Microlens-array-based exit-pupil expander for full-color displays

Hakan Urey and Karlton D. Powell

Two-dimensional arrays of microlenses can be used in wearable display applications as numerical aperture expanders or exit-pupil expanders (EPEs) to increase the size of the display exit pupil. A novel EPE approach that uses two microlens arrays (MLAs) is presented. The approach is based on cascading two identical microlens arrays spaced precisely at one focal-length distance with submicrometer registration tolerances relative to each other. The ideal MLA for this application requires a 100% fill factor, sharp seams between microlenses, and a perfect spherical profile. We demonstrate a dual-MLA-based EPE that produces excellent exit-pupil uniformity and better than 90% diffraction efficiency for all three wavelengths in a color-display system. Two-MLA registration is performed with submicrometer precision by use of far-field alignment techniques. Fourier optics theory is used to derive the analytical formulas, and physical optics beam propagation is used for numerical computations. Three MLA fabrication technologies, including gray-scale lithography, photoresist reflow, and isotropic etching, are evaluated and compared for an EPE application. © 2005 Optical Society of America


1. Introduction

Retinal scanning display (RSD) systems operate by scanning a light beam onto the viewer’s retina in a two-dimensional (2-D) raster format.1 A RSD system typically creates a small exit pupil that is approximately the size of the eye pupil. A large exit pupil (of the order of 15 mm) is desired in wearable display systems to allow for tolerance in the placement of the display relative to the user’s eye and to allow for eye rotation without loss of the image. To obtain a large exit pupil, one needs to place an exit-pupil expander (EPE), which is essentially a numerical aperture (NA) expander, at an intermediate image plane in the optical train. Exit-pupil expansion can be achieved by use of 2-D binary gratings,2 multilevel diffraction gratings, single microlens arrays, 2-D holographic gratings, fiber optic face plates, and a number of other diffuser approaches. However, none of these approaches produces a high degree of uniformity while maintaining high efficiency in multicolor displays.3

In this paper we discuss a novel solution to the color exit-pupil expansion problem that uses two identical microlens arrays (MLAs) that are precisely registered relative to each other and separated by one focal-length distance. The EPE application in display systems provides a novel use for microlens arrays. Recently, a similar dual-MLA configuration was employed as a beam homogenizer for illumination applications.4 This paper expands our work presented in Ref. 5 and adds analytical results and a comparison of micro-optics fabrication technologies. In Section 2 we discuss RSD optics and EPE functionality as well as diffractive optical approaches to the color EPE problem. In Section 3 we present the novel dual-MLA concept and present analytical, numerical, and experimental results. In Section 4 we discuss a technique for alignment and bonding of two MLAs and describe EPE fabrication and performance with three different microfabrication technologies: gray-scale lithography, photoresist reflow, and isotropic etching.

2. Color Exit-Pupil Expander Problem and Diffractive Approaches

Figure 1 illustrates RSD optics and the EPE operation.1,2 The EPE is placed at an intermediate image plane between the scanners and the display exit pupil. A focused spot scans across the EPE and produces multiple diffraction orders at the exit pupil, where
each diffraction order contains the full image information. The eye pupil samples a few such diffraction orders and forms a retinal image. Overall luminance and color uniformity across the exit pupil perceived by the viewer is a function of the uniformity of the diffraction-order relative intensities, focused spot size, microlens diameter, scanning beam profile, and the viewer's eye-pupil size.\textsuperscript{3}

In typical RSDs, the exit-pupil size without an EPE is of the order of 1 to 3 mm. To enlarge the NA of the incoming beam for the required exit-pupil size, we place the EPE at the intermediate image plane between the scanner and the exit pupil. Note that the optical invariant before and after the EPE plane does not remain constant in the presence of an EPE. A number of technologies such as controlled angle diffusers that are used for screens appear to be good EPE candidates. However, to reduce system size while maintaining resolution, the display pixel size and the size of the focused scanned spot incident on the EPE are designed to be smaller than 20 \mu m. The small spot size prohibits the use of diffuser materials, as they are typically dependent on the illuminating beam's being much larger than the features of the diffuser material.

