CURRENT STANDARDS AND TECHNIQUES IN AR/VR/MR NEAR-TO-EYE DISPLAYS

by Luke Turner


_____


A Report Submitted to the Faculty of the

JAMES C. WYANT COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2021

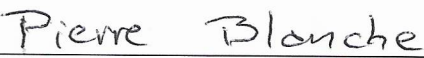## THE UNIVERSITY OF ARIZONA

## GRADUATE COLLEGE

The master's report title "Current Standards and Techniques in AR/VR/MR Near-To-Eye Displays" prepared by Luke Turner, has been submitted in partial fulfillment of requirements for a master's degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

As members of the master's Committee, we certify that we have read the report prepared by Luke Turner, "Current Standards and Techniques in AR/VR/MR Near-To-Eye Displays" and recommend that it be accepted as fulfilling the report requirement for the master's degree.
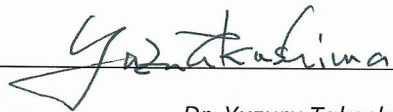
_____                    12/6/2021
Dr. Russell Chipman, Chair                                         Date

_____                    12/6/2021
Dr. Pierre Blanche, Member                                        Date

_____                    12/6/2021
Dr. Yuzuru Takashima, Member                                   Date

### APPROVAL BY REPORT ADVISOR

This report has been approved on the date shown below:

_____                    12/14/2021
Dr. Russell Chipman, Chair                                         Date

# Acknowledgements

I'd like to thank my wife, Renae, for always believing in me and pushing me to pursue my dream. I would not have been able to finish this program without her encouragement. I would also like to thank my advisor, Russell Chipman, for giving me this opportunity to work underneath him and be a guiding presence in my pursuit of knowledge.

# Table of Contents

# List of figures

# Abstract

Augmented reality (AR), virtual reality (VR), and mixed reality (MR) near-to-eye (NTE) displays are complex technologies where a wide field of view, high resolution, and a large eyebox are desired. Special consideration must be taken during the design of these systems to achieve a high-quality image, however optimizing the design for all three is difficult to achieve. Specifically, methods of enlarging the eyebox, also called exit pupil expansion (EPE), have been a focus of research in recent years. In this review, options are discussed evaluating the user experience,  human eye limitations , and survey design techniques used to create the optical subsystems, which image the display into the user's eyes, namely the display engine, imaging optics, and optical combiners. A focus of this report is methods of exit pupil expansion for mixed reality displays and advancements in polarization-based optical folding used to reduce the imaging system of a VR display to under 10 millimeters.

## Chapter 1. Introduction

Virtual reality (VR), augmented reality (AR) and mixed reality (MR) are emerging

technologies based on the integration of virtual digital elements into the real world. There are

many levels of immersion that fall on the mixed reality spectrum, figure 1, with augmented

reality being closer to a real environment and virtual reality being closer to a virtual

environment. The optics used in these three groups are similar because all image a display into

the user's eyes. Thus, techniques used in MR design will often apply to both AR and VR. For the

purposes of this paper, only the optical components related to the display illumination and

image formation elements are discussed. These elements contain current hardware designs,

issues, and mitigation techniques used in mixed reality head-mounted displays (HMDs).
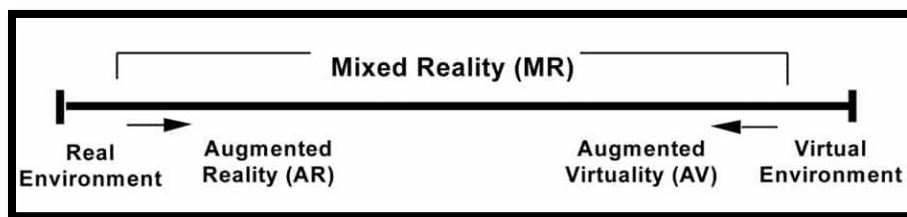


*Figure 1: Mixed Reality Spectrum Continuum [Referenced from*

*https://www.aniwaa.fr/guide/vr-ar/vr-ar-mr-guide-ultime/]*

VR refers to the total immersion into a virtual universe, blocking out the real world,

generating realistic images, sounds and other sensations that simulate a user's physical

presence in the virtual environment. AR overlays digital images into the real environment. An

image from a display is coupled into the eye using beamsplitters, lenses, prisms, waveguides, or

other optical elements. MR refers to everything besides the two extremes, but generally refer

to an interactive experience of a real-world environment that is enhanced by computer

generated information. This has a broad meaning and can be used for many kinds of systems. Some MR systems, like the Microsoft HoloLens, allow the user to interact with digital objects while having full view of the real world using a see-through visor. Other systems produce a video see-through (VST) device that can display the real world as a virtual environment but blocks direct view outside the headset. This can be done using cameras on the headset to produce a digital image to the user.

Virtual reality systems are mostly used for gaming entertainment, but they can also be used as an educational tool and a virtual social environment. These often require a controller to interact with the virtual environment. Augmented reality eyewear can be used as a tool to display information, around the house, in a medical environment, or in an office to speed up work production. Amazon has over 420 hits when searching for augmented reality glasses. Many of the cheaper designs use a smartphone as the display, which reduces the need to include the display with the headwear. More expensive headsets will generally have a display already installed into the device and use a computer to input the programs used. Mixed reality headsets may use a controller as well, but if a headset has external cameras, hand tracking is a possibility. Mixed reality is a relatively new concept, but augmented and virtual reality have been an idea in progress since 1962 [20]. U.S. patent 3050870 is seen providing a stereoscopic pair of separated images corresponding to parallactically displaced view of the same object, figure 2. Since then, many new inventions and technological strides that increased the performance of these machines including the Virtual Boy (Nintendo 1995), Kinect (Microsoft 2010), Glass (Google 2014), HoloLens(Microsoft 2016), Oculus Rift (Oculus 2016), and many

more. These products seek to improve the current standards of mixed reality to provide users with the best experience possible.
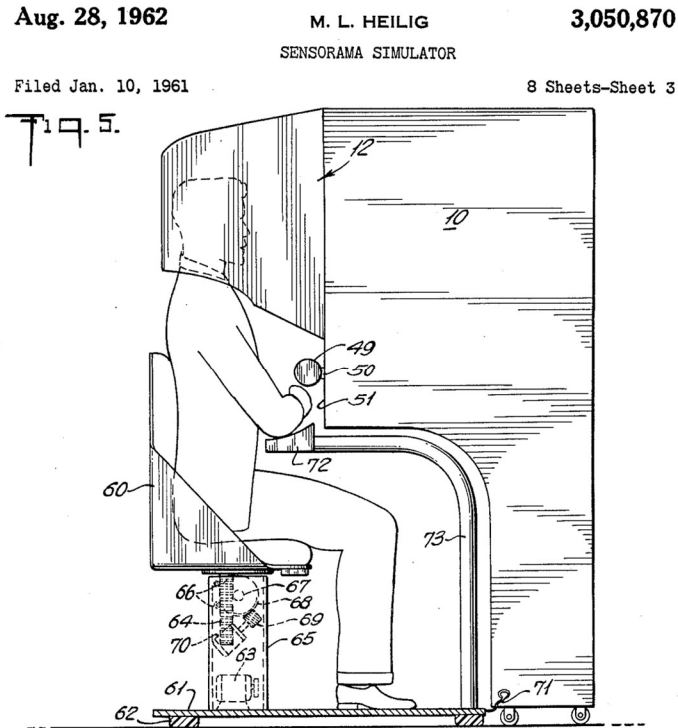
Figure 2: The Sensorama [US patent 3050870]

## Chapter 2. Objectives for the Best User Experience

Comfort plays a major role for the headset. One cannot remain totally immersed in a virtual world long if there is discomfort in wearing the HMD, because that discomfort brings the user back to the real world. There are three main areas of comfort that need to be addressed in a mixed reality HMD.

First, MR should be physically comfortable to wear for longer periods of time. Common issues include weight, temperature, center of gravity, motion, nausea, a cordless experience, and more. The weight of device is related to the size of the optics required, and so making the

systems shorter or folded will reduce the size and weight of the headsets. Lessening the weight and moving the center of gravity closer to the human head natural position also alleviates the neck strain felt by the user. Many of these features are inherent to the mechanical structure of the headset, but also are limited by the optical design of the eyepieces. Thicker lens elements and bulky reflective and refractive freeform designs will inevitably lead to a larger headset to mount these optics properly.

Comfort is a subjective concept. The use of EEG (electroencephalography) sensors has been used to take measurements of mental activity while subjects are immersed in VR [20]. Another study compares the EEG measurements from camera motion effects such as yaw, roll, pinch, zoom-in and zoom-out in 16 different participants [21] to provide quantitative and objective measurements for VR sickness, nausea induced by the display in the VR system. This nausea is a pattern in MR systems but is most known for VR HMDs.

The second area of comfort is related to displaying a clear image with minimal aberrations and a large FOV. High quality images in HMDs need to be able to compete with current television, phone and watch resolutions to be viable in the marketplace. It is the work of the optical designer to limit image aberrations to acceptable levels.

Lastly, social comfort influences the design of the aesthetics, form factor of the headset, and use in a public setting. An important feature of AR design is to include an unaltered eye view of the HMD wearer, allowing for continuous eye-contact and eye expressions discernment. Having eyewear that does not distort or block the eye will increase the normalcy

in day-to-day interactions. While social comfort and wearability are important factors, the visual features are more pressing to an optical engineer.

## 2.1 Visual Comfort

Visual comfort refers to the optical features that allow a clear image to be formed on the user's retina. These features include having a large and uniform eyebox (larger than the human pupil) to allow for a wide range of interpupillary distances (IPDs) and motion of the eye, an angular resolution close or greater than 20/20 visual acuity, no screen door or Mura effects, high brightness, high contrast, minimized ghost images, a wide FOV for the display, and a 200+ degree see-through peripheral vision. These features allow the user to comfortably see what is in front of them, whether it be virtual or real.

Consider the large and uniform eyebox requirement. There are many different definitions for the eyebox. Some optical designers refer to the unvignetted pupil dimensions, while other refer to the amount of motion a pupil can have [2]. A simple definition of the eyebox is the 3D volume located between the combiner and the human eye pupil over which the entire FOV is visible for a typical pupil size. The 3D shape is dependent on the type of optics used and the emission cone of the display, figure 3a-3b. According to the American Academy of Ophthalmology, typical eye pupils have a diameter of 2-8mm. Having a large eyebox allows more tolerance to accommodate the user's interpupillary distance and rotations of the eye when looking away from the optical axis. Alternately, having a smaller eyebox will lead to vignetting, or a dimming of the image.

A uniform eyebox is desired to reduce changes in brightness when viewing the display at different locations in the eyebox. Nonuniformity can be defined using the Michelson contrast: (max-min)/(max+min). A nonuniformity value of 1 indicates that at least one of the source points cannot be seen for a particular eye position. A nonuniformity value of 0 indicates that all points take in the nonuniformity calculation have the same retinal illuminance value. Lower nonuniformity values are desirable [2]. Eyebox nonuniformity causes undesired attention to the variation of brightness of the virtual images, causing safety issues [36]. According to Clause 4.3.2 of SAE AS8055 issued by the International Association of Automata Engineers, the luminance variation of the displayed symbols shall meet the following requirements for HUDs: 1) when measured from the center of the HUD eyebox, the luminance variation of symbols contained in the monocular instantaneous field shall be less than 30%; 2) when measured over
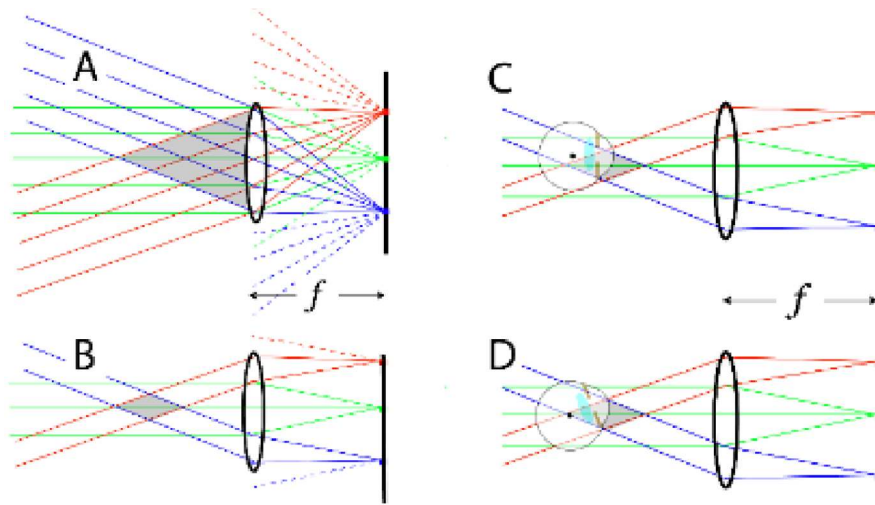


*Figure 3: (A) A broad emission cone forms a triangular eyebox. (B) A narrow emission cone forms a diamond shaped eyebox. (C) The eye located within the eyebox. (D) The eye rolled up to view the top edge of the image [Cakmakci, 2]*

80% of the HUD eyebox, the luminance variation of any symbol shall not vary by more than 30% [35]. Even though these standards are in place for HUDs in airborne vehicles, the same reasoning applies to HMDs.

