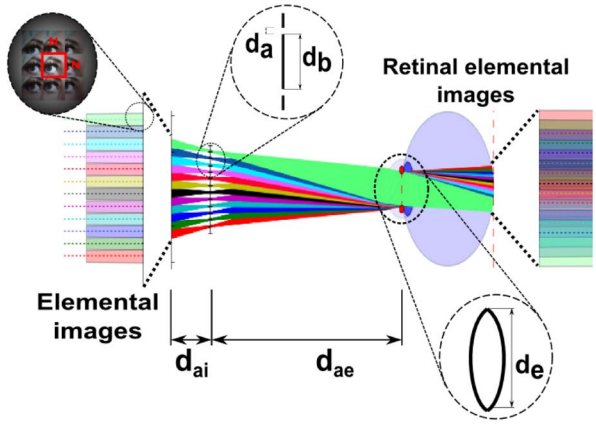

Fig. 1. 2D sketch showing the image formation on the retina by using a pinhole aperture array in front of a screen. Each elemental image consists of $N \times N$ pixels. Depending on the pinhole aperture array's distance to the screen (d_{ai}), eye pupil size (d_e), cornea to pinhole aperture array distance (d_{ae}), pinhole diameter (d_a), and the spacing of the pinholes (d_b), the retinal elemental images can be made to abut or overlap (as shown here) on the retina.

A constant in our design is the eye pupil size d_e, which is equal to the size of the eyebox in our solution. We have chosen d_e as 8 mm, which varies between 2 and 8 mm in reality [8]. Choosing an eyebox bigger than the actual size of an eye pupil will provide the freedom to compensate for change in gaze without requiring a pupil tracker. For example, a 4 mm eye pupil size will have enough freedom to gaze at different parts of the screen, when an 8 mm eyebox is designed.

As highlighted earlier, we have chosen $d_{ae} + d_{ai}$ to be 45–50 mm. Note that the choice of d_{ae} and d_{ai} directly affects the spacing of the pinholes d_b as shown in Fig. 2(a). Using Fig. 2(a), the spacing of the pinholes can be formalized as

$$d_b \geq \frac{d_e d_{ai}}{d_{ai} + d_{ae}}. \tag{1}$$

The size of a single elemental image d_1 can be calculated using Fig. 2(b) as

$$d_1 = \frac{d_e(m + d_{ai})}{d_{ae} - m}, \qquad \text{where} \quad m = \frac{d_a d_{ae}}{d_a + d_e}. \tag{2}$$

Equation (2) contains the pinhole size d_a as a variable; d_a is directly correlated with the angular resolution of the system. d_a has to be selected in a way that the maximum possible angular resolution is provided. According to Fourier optics, the angular

