Prospects and challenges in augmented reality displays

Yun-Han LEE, Tao ZHAN, Shin-Tson WU

College of Optics & Photonics, University of Central Florida, Orlando, FL 32816, USA

* Corresponding author, swu@creol.ucf.edu

Received: 5 December 2018 Accepted: 17 December 2018

Supported by Air Force Office of Scientific Research (FA9550-14-1-0279); Goertek Electronics.

Citation: Yun-Han LEE, Tao ZHAN, Shin-Tson WU. Prospects and challenges in augmented reality displays. Virtual Reality & Intelligent Hardware, 2019, 1(1): 10—20

DOI: 10.3724/SJP.2096-5796.2018.0009

Abstract Augmented reality (AR) displays are attracting significant attention and efforts. In this paper, we review the adopted device configurations of see-through displays, summarize the current development status and highlight future challenges in micro-displays. A brief introduction to optical gratings is presented to help understand the challenging design of grating-based waveguide for AR displays. Finally, we discuss the most recent progress in diffraction grating and its implications.

Keywords Near-eye displays; See-through display; Augmented reality; Micro-displays; Bragg gratings

1 Introduction

Augmented reality (AR) displays are highly anticipated as the next-generation display devices. In a glasses-like form-factor, the seamless overlay between computer-generated information and the physical world will enrich human interaction with both physical and digital world. In this paper, we first review the major configurations of AR displays adopted in present commercialized or near-commercialized products, and then discuss the advances and challenges on the micro-displays and grating couplers.

2 Construction of major AR displays

The basic construction of major AR displays includes a display unit, magnifying optics, and a combiner to overlay displayed content and physical world[10]. Based on the type of combiners, AR devices can be categorized into partial-mirror and grating-based configurations (Figure 1).

A typical example of partial-mirror based AR devices is the Google Glass™[2,3], where a beam splitter cube is utilized as the combiner in Figure 1(a). The magnifying optics is a reflective concave mirror disposed directly on the beam-splitter cube. This configuration provides the simplest solution to AR glasses with larger form-factor. Figure 1(b) shows another configuration utilizing a free-form partial reflector[4,5] both as a magnifying optics and a combiner, eliminating the need for additional magnifying optics (e.g., Mira™, Meta 2™ and DreamGlass™). An interesting configuration adopted by NED AR™ utilizes a free-form prism for achieving glasses-like form-factor in Figure 1(c). This configuration sophisticatedly incorporates two refraction surfaces, a total-internal-reflection surface, and a partial reflection surface into one element, and therefore allows extra design freedom[6–9], reducing device size with minimal sacrifice on image quality and field of view (FOV). A correcting prism is necessary to cancel the refraction of the ambient light from the main free-form prism. To further reduce size, partial reflector
array in Figure 1(d) is devised (e.g., Lumus™[10–15]). In this configuration, the thickness of combiner optics can be further reduced. The see-through transmittance is also improved as the reflectance of each partial mirror is reduced to distribute light over a larger eye box. This design sacrifices the FOV due to the limited angular range allowed in the waveguide. The image quality would also be decreased as multiple reflections could result in stray light and non-uniform intensity. In general, the optical quality and efficiency in partial-reflector based AR are higher, and yet the form-factor tends to be larger than a regular pair of glasses, except for the partial mirror array configuration.

Grating-based waveguide AR devices[16–18] utilize diffraction to guide the display light through a thin glass plate (the "waveguide") whose thickness ranges from 0.3mm to 1mm. As shown in Figure 1(e), in general, an input-coupling grating is deployed at the display end to diffract input light to an angle larger than the critical angle. The diffracted light will then travel toward viewing region through total internal reflection (TIR). An output-coupling grating is deployed in front of the eye to diffract TIR-guided light such that it travels toward the viewer. To distribute light over a larger eye box, low-efficiency gratings (leaky gratings) with gradient efficiency is usually employed; the process of distributing light across wide eye box is referred to as exit-pupil expansion or pupil replication. Both the dispersive nature of gratings and multiple diffractions post challenges for achieving large FOV and uniform output intensity. This manifests as obvious non-uniformity in color across FOV typically when a white image is displayed. To achieve better uniformity, the overall efficiency is often compromised. Despite the inferior display quality, the pursuit for smaller form-factor still draws wide attention for grating-based waveguide AR devices.

