

Optical architectures for Augmented and Mixed Reality headsets

Bernard C. Kress

Microsoft Corp. HoloLens Team
1 Microsoft Way, Redmond, 98052 WA. USA

ABSTRACT

This paper is a timely review of the main optical architectures and technologies used today in enterprise and consumer Head Mounted Displays (HMDs), over a range of implementations including smart glasses, Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) headsets.

In addition to the optical architecture review, we discuss the human immersive experience and the need for a human centric optical design perspective to build the most comfortable headset without compromising the user's sensory experience (display, immersion, interaction). We discuss the major optical challenges to overcome in order to provide the user with the visual and sensory experience that will eventually enable the market analysts' optimistic expectations for the coming years. These challenges range from wearable and visual comfort to sensory and display immersion experience. We present the most appropriate optical technologies to address those challenges, and discuss the latest product implementations attempting to solve such challenges.

1. Introduction

Defense has been the first application sector for AR and VR, as far back as the 1960s [1]. Based on such early developments, the first consumer VR/AR boom (which included also the smart glass concept) has expanded in the early 90s and contracted a decade later, as a poster child of a technology endeavor ahead of its time, and also ahead of its markets [2]. Due to the lack of available consumer display technologies and related sensors, novel optical display concepts had been introduced throughout this decade [3], [4], which are still