Each headset could have the ability to mechanically adjust for the individual's interpupillary distance and be set at that location for future use. However, the need to change the IPD on the headset when switching users would become a nuisance and require a setup process each time. Current MR systems usually have some form of IPD adjustment, whether that is a mechanical adjustment (used in the HTC Vive) or a digital adjustment (used in the PlayStation VR). Increasing the eyebox will allow a greater range of people to wear the headset without needing to change the interpupillary distance. There are many methods being used today to increase the eyebox dimensions including pupil replication and pupil steering.

Many people have some form of vision impairment, so designing an active visual impairment correction would be beneficial for people who wear prescription glasses. When considering visual impairment, the far-field is the primary concern for correction because virtual images are projected at infinity. However, there are also reasons to correct for hyperopia (farsightedness), especially in see-through MR headsets. If the combiner is not located on the glass, like used in Google Glass, a simple fix could be to substitute the original glass for the user's required prescription. This idea can also be used in VR by changing the distance between the lens and the display. This method can compensate for myopia (nearsightedness) and hyperopia, but not for astigmatism. The human eye is the sensor for these imaging systems, so understanding the optical properties of the eye will give insight into some of the specifications and requirements of MR systems.

## Chapter 3. The Human Visual System

Optical properties of the imaging sensor, in this case the eye, are very important in the design of optical systems. The monocular FOV of the human eye is about 160 degrees (horizontal) by 130 degrees (vertical). The combined binocular FOV is about 200 by 130 degrees, with an overlapped region of 120 degrees horizontally [7]. These values are averages based on one study, but an individual can have slightly different FOV abilities. For this reason, there are some discrepancies in the exact numbers that should be used to define a human's FOV. The horizontal FOV is shown in figure 4, with the white circle showing the fixed foveated region where sustained eye gaze is possible. This region defines specification for high-end AR/MR devices. VR seeks to expand the FOV to be closer to how the eye functions but reaching a 90-degree diagonal FOV is sufficient to cover most of the binocular vision and is the benchmark standard for VR stereoscopic systems.

The ability of the eye to resolve small features is referred to as visual acuity. The resolution limit of the human eye is determined by the average spacing of cone cells in the fovea, given as 0.5 arcminutes [22]. According to the National Institute for Standards and Technology, good vision in young adults is usually 20/20 or better, meaning that at 20 feet away, they can see alternating black and white lines that subtend an angle of 1 arc minute [23]. This is the common standard when designing optical systems for the human eye.

The human eye has a spectral response which must be accounted for during the design process. A light source radiating 1 Watt of green light will appear much brighter than another source radiating the same amount of power of red light because the eye is more sensitive in the
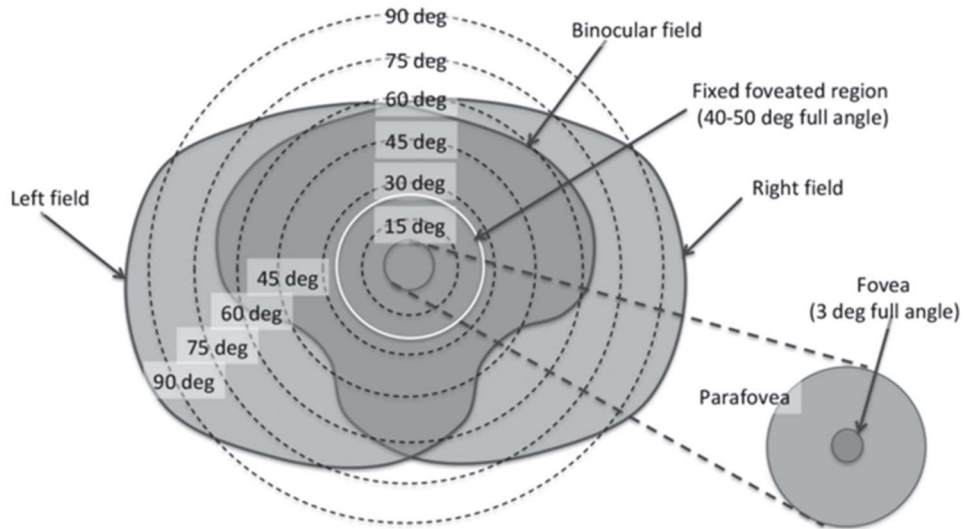
*Figure 4: Human binocular field of view [Kress, 1]*

green region. There are two different luminous sensitivities for low (scotopic) and high

(photopic) illumination levels, figure 5 [24]. Scotopic corresponds to the sensitivity of the rods

and photopic corresponds to the sensitivity of the cones.
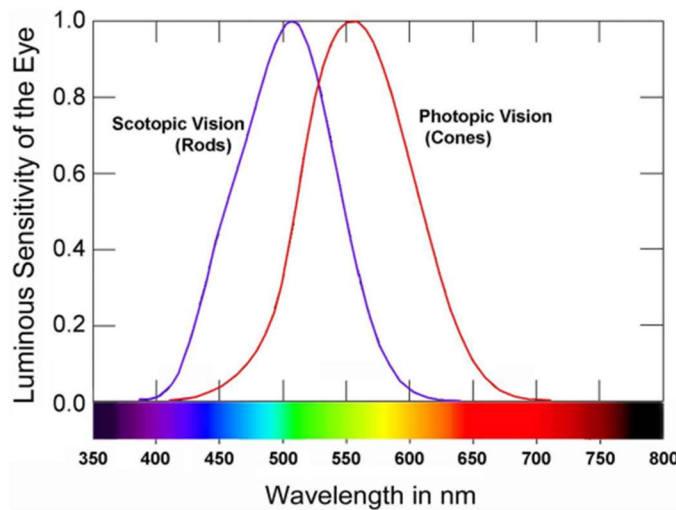


*Figure 5: Luminous sensitivity of the eye [Malacara, 24]*

Based on the luminous response of the eye, how bright should the display be?

Luminance is the power of the light per unit area per unit projected solid angle weighted by the

spectral response of the eye [25], often referred to as nits or candelas per square meter(cd/m$^2$).

Some typical luminance values are shown in figure 6. Illuminance is the power per unit area

that is incident on a surface weighted by the spectral response of the eye. Its most common

units are the lux (lx = lm/m$^2$) or the footcandle (fc = lm/ft$^2$) depending on if imperial or metric

units are used. To determine the proper light efficiency of the system, two metrics should be

used: one based on the luminance of the system in nits, and the other based on the luminous

throughput of the system spread over the entire eyebox in lumens. So, an effective way to

measure the efficiency of a system is to measure the brightness in nits at the eye as a function

of lumens produced by the light engine. Typical efficiency levels for current waveguide

combiners in AR systems range from 50-1000 nits/lumen [1].

| Source | Luminance (cd/m$^2$) | Source | Luminance (cd/m$^2$) |
|---|---|---|---|
| Night sky, cloudy, no moon | $10^{-4}$ | Minimum visible level (human) | $3 \times 10^{-6}$ |
| Darkest sky | $4 \times 10^{-4}$ | Scotopic vision valid (human) | $< 0.003$ |
| Night sky, clear, no moon | $10^{-3}$ | Photopic vision valid (human) | $> 3$ |
| Night sky, full moon | $10^{-2}$ | Green electroluminescent | 25 |
| Clear sky 0.5 hr after sunset | 0.1 | T8 cool white fluorescent | $10^4$ |
| Clear sky 0.25 hr after sunset | 1 | Acetylene burner | $10^5$ |
| Cloudy sky at sunset | 10 | 60 W inside frosted lamp | $1.2 \times 10^5$ |
| Gray sky at noon | $10^2$ | Candle | $6 \times 10^5$ |
| Cloudy sky at noon | $10^3$ | Sodium vapor lamp | $7 \times 10^5$ |
| Moon | $2.5 \times 10^3$ | High-pressure Hg vapor lamp | $1.5 \times 10^6$ |
| Average clear sky | $8 \times 10^3$ | Tungsten lamp filament | $8 \times 10^6$ |
| Clear sky at noon | $10^4$ | Plain carbon arc crater | $1.6 \times 10^8$ |
| Solar disk | $1.6 \times 10^9$ | Cored carbon arc crater | $10^9$ |
| Lightning | $8 \times 10^{10}$ | Atomic fusion bomb | $10^{12}$ |

*Figure 6: Typical luminance levels [Palmer, 25]*

The light sensitive elements in the retina are of two types, rods and cones. The rods are

responsible for night vision, while the cones' function is daylight vision. The sharpest vision

occurs in the central part of the retina, called the fovea, where most cones are located, and

there are no rods [24], figure 7. Displays used in HMD devices are not bright enough to damage

the human eye and work in the visible wavelength range, so the total luminance of the source

will not be a limitation in the safety of the designed system. Some systems use LED or laser

illumination, and that could be more dangerous, so it is best to refer to laser safety guidelines

regarding safe illumination of the human eye. One concern is UV LEDs leaking into the

transmitted beam and causing retinal damage. Products should meet the limitations set in

IEC/EN 62471 to avoid damaging the eye.
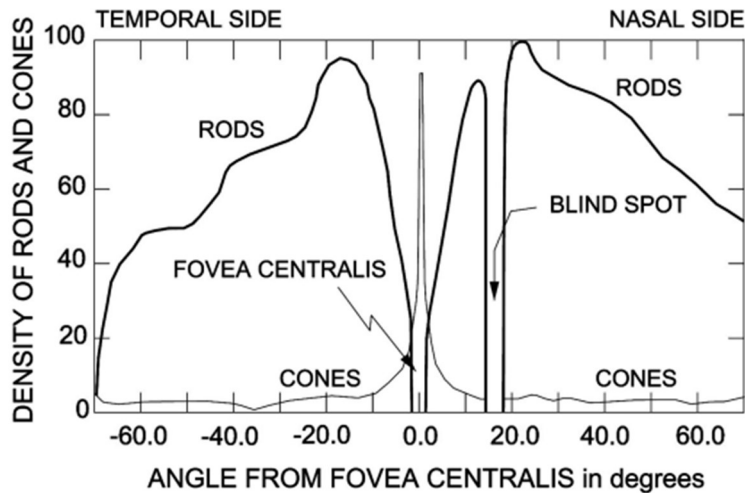


*Figure 7: Distribution of rods and cones in the retina along a horizontal plane through the fovea [Malacara ,24]*

## 3.1 Resolution Enhancement

A system that reaches the retinal resolution limit is desired in HMDs. A straightforward

solution to increasing the resolution is to increase the pixel count of the display source.

However, smaller displays are desired because this reduces the size and weight of the device.