3 Advances in micro-displays

As a see-through device, AR displays post strict requirement on display brightness. For partial-mirror based ARs, the efficiency can be as high as 50%, while for grating-based waveguide ARs, the efficiency can be as low as 10%. To establish a quantitative guideline, ambient contrast ratio (ACR) is defined as[19,20]:

\[
ACR = \frac{L_{\text{on}} + L_{\text{ambient}} \cdot T}{L_{\text{off}} + L_{\text{ambient}} \cdot T}
\]

where \(T\) is see-through transmittance, \(L_{\text{on}}\) (\(L_{\text{off}}\)) represents on- (off-) state luminance (nits) while \(L_{\text{ambient}}\) is the
ambient luminance. In general, ambient lighting condition is measured in illuminance (lux); for comparison, we translate illuminance to luminance by dividing a factor of \( \pi \) (i.e., ambient light scattered from a perfectly white paper). For instance, in a normal living room, the illuminance is around 100lux, and thus it is translated to \( L_{\text{ambient}} \approx 30 \) nits, while in a regular office, \( L_{\text{ambient}} \approx 150 \) nits. For outdoors, on an overcast day, \( L_{\text{ambient}} \approx 300 \) nits, and \( L_{\text{ambient}} \approx 3000 \) nits on a sunny day.

Even for a high contrast display \( (L_{\text{on}}/L_{\text{off}} > 100) \), ambient light can wash out the content and renders the image unrecognizable. A general rule requires a minimum ACR of 3:1 for recognizable images, an ACR of 5:1 for adequate readability and an ACR over 10:1 for appealing quality\(^{[29]} \). Therefore, for an AR glass with \( T=90\% \), in an office lighting condition, the target luminance level through the device should be at least 550 nits in order to achieve ACR=5:1.

Due to the maturity of projection displays, liquid-crystal-on-silicon (LCoS) is the first choice for micro-displays in AR glasses. LCoS is a high-resolution reflective display. In this type of displays, light emitting diodes (LEDs) are used as light source, and the light is reflected by a polarizing beam splitter (PBS). Upon reflection toward LCoS, the polarization is modulated and reflected toward the PBS, which translates the phase retardation of each pixel into intensity and therefore generates images. Because the driving circuit is hidden beneath the mirror, small pixel sizes (<5\( \mu \)m) with a large aperture ratio (>93\%) can be achieved. Also, the advancement in LED technology enables high-luminance LCoS (3\( \times \)10\(^{n}\) nits)\(^{[21,22]} \), these advantages lead to their early adoption in Google Glass, HoloLens and Magic Leap One AR glasses. The major disadvantage is the requirement for a polarizing beam splitter, preventing further decrease in form-factor. Similar to LCoS, digital light processing (DLP) is also a reflective display utilizing the switching of micro-mirrors to generate images. They can also provide high brightness with a reflective configuration\(^{[23]} \). Although a polarizing beam splitter is not needed, the reflective optical path still has a relatively large form factor. This type of technology is adopted in Vuzix Blade\(^{\text{TM}} \).

To further reduce form factor, emissive displays are being evaluated for their potential in replacing LCoS or DLP. Although organic light emitting diodes (OLEDs) have been widely implemented in smartphones, smartwatches, and several virtual reality headsets, their lifetime degrades significantly at higher brightness level, making them unsuitable to drive at the same luminance level as the reflective displays. At the current stage, 5000-nit luminance level can be attained for full-color OLED micro-displays\(^{[24,25]} \).

To achieve high brightness, micro-LED displays are being heavily investigated. In recent advances, over 10\(^n\) nit peak luminance has been reported\(^{[26-28]} \). The major challenges with micro-LEDs include mass-transfer yield, non-uniform pixel brightness, high cost and decreased external quantum efficiency at smaller die size (a drop of 50\% when reducing pixel size to 5\( \mu \)m\(^{[29]} \)). If these technical challenges can be successfully overcome, micro-LED could be a disruptive display technology for AR glasses.

To compare ACR, let us assume a partial mirror of 50\% transmittance and a waveguide device of 10\% or 20\% overall efficiency. If a micro-display has 4X extension in eye-box along both directions, then the brightness would be divided by a factor of 16. With these assumptions, the ambient contrast ratio in office ambient (150 nits) is compared in Table 1.

In this estimation, the current OLED brightness in any case cannot achieve an ACR of 5:1. This makes OLED displays unsuitable for compact AR glasses without significant improvement in brightness. Using LCoS/DLP as display modules, for partial-mirror AR glasses, the brightness is more than enough due to the high overall efficiency (~50\%) and 50\% blockage of the ambient light. For waveguide ARs, at an overall efficiency of 10\%, the ACR is quite insufficient with the inherent high see-through transmittance. An additional light shield is required to reduce ambient transmittance for readability. At 30\%
transmittance, the brightness is reasonably sufficient. This results in sunglasses-like appearance. If the overall efficiency is increased to 20%, higher see-through transmittance over 50% can be allowed. Apparently, before micro-LED displays become widely available, boosting overall efficiency is a crucial subject for grating-based waveguide AR glasses.