Diffractive optical elements (DOEs) can be used to solve the NA expansion problem in monochrome displays. Efficiency is defined as the fraction of incident power transferred into the useful output NA of the EPE, and uniformity is defined as the variation in brightness within the useful NA of the EPE. Efficiency of the order of 65\%, which is within 2\%–3\% of the theoretical limit, and uniformity better than 20\% have been achieved in red monochrome and green monochrome displays by use of a diffractive EPE.\textsuperscript{2} Diffractive EPEs work well at the design wavelength, but the intensity of the central 0th order grows rapidly if the wavelength changes or if the etch depth of the grating does not meet the design specifications as a result of manufacturing tolerances. Figure 2 illustrates another problem with diffractive EPEs: the number of diffraction orders produced at each wavelength is the same but the diffraction angle for each order is proportional to the input wavelength. If the display's useful NA is defined as the overlapping white exit pupil formed by the smallest wavelength in the system, typically blue, then the theoretical efficiencies of green and red reduce from \textasciitilde70\% to 50\% and 35\%, respectively, which are significant losses in light power.

Even-orders-missing binary DOEs can be used to reduce the bright central diffraction-order problem in microdisplays, but the scaling of NA with wavelength remains an issue, limiting efficiency. Likewise, multilevel DOEs can be used to achieve reduced 0th order and better uniformity than binary gratings. Analysis...
shows that an eight-level (three-mask) DOE is needed to achieve good exit-pupil uniformity. An exit-pupil profile corresponding to an eight-level DOE is shown in Fig. 2(b). A three-level (two-mask) DOE is designed and fabricated to balance the central 0th order for red, green, and blue laser illumination. Figure 2(a) is a scanning-electron microscope picture of the three-level DOE. In theory, a three-level mask should do fairly well in balancing the 0th order. However, as illustrated in Fig. 2(c), alignment tolerances between different level masks even for a two-level mask preclude achieving acceptable uniformity and efficiency when the unit cell size is of the order of 20 μm and the output NA is of the order of 0.25.

The reasons that color exit-pupil expansion is a difficult problem can be summarized as follows:

- High efficiency and good uniformity are desired at the exit-pupil plane.
- Use of DOEs is ruled out owing to poor efficiency and uniformity in the presence of multiple wavelengths.
- The EPE surface should be cosmetically defect free as it is placed at an intermediate image plane.
- The unit cell size of the periodic structure is small (10 to 20 μm) and the NA required after the EPE is large (0.15 to 0.25), which pushes the limits of micro-optics fabrication technology.

3. Microlens Exit-Pupil Expander Theory of Operation

A single microlens array functions in a way similar to DOEs and expands the NA of the beam. A novel approach to exit-pupil expansion is the use of a dual-microlens array (DMLA), in which two identical MLAs are cascaded with one focal-length distance between them. Figure 3 compares NA expansion using one microlens array and the DMLA. Figures 3(b) and 3(e) illustrate beam propagation along the optical axis when a focused, approximately Gaussian, spot illuminates the microlens array. The exact form of the focused spot is determined by aperture clipping effects. The simulation results are obtained by physical optics beam propagation with the Fresnel diffraction integral. The Gaussian spot covers one microlens fully and two microlenses partially and forms an exit-pupil pattern in the far field, which consists of a 2-D array of diffraction orders. Notice that both a single MLA and a DMLA form a focused spot at the array’s focal plane and then expand the beam to the NA defined by the $f_{\text{MLA}} = f_{\text{MLA}}/d$ value of the MLA.

An array of microlenses can be represented as a convolution of a quadratic phase function, corresponding to a spherical lens profile, with an impulse train function:

$$t_{\text{MLA}}(x, y) = \exp[-i\pi(x^2 + y^2)/\lambda f_{\text{MLA}}] \text{rect}\left(\frac{x}{d}, \frac{y}{d}\right)$$

$$\times \sum_n \sum_m \delta(x - nd, y - md), \quad (1)$$

where $\lambda$ is the illumination wavelength, $d$ is the MLA pitch on a square grid, and $f_{\text{MLA}}$ is the focal length of the microlenses. In Eq. (1), rect represents the rectangular aperture function with width $d$, ** denotes 2-D convolution, and $\delta$ is the Dirac delta function. Assume that a focused beam represented by a wave amplitude $h(x, y)$ illuminates a few microlenses in the array. If microlenses in the array are indexed in a matrix form with indices $(i, j)$, the illumination function for each microlens, represented by $s_i(x, y)$, can be expressed in its own coordinate system centered at the corresponding microlens by

$$s_i(x, y) = h(x - id, y - jd) \text{rect}\left(\frac{x}{d}, \frac{y}{d}\right). \quad (2)$$