These competing factors require the optical engineer to make a choice between high resolution

and size, and so methods to increase the resolution without using an extremely dense pixel

array are being researched. One common approach to simulate a higher pixel density, (or

smaller pixel pitch), is to use the knowledge that the human eye has a higher resolution

perception when light focuses on the fovea [24]. This foveated view covers only 2-3 degrees

and is set off axis from the optical axis of the eye by about 5 degrees [1]. Because humans have

increased perception when we look straight at something, we can use a non-uniform pixel

density in the display panel that has higher resolution at the center, figure 8 (bottom). This

fixed foveation method give the user an increased resolution in the center of the display. This

can be achieved by using a beam splitter to combine images displayed from a low-resolution

panel and a high-resolution one [8]. This has the drawback of a larger device volume from the

two displays. Miniaturizing this optical layout is an essential task for the future development of

foveated devices.
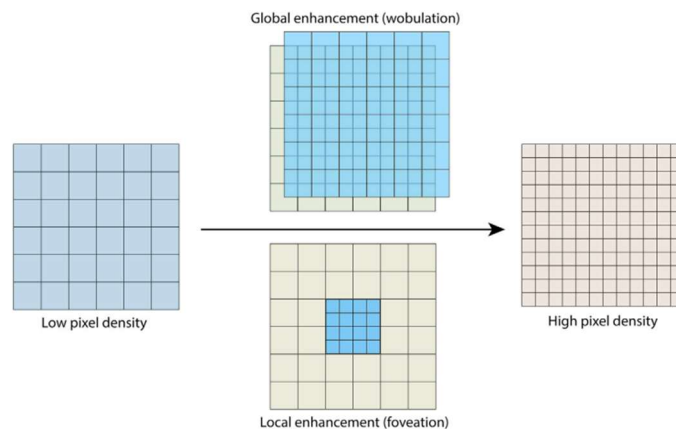


*Figure 8: Various pixel densities and enhancements of apparent pixel densities including foveation and wobulation [Zhan, 8]*

The second option uses the wobulation method to increase the angular resolution,

figure 8 (top). The idea here is to slightly shift the pixels so that they display between other

pixels either mechanically, or electrically using liquid crystals. This technique originally used a

spinning wedge to produce a slight angular shift, synchronized with the display refresh rate, to shift the apparent location of the pixels in a circular motion, figure 9. There are many new methods of wobulation including the use of liquid crystals [26] and piezoelectric actuators to mechanically shift the location of the display [9].
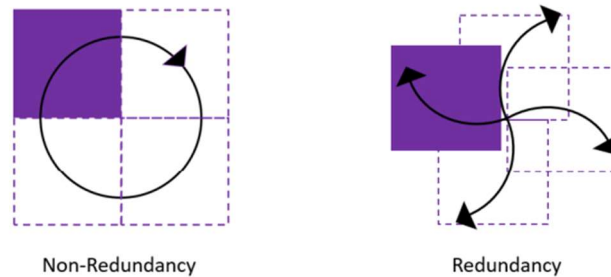


Figure 9: Non-redundancy pixel trace (left), and redundancy pixel trace (right) [Nguyen, 9]

## 3.2 Vergence Accommodation Conflict

Because of human's stereoscopic vision, we see a different image in each eye. A common issue in HMDs for stereoscopic viewing is allowing the user to view virtual objects as if they were real object with depth. The vergence accommodation conflict (VAC) is described as the mismatch of visual depth cues when trying to focus on an object with both eyes, visualized in figure 10.

Vergence is the simultaneous rotation of the two eyes to maintain binocular vision. In other words, it is the changing viewing angle of the two eyes such that the depth of the objects is fitted properly. Accommodation is the oculomotor response to the distance of the object, or how much the eye muscles contract to place the focus of the object onto the retina properly. Accommodation is driven by the retinal blur. This blurring occurs due to the reasons discussed above with the fovea providing the sharpest vision as well as further objects being blurred

when focusing on closer objects [4]. This issue is present in HMDs because the fixed display plane forces the viewer to accommodate to a single distance to maintain a sharp image, regardless of their vergence state. This contributes to visual fatigue during prolonged use, and in some cases, nausea lasting well after the headset is taken off. Therefore, many headsets limit the virtual objects depth range, keeping the VAC discomfort to within acceptable limits.
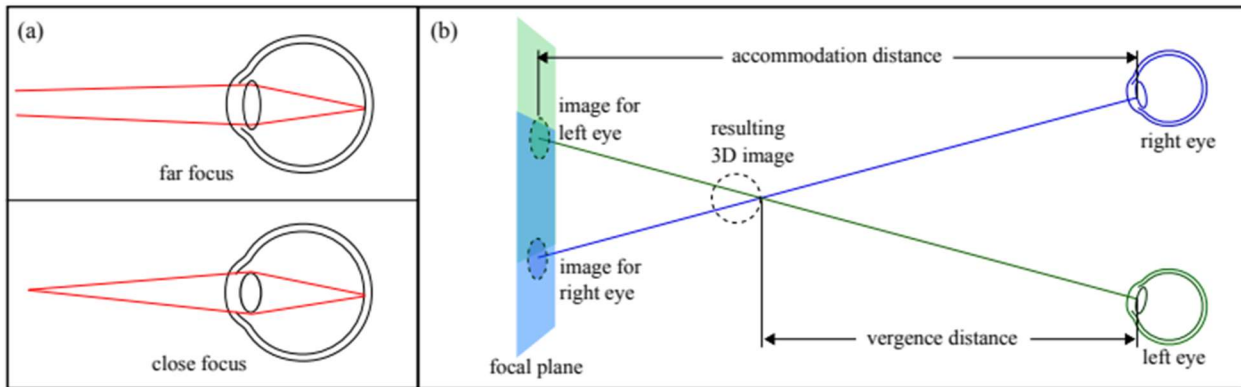


*Figure 10: Sketch of the VAC issue. (a) Accommodation of the eye for close and far field objects. (b) Conceptual representation of the VAC. [Kramida, 4]*

Since the convergence cue in properly configured stereo displays mostly corresponds to natural world viewing, but the accommodation does not, vast majority of the effort on resolving VAC gears towards adjusting the retinal blur cue to the virtual depth of the content [4]. Many mitigation techniques have been utilized to produce multiple focal plane distances such as sliding relay lenses [28], deformable mirrors [29], electroactive liquid lens displays [30]. A more complete list of techniques can be found in a paper by Kramida and Varshney [4]. If the optics are not designed to mitigate VAC, digital rendering of blur can be added to increase the available cues.

### 3.3 Pupil Swim

Active pupil swim is another feature that is commonly present in HMDs. For optical systems with an external pupil, when the eye's pupil moves around, there is a change in the aberrations present, specifically distortion and lateral chromatic aberration that can result in a "swimming" sensation of the image [5]. This means that the perceived location of each point in the image shifts in a non-uniform way. For example, if you look at a rectilinear grid in a VR headset, the shapes and color shift of the grids might change as you rotate your head and eyes. If the MR display has eye-tracking capabilities, images can be rendered that compensate for the distortion and chromatic shift on a pixel-by-pixel basis. One method for self-calibrating eye tracking is defined by Huang and Bulling [31] which uses only the motion of the eye to produce a distorted gaze projection plane. Eye tracking is becoming a standard feature for many headsets today, providing uses such as optical foveation, vergence tracking, distortion mapping, and pupil swim compensation [1].

### Chapter 4. Near-to-Eye Displays

Near-to-eye displays consist of three main components: a display engine, imaging optics, and an optical combiner. The display engine forms the original image, which can be a light-emitting display such as an array of organic light-emitting diodes (OLEDs) or a light-modulating device illuminated by one or more light sources. The imaging optics will image the display either by forming a pupil to produce an intermediate image that can be processed and modified, or not forming a pupil to display the image at optical infinity. Optical combiners will

take the image of the display and move it to the user's pupil through beamsplitters, freeform prisms or waveguides.

Near-eye displays often suffer from image artifacts that degrade the system such as pupil swim (discussed earlier), image contrast, stray light, and ghost images [5]. One significant factor to all these performance metrics is the performance of the imaging optics. These can come in many different forms including smooth refractive lenses, Fresnel lenses, polarization-based pancake optics systems, freeform prisms, waveguides, and holographic optical elements.

## 4.1 Imaging Optics

Due to ergonomic requirements, the imaging optics should be compact and light, which generally brings a significant sacrifice in imaging quality. A conventional singlet lens has limited stray light and ghost images, but also has a comparably large volume and weight. These lenses are usually limited by field curvature but can be mitigated by use of aspherics, field lenses, and multiple elements [8]. Because of the weight and similar optical performance, Fresnel lenses
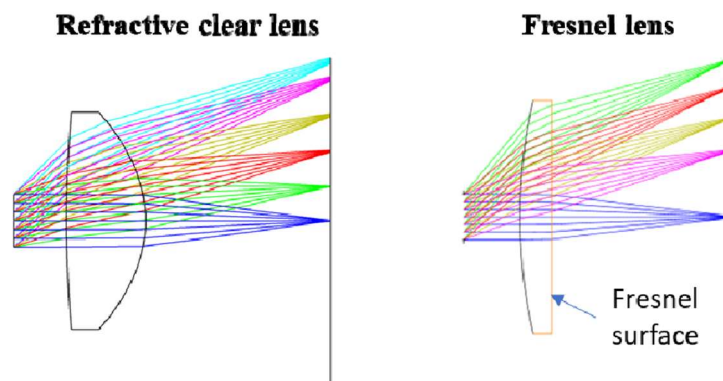


*Figure 11: (Left) A standard refractive lens used for imaging a display to an external pupil. (Right) A Fresnel lens using the same imaging [Geng, 5]*

are an alternative to traditional refractive lenses in a VR display, but at the cost of possible stray light and image shifts.

Fresnel lenses are incorporated in many commercial VR systems (Oculus CV1, HTC Vive) due to their potential of being light weight and low volume compared to standard refractive glasses [1]. Fresnel lenses are also popular because the field curvature behaves differently compared to normal smooth refractive lenses and allows better performance at the edges of the field with a similar number of elements [5]. A few drawbacks of using Fresnel lenses include increased diffraction artifacts from the Fresnel structure limiting the diffraction MTF and the stray light artifacts from the facets etched into the Fresnel structure. Because VR HMDs are often limited by the available display panel resolution, the optics do not need to be diffraction limited and some diffraction effects can help to mask the gaps between pixels, reducing the "screen-door" effect.

Using polarization-based pancake lenses, or catadioptric lenses, is a promising solution for use in VR displays. Catadioptric refers to optical systems that have reflective and refractive components. These reflective components can be mirrors (like used in a Cassegrain telescope) or can be optical thin film coatings that produces a partial reflectivity over the desired angle and wavelength range as shown in the singlet lens in figure 12 (right). These systems are extremely useful in reducing field curvature because the Petzval curvature is inversely proportional to the refractive index and is opposite for positive and negative elements. Thus, by including reflective components with positive power, and refractive components with positive power, these catadioptric lenses can cancel the Petzval curvature. Using these reflective surfaces, systems with shorter focal lengths and a larger FOV can be achieved more easily.
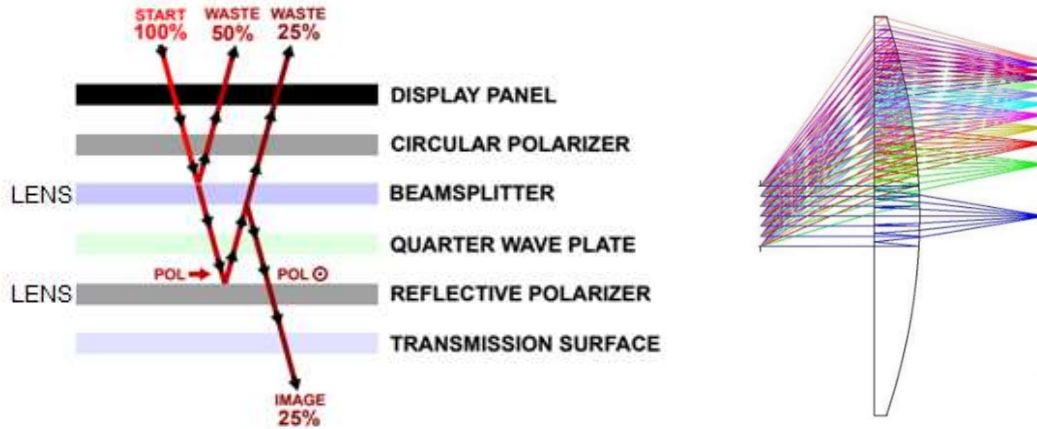
*Figure 12: (Left) Pancake Optics allow the cavity between a beamsplitter and a reflective polarizer to be traversed three times, which reduces space and improves performance [Maimone, 11]. (Right) A design of a catadioptric low curvature singlet lens [Wong, 12].*

Metallic wire-grid polarizers could also be used to create the reflective polarizers, but polymetric, birefringent, multi-layer reflective polarizers offer numerous advantages such as low scattering and absorption and high polarized reflectivity [32]. To utilize the properties of catadioptric systems properly, optical power is placed on surfaces that have multiple interactions, like the beamsplitter and the reflective polarizer surfaces in figure 12 (left). This allows less power to be imprinted on the surfaces while maintaining a similar focal length compared to non-catadioptric systems.