4 Advances in grating couplers

Through magnification optics, the display light will be collimated and then coupled into a waveguide through a diffraction grating. Different from the convention in integrated photonics, the term "waveguide" here refers to a transparent, flat glass or plastic whose thickness ranges from 0.3mm to 1mm.

The critical angle of TIR posts a limit on the shallowest angle light may travel within the waveguide. Also, it is preferred to maintain the steepest angle less than 80° to prevent gaps between each two TIRs. As such, material for high-index waveguide is also a critical topic to attain a wider angular range. The current commercially available glasses have a refractive index as high as 2.0, while plastics can have a refractive index around 1.75. Although high index materials can essentially increase the FOV to over 50° × 50°, the pursuit of large-FOV posts critical challenges in output uniformity due to the diffractive nature of gratings.

To understand the difficulties in achieving high-quality grating-based ARs, it is worthwhile to investigate the basics of diffraction gratings. Gratings represent periodic modulation of refractive index. When light passes through a typical grating, it will diffract into multiple orders (directions). The first-order diffraction angle can be described by

$$\sin \theta_{\text{in}} - \sin \theta_{\text{out}} = \frac{m\lambda}{n\Lambda},$$

where $\theta_{\text{in}}$ represents the incident angle, $\theta_{\text{out}}$ represents the diffracted angle, $m$ is the order of diffraction, $n$ is the refractive index and $\Lambda$ is the horizontal periodicity.

Equation (2) states the wavelength dependency on diffraction angle. If the same grating is used for RGB colors, they will be diffracted into different angular range, and yet the allowable angular range limited by the waveguide is the same for RGB. Therefore, if the grating is optimized for the green wavelength to have the widest FOV, then the red and blue colors will be cut and have narrower FOV. To have the same FOV, three different gratings are needed to diffract RGB lights into the same nominal angle. A three-grating single-waveguide design is reported by Mukawa et al.\cite{15}. However, in this design, cross-talk between different primaries (e.g., red light being diffracted by the grating designated for green light) become an

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>LCoS/DLP</th>
<th>micro-OLED</th>
<th>micro-LED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminance at display output/nits</td>
<td>$3 \times 10^4$</td>
<td>$5 \times 10^3$</td>
<td>$10^5$</td>
</tr>
<tr>
<td>Partial-mirror ARs, $\eta$=50%</td>
<td>937.5</td>
<td>156.3</td>
<td>31250</td>
</tr>
<tr>
<td>ACR@500 lux, $T$=50%</td>
<td>13.5</td>
<td>3.1</td>
<td>417.7</td>
</tr>
<tr>
<td>Waveguide ARs, $\eta$=10%</td>
<td>187.5</td>
<td>31.3</td>
<td>6250</td>
</tr>
<tr>
<td>ACR@500 lux</td>
<td>$T$=90%</td>
<td>2.4</td>
<td>47.3</td>
</tr>
<tr>
<td></td>
<td>$T$=50%</td>
<td>3.5</td>
<td>84.3</td>
</tr>
<tr>
<td></td>
<td>$T$=30%</td>
<td>5.2</td>
<td>139.9</td>
</tr>
<tr>
<td>Waveguide ARs, $\eta$=20%</td>
<td>375.0</td>
<td>62.5</td>
<td>12500</td>
</tr>
<tr>
<td>ACR@500 lux</td>
<td>$T$=90%</td>
<td>3.8</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td>$T$=50%</td>
<td>6.0</td>
<td>167.7</td>
</tr>
<tr>
<td></td>
<td>$T$=30%</td>
<td>9.3</td>
<td>278.8</td>
</tr>
</tbody>
</table>

*Table 1 The ACR comparing different AR configurations.

*micro-LED is still under active development. The listed luminance level only serves as a reference. \( \eta \): Overall efficiency.
issue. Low cross-talk single waveguide design is often achieved by reducing spectral bandwidth of grating
couplers and thereby sacrificing the overall efficiency. To maintain high efficiency, two- and three-waveguide
designs are being utilized\(^{[30]}\); however, in this case, the thickness is compromised.

