A foveated contact lens display for augmented reality

Chen, Jie, Mi, Lantian, Chen, Chao Ping, Liu, Haowen, Jiang, Jinghui, et al.


Event: SPIE AR VR MR, 2020, San Francisco, California, United States
A foveated contact lens display for augmented reality

Jie Chen, Lantian Mi, Chao Ping Chen*, Haowen Liu, Jinghui Jiang, Wenbo Zhang, and
Yuan Liu
Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University,
Shanghai, China
*Corresponding author: ccp@sjtu.edu.cn

ABSTRACT

We present a design of a contact lens display, which features an array of collimated light-emitting diodes and a contact lens, for the augmented reality. By building the infrastructure directly on top of the eye, eye is allowed to move or rotate freely without the need of exit pupil expansion nor eye tracking. The resolution of light-emitting diodes is foveated to match with the density of cones on the retina. In this manner, the total number of pixels as well as the latency of image processing can be significantly reduced.

Keywords: augmented reality, foveated imaging, contact lens display, near-eye display, retinal projection, visual acuity

1. INTRODUCTION

Why is making a good near-eye display (NED) for augmented reality (AR) so hard? A blog recently posted on LinkedIn by Daniel Wagner [1], chief technology officer of DAQRI, truly strikes a chord among the AR community. To digest this blog, there are a lot to take in. Yet we can follow a Liebig’s barrel method [2] to sift out the most decisive factors that hold back the current NEDs from being good. According to the blog, there are two types of NEDs, i.e. the free-space-combiner-based and waveguide-based. For free-space-combiner-based NEDs [3–5], the trade-off between the field of view (FOV) and exit pupil is identified as its shortest stave. To have a large FOV, the exit pupil will become unacceptably small. To have a big exit pupil, either the FOV needs to be decreased or the size of the device will be increased to be even uncomfortable to wear. For waveguide-based NEDs [6–8], the good news is FOV is no longer coupled with exit pupil as the latter can be expanded via a number of techniques. The bad news is FOV will be limited by the condition of total internal reflection. For most consumers, a small FOV is definitely a deal breaker. Although not mentioned in the blog, it is necessary to be aware of another type of NEDs [9–11], which exploit the direct retinal projection without the need of a combiner or waveguide to form an intermediate virtual image. The minimal design aside, such displays are often associated with a fairly large FOV of more than 100°. That being said, they are not very successful in making real products. For example, virtual retinal display [9]—also known as retinal scanning display—requires that all beams be converged to the center of lens of eye so that the image would be unaffected no matter how the eye accommodates its diopter. Unfortunately, this condition can be only met if the eye does not move or rotate. Pinlight display [10]—a mimic of the pinhole camera—allows for certain eye motion. But its image is subject to diopter of eye, which is a variant. From the above, it can be seen that each type of NEDs has its short staves. To patch up the stave, or to build a new barrel is a question.

Customarily, NEDs are mounted on a helmet or eyewear. In this work, we attempt to shift this paradigm by building the infrastructure directly on top of the eye. Since the jargon NED is not applicable or at least inaccurate for our case, we shall re-name this type of display as contact lens display (CLD). Hereinafter, the principle as well as the potential of CLD is to be unfolded in full.

2. DESIGN PRINCIPLES

2.1 Proposed structure

The proposed CLD consists of two components, i.e. a contact lens and a collimated light-emitting diode (LED) array, as shown in Fig. 1. For the sake of symmetry, the contact lens, LED array and eye are center-aligned. Adjacent to the cornea is a thin layer of contact lens for fixing the refractive errors. On top of the contact lens is an array of LEDs, each pixel of which is able to emit a collimated beam of light towards the center of lens of eye. In this regard, our CLD is analogous to...
the said virtual retinal display or retinal scanning display. The major difference is that in our CLD, eye can rotate freely without losing the image, thanks to the surface tension of tear that tightly adheres the whole device to the eye.

Fig. 1. Cross-section view of the proposed contact lens display. Adjacent to the cornea is a thin layer of contact lens for fixing the refractive errors. On top of the contact lens is fabricated an array of LEDs, each pixel of which is able to emit a collimated beam of light towards the center of lens of eye. For the sake of symmetry, the contact lens, LED array and eye are center-aligned.