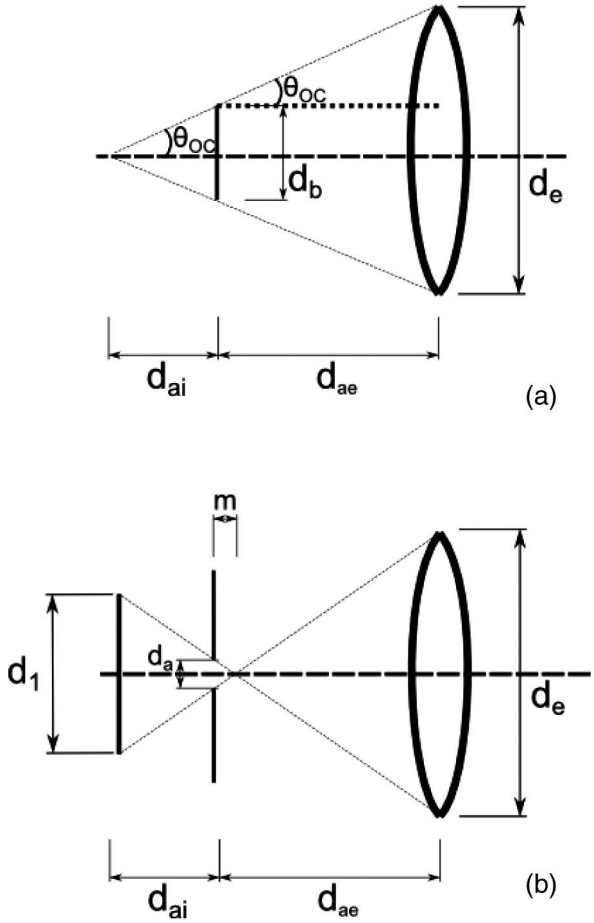

Fig. 2. 2D sketch showing (a) effect of the changing d_{ae} and d_{ai} over d_b and (b) effect of changing d_a over a single elemental image's size (d_1) on screen. Both sketches are valid for a fixed pupil size (d_e).

resolution of such a system can be calculated using the Rayleigh resolution formula ($\theta \approx \lambda/d_a$); on the other hand, geometric optics predicts a growing angular spot size on the cornea with increasing d_a ($\theta \approx d_a/(2d_{ai})$). The case in which two estimations meet is the optimal pinhole size in terms of angular resolution, which is given as

$$d_a \approx \sqrt{2\lambda d_{ai}}. \quad (3)$$

Note that Eq. (3) is valid as long as paraxial approximation holds. It is beneficial to select d_a accordingly in order to achieve the best possible angular resolution. A sample design space is scanned as in Fig. 3 using the derived set of equations [Eqs. (1)–(3)]. Figure 3 provides an atlas for possible designs with $d_e = 8$ mm and optimum d_a derived using Eq. (3).

4. CONTENT CREATION

Our system requires a content creation routine similar to other lightfield displays. The first task in content creation is to find out how many pixels fit inside a single elemental image. The pixel pitch of a display can be calculated by using the ppi (n_{ppi}) value of the display. Using Eq. (2) or Fig. 3(b), a designer can estimate size of a single elemental image (d_1). Once d_1 is calculated, the amount of pixels in a single elemental image can be calculated as $N = n_{ppi}d_1/25.4$. The location and size of each elemental image on the retina plane play an essential role in finding how much each elemental image overlaps on the retina. Overlap ratios on the retina can be found through ray tracing the whole system. This information is useful for mapping any content on each elemental image. The pseudo-code to map a sample image to a single elemental image is provided as

$$i = 5; \quad j = 3$$

$$\text{elemental} = \text{image.get\_region}(i * N * o, j * N * o, N, N),$$

where i and j are the numbers of elemental images at the x and y axes of the screen, N is the size of a single elemental image, o is the result of subtracting the overlap ratio in between retinal elemental images from 100%, elemental is the result of the mapping, and image is the source image, which is desired to be projected on to the retina. In this pseudo-code, the fifth in x axis, third in y axis elemental image is filled with an N-by-N rectangular region cropped from a given image. The cropping operation started with the offset coordinates of $i * N * o$ in the x axis, and $j * N * o$ in the y axis.

5. EXPERIMENT AND DISCUSSION

Each retinal elemental image can be tiled on the retina by either abutting each image with a minimum overlap, which provides higher effective resolution but requires precise eye tracking, or heavily overlapping the retinal elemental images to sacrifice effective resolution while providing a bigger self-repeating eyebox without eye tracking [3]. We provide samples from each scenario with multiple prototypes.

Prototype I: We first demonstrate a prototype with overlapping retinal elemental images as shown in Fig. 4. The prototype uses an active matrix organic light emitting display (AMOLED) found in a smartphone, with 386 ppi and a 5.7 in. diagonal screen size. The screen is positioned at