Some major pitfalls of using these systems exist fundamentally in their design. Using the example in figure 12 (left), the signal needs to hit the beamsplitter twice, so the system throughput is limited to 25% if the initial display panel is polarized, or 12.5% if the display is unpolarized. This wasted power has a direct correlation to less battery life in the system. Additionally, due to the various individual components and air-material interfaces, the system is subject to ghost images and stray light paths. Depending on the performance of the quarter wave plate, there will be polarization-based light leakage due to the retardance difference at

varying angles of incidence and wavelengths. Lastly, any lens materials that are inside the polarization cavity need to have low birefringence to minimize any unwanted polarization changes. Using plastic injection molding will have lighter weight but will also have significant residual stress which results in unwanted birefringence [5]. As a result, molded plastic generally leads to relatively low image contrast.

## 4.2 Display Engine

An HMD is a very complex system with a couple optical subsystems. Particularly in see-through HMDs in MR and AR systems, the display and the combiner are the two most important engines that need to be designed together to ensure proper eyebox size, illumination, and form factor. Displays commonly used can be broken down into two categories: panel-based displays and scanning displays.

### 4.2.1 Panel-Based Displays

Panel-based displays are commonly used in electronics such as smartphones, smart watches, and televisions. These displays come in many different forms: liquid crystal display (LCD), light emitting diode (LED), organic light emitting diode (OLED), digital light processing (DLP), and liquid crystal on silicon (LCoS). The illumination engine for AR headsets is often half of the display engine volume. LCDs, LCoS-based displays and digital micromirror devices (DMDs) require some illumination source, either front-lit or back-lit. One advantage of LED or OLED displays is the reduction in size and weight due to the diodes being self-illuminated rather than through an external illumination source. Each of these displays come with their own size and brightness, which is an important factor in designing the system. In some cases, brightness is

more important, but for others, size and weight might be prioritized (such as for smart glasses and other consumer AR headsets). Figure 13 emphasizes the balance between the display engine size and the brightness at the panel depending on if back-lighting or front-lighting illumination is used. Listed in the figure are a few other displays such as micro-OLED and micro i-LED, which are still evolving technology based on creating smaller pixel sizes, as well as laser beam scanners (LBS), which will be discussed in the next section.

Panel displays have some inherent issues because of the grid structure of the pixels. Firstly, if the MTF of the display system is well resolving the pixels, the user might be able to see
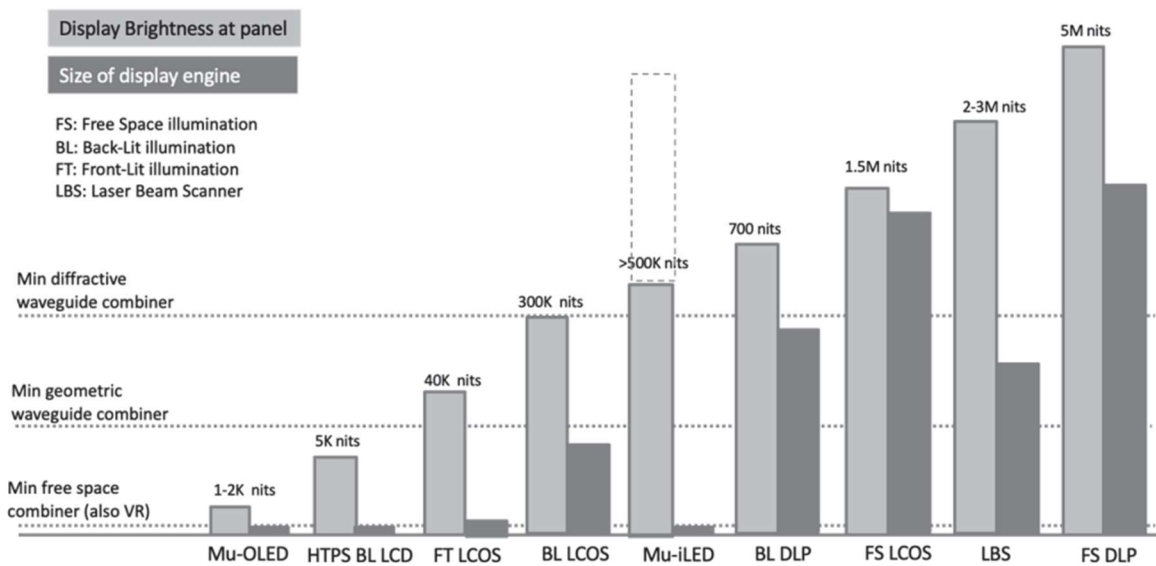


*Figure 13: Brightness and size of display engines for various architectures [Kress, 1].*

the gaps between the pixels, coined the "screen-door" effect. This effect can be mitigated by reducing the pixel spacing, or by intentionally reducing the MTF of the system so the gaps cannot be resolved. This mitigation technique can be seen in figure 14 below. This, however, also affects the apparent display quality.

27

Reducing this effect is still a topic of research. One method employed by Nguyen et al. is to mechanically shift the pixels using piezo actuators to increase the apparent size of the pixels and to reduce the spacing between the pixels [9]. This has shown good qualitative results but has yet to completely mitigate the effect.

A second parasitic effect of these display panels is the so-called Mura effect, from the Japanese meaning unevenness or nonuniformity. Displays are typically composed of multiple layers of material and substrates bonded together. It is virtually impossible to combine these layers with high precision every time resulting in imperfect illumination of the entire display. Mura defects are contrast defects, where one or more pixels is brighter or darker than surrounding pixels when they should have uniform luminance [40]. An immersive display such as VR or a large television increases the perception of the Mura effect. AR systems have a
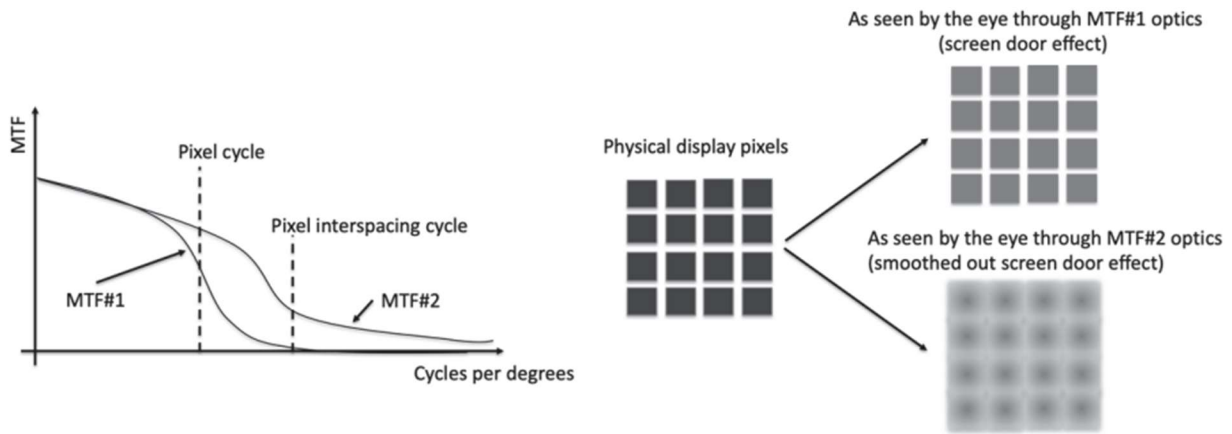


*Figure 14: Tuning the MTF of the optics to smooth out the screen door effect [Kress, 1]*

lessened perception of the effect because the background color and uniformity change constantly as the head moves around the scene. The Mura effect is a fundamental design flaw of current panel-based displays, but U.S. patent 8049695B2 shows the workings of a correction

scheme that uses a gray level display and interpolated corrective data to reduce the Mura effects generally visible by the human eye [40].

## 4.2.2 Scanning Displays

Scanning displays are implemented in many HMDs. The main advantages of these systems are their small size, high brightness and efficiency with laser illumination, high contrast, and optical foveation. Scanning is usually accomplished with either cascaded 1D MEMS (microelectromechanical systems) mirrors (figure 15a) or a single 2D MEMS mirror (figure 15b). Due to the scanning nature of LBS, individual pixels are not confined to a pre-defined grid like in panel-based displays. Using two cascaded mirrors can help with the angular swing amplitude and speed. One major drawback of using two 1D mirrors is the alignment of both mirrors within the opto-mechanical tolerances of the system [33]. The driving electronics of this architecture is also typically more complex and larger. Using a single 2D MEMS mirror, where the mirror can



*Figure 15: LBS architecture based on two separate 1D MEMS mirrors (a) and architecture based on only one 2D MEMS mirror capable of moving independently in two axes (b). [Reitterer, 33]*

be scanned in both axes, requires a careful design because of the mirror deformations when in motion, and the mechanical suspension of the mirrors. The main advantage of this system is that there is only the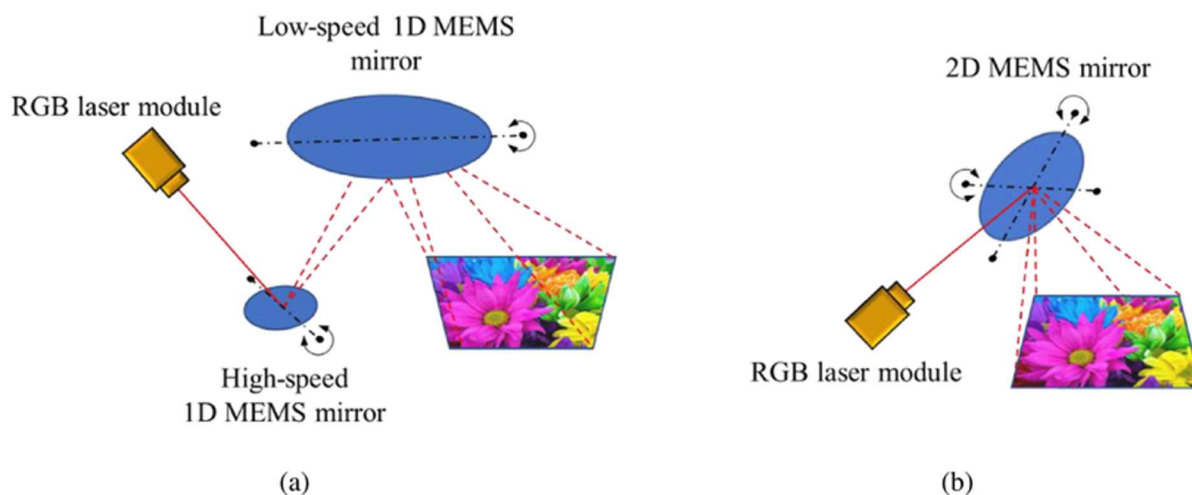 need for one mirror, thereby having fewer alignment steps, smaller driving electronics and lower power consumption [33]. The disadvantages include the possible crosstalk between the two axes, and difficulty in achieving higher frequencies because of the larger mirror required compared to a 1D MEMS mirror. In many cases, an intermediate image and a diffuser are required to form a decent eyebox [41]. This reduces the appeal of a MEMS display because relay lenses are required and the presence of a diffusing element in the optical path may cause speckle to appear.

## 4.3 Optical Combiners

Optical combiners are a complex and costly optical element in HMD architecture. Optical combiners image the exit pupil of the display optics into the pupil of the eye, while also providing a see-through view into the real world in AR applications. They often define the size and aspect ratio of the entire headset. The main types of combiners used in AR NTE systems include free-space, freeform TIR, waveguide, and holographic combiners.

### 4.3.1 Free-Space Combiners

Free-space combiners have been used extensively in defense applications, especially for HMDs in aircrafts. These systems can be very simple to create and understand but become more complex as the FOV increases. The most straightforward free-space combiner would be a tilted flat 50/50 mirror or a beamsplitter, figure 16 (left). These simple architectures typically

have a FOV of less than 30 degrees as tight limitations need to be present to keep the eyebox size reasonable.

A simple extension of this combiner is to add a second fully reflective mirror (figure 16 right). This is used in minimalist AR displays that implement the user's smartphone as the display panel. This shape allows for low cost (there is a cardboard DIY kit produced by Aryzon for 30 Euros) and a relatively large FOV of 45 degrees because of the large display.