2.2 Photoreceptor cells and visual acuity

Photoreceptor cells are neuroepithelial cells found in the retina that are capable of converting the photons into biological signals [12]. One type of photoreceptor cells, cones, are less brightness sensitive than others, but can discern the colors. The human retina contains about 6 million cones. Number-wise, we are particularly interested in the number of cones as it is closely linked to the visual acuity, the ability to resolve the spatial details [13]. By retrieving the data from Curcio et al.’s study on the photoreceptor cells [14], the density of cones is plotted with respect to the eccentricity—a rotation angle $\varepsilon$ about the axis connecting the center of fovea and the center of eye (see Fig. 2)—as shown in Fig. 3. The positive sign of eccentricity signifies the temporal side, while the negative sign the nasal side. It can be seen that cones are vastly concentrated at the fovea, whose eccentricity is up to 3.3° [15]. Between −13.6° to −21.6° at the nasal side is a photoreceptor-free region called blind spot [16].

Fig. 2. Illustration of eccentricity and angular resolution. The eccentricity is a rotation angle $\varepsilon$ about the axis connecting the center of fovea and the center of eye. Suppose two LEDs A and B on the cornea are just resolvable and coaxial with two cones $A'$ and $B'$ on the retina. Relative to the center of lens, the angle subtended by A and B is equivalent to the angle $\alpha$ subtended by $A'$ and $B'$.
Visual acuity (VA) refers to the clarity of vision, which can be quantitatively described as the reciprocal of angular resolution [17], i.e.

\[ VA = \frac{1}{\text{angular resolution}} \]  

(1)

Suppose two LEDs A and B on the cornea are just resolvable and coaxial with two cones on the retina, as shown in Fig. 2. Relative to the center of lens, the angle subtended by A and B is equivalent to the angle \( \alpha \) subtended by \( A' \) and \( B' \). With the assumptions that the eye is perfectly healthy having no refractive errors, retinal detachment, macular degeneration etc. and that the beam size of LEDs is small enough compared to the pupil that no diffractions would occur, then angular resolution \( \alpha \) will be solely dependent on the density of cones \( \rho \) as

\[ \alpha = 2\sin^{-1}\left[ \sin\left( \frac{\sqrt{1/\rho}}{2\sqrt{[r^2\sin(\varepsilon)]^2 + [r\cos(\varepsilon) + d_{el}]^2}} \right) \right] \]  

(2)

where \( r \) is the radius of eyeball, \( \varepsilon \) the eccentricity, and \( d_{el} \) the distance from the center of eye to the center of lens. With the above equations and the densities of cones given in Fig. 3, VA can be calculated as a function of eccentricity. Within the fovea, the best VA can be as high as 2.6. Only a small portion, about 1% of the population, can have a VA above 2.0 [18].

### 2.3 Contact lens

The design of contact lens should conform to the prescription acquired from an optometrist or ophthalmologist. It is important to be aware that a prescription for contact lens is not the same as a prescription for eyeglasses. This is because the working distance of contact lens is obviously much shorter than that of eyeglasses. Say a user is nearsighted and his/her eyeglasses have a diopter \( P_g \) of \(-3.00 \) m\(^{-1}\) for both eyes. Then, his/her contact lens shall have a diopter \( P_c \) [19], i.e.

\[ P_c = \frac{P_g}{1 - d_g P_g} \]  

(3)

where \( d_g \) is the distance between the eyeglasses and eye. When \( d_g = 12 \) mm, \( P_c = -2.896 \) m\(^{-1}\). Treating the contact lens as a thin lens [20], we have

\[ P_c = (n - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right) \]  

(4)

where \( n \) is refractive index of contact lens, \( R_1 \) the radius of curvature of front surface, and \( R_2 \) the radius of curvature of back surface or anterior cornea. In order to be compatible as an encapsulation layer for the LED array [21], contact lens is supposed to be air impermeable. To meet this purpose, polymethyl methacrylate (PMMA)—the raw material for the old-fashioned hard contact lens [22]—is selected as the lens material. Besides PMMA, other plastic polymers, e.g. polysiloxane (silicone) and polyethylene terephthalate (PET), might work as well [23]. Per the above rules, a contact lens can be tentatively designed with the parameters listed in Table 1, which also includes a couple of physical dimensions, such as the thickness \( t_c \).
Table 1. Parameters of contact lens.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>refractive index (532 nm)</td>
<td>$n_c$</td>
<td>1.4937</td>
</tr>
<tr>
<td>radius of curvature of front surface</td>
<td>$R_1$</td>
<td>8.1301 mm</td>
</tr>
<tr>
<td>radius of curvature of back surface</td>
<td>$R_2$</td>
<td>7.7600 mm</td>
</tr>
<tr>
<td>diopter</td>
<td>$P_c$</td>
<td>-2.896 m$^{-1}$</td>
</tr>
<tr>
<td>thickness</td>
<td>$t_c$</td>
<td>0.05 mm</td>
</tr>
</tbody>
</table>