Several factors in the coupling process make the uniform distribution of outcoupled intensity over large
FOV a tough challenge. Firstly, light from different pixel travels at different angles in the waveguide. This
translates to different travel distances between each TIR, and thus different total number of TIR across the
out-coupler. Also, the spectral bandwidth results in an angular spread for each pixel. Finally, for gratings
with lower index modulation (\(\delta n < 0.1\)), diffraction spectra for different pixels are different. Due to these
reasons, designing a uniform exit-pupil expansion at all angles is exceedingly difficult, especially when the
desired FOV is large. Although reducing overall efficiency will contribute to uniformity, low optical
throughput is undesirable as mentioned in the previous section.

4.1 Volume bragg gratings

Two major types of gratings are utilized as input and output couplers: volume Bragg gratings (VBGs) and
surface-relief gratings (SRGs). Sony\(^{TM}\) promoted VBGs in Figure 2(a) for application as grating
combiners\(^{[31-33]}\). VBGs are fabricated by differential polymerization of monomer precursors under
interference exposure, and they exhibit small sinusoidal index modulation (\(\delta n < 0.05\)) across the volume.
Due to the high spectral selectivity, reflective VBGs allow for single-waveguide design\(^{[18]}\), and they exhibit
superior see-through quality. However, the spectral selectivity leads to lower efficiency as only a minor
portion of the display light is utilized. Although VBG materials with a large \(\delta n\) were demonstrated\(^{[34]}\), they
are comparatively more sensitive to environmental stimuli such as temperature or humidity.

![Figure 2](image_url)  Three major types of gratings. (a) Volume Bragg grating; (b) Surface-relief grating; (c) Polarization volume grating.

Another interesting type of VBGs promoted by Digilens\(^{TM}\) is based on holographic polymer-dispersed
liquid crystals (HPDLCs; Figure 3)\(^{[35-39]}\), which employs phase separation of LCs and monomers to create
index modulation with interference exposure. After exposure, the LC nano-droplets formed in the grating
structure can be driven by an applied voltage, allowing the realization of electrically switchable VBGs.

4.2 Surface-relief gratings

Nokia\(^{TM}\) and Microsoft\(^{TM}\) promoted the use of SRGs as combiners in Figure 2(b)\(^{[40-44]}\). In this type of
gratings, a master mold is first defined by interference and then subjected to ion-beam etching at an
oblique angle to create slanted binary features\(^{[41]}\). The feature on the prepared master mold is replicated
onto polymer materials by nanoimprint for mass production. As a result of high index contrast between air
and polymer, large binary index modulation can be achieved (\(\delta n > 0.5\)). To reduce \(\delta n\) for design purpose,
other polymer material can be utilized to substitute air. In Figure 4(a) and 4(b), we compare the performances of two input-coupling gratings: transmissive SRG ($\delta n=0.7$, $d=0.27\mu m$) and reflective VBG ($\delta n=0.05$, $d=15\mu m$). As discussed before, the high angular/spectral selectivity makes it possible to adopt the single-waveguide design with low-$\delta n$ VBGs. On the other hand, high $\delta n$ SRGs exhibits ultra-wide spectral and angular bandwidth, making it advantageous for high overall efficiency comparing to VBGs. The wide bandwidth also makes it necessary to adopt three-waveguide design to prevent crosstalk between primary colors. Due to the advancement in nanoimprint, mass replication of high-efficiency transmissive SRGs and low-efficiency reflective SRGs is proven viable\cite{30}.

While VBGs and SRGs are well-developed and adopted for commercial products, a different type of Bragg grating based on liquid crystals (LCs), called polarization volume gratings (PVGs; Figure 2(c)), is demonstrated in recent years\cite{45-52}. Two methods can be used for fabricating PVGs. The first method utilizes photo-alignment and self-assembly of the LC material\cite{46,47,50}: a photo-alignment layer is first coated onto the desired substrate, and then the coated substrate is subjected to interference exposure to define the in-plane crystal axis. Liquid crystal polymer precursor is then coated onto the exposed photo-alignment layer. The chiral agent in the precursor promotes the self-assembly of LCs into a helical structure following the recorded grating periodicity on the photo-alignment layer and formed PVGs. The other method\cite{52} utilizes the photocycloaddition of cinnamate moieties to record the volumetric polarization field into the material: a film of LC precursor containing cinnamic ester groups is first coated onto a substrate. The coated sample is subjected to interference exposure of left-handed and right-handed polarized lights, and the three-dimensional polarization field is recorded. After annealing at elevated temperature, the birefringence increases and a PVG is formed. With either method, instead of index variation, Bragg diffraction of PVGs is based on the helical rotation in LC optical axis. This means the effective index modulation $\delta n$ is essentially the birefringence ($\Delta n$) of the LC material. Due to the heavy investment of LCD industry, wide
selection of $\Delta n$ ranging from 0.05 to 0.4 is available. Besides high-efficiency and large diffraction angle, reflective PVGs show unique polarization selectivity regardless of the modulation factor (product of $\delta n$, wave vector $k$, and grating thickness $d$). As illustrated in Figure 5, only circularly polarized light with the same handedness to the helical rotation will be diffracted (in this case, right-handed), while light with the opposite handedness will pass through.