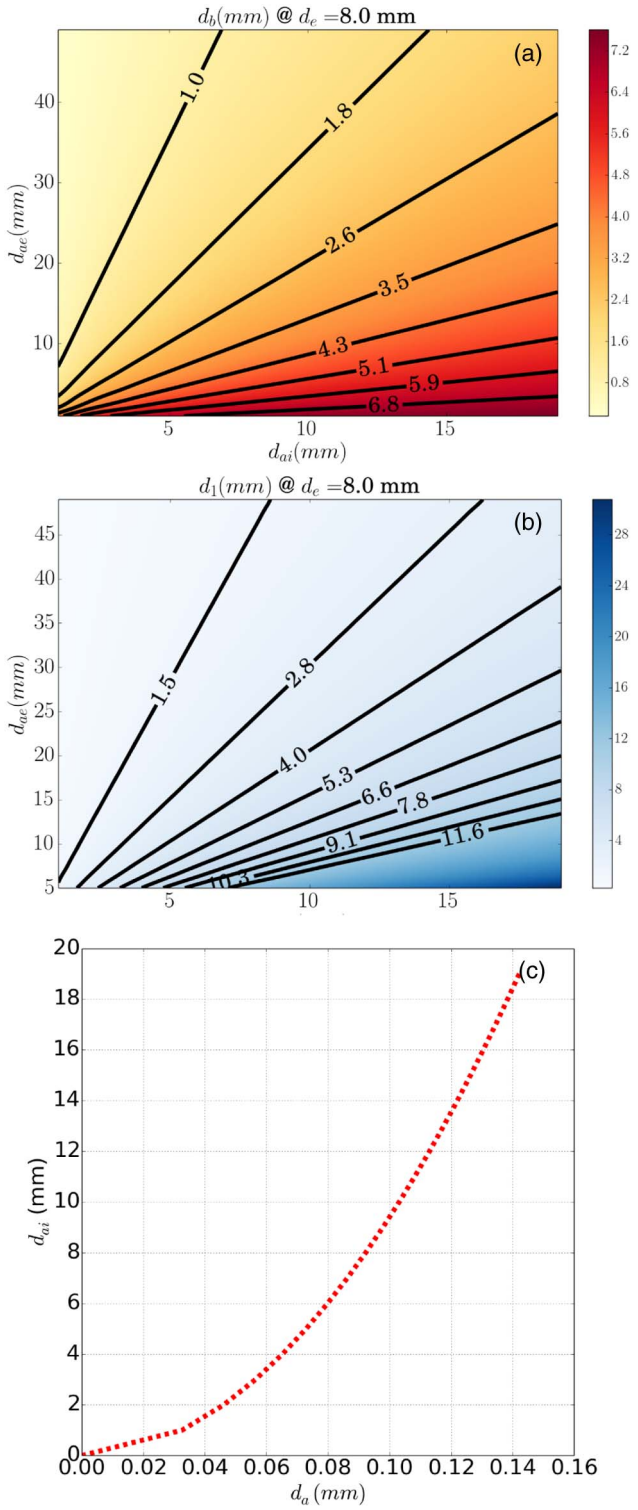

Fig. 3. Simulation results showing (a) minimum spacing between pinholes d_b in order to use screen without any crosstalk; (b) size of a single elemental image d_1, when d_a is chosen as in Eq. (3); and (c) optimum pinhole size with the changing screen to pinhole aperture array distance d_{ai}. Note that (a) and (b) are plotted using a constant eye pupil size ($d_{eye} = 8$ mm).

45–50 mm from the viewer's eye, and the distance varies roughly with the varying facial structure from user to user. The same screen is used for both eyes, with the overall system