Other free-space combiners include using one or multiple mirrors to reflect the light to the eye. One such example is called a birdbath which combines a beamsplitter with a curved mirror. These are commonly used because they can reach relatively large FOV (AHMD with a
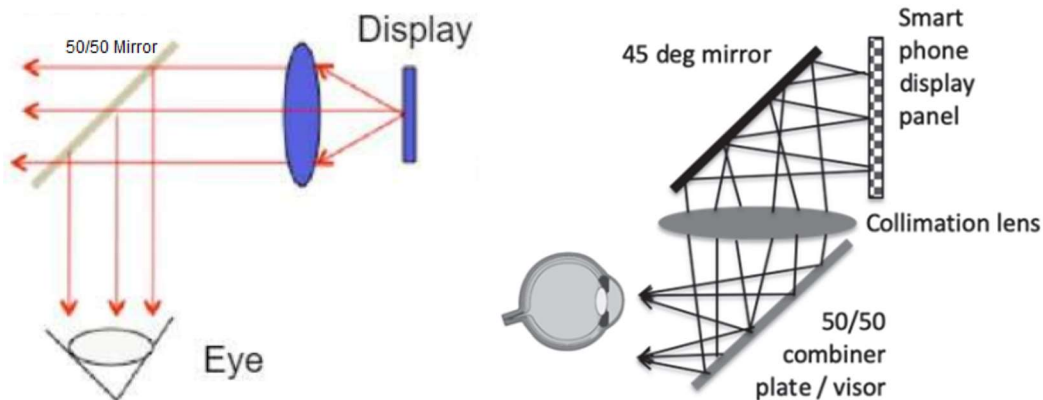


*Figure 16: One [Kress, 42] and two-mirror free-space combiners [Kress, 1]*

65-degree HFOV and BAE with a 40-degree DFOV) [34] and can better correct for distortion and pupil swim since there are more optical surfaces to work with. Seen in figure 17, these birdbath architectures can use a partially reflective mirror, or a fully reflective mirror with a beamsplitter as the combiner. A fully reflective mirror would lead to higher efficiency overall, but also has a bulkier form factor. When choosing the correct free-space combiner, the optical engineer must

consider the power output and size of the display to have a bright enough image for the user. In these examples, the overall efficiency of the combiner design architecture must be chosen properly, since there will be about 50% of light lost at each pass of a 50/50 partial reflector,
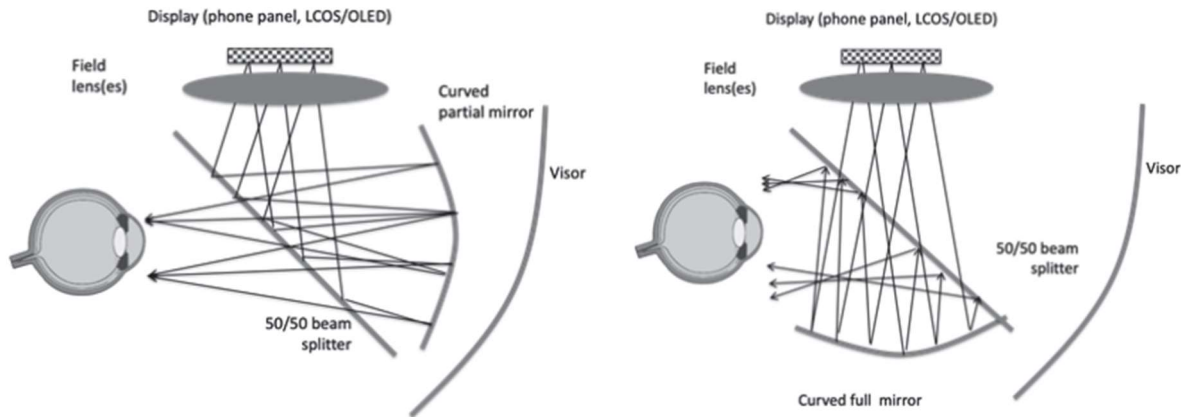


*Figure 17: Partially and fully reflective mirror birdbath design for use in AR glasses [Kress, 1]*

leading to only 12.5% of the light reaching the pupil in figure 17 (left) and 25% of the light reaching the pupil in figure 17 (right).

## 4.3.2 Freeform TIR Combiners

Total internal reflection (TIR) is the Fresnel reflection of light from a higher index to a lower index when the incident angle is larger than the critical angle, figure 18 (right). When this happens, there is no transmission of the light through the interface and all the light gets reflected following Snell's Law. To find the critical angle, we need to use Snell's Law (below) to determine the critical angle that TIR occurs shown in the equation below, where n is the index of refraction, and $\theta$ is the angle of the ray inside the medium with respect to the surface normal. For TIR, we require that $n_1 > n_2$, where $n_1$ is the index inside the propagating medium.

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)$$
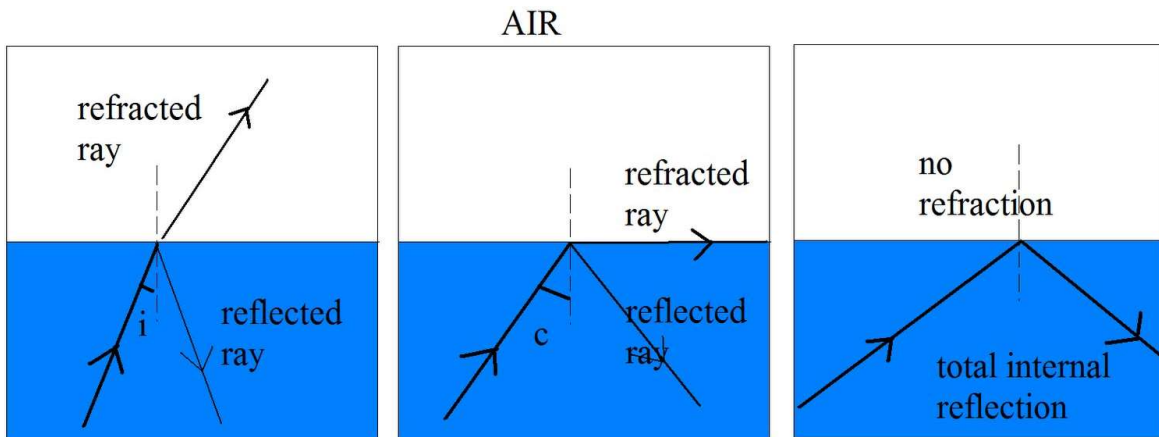


*Figure 18: Illustration of Snell's Law and TIR [https://4.bp.blogspot.com/-ZTvAeZX87Ug/TiWENZNP5XI/AAAAAAAAAEw/6Uvhg-s8ph8/s1600/fig1.bmp]*

We find that $\theta_c$ is the critical angle, where any angle larger than this will result in TIR. This value ends up being about 41.8 degrees for indices of 1.5 (standard glass) and 1 (air).

Typical freeform prism combiners include a semi-transparent coated surface and a TIR surface. Figure 19 shows a design taken from U.S. Patent 20140009845A1. The use of the TIR surface in this design is interesting because it only reflects the light incident from the microdisplay, not from the outside world. This design allows real objects to transmit through the eyepiece to be combined with the microdisplay at the eye pupil. The compensator (second piece of glass attached to the half-mirror surface in figure 19) is used here eliminate the power from the curved surface of the prism, so that real objects can pass through the prism unperturbed. Compensators will generally be another prism bonded to the TIR prism or

separated by an air gap if TIR is required in that section of the freeform. The TIR structure could allow for a single TIR reflection, or multiple (figure 19 middle). With the emergence of freeform diamond-turning machines, freeform prism combiners have become more popular and easier to manufacture [1].

Figure 19 (right) shows a custom freeform optical design which includes a tunable lens and an aperture array created by Huang and Hua in a recent paper [14]. A compensator was later added to the eyepiece to correct for the see-through aberrations present when looking into the real world. This design was able to create a 3D display view of 30 degrees by 18 degrees, with a spatial resolution of 3 arc minutes and provides a see-through field of 65 degrees by 40 degrees. The use of the tunable relay group, just before the eyepiece, allows for focusing to different diopters depending on the user's visual acuity.
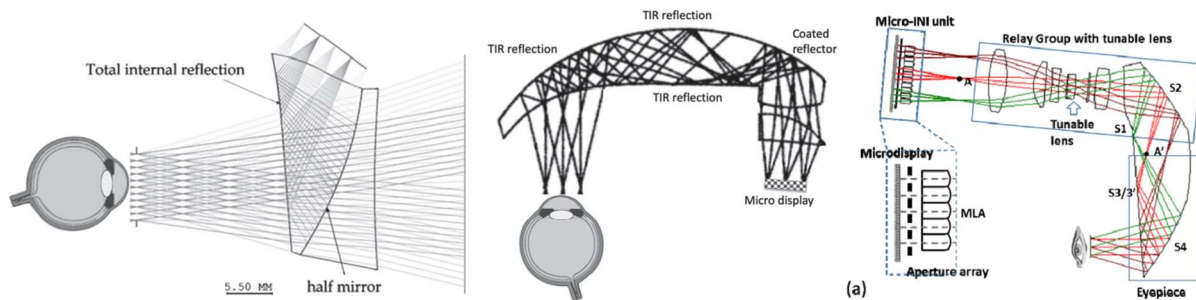


*Figure 19: (Left) TIR combiner prism with a compensator [US patent 20140009845A1]. (Middle) A 5 bounce curved TIR prism combiner [US patent 201201162549A1]. (Right) TIR prism that includes part of the prism with the relay group [Huang, 14]*

An issue with using freeform combiners is their thick and heavy components. These freeforms can be diamond turned from glass or injection molded with plastics, but both need large optics to establish a wide FOV of over 30-40 degrees. Plastic materials will help to reduce the weight of the components but, there is usually some residual birefringent stress in the

elements that could contribute to image doubling or a general degradation of the MTF. The advantage of these devices are the many degrees of freedom in the design. These structures can give a high-quality image because each TIR bounce can be specifically shaped to provide a reduction in aberrations. As the number of TIR reflections increase, the prism can become smaller and will behave more like a waveguide.

### 4.3.3 Waveguide Combiners

Waveguide combiners are based on TIR propagation of the entire field in an optical guide with a single entrance pupil and often many exit pupils. These consist of the input and output couplers that can be either simple prisms, micro-prism arrays, embedded mirror arrays, surface relief gratings (SRGs), thin or thick analog holographic gratings, metasurfaces, or resonant waveguide gratings. 1D or 2D pupil expansion may be necessary to keep the eye pupil within the eyebox properly during rotation when looking around. Pupil expansion may also be necessary in order to cover a wide range of interpupillary distance as discussed earlier in this paper. Figure 20 shows a few examples of flat waveguides with a single exit pupil outcoupler.



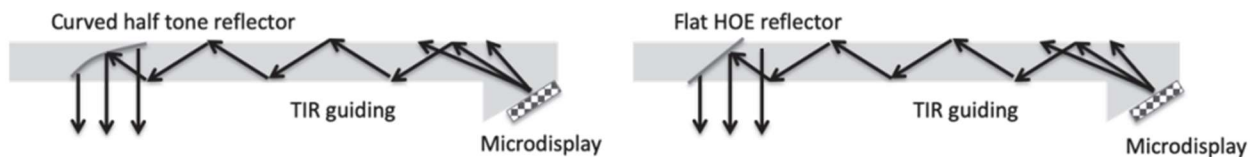*Figure 20: Single exit pupil flat waveguide combiners with a curved reflector (left) or a flat HOE reflector (right) out-coupler [Kress, 1]*

These have been implemented into products such as the Espon Moverio BT-300 (23-degree DFOV). Waveguides don't have to be flat, however. Adding curvature to the waveguides makes the construction much more complicated but improves the form factor by contouring the

waveguide to the user's face, similar to eyeglasses. This has been seen in products by Tooz technologies and in many patents owned by Oscar A. Martinez and Ozan Cakmakci, such as U.S. patent 9366869B2 [43]. One needs to take into consideration the power imprinted on the TIR field at each reflection when designing the outcoupler.