### 2.4 Field angle versus eccentricity

Picture a ray (yellow line) is incident to the contact lens at a field angle $\theta$, then refracted towards the center of lens at an angle $\beta$, and finally hits the retina at an eccentricity $\varepsilon$, as shown in Fig. 4. We could derive

$$
\theta = \beta + \sin^{-1} \frac{n_c \sin \beta}{R_1} - \sin^{-1} \frac{d_{cl} \sin \beta}{R_1} \quad (5)
$$

where

$$
\beta = -\sin^{-1} \frac{(r \sin \varepsilon)^2}{\sqrt{(r \sin \varepsilon)^2 + (r \cos \varepsilon + d_{cl})^2}} \quad (6)
$$

in which $d_{cl}$ is the distance from the vertex of contact lens to the center of lens. For the human eye, its eccentricity $\varepsilon$ varies from $-96^\circ$ (nasal) to $+80^\circ$ (temporal) [14]. The corresponding range of field angle is therefore $-66^\circ$ (temporal) to $+81^\circ$ (nasal). To avoid confusion about the plus/minus sign, it shall be noted that, due to the mirror imaging of lens, the orientation of field angle $\theta$ is opposite to that of eccentricity $\varepsilon$.

![Diagram of field angle and eccentricity](https://via.placeholder.com/150)

Fig. 4. Picture a ray (yellow line) is incident to the contact lens at a field angle $\theta$, then refracted towards the center of lens at an angle $\beta$, and finally hits the retina at an eccentricity $\varepsilon$. $d_{cl}$ is the distance from the center of lens to center of eye, $d_{el}$ the distance from contact lens to center of lens, $r$ the radius of eye, and $R_1$ the radius of curvature of front surface of lens.

### 2.5 Foveated LED array

The concept of foveated imaging originates from the computer graphics, which is intended to speed up the rendering of high-resolution images [24]. However, the physical resolution of the image remains intact, and an additional device to track the eye is needed [25,26]. Alternatively, this concept can be realized by re-arranging the pixels to match with the distribution of cones. Under the presumption that, upon the arrival at the retina, the collimated beam of each LED is received by one single cone only, it is legitimate to equate the number of LEDs to that of cones. Let the area of LEDs be decomposed into $M$ rings and the width of each ring be equal to that of LED. For the $i^{th}$ ring, the number of LEDs $N_i$ it contains can be roughly estimated by

$$
N_i = \frac{\pi \sin \beta}{\sin \left( \frac{\beta}{2} \right)} \quad (7)
$$
where \( \alpha_i \) is the angular resolution on the \( i^{th} \) ring. Hence, the total number of LEDs \( N \) is

\[
N = \sum_{i=1}^{M} N_i
\]

Substituting Eq. (5) into Eq. (7), the number of LEDs on each ring can be calculated as a function of field angle. To better appreciate the benefit of foveated pixel arrangement, the minimal number of pixels—at an aspect ratio of 16:9—required to yield an angular resolution of 1’ is computed with respect to the diagonal FOVs, as shown in Fig. 5. Take a FOV of 100° as an instance. The minimal numbers of pixels for the non-foveated and foveated displays are 15.38 and 3.20 million, respectively. The latter is merely about 1/5 of the former.

Fig. 5. Number of pixels—at an aspect ratio of 16:9—required to yield an angular resolution of 1’ is computed with respect to the diagonal FOVs. Take a FOV of 100° as an instance. The minimal numbers of pixels for the non-foveated and foveated displays are 15.38 and 3.20 million, respectively. The latter is merely about 1/5 of the former.

3. RESULTS AND DISCUSSION

3.1 Simulation setting

The whole idea of CLD is validated through the numerical simulation on Code V (Synopsys). The design wavelength is 532 nm. Fig. 6 outlines the optical surfaces defined in Code V, which are in turn (1) contact lens, (2) anterior cornea, (3) posterior cornea, (4) pupil, (5) anterior lens, (6) virtual plane, (7) posterior lens, and (8) retina. The initial parameters of eye are adapted from our previous eye model [27], in which the lens has gradient refractive indices and it is split into anterior and posterior lenses with a virtual plane being inserted in between. The total length of eye is 24 mm. For the real image, the object is positioned at 3 m ahead of the eye. For the virtual image, the object coincides with the contact lens. During the optimization, the radii of anterior and posterior lenses are set as the variables. Table 2 summarizes the as-optimized parameters for each surface.