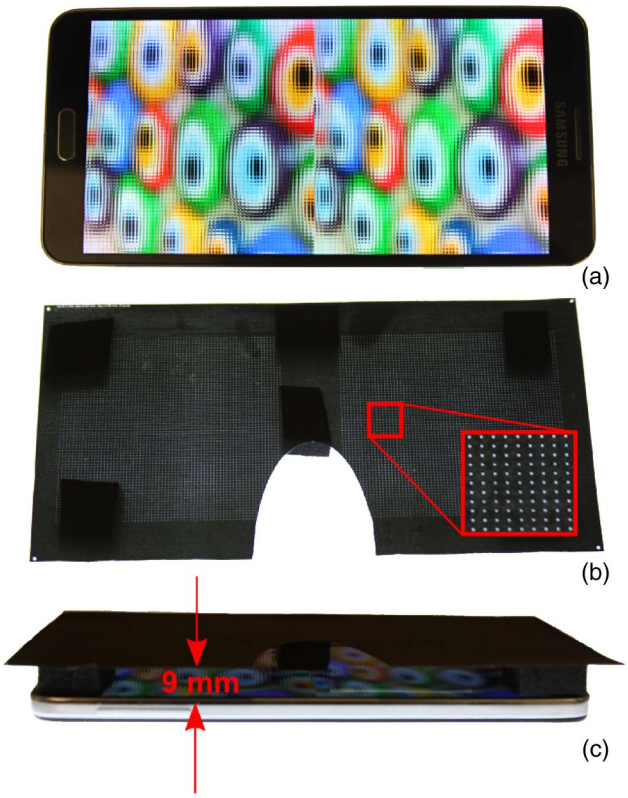

Fig. 4. Components of our system: (a) display, which is part of a mobile device, with 386 ppi and 5.7 in. diagonal screen size. (b) Aperture array with 0.1 mm pinhole diameter and 1.4 mm center-to-center spacing between pinholes. (c) Complete prototype with the thick spacers under the array, which has 9 mm thickness.

providing a binocular FOV of 70° × 45°. Our design uses 9 mm rubber spacers between the pinhole aperture arrays and the front surface of a smartphone. Note that the protective cover glass on such devices has some additional thickness, which typically varies between 0.5 and 2 mm. The overall spacing between the display and the aperture array is 9.5–11 mm.

We have manufactured our aperture arrays by photoplotting. The aperture array contains pinholes with 0.1 mm diameter and 1.4 mm center-to-center spacing. This leads to overlapping elemental images on the retina of about 76%, which decreases the overall effective resolution of the screen from 1920 × 1080 to 460 × 260.

We select the pinhole diameter by finding the angular resolution limits predicted by geometric as well as diffraction optics; the optimum pinhole diameter lies at their intersection and should provide the highest angular resolution possible. Simulations of our system using both ray tracing and Fourier optics showed this optimum diameter to be ~0.1 mm. Figure 5 shows our initial estimations of retinal spot size with this specific configuration. Retinal features (cones and rods) at the optimum pinhole diameter approach the spot size levels, which nearly matches with retinal features.

Images to be viewed through an aperture array display must first be processed computationally. The sample input content in Fig. 6(a) was rendered as in Fig. 6(b). The rendered content displayed on the screen was first observed without a pinhole

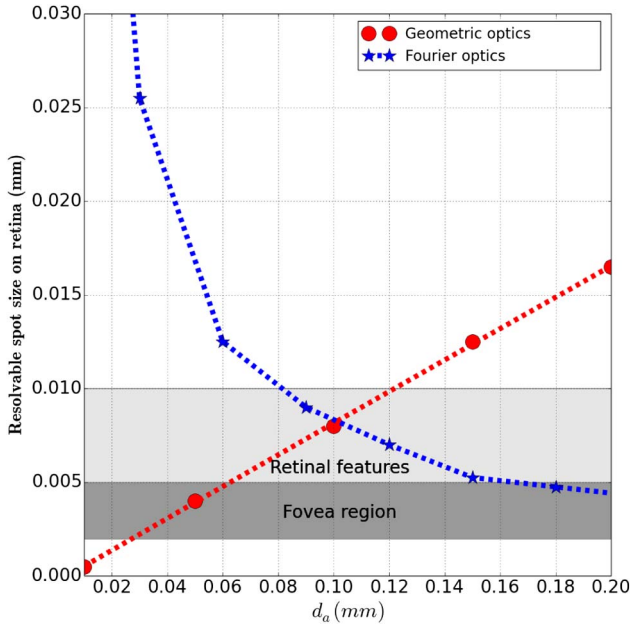

Fig. 5. Sketch showing spot size on the retina with the changing pinhole diameter. The system is simulated using a custom in-house built ray optics and Fourier optics simulator ($d_{ai} = 9$ mm, $d_{ae} = 45$ mm; see Fig. 1).

aperture array at a distance of 45 mm from pinhole aperture arrays; the observed image was blurry as expected as in Fig. 6(c). Later, the resultant image was observed with a pinhole aperture array as in Fig. 6(d). Both of the photographs were captured with a Point Grey CCD (FL3-U3-32S2C) equipped with an objective lens (YV2.8 × 2.8SA-SA2, focal length of 2.8–8 mm and F-number of 1.2), chosen to roughly match