### 4.3.4 Waveguide Couplers

Various techniques to couple light into a waveguide include cutting the waveguide at an angle (Lumus LOE and Optivent Prisms, figure 21 incoupler), mirrors (Google Glass and Espon Moverio BT300), HOEs [44], and SRGs [45]. The coupler elements are the key feature of the waveguide combiner and must be chosen properly based on the type of waveguide used and other constraints in the system (i.e., wavelength selection, index of refraction and cost). A general protocol for coupler design is that the in-coupler should have a high efficiency (thereby reducing the wasted illumination from the display) and the out-coupler should have a uniform light output, creating a uniform eyebox (as described in section 2.1). Both couplers should have an adequate angular bandwidth to accommodate a reasonable FOV and be optimized to avoid undesired diffraction orders if using diffractive couplers. Waveguide couplers can be broken down into two main categories: refractive/reflective and diffractive/holographic.

The simplest TIR coupler is a prism. This prism can be bonded to the waveguide, or the waveguide can be cut at an angle (figures 20 and 21 at the microdisplay location) to allow normally incident light to be refracted into the waveguide and guided by TIR. This is done by using a prism of the same or different index of refraction to insert the light into the waveguide at the proper angle so that TIR occurs in the waveguide-air interface.
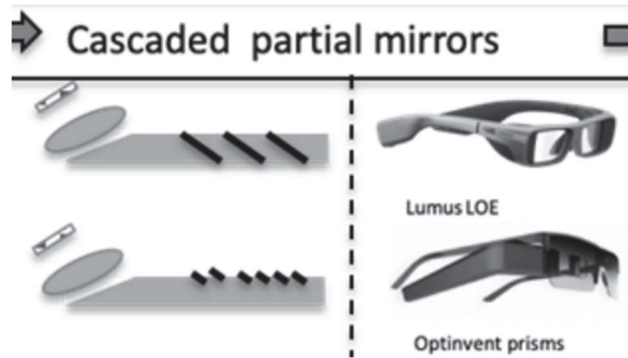
*Figure 21: Examples of cascaded partial mirrors [Kress, 1]*

Using cascaded partial mirrors is another approach for out-coupling the light from a waveguide (used in the Lumos LOE) [1]. The multiple mirrors contribute to an increased eyebox size through exit pupil expansion. The construction of these partial mirrors generally involve coating mirrors on cut-apart pieces and integrating them back together, which may result in a high cost. In figure 21, there are two examples of using cascaded mirrors with references to the product they are used in. In the top image, partial mirrors span the entire waveguide and in the bottom image, microprisms are used which can all be coated with a partially reflective coating mirror layer, or can have alternatively fully reflective and transmissive facets to provide a 50% transmission see-through experience.

## 4.3.5 Holographic and Diffractive Couplers

HOEs and other diffractive optics working in transmission or reflection are very effective couplers for AR architecture. A full-color hologram based on multiplexing can allow for a single plate waveguide architecture which reduces the weight, size and costs while increasing yield (no plate alignment required), however a main issue with this design is the reduced efficiency, possible crosstalk and a limited FOV with a single TIR waveguide. As a general rule, the

efficiency of each hologram during multiplexing follows an inverse square law $\frac{1}{N_h^2}$, where $N_h$ is the number of holograms. However, in a material where a large index modulation can be achieved, this rule can be broken and a larger efficiency per hologram can be achieved [37]. A material must also be chosen properly so that it is panchromatic, or sensitive to all wavelengths chosen.

In figure 22, three separate waveguides and holographic couplers are utilized for each wavelength and recombined at the exit pupil. This is possible due to the wavelength selectivity inherent in holographic optics. Like holographic gratings formed from interferometric exposure, surface relief gratings formed using other methods are also very effective in coupling the light into waveguides. The structures of these gratings come in many forms including blazed, slanted, binary and multilevel. Asymmetric gratings are more desirable, especially blazed gratings, but typical periods for TIR grating couplers in the visible spectrum are below 500 nm and are difficult to manufacture with few errors [46].
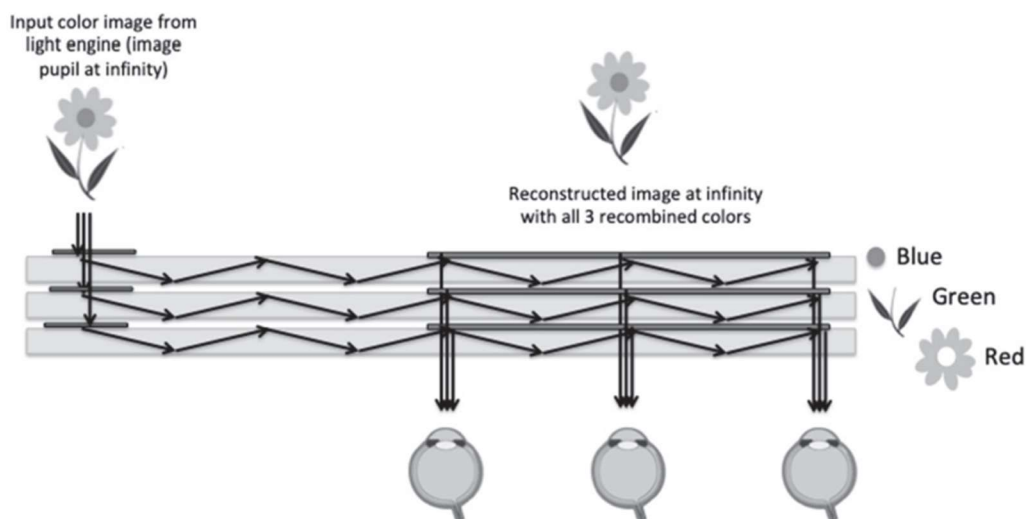


*Figure 22: Stacked waveguide combiners using holographic couplers separately for each wavelength [Kress, 1]*

Figure 23 shows various SRGs and how they can be integrated as in-coupling and out-coupling elements. SRGs can be chosen properly to diffract specific spectral bands and pass other bands to be processed by the next coupler area located on the following guide. This allows a similar architecture between holographic recordings (or photopolymer gratings, PPGs) and SRGs as seen in figure 24 below.
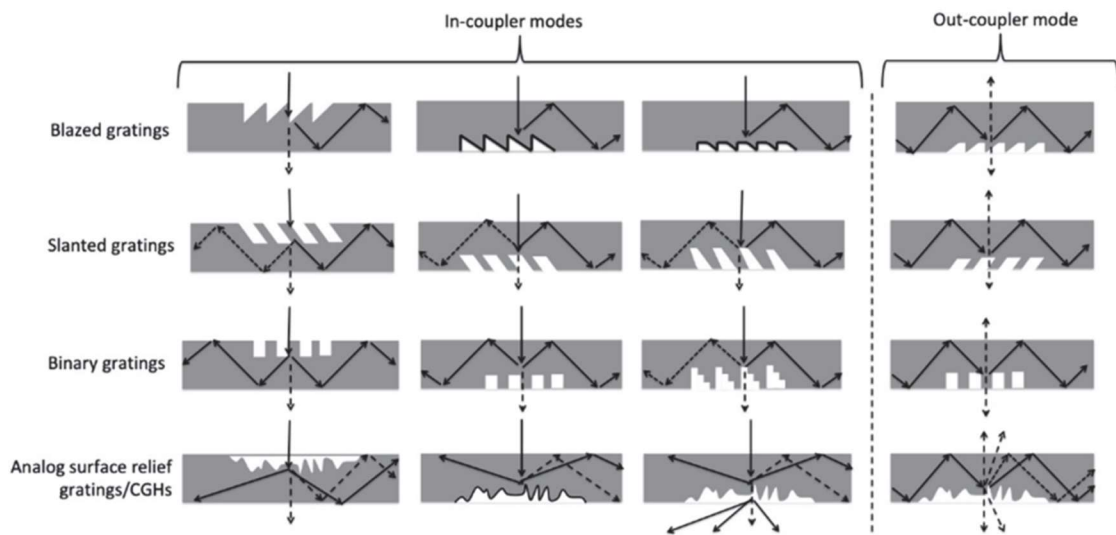


Figure 23: Surface relief grating types used as waveguide combiner couplers [Kress, 1]
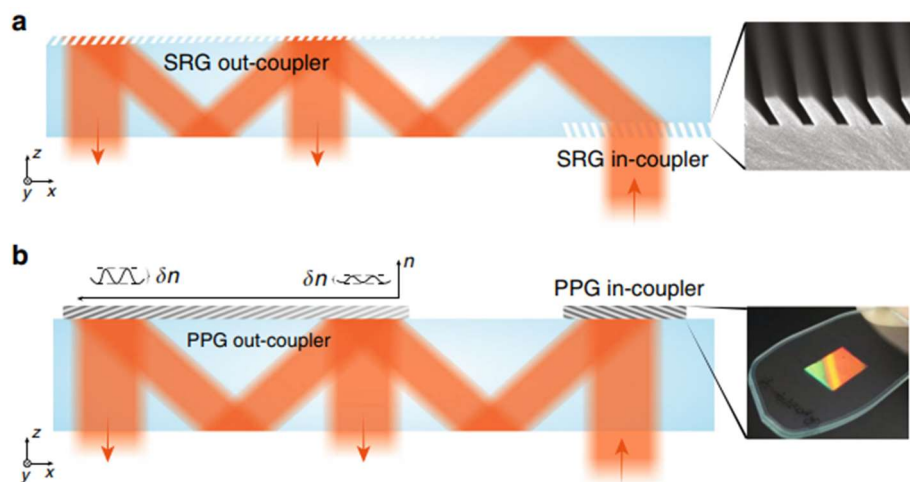


Figure 24: Schematics of waveguide combiners based on (Top) SRGs, and (Bottom) PPGs [Xiong, 6]

In general, the most common processes for fabrication of these SRGs involve photolithographic methods, direct machining or replication using molds and casts. Holographic interferometric exposure is often used when sub-wavelength period structures are desired. This method is also useful when generating a large volume of these elements. Binary optics is a method of lithographically exposing a photoresist layer on the substrate and is a very common method of making diffractive optics. This process can produce $2^N$ phase levels from $N$ masks which gives a very high flexibility in comparison to interference recording [38], however due to the standard grating period required (less than 500 nm), this process becomes very fragile and difficult to manufacture properly [46]. One other method of creating these SRGs is to use mechanical ruling, which uses a sharp stylus tip to scrape away the substrate material in a highly controlled manner, capable of thousands of grooves per millimeter. Because of the finite size of the tip, it is necessary to achieve a balance between the fineness of the cut and accuracy of the surface profile. For more information on SRG fabrication techniques, see reference [38].

## 4.4 Exit Pupil Expansion

Seen in previous examples, pupil expansion is commonly needed when the FOV for the AR/MR display is large, or binocular vision displays are used in order to match the exit pupil of the imaging system to the entrance pupil of the eye. A quick method without any pupil expansion is to mechanically adjust the location of the pupil by moving the entire system. This is a popular method in monocular AR glasses and lower resolution binocular MR systems. When these mechanical adjustments are ruled out, the next option would be increasing the size of the exit pupil.

## 5.4.1 Exit Pupil Replication

One popular method for increasing the eyebox is using exit pupil replication, which can be done in 1D or 2D. In the exit pupil expansion (EPE) process, a portion of the trapped light is repeatedly coupled out of the waveguide in each TIR, creating multiple exit pupils. This increases the effective eyebox size [6]. This is done so that the image field appears at least once over the full human pupil. As the eye moves around, it can move out of the original pupil, but through pupil expansion, a second pupil can be seen by the eye. This is shown in figure 25 below, which uses a spatially varying surface relief grating to outcouple a uniform eyebox. This is important so that each pupil replication has the same intensity distribution. If a uniform SRG is used to outcouple the light, then each pupil would have a dimmer output compared to the previous pupil.

For instance, if we assume the model in figure 25, such that light gets out-coupled four times, then the diffraction efficiency of each TIR bounce should be 1/4, 1/3, 1/2, and 1 respectively. Suppose we have 100 units of power in the waveguide. For a uniform pupil replication, ideally each outcouple should have 25 units and pass the rest, resulting in 25% diffraction efficiency. This leaves 75 units for the second TIR bounce. If 25 units are to be outcoupled, then we need 25/75 or 33.3% diffraction efficiency and so on. This efficiency distribution can be written explicitly in the following equation:

$$\eta_i = \frac{1}{N_T - N_i + 1}$$

where $\eta_i$ is the diffraction efficiency of the $i$th pupil, $N_T$ is the total number of replication pupils and $N_i$ is the individual replication pupil (where $N_i = i = 1$ for the first outcoupled pupil).