In this representation, the microlens that is at the center of the axis has an index $(i, j) = (0, 0)$. Each portion of $h(x, y)$ in Eq. (2) goes through the first MLA, propagates a distance $f_{\text{MLA}}$, and then goes through the second MLA. The second MLA is registered with respect to the first MLA and adds a quadratic phase factor, which makes each microlens pair separated by one focal length a perfect Fourier transformer, limited only by cutoff of high spatial frequencies due to the finite microlens numerical aperture and pitch. An optical Fourier transform (FT) of each segment, $s_i$, spatially separated from each other, is obtained after the second MLA. The complex wave amplitude after the second MLA can be described as the optical FT of Eq. (2) rewritten in an array format similar to Eq. (1):

$$t_{\text{MLA}}(x, y) = \frac{1}{f_{\text{MLA}}} \sum_{n} \sum_{m} \text{FT}[s_{nm}(x, y)] \text{rect}\left(\frac{u}{d}, \frac{v}{d}\right)$$

$$\times \delta(x - nd, y - md). \quad (3)$$

The constant scaling factor and the coordinate definition for the FT spatial frequency variables $u, v$ outside the FT operation are due to the optical FT operation. Exit-pupil-forming optics (i.e., the ocular (viewing optics) illustrated in Fig. 1) with focal length $f_{\text{oc}}$ that follow the EPE act as another optical Fourier transformer across the entire array and produce the following wave amplitude, which is the FT of Eq. (3):
As a special case, assume that a uniform beam [i.e.,
\( h(x, y) = 1 \)] illuminates a \( (2N + 1) \times (2N + 1) \) array of microlenses. The solution is somewhat similar to that for the pinhole array diffraction problem. As the exit-pupil pattern will be symmetrical in \( x \) and \( y \), we can write the \( y = 0 \) cross section of \( t_{\text{exp}} \) as

\[
t_{\text{exp}}(x, y) = -\frac{1}{\lambda^2 f_{\text{MLA}} f_{\text{ocu}}} \sum_{n=-N}^{N} \sum_{m} s_{nm} \left( \frac{f_{\text{MLA}}}{f_{\text{ocu}}} x, \frac{f_{\text{MLA}}}{f_{\text{ocu}}} y \right) \times \exp \left[ -\frac{2\pi}{f_{\text{ocu}}} (xnd + ymd) \right].
\]

(4)

Notice that the resultant waveform is a superposition of wave amplitudes illuminating each MLA at the input side multiplied by a linear phase factor, where the slope of the phase is proportional to the distance of the particular MLA from the optical axis. Because of the periodicity of the microlenses, coherent beam illumination creates interference fringes at the exit-pupil plane. If the illumination has low spatial coherence, as does the light from LEDs, the interference fringe visibility will be lower at the exit-pupil plane.

Note that, unlike for diffractive EPEs, exit-pupil size is independent of wavelength and determined only by MLA and ocular parameters:

\[
D_{\text{exp}} = df_{\text{ocu}}/f_{\text{MLA}}.
\]

(5)

Microlenses act as diffraction gratings with period \( d \), creating diffraction orders spatially separated by \( \lambda f_{\text{ocu}}/d \). The number of interference orders \( N_{\text{order}} \) within the exit pupil is then given by

\[
N_{\text{order}}(\lambda) = d^2/\lambda f_{\text{MLA}}.
\]

(6)

As a special case, assume that a uniform beam [i.e., \( h(x, y) = 1 \)] illuminates a \( (2N + 1) \times (2N + 1) \) array of microlenses. The solution is somewhat similar to that for the pinhole array diffraction problem. As the exit-pupil pattern will be symmetrical in \( x \) and \( y \), we can write the \( y = 0 \) cross section of \( t_{\text{exp}} \) as

\[
t_{\text{exp}}(x) = \sum_{n=-N}^{N} \exp \left( -\frac{2\pi xnd}{\lambda f_{\text{ocu}}} \right) = 1 + 2 \sum_{n=1}^{N} \cos \left( \frac{2\pi xnd}{\lambda f_{\text{ocu}}} \right).
\]

(7)