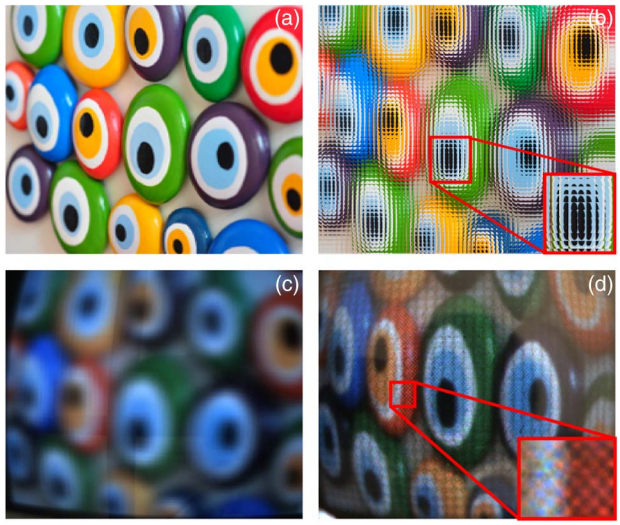

Fig. 6. (a) Sample desired image, (b) rendered content to be displayed on a screen, (c) blurry image observed on the screen without any pinhole aperture array, shot to mimic the percept of an unaided human observer (camera focal length 2.8–8 mm and F-number 1.2), and (d) image observed through our HMD prototype with a close-up region demonstrating the observability of PenTile OLED structure (386 ppi) (Media 1).

the aperture and FOV of the human eye. Figure 6(d) also provides a close-up view; note that the PenTile structure of the OLED display can be observed. Each pixel in the display has 65 μm pixel size, so resolving the PenTile structure requires 20–30 μm resolution on the screen plane or ∼1.4–2.1 arcmin of angular resolution. This resolvable spot size on the screen plane matches well with our initial resolvable spot size estimate on the retina, which can be found in Fig. 5.

In informal subjective tests, viewers find our prototype to have satisfactory resolution and FOV. The subjects also indicate that the brightness level of the display is sufficient when screen brightness is set to maximum.

Prototype II: The overall screen resolution can be improved by changing the center-to-center spacing between pinholes, and the distances between the cornea, display, and aperture. We also demonstrate alternative designs. The second design has a thickness of 35 mm, 0.3 mm pinhole diameter, and 20 mm cornea-to-aperture distance. Figure 7 shows the desired target image, and the two images as observed through our two different prototypes. Figure 7(c) shows the observed image through another, thicker prototype with less overlap between retinal elemental images (30%). The overall effective resolution is enhanced in this second sample design, since the content displayed on the screen contains fewer overlapping pixels representing the same image region.

The toning artifacts caused by overlapping retinal elemental images become more visible as overlaps between regions are decreased in Fig. 7(c). The dark edges are a result of vignetting; the brightness variation through the image is a result of overlapping regions on the retina plane. It should be possible to adjust the brightness level according to photometric values to have a smoother observed image. Note that this would require precise eye pupil size and gaze information to render the content correctly as proposed in [3].

Our proposed approach can support overall thickness smaller than 9 mm. Existing mobile displays generally come with a protective glass, which has a typical thickness around 0.55–2 mm. Such distances can be enough to get sharp images at the final plane. Thus, a pinhole aperture array can also be placed in direct contact with the front surface of a display in many cases. Figure 8 shows a smartwatch screen with 277 ppi observed through a single pinhole from 10 mm distance. With such a thickness setting, high ppi display, and

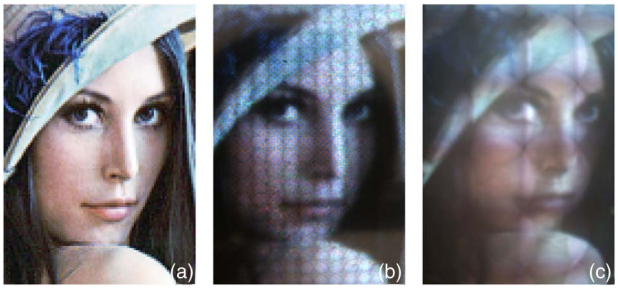

Fig. 7. Picture and two photographs showing (a) desired target image, (b) image observed through an HMD prototype with overlap ratio of 76% between retinal elemental images, and (c) image observed through another HMD prototype with a lower overlap of 30%.