There are many ways to modify the grating including changing the depth, duty cycle, slant angle, and blaze angle. Modeling the efficiency of SRGs can be performed accurately using the rigorous coupled wave analysis (RCWA), the Fourier modal method (FMM) or the finite difference time domain (FDTD) method. The FDTD can model non-periodic nanostructures and show all the diffracted fields, polarization conversions and the entire complex field [1] but require a time costly calculation process and a large amount of memory. RCWA can accurately model quasi-periodic structures but will only show the diffraction efficiencies of each order.
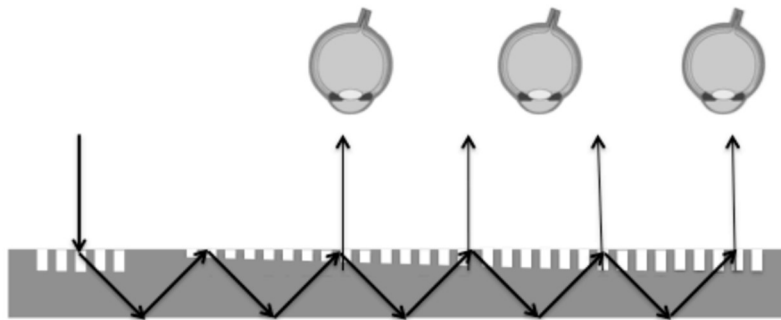


*Figure 25: Modulation of the outcoupling efficiency to build up a unform eyebox [Kress, 1]*

## 5.4.2 Pupil Switching

Another method for placing the eyebox in the correct location is called pupil switching. This is accomplished using a pupil tracker. Instead of providing a large eyebox, the exit pupil is shifted to cover the expanded eyebox area. A proposed practical method is described by Jang et al. in which the main diffractive HOE takes an oblique point source and diffracts it into the eye [16]. As seen in figure 26 below, the hologram provides maximum diffraction efficiency with a certain angle of incidence and wavelength. Additionally, within a certain degree of tolerance,

the HOE still behaves as expected without excessive efficiency degradation which is shown in the figure (b,c). Within this tolerance, and in accordance with pupil tracking, it is possible to create a shifted exit pupil. Thus, the eyebox can be effectively increased to cover a larger area.
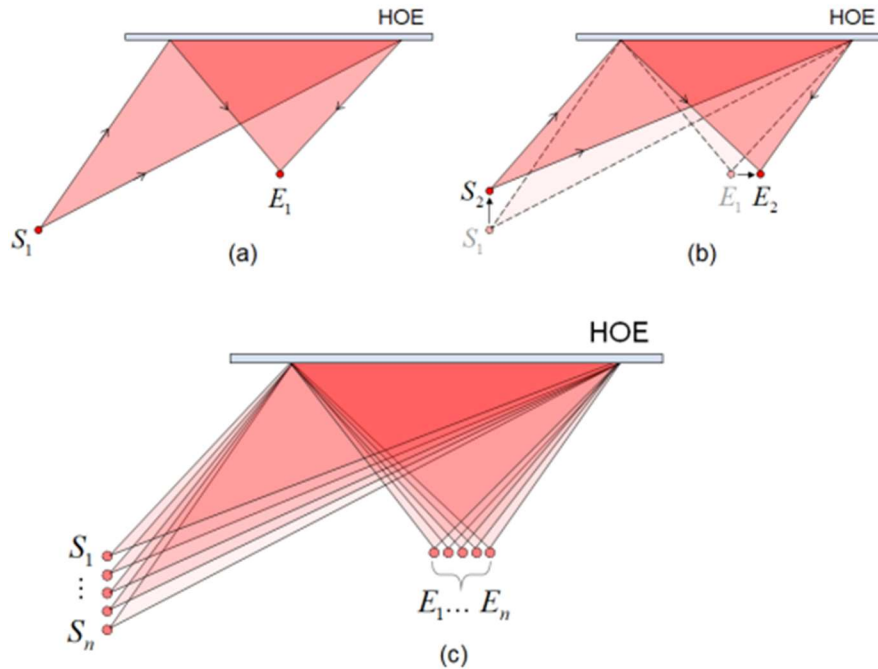


*Figure 26: Basic principle of exit-pupil switching for a lens HOE [Jang, 16]*

To implement this switchable point source, a second HOE is used which they define as a pupil-shifting HOE (PSHOE). This is accomplished by using a lens array to focus the signal wave, and a spherical wave from a point source as the reference wave to create the hologram. In this way, numerous lenslet hogels can be formed that generate point light sources with shifted locations. These can be seen in figure 27 below. Using this design, the HOE can create a light source array within a narrow space by using a single light source and only changing the direction of the laser beam. The additional lens is used so that each point source is focused on a certain distance to use the maximum area of the SLM in the prototype design shown in the figure below.
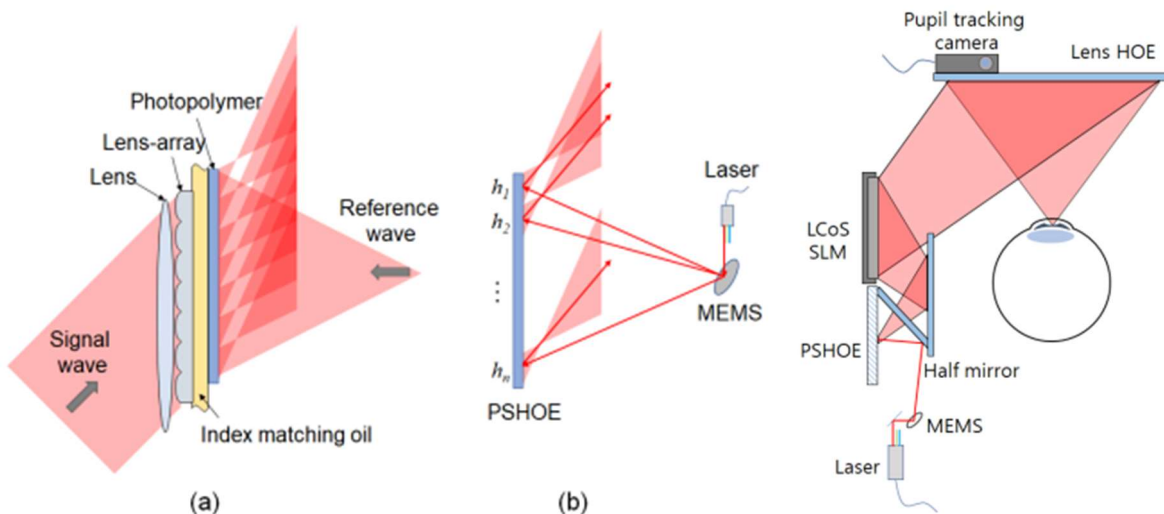
*Figure 27: Configuration of proposed pupil shifting HOE: (a) Optical Fabrication stage, (b) operation of PSHOE, (right) design example of a compact prototype [Jang, 16]*

Using this method, the eyebox is expanded from 2mm to 9mm, which is close to the desired minimum of 10mm for practical use [1,6,16]. This minimum is based on accommodation of different interpupillary distances and the average size of the human pupil. If the shift is larger than a certain level, distortion becomes too large, and the exit pupil cannot operate properly due to vignetting of the SLM and the half mirror. Thus, the resolution of the image will also decrease because less pixels on the SLM are being used. This paper shows a practical use of pupil switching for the purposes of pupil expansion.

## Chapter 5. Case Study: Pancake Lens VR Prototype

Many current commercially available VR systems still have box-like form factors and provide only a fraction of the resolution of the human eye. In the following work by Andrew Maimone and Junren Wang at Facebook Reality Labs, a pancake based optical system is designed for VR use which has a form factor nearing sunglasses [11]. They combine the

techniques of polarization-based optical folding and holographic optics to gain a better performance than if they were used separately. They augment their solution with a laser illuminated LCD, directional backlighting, and color-multiplexing to achieve a wide FOV and high



*Figure 28: (Left) Sunglasses prototype. The driving electronics and light sources are mounted externally. (Right) Display module for this prototype. Consists of a backlight, display panel, and eyepiece. The total thickness of the module is 8.9mm [Maimone, 11]*

resolution while keeping the thicknesses to under 10 millimeters. A prototype of this model is shown in figure 28. I want to discuss some of the techniques they used in their designs and the effectiveness in the results of their prototypes.

Generally, in NTE displays, there needs to be some focusing power to collimate the display source or focus if the object is not at infinity. This is accomplished by through the combination of adding curvature to the beamsplitter surface, adding curvature to the reflective polarizer surface, or by adding refractive transmissive elements. The focusing power is limited by the physical space of the optics and can add significant weight to the display. The approach

45

here is to use HOEs to provide all the focusing power such that the elements in the pancake can remain flat. Most of the weight from the HOEs come from the substrate they are attached to, which means that very little weight would be added to the system if these holograms were placed on the beamsplitter or the reflective polarizer. Flat holographic surfaces can also have high focusing power, which is only limited to optical aberrations and fabrication constraints [11]. Also, color multiplexing can be used such that each wavelength can be individually controlled by the hologram, giving more degrees of freedom than refractive glasses.

There are some inherent issues present in this design. Namely, holographic optics are highly dispersive over a range of wavelengths and angles. This drives the need for laser illumination, which has a small spectral bandwidth compared to LEDs and will provide a higher efficiency. However, using laser illumination introduces the additional challenge of eliminating speckle. The term speckle refers to the random interference of a highly coherent source with itself when diffusely reflected at a surface. This generates a granular pattern that severely diminishes image quality of laser projection displays seen in figure 29 (left). There are many known ways to
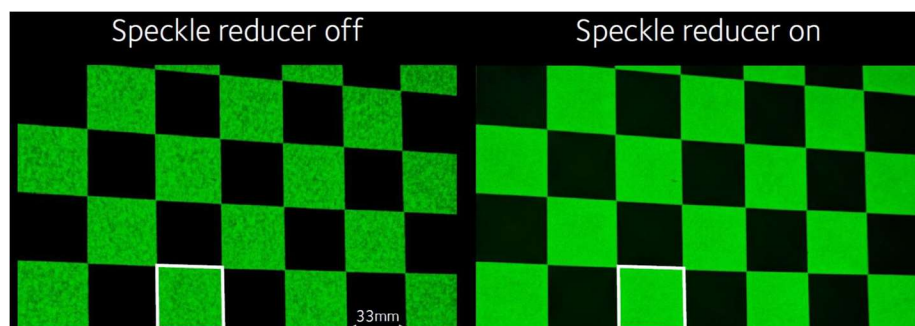


*Figure 29: Contrast levels without (left) and with (right) a laser speckle reducer [Oxford University Innovation Ltd.]*

reduce the effects of speckle (figure 29 right) including polymer dynamic diffraction gratings, MEMS deformable mirrors, MEMS in-plane vibrations, MEMS oscillating mirrors, sinusoidal vibrating random diffuser and more [39]. This group decided on a laser speckle reducer from Optotune which uses four electro-active polymers to move the central diffuser along the X and Y axes, forming a circular pattern of motion. This helps the collimated source to appear as a uniform distribution of light.