For very large \( N \), an impulse-train-like waveform would be obtained with a period of \( \lambda f_{\text{ocu}}/d \) within the exit-pupil envelope. We can rewrite Eq. (7) without the summation by using the following notation:

\[
\sum_{n=0}^{N} \exp(j\delta n) = \exp(jN\delta/2) \left( \frac{\sin((N + 1)\delta/2)}{\sin(\delta/2)} \right).
\]

(8)

After some algebra, \( t_{\text{exp}} \) in Eq. (7) can be rewritten as

\[
t_{\text{exp}}(x) = \sum_{n=-N}^{N} \exp(j\delta n)
= 2 \cos(N\delta/2) \left( \frac{\sin((N + 1)\delta/2)}{\sin(\delta/2)} \right) - 1,
\]

(9)

where \( \delta = 2\pi xd/\lambda f_{\text{ocu}} \).

There are two important advantages to the DMLA approach to exit-pupil expansion: (1) As illustrated in Figs. 4(a) and 4(b), for DOEs the expanded NA scales with wavelength, whereas the expanded NA for the DMLA is determined only by \( f \)-number \( f_{\text{ocu}} \) of the microlens and is independent of the wavelength, thus allowing all light to contribute to the useful NA of the microlens, which results in high efficiency for all colors. (2) Diffraction orders are formed within the useful NA of the MLA owing to the periodicity of microlenses by overlap interference of light emanating from neighboring discrete microlens locations. The peak intensities of the resultant diffraction orders are uniform within \( 15\% \) for small Gaussian spot illumination (i.e., when the full width at 50% intensity spot diameter is equal to the MLA’s diameter). Thus the DMLA approach provides better efficiency and exit-pupil uniformity than DOEs.

4. Microlens Array Alignment and Integration

Figure 5 illustrates the computer-controlled setup designed to align and bond two identical microlenses with six degrees of freedom: \( x, y, z, \) tip, tilt, and rotation. Alignment tolerances are \( \pm 0.5 \) \( \mu \)m for in-plane registration \( (x, y) \), \( \pm 1 \) \( \mu \)m for MLA intermediate spacing \( (z) \), and better than \( \pm 3 \) mdeg for rotation, tip, and tilt. Alignment by use of near-field techniques to such tight tolerances requires sophisticated aligners with long working distances and high-magnification objectives. Even then, an alignment procedure using near-field techniques would be difficult. The alignment procedure is simplified by observation of the far-field diffraction pattern, which is the expanded exit-pupil pattern. Different types of alignment error produce interesting moirélike artifacts when they are imaged. Using visual feedback from the far-field interference patterns and uniformity variations, one can correct all rotational and translational errors. The setup illustrated in the figure is not well suited for high-volume manufacturing, but it can be used for low- and mid-volume manufacturing. UV-curable epoxy is placed between the MLAs and, once the alignment procedure is completed, the MLAs are fixed in place by UV exposure. The alignment and bonding of two microlens arrays by UV-curing epoxy can be achieved with the required precision within a relatively short time.

Fig. 5. D-DMLA alignment and bonding setup by far-field alignment techniques.
5. Microlens Array Fabrication Technologies and Comparison

There are a number of fabrication technologies for making microlenses for a variety of applications. In our design the microlens’s diameter is 15 µm and the focal length is approximately 30 µm. Once a good master is obtained, injection molding, stamping, and other techniques can be used to replicate the microlens arrays. There are always limitations to fabrication techniques determined by process types and inherent capability. In the process of implementing a usable DMLA color EPE, a handful of issues were encountered that had to be overcome. In this paper we discuss three of the principal issues, which are related primarily to obtaining good far-field uniformity from the EPE element. They involve seam sharpness, or seam radius, encountered in gray-scale lithography; flat gaps in the seam areas encountered in the photoresist reflow method; and flat-topped profiles encountered in subtractive isotropic etching.

A. Gray-Scale Lithography

As illustrated in Fig. 6(a), gray-scale lithography can involve imaging a gray-scale mask through a projection aligner or direct exposure through contact printing, or it can be achieved by using laser writing on a substrate with an analog photoresist. The photoresist is hardened in varying amounts by adjustment of the laser exposure’s duration, intensity, or both. After development and rinsing, three-dimensional surface relief features of varying depth profiles can be produced on the photoresist. Reactive ion etching (RIE) of the photoresist transfers the gray-scale pattern onto the glass substrate underneath the resist. As illustrated in Fig. 6(c), edges and corners of the microlenses are not sharp. Furthermore, step-and-repeat exposure of the substrate results in stitching errors on the surface, which are easily seen with the naked eye.