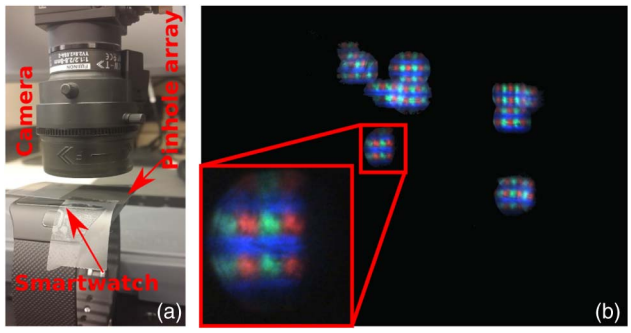

Fig. 8. Two photographs showing (a) smartwatch's screen with 277 ppi observed from 10 mm distance through direct contact pinholes with 0.1 mm diameter and 1.4 mm center-to-center spacing, and (b) what a human observer sees through the pinholes (camera's focal length range, 2.8–8 mm; F-number, 1.2). A single pinhole's image is magnified in the lower left. The content shown on the display consists of sparsely placed white pixels.

correct sized pinholes, it is possible to build a very thin near-eye HMD.

Prototype III: We improved the prototypes discussed previously with a new prototype using the same display. The general aim in this extended work is to provide a better user experience overall, and to enhance the capabilities of the previous prototypes.

The new prototype has a different pinhole pattern as in Fig. 9(a), which contains pinholes with a diameter size of

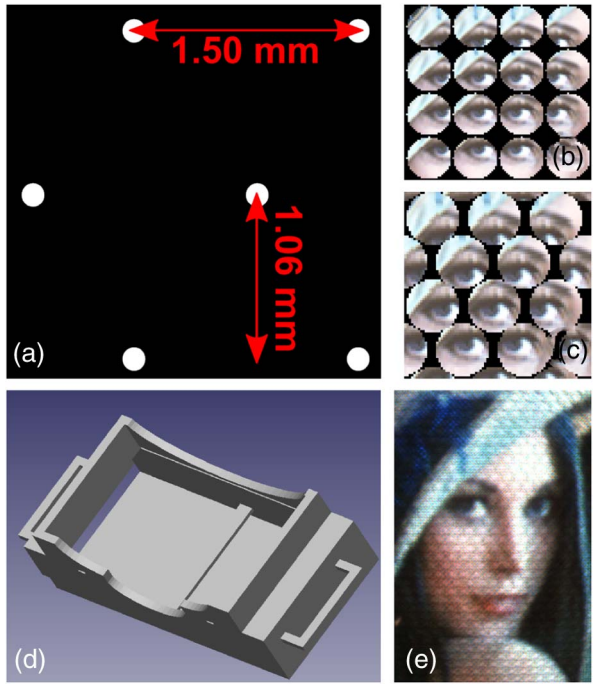

Fig. 9. (a) Sketch showing the pinhole pattern of the new prototype. (b) Sample content on display plane from previous demonstrators. (c) Sample content on display plane from the new prototype. (d) Images showing housing design for the new prototype from different perspectives. (e) Photograph showing an image captured from the new prototype (Media 2).

0.15 mm. Thus, it provides ×2.25 brightness than the previous state. The center-to-center spacing of pinholes is 1.50 mm at the vertical axis, and the horizontal row-to-row spacing between pinholes is 1.06 mm in this configuration. The rest of the distances were kept the same with previous prototypes.

The main intention in this type of design is to provide a denser pinhole array pattern, which decreases the visible effect of the pinhole pattern, and uses more pixels of the display. Such a design requires a different arrangement of the content: in the previous state, the visible part of the content on the display through pinholes was as in Fig. 9(b); however, the new design requires sample content as shown in Fig. 9(c).