Fig. 6. Optical surfaces defined in Code V, which are in turn (1) contact lens, (2) anterior cornea, (3) posterior cornea, (4) pupil, (5) anterior lens, (6) virtual plane, (7) posterior lens, and (8) retina. For the real image, the object is positioned at 3 m ahead of the eye. For the virtual image, the object coincides with the contact lens.
Table 2. Parameters used in the simulation.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Surface type</th>
<th>Radius (mm)</th>
<th>Thickness (mm)</th>
<th>Refractive index*</th>
<th>Semi-aperture (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>real/virtual object</td>
<td>sphere</td>
<td>∞/8.1301</td>
<td>3000/0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (contact lens)</td>
<td>sphere</td>
<td>8.1301</td>
<td>0.0500</td>
<td>1.4937</td>
<td>3.3903</td>
</tr>
<tr>
<td>2 (anterior cornea)</td>
<td>asphere</td>
<td>7.7600</td>
<td>0.5500</td>
<td>1.376</td>
<td>3.3239</td>
</tr>
<tr>
<td>3 (posterior cornea)</td>
<td>asphere</td>
<td>6.5200</td>
<td>3.0500</td>
<td>1.336</td>
<td>2.9574</td>
</tr>
<tr>
<td>4 (pupil)</td>
<td>sphere</td>
<td>infinity</td>
<td>0</td>
<td>1.336</td>
<td>1.2078</td>
</tr>
<tr>
<td>5 (anterior lens)</td>
<td>asphere</td>
<td>11.8105</td>
<td>1.7136</td>
<td>AL</td>
<td>1.1639</td>
</tr>
<tr>
<td>6 (virtual plane)</td>
<td>sphere</td>
<td>Infinity</td>
<td>1.3522</td>
<td>PL</td>
<td>0.0013</td>
</tr>
<tr>
<td>7 (posterior lens)</td>
<td>asphere</td>
<td>−5.9391</td>
<td>17.3213</td>
<td>1.336</td>
<td>0.8930</td>
</tr>
<tr>
<td>8 (retina)</td>
<td>sphere</td>
<td>−13.0000</td>
<td>0.0000</td>
<td>1.336</td>
<td>10.1176</td>
</tr>
</tbody>
</table>

*Refractive indices are left empty when the medium is air.  
\[\beta\] is the gradient refractive indices of anterior lens [27].  
\[\text{PL}\] is the gradient refractive indices of posterior lens [27].  
The semi-aperture of virtual plane is 1.3 µm, i.e. half the diameter of collimated beam of LED.

3.2 Field of view

Referring to Fig. 7, the following geometric relationship among FOV and other parameters can be obtained

\[
FOV = 2\left(\tan^{-1}\frac{D}{2d_{pl}} + \sin^{-1}\frac{nDd_{cl}}{R_1\sqrt{D^2+4d_{pl}^2}} - \sin^{-1}\frac{Dd_{cl}}{R_1\sqrt{D^2+4d_{pl}^2}}\right)
\]  

(9)

where \(d_{pl}\) is the distance from the center of the lens to the pupil. Assigning the above parameters with the values provided in Table 3, \(\text{FOV}\) is calculated as 82°. A careful examination of Eq. (9) implies that \(\text{FOV}\) is proportional to the pupil size. As can be seen in Fig. 8, when the pupil dilates to 8 mm in diameter, \(\text{FOV}\) reaches up to 142°. Unlike the rectangular NED, our CLD is round in shape. If both are of the same \(\text{FOV}\), the round one will indisputably have a bigger image size.

Table 3. Parameters for calculating FOVs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pupil size</td>
<td>(D)</td>
<td>2.4 mm</td>
</tr>
<tr>
<td>distance from the center of the lens to the pupil</td>
<td>(d_{pl})</td>
<td>1.7136 mm</td>
</tr>
<tr>
<td>refractive index of contact lens</td>
<td>(n)</td>
<td>1.4937</td>
</tr>
<tr>
<td>distance from contact lens to center of lens</td>
<td>(d_{cl})</td>
<td>2.766 mm</td>
</tr>
<tr>
<td>radius of curvature of front surface of lens</td>
<td>(R_1)</td>
<td>8.1301 mm</td>
</tr>
</tbody>
</table>

Fig. 7. Geometric relationship among FOV and other parameters. \(D\) is the pupil size, \(d_{pl}\) the distance from the center of the lens to the pupil, \(d_{cl}\) the distance from contact lens to center of lens, and \(R_1\) the radius of curvature of front surface of lens.
Fig. 8. FOV versus the pupil size. When the pupil is 2.4 mm in diameter, FOV is 82°. When the pupil dilates to 8 mm in diameter, FOV reaches up to 142°.