The effect of seam smoothing on the exit pupil pattern can be accurately simulated by physical optics beam propagation. Our numerical results and the experimentally observed far-field patterns agree well. Figure 7 illustrates the microlens profiles and the corresponding far-field patterns for three values of seam radius. In the simulations the areas where the MLA profile deviates from a perfect one is retrofitted with a seam profile, which has a radius that is curved up and centered along the seam border between the microlenses within the MLA. Based on the far-field uniformity requirement of 20%, the maximum acceptable seam radius was determined to be 1.5 µm and has proved difficult to achieve.
B. Photoresist Reflow

The photoresist reflow technique is a simple process whereby a master is made by patterning a photoresist into an array of pillars and then melting the photoresist pillars to form a hemispheric surface profile by surface tension.\(^{12,13}\) The resist volume controls the \(f_0\) of the microlenses. Microlenses of this type show excellent surface profile and surface quality; however, a small gap remains between the microlenses, and the light that goes through the gap creates a hot spot at the center of the exit pupil, degrading the far-field uniformity. Figure 8 shows the microlens profiles and the exit-pupil profiles as a function of gap spacing. For the simulations, first a perfect MLA profile is created, then the gap spacing is simulated by setting the phase function to a constant level in the areas corresponding to the gap such that the gap value corresponds to the flat region between nearest neighboring microlenses. The technology limit for the minimum gap is \(\sim 1 \mu m\), which yields an even larger gap at the corners. Adding an absorptive masking material in the gap areas can block the light from exiting through the gap and result in output uniformity better than that provided by any other technology, but the masking reduces the efficiency to \(\sim 50\%\) for an approximately \(2 \mu m\) gap. Addition of an absorptive layer also complicates the process by adding one more lithography step. An exit-pupil pattern produced by photoresist reflow MLAs with masking of the gap areas is illustrated in Fig. 8 (bottom).

C. Isotropically Etched Subtractive Process

A substrate is etched isotropically through an array of pinholes and forms a master mold. Microlens arrays can then be produced on plastic by any of a
number of replication technologies. The microlens profile is determined by etch selectivity in horizontal and vertical directions. The size of the etch hole results in a flat top at the center of each microlens. Figure 9 illustrates the effect of the flat top on the exit pupil’s uniformity. The flat top is simulated by setting the phase function to a constant value within the flat-top diameter, $d_{ef}$, at the center of the MLA while fitting a profile radius constrained to be tangent to the edge of the flat diameter disk and the ideal microlens profile at its edge. Once we achieved flat-top sizes smaller than 1.5 μm (~10% of the microlens diameter), the efficiency of a DMLA produced with this approach was measured to be better than 90% for all colors, and uniformity was better than 20%. The excellent far-field performance, excellent surface quality, and suitability for high-volume manufacturing offered by this approach make it attractive for fabrication of MLAs.

6. Exit-Pupil Uniformity and System Performance

To give an indication of the full array size, we show a bonded DMLA in Fig. 10. One has to keep in mind that there are more than $2 \times 10^6$ microlens sets in this array. The diffraction efficiency is of the order of >90% into the desired exit pupil pattern. The particular arrays that we used to fabricate this EPE were replicated by UV casting and can be fabricated in a fairly cost-effective manner.

7. Conclusions

The dual-microlens array is a novel approach to solving the difficult problem of exit-pupil expansion in full-color displays. Many other technologies such as various types of diffraction grating and holographic approaches were tested, and the DMLA approach proved to be superior. We also investigated a number of MLA fabrication technologies. Each approach exhibits an associated artifact, which is illustrated by experiments and numerical simulations. The technology best suited for our requirements proved to be the isotropically etched subtractive process for master making, and replication can be made by use of any of a number of plastic replication technologies. Using the DMLA approach yields efficiencies for red, green, and blue of >90%, and the uniformity is better than 20% across the display exit pupil. A challenge that we could present to microlens makers is to develop a replication technology that permits stamping microlenses with submicrometer registration tolerances from both sides onto thin substrates (50 to 100 μm thick). We could then produce the DMLA on a single substrate without requiring separate alignment and bonding.

We are grateful to Peggy Lopez (Microvision, Inc.) for her help with microlens testing.

References