Another improved aspect in this new prototype is a new housing for the display as shown in Fig. 9(d). This design is 3D printed in-house. Figure 9(e) shows a sample image from the new prototype, which is captured using the same camera with the same settings as in Fig. 7.

We have conducted an informal subjective experiment on the overall performance of the display with 31 participants; the participants were both shown static and moving scenes. Below you can find a summary of the negative feedback from the participants of this experiment.

- Overall resolution on retina found to be low as in the case of all other lightfield displays.
- People with different eye prescriptions detected image distortions, due to the shape of their eyes' point spread function (PSF); however, the majority of people were able to perceive clearly.

On the other hand, the positive feedback from the participants of the experiment is as follows:

- Large FOV was well received.
- People liked the idea of having a cost-effective, simple solution without requiring much optics.
- Nobody complained about the form factor or ergonomics during the experiments.
- Nobody complained about brightness during the experiments. We believe housing helped in this case.

6. FUTURE WORK

A major trade-off in our proposal is the low light efficiency of the overall system. For our HMD prototypes, it is possible to have larger pinhole diameters and give up some angular resolution, but overall system light efficiency would remain low. Our prototype also does not directly address the accommodation-vergence conflict. However, by overlapping retinal elemental images with higher percentages, it is possible to address both the accommodation-vergence conflict and the toning effect on the image [1,3]. The amount of overlap could be increased further using wavelength selective pinholes as in [5]; thus a better representation of a lightfield can be achieved. The pinhole pattern can be improved in our future prototypes; further analysis on sampling as discussed in [9] can help us to design different pinhole patterns to improve sampling on the screen plane, and to overcome toning-related issues. Elliptical distortions (astigmatism) of a viewer's eye can be corrected by modifying the content accordingly as in [4]. Another aspect of this type of display is dependency on the smartphone's computational capability; our aim is to stream content to the smartphone to have more processing power for smarter content generation, and to address some of the highlighted issues. We intend to tackle these issues in the near future with new designs based on this work. We believe this is a basis of a framework to build simple HMDs with too few optical components.

7. CONCLUSION

In this paper, we proposed a pinhole aperture array-based approach to build a wide FOV and high optical angular resolution HMD using mobile displays. We also demonstrated multiple proof-of-concept prototypes using a photo-plotted mask on a transparency film together with an existing mobile phone's display, or a smartwatch's screen. We believe this is the simplest HMD setting proposed so far in the domain of computational near-eye display.

The authors thank Andrew Maimone for fruitful discussions and useful insights.

REFERENCES

1. D. Lanman and D. Luebke, "Near-eye light field displays," ACM Trans. Graph. **32**, 220 (2013).
2. A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (IEEE, 2013), pp. 29–38.
3. 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* (ACM, 2014), p. 20.
4. F.-C. Huang, G. Wetzstein, B. A. Barsky, and R. Raskar, "Eyeglasses-free display: towards correcting visual aberrations with computational light field displays," ACM Trans. Graph. **33**, 59 (2014).
5. K. Akşit, A. H. G. Niaki, E. Ulusoy, and H. Urey, "Super stereoscopy technique for comfortable and realistic 3D displays," Opt. Lett. **39**, 6903–6906 (2014).
6. R. Sprague, A. Zhang, L. Hendricks, T. O'Brien, J. Ford, E. Tremblay, and T. Rutherford, "Novel HMD concepts from the DARPA SCENICC program," Proc. SPIE **8383**, 838302 (2012).
7. W. Song, Y. Wang, D. Cheng, and Y. Liu, "Design of light field head-mounted display," in *International Optical Design Conference* (Optical Society of America, 2014), paper ITh4A-3.
8. S. De Groot and J. Gebhard, "Pupil size as determined by adapting luminance," J. Opt. Soc. Am. **42**, 492–495 (1952).
9. S. Dammertz and A. Keller, "Image synthesis by rank-1 lattices," in *Monte Carlo and Quasi-Monte Carlo Methods* (Springer, 2008), pp. 217–236.