3.3 Angular resolution

For the resolution of our CLD is foveated, so is its angular resolution. If measured in arcminute (′), the angular resolution for the $i^{th}$ ring of the field angle $\theta$ can be calculated with

$$\text{Angular resolution} = \frac{21600 \sin \theta}{N_i} \quad (10)$$

As shown in Fig. 9, the best angular resolution is 0.38′ at 0°, whereas the worst is 3.11′ at −41°.

Fig. 9. Angular resolution versus the field angle. The best angular resolution is 0.38′ at 0°, whereas the worst is 3.11′ at −41°.

3.4 Modulation transfer function

Modulation transfer function (MTF) is calculated for both real and virtual images, as shown in Fig. 10. For the real image, MTF of the field of 0° is 0.669757 at 30 cycle/degree. For the virtual image, MTFs of all fields are above 0.999999 at 30 cycle/degree.
For the real image, MTF of the field of 0° is 0.669757 at 30 cycle/degree. For the virtual image, MTFs of all fields are above 0.999999 at 30 cycle/degree.

### 3.5 Contrast ratio

Contrast ratio (CR) of image is defined as

$$CR = \frac{CR_0 + 1 + MTF \cdot (CR_0 - 1)}{CR_0 + 1 - MTF \cdot (CR_0 - 1)}$$

where CR₀ is the CR of object [28]. For the real image, CR of object can be infinitely large. For the virtual image, CR of object or LED, is set as 5000. For the field of 0° at a spatial frequency of 30 cycle/degree, CRs of real and virtual images are calculated as 5 (MTF= 0.669757) and 4988 (MTF= 0.999999), respectively.

### 3.6 Distortion

Distortion is defined as

$$\text{Distortion} = \frac{h_a - h_p}{h_a}$$

where \( h_a \) is the height of actual image, and \( h_p \) the height of paraxial image calculated with the first-order approximation [27]. As shown in Fig. 11, distortions of real and virtual images are 10% and 6%, respectively.
3.7 Simulated imaging

Fig. 12 shows the original image (Snellen chart [29]) alongside the real and virtual images formed on the retina. As expected, the virtual image is sharper and less distorted than the real image especially at the marginal fields, which agrees with the foregoing MTF and distortion. However, the foveated effect of the virtual image is not visible. The reason is that, in Code V and other simulation tools, the influence of photoreceptor cells on the image has yet to be factored into.

4. CONCLUSIONS

As an alternative to the established NEDs, a foveated CLD has been proposed. To justify this concept, the working principles for each component have been explained. To evaluate its performance, numerical simulations have been carried out. For the real image, MTF is 0.669757 at 30 cycle/degree, CR is 5, and distortion is 10%. For the virtual image, FOV is 82°, best angular resolution is 0.38′, MTF is above 0.999999 at 30 cycle/degree, CR is 4988, and distortion is 6%. Compared to the retinal-projection-based NEDs, our CLD has its own pros and cons. First (pro), eye is allowed to move or rotate freely without the help of exit pupil expansion nor eye tracking. This is a big advantage over its counterparts, especially those with small exit pupils [9,10,27,30]. Second (pro), the physical resolution is foveated to match with the distribution of cones. This will significantly reduce the total number of pixels as well as the latency incurred by the image processing. Third (pro), no burden or weight on the head and no worries about head-related human factors, such as the shape of head, interpupillary distance etc. Fourth (con), the fabrication and popularization of CLD will face a bunch of challenges. Among others, the safety of CLD is arguably the number one issue. In addition, the power supply of CLD necessitates either the standalone built-in battery—e.g. glucose biofuel cell—or the wireless charging by means of electromagnetic induction [31]. Sadly, those issues will bring us back to the question raised at the beginning. Why is it so hard to make a good AR display? In Wagner’s closing remarks, he expressed a kind of pessimistic mood towards the possibility of breakthroughs in the short term. Still, we believe if we keep thinking outside the box, a dream solution might be just around the corner.

ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (61831015), Science and Technology Commission of Shanghai Municipality (19ZR1427200, 1801H163000, 1701H169200), and Shanghai Shadow Creator Inc. (18H100000559).
REFERENCES