High polarization selectivity introduces new designs\cite{53}. It may be utilized for boosting efficiency\cite{54}, increasing FOV\cite{55} and alternative exit-pupil expansion design\cite{56}. As shown in Figure 6, instead of utilizing gradient efficiency, an alternative exit-pupil expansion design utilizes polarization management and polarization selectivity to distribute the output light\cite{56}. In such design, a liquid crystal layer is disposed to control the polarization state, as the optical axis and thickness of LC layer can be easily controlled, the polarization state can also be controlled such that at every interaction with the out-coupling PVG, a similar amount of light is diffracted across the eye box. This allows distribution of light without introducing non-uniform see-through transmittance and adds new free parameters for optimization of exit pupil expansion.

![Figure 5](image1.png)

**Figure 5** Efficiency for RCP and LCP of a PVG with $\Delta n=0.15$, horizontal period of 440 nm, vertical period of 205nm and central wavelength of 532nm.

![Figure 6](image2.png)

**Figure 6** An illustration of exit-pupil expansion based on (a) gradient output grating and (b) polarization control. Instead of gradient output efficiency, an additional LC layer is utilized to control the polarization state and thereby distribute light evenly at each interaction with the polarization selective PVG.

## 5 Conclusion

In conclusion, we have reviewed major configurations for AR displays. While grating-based ARs provide the slimmest geometry, a major improvement needs to be made in the optical throughput and display brightness. The dispersive nature of gratings and the consequent intensity/color non-uniformity is also a critical issue to be addressed. We also reviewed the optical properties and advantages of VBGs and SRGs and presented a newly developed PVG. With unique polarization selectivity, PVGs open new design space and serve as a potential candidate for achieving efficient and uniform throughput AR displays.
References

   DOI:10.1109/JDT.2006.879846


   DOI:10.3807/JOSK.2015.19.6.614


   DOI:10.1117/12.349411


   DOI:10.1364/AO.48.002655

   DOI:10.1889/1.2036446


   DOI:10.1364/OE.22.020705

   DOI:10.1364/AO.54.008354

   DOI:10.1364/AO.57.009246

   DOI:10.1002/sdtp.11589

   DOI:10.1364/AO.34.001352

   DOI:10.1889/1.2206112

   DOI:10.1889/JSID17.3.185
DOI:10.1364/AO.31.005988

DOI:10.1364/OE.25.033643

DOI:10.1117/12.668481

22 Huang Y, Liao E, Chen R, Wu S-T. Liquid-Crystal-on-Silicon for Augmented Reality Displays, 2018, 8(12): 2366
DOI:10.3390/app8122366

DOI:10.1002/sdtp.10269

DOI:10.1002/sdtp.12445

DOI:10.1002/sdtp.11674

DOI:10.1002/sdtp.10255

DOI:10.1002/jsid.516

DOI:10.1002/jsid.649

DOI:10.1002/sdtp.11615

DOI:10.1002/sdtp.11586


DOI:10.1088/1464-4258/11/2/024008

DOI: 10.1117/12.2191587

32–39
DOI:10.1021/cm00034a025
DOI:10.1063/1.110936
DOI:10.1364/JOSAB.19.002995
DOI:10.1155/2008/684349
DOI:10.1002/sdt.p.12523
DOI:10.1364/JOSAA.12.001077
DOI:10.1364/OE.15.002067
DOI:10.1117/12.768666
DOI:10.1889/JSID17.8.659
DOI:10.1364/AM.49.005454
DOI:10.1364/OE.24.017746
DOI:10.1364/OE.25.006283
DOI:10.1364/OE.25.019298
DOI:10.1038/nphoton.2016.66
DOI:10.1515/odps-2017-0008
DOI:10.1515/odps-2017-0010

DOI:10.1007/s0034


DOI:10.1117/12.2315635

DOI:10.1002/jsid.635