Because of the angular selectivity of the holograms, directional backlighting is employed to limit the angle of light incident on the holographic surface. A typical display panel emits light in all directions, approaching a solid angle of $2\pi$ steradians, but in a typical near-eye display, only a fraction of this light contributes to the viewing eye box and the rest is wasted and can scatter inside the display and cause a reduction of image contrast. Thus, there is a need to control the emission angles of the display panel. One such technique uses a microstructure light guide plate with three micro-prism arrays to direct the light. This method achieved high uniformity and a FWHM at $\pm 4°$ with low power consumption [17]. The results for this device are shown in figure 30. Directional backlighting is not always necessary in MR systems, but it does provide the capability for a brighter image which is necessary for outdoor use. In these pancake holographic lenses, it also allows for better angle control which increases the efficiency of the hologram and reduces stray light entering the eyebox. Due to the pancake design, the system here is prone to ghost images which are seen in the benchtop model, but not as noticeable in the sunglasses prototype due to the use of anti-reflection coatings on the surfaces.
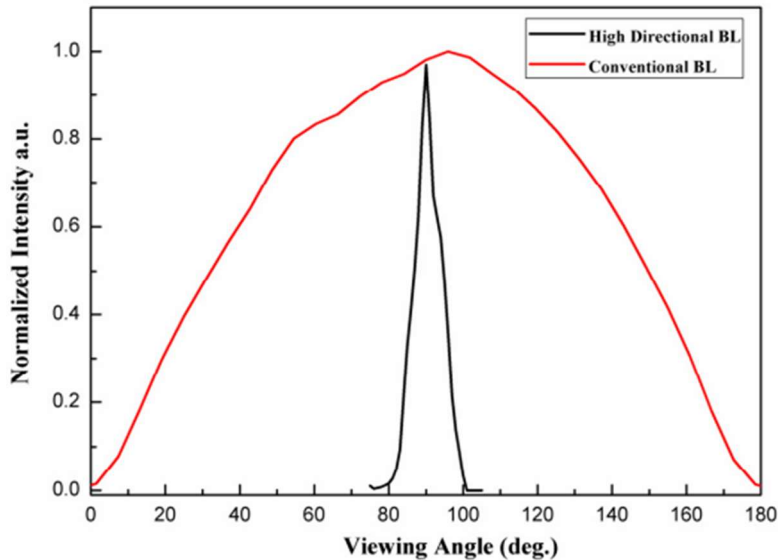
*Figure 30: Light intensity distribution of conventional and high directional backlighting [Wang, 17]*

Three separate prototypes were created using the techniques listed above. One benchtop full-color model, a scaled down monochromatic version using one holographic element, and a high-resolution monochromatic prototype that uses two holographic elements. All three versions had a diagonal FOV of over 90 degrees, and resolution ranging from 1-5 arcminutes. These values are comparable to current VR headsets (Oculus Quest 2 with an 89x93 (HxV) FOV and Sony PSVR with a 100-degree HFOV), and therefore was a successful step in flattening the VR headset into a smaller form factor.

## Chapter 6. Conclusion

The motivation behind this paper was to aggregate a collection of articles and papers that showcased the current specifications and requirements for creating MR architectures. This report is by no means a complete collection, as there is ongoing research on the topic, both by

industry and through educational research. This report shows the main approaches of creating

AR/VR/MR head mounted displays from an optical engineering point of view. This deals

specifically with the display engine, combiner optics and imaging optics of the system. Various

pitfalls of the design process are selected, and mitigation techniques revealed. There are many

aspects to designing a sophisticated head mounted display, but there are still many more

options and techniques that haven't been tried yet. As the development of better and smaller

display panels and pixels are created, the mixed reality world will begin to integrate more into

the real world, but we are still years away from having truly immersive systems.

# References:

[1] Kress, B. C. (2020). *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets*. SPIE.

[2] Cakmakci, O., Hoffman, D. M., & Balram, N. (2019). 31-4: invited paper: 3D eyebox in Augmented and Virtual Reality Optics. *SID Symposium Digest of Technical Papers*, *50*(1), 438–441. https://doi.org/10.1002/sdtp.12950

[3] Draper, C. T., Bigler, C. M., Mann, M. S., Sarma, K., & Blanche, P.-A. (2019). Holographic waveguide head-up display with 2-D pupil expansion and longitudinal image magnification. *Applied Optics*, *58*(5). https://doi.org/10.1364/ao.58.00a251

[4] Kramida, G. (2016). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, *22*(7), 1912–1931. https://doi.org/10.1109/tvcg.2015.2473855

[5] Geng, M., Bryars, B. J., Wheelwright, B. M., Peng, F., Lam, W. S., Fu, Y., Sohn, A., Sulai, Y., Gollier, J., McEldowney, S., Lewis, B., Chan, N., Fix, A., Lanman, D., Cardenas, N., & Yoon, Y. (2018). Viewing optics for immersive near-eye displays: Pupil swim/size and weight/stray light. *Digital Optics for Immersive Displays*. https://doi.org/10.1117/12.2307671

[6] Xiong, J., Hsiang, EL., He, Z. *et al.* (2021). Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light Sci Appl* **10,** 216 https://doi.org/10.1038/s41377-021-00658-8

[7] Sulai, Y., Geng, M., Luanava, S., Gao, W., Gollier, J., Wheelwright, B. M., & Choi, S. (2018). Field of view: Not just a number. *Digital Optics for Immersive Displays*. https://doi.org/10.1117/12.2307303

[8] Zhan, T., Yin, K., Xiong, J., He, Z., & Wu, S.-T. (2020). Augmented reality and virtual reality displays: Perspectives and challenges. *IScience*, *23*(8). https://doi.org/10.1016/j.isci.2020.101397

[9] Nguyen, J., Smith, C., Magoz, Z., & Sears, J. (2020). Screen door effect reduction using mechanical shifting for virtual reality displays. *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR)*. https://doi.org/10.1117/12.2544479

[10] Kostuk, R. K. (2021). *Holography: Principles and applications*. CRC PRESS.

[11] Maimone, A., & Wang, J. (2020). Holographic optics for thin and lightweight virtual reality. *ACM Transactions on Graphics*, *39*(4). https://doi.org/10.1145/3386569.3392416

[12] Wong, T. L., Yun, Z., Ambur, G., & Etter, J. (2017). Folded optics with birefringent reflective polarizers. *Digital Optical Technologies 2017*. https://doi.org/10.1117/12.2270266

[13] Wu, T., Sher, C.-W., Lin, Y., Lee, C.-F., Liang, S., Lu, Y., Huang Chen, S.-W., Guo, W., Kuo, H.-C., & Chen, Z. (2018). Mini-led and micro-led: Promising candidates for the Next Generation Display Technology. *Applied Sciences*, *8*(9), 1557. https://doi.org/10.3390/app8091557

[14] Huang, H., & Hua, H. (2018). High-performance integral-imaging-based light field augmented reality display. *Digital Optics for Immersive Displays*. https://doi.org/10.1117/12.2315661

[15] Taya, S., El-Agez, T. (2012). Slab waveguide sensor based on amplified phase change due to multiple total internal reflections. *Turkish Journal of Physics*, 36. https:// doi.org/10.3906/fiz-1103-23

[16] Jang, C., Bang, K., Li, G., & Lee, B. (2019). Holographic near-eye display with expanded eye-box. *ACM Transactions on Graphics*, *37*(6), 1–14. https://doi.org/10.1145/3272127.3275069

[17] Wang, Y.-J., Ouyang, S.-H., Chao, W.-C., Lu, J.-G., & Shieh, H.-P. D. (2015). High directional backlight using an integrated light guide plate. *Optics Express*, *23*(2), 1567. https://doi.org/10.1364/oe.23.001567

[18] Song, W., Li, X., Zheng, Y., Liu, Y., & Wang, Y. (2021). Full-color retinal-projection near-eye display using a multiplexing-encoding holographic method. *Optics Express*, *29*(6). https://doi.org/10.1364/oe.421439

[19] Heilig M. (1962). US Patent #3,050,870

[20] Dorothy Strickland, Dan Chartier; EEG Measurements in a Virtual Reality Headset. *Presence: Teleoperators and Virtual Environments* 1997; 6 (5): 581–589. doi: https://doi.org/10.1162/pres.1997.6.5.581

[21] Lim, H. K., Ji, K., Woo, Y. S., Han, D.-uk, Lee, D.-H., Nam, S. G., & Jang, K.-M. (2020). Test-retest reliability of the virtual reality sickness evaluation using electroencephalography (EEG). *Neuroscience Letters*, *743*, 135589. https://doi.org/10.1016/j.neulet.2020.135589

[22] Curcio, C.A., Sloan, K.R., Kalina, R.E., and Hentrickson, A.E. (1990). Human photoreceptor topography. J Comp. Neurol. *292,* 497-523

[23] Evans, J. M. (2006, June 15). *Standards for Visual Acuity*. National Institute for Standards and Technology. Retrieved November 29, 2021, from https://www.nist.gov/system/files/documents/el/isd/ks/Visual_Acuity_Standards_1.pdf

[24] Malacara, D. (2011). *Color vision and colorimetry: Theory and applications*. SPIE Optical Engineering Press.

[25] Palmer, J. M., & Grant, B. G. (2010). *The art of radiometry*. SPIE Press.

[26] Zhan, T., Lee, Y.-H., & Wu, S.-T. (2018). High-resolution additive light field near-eye display by switchable pancharatnam–berry phase lenses. *Optics Express*, *26*(4), 4863. https://doi.org/10.1364/oe.26.004863

[27] A. Sherstyuk and A. State, "Dynamic eye convergence for head-mounted displays," Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology - VRST '10, p. 43, 2010. [Online]. Available: http://portal.acm. org/citation.cfm?doid=1889863.1889869

[28] S. Shiwa, K. Omura, and F. Kishino, "Proposal for a 3D display with accommodative compensation: 3DDAC," Journal of the Society for Information Display, vol. 4, no. 4, pp. 255–261, Dec. 1996. [Online]. Available: http: //onlinelibrary.wiley.com/doi/10.1889/1.1987395/abstract

[29] S. C. McQuaide, E. J. Seibel, J. P. Kelly, B. T. Schowengerdt, and T. A. Furness, "A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror," Displays, vol. 24, no. 2, pp. 65–72, Aug. 2003. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S0141938203000167

[30] ] S. Liu and H. Hua, "Time-multiplexed dual-focal plane head-mounted display with a liquid lens," Optics Letters, vol. 34, no. 11, p. 1642, May 2009. [Online]. Available: http://www.opticsinfobase.org/viewmedia.cfm? uri=ol-34-11-1642&seq=0&html=true

[31] Huang, M. X., & Bulling, A. (2019). Saccalib. *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. https://doi.org/10.1145/3317956.3321553

[32] Nevitt, T.J. and Weber, M.F., "Recent advances in multilayer polymeric interference reflector products," Thin Solid Films 532, 106-112 (2013)

[33] Reitterer, J., Chen, Z., Balbekova, A., Schmid, G., Schestak, G., Nassar, F., Dorfmeister, M., & Ley, M. (2021). Ultra-compact micro-electro-mechanical laser beam scanner for augmented reality applications. *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) II*. https://doi.org/10.1117/12.2576704

[34] C. E. Rash, "Helmet-mounted Displays: Design Issues for Rotary-wing Aircraft," SPIE Press, Bellingham, WA (2001).

[35] S. A. E. Aerospace Standard, AS8055, "Minimum Performance Standard for Airborne Head Up Display (HUD)," (2008).

[36] V. K. Vinod Karar and S. G. Smarajit Ghosh, "Estimation of tunneling effect caused by luminance non-uniformity in head-up displays," Chin. Opt. Lett. **12**(1), 013301 (2014).

[37] Blanche, P.-A. (2014). *Field Guide to Holography*. SPIE Press.

[38] O'Shea, D. C., Suleski, T. J., Kathman, A. D., & Prather, D. W. (2015). In *Diffractive optics: Design, fabrication, and test* (pp. 115–166). essay, SPIE Press.

[39] Akram, M. N., & Chen, X. (2015). Speckle reduction methods in laser-based picture projectors. *Optical Review*, *23*(1), 108–120. https://doi.org/10.1007/s10043-015-0158-6

[40] Y. Ozawa, S.J. Daly, X. Feng, "Correction of visible mura distortions in displays by use of flexible system for memory resources and mura characteristics," U.S. Patent 80409695B2, Nov. 01, 2011.

[41] Kress, B. C. (2014). *Field Guide to Digital Micro-optics*. SPIE Press.

[42] Kress, B., & Shin, M. (2013). Diffractive and holographic optics as optical combiners in head mounted displays. *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*. https://doi.org/10.1145/2494091.2499572

[43] O.A. Martinez, O. Cakmakci, "Thin curved eyepiece for see-through head wearable display," U.S. Patent 9366869B2, Jun. 14, 2016.

[44] H.J. Caulfield, Q. Huang, A. Putilin, V. Morozov, J. Shamir, "Waveguide hologram illuminators," U.S. Patent 5854697A, Dec. 29, 1998.

[45] J. Upatnieks, "Compact head-up display," U.S. Patent 4711512A, Dec. 08, 1987.

[46] Fleming, M. B., & Hutley, M. C. (1997). Blazed Diffractive Optics. *Applied Optics*, *36*(20), 4635. https://doi.org/10.1364/ao.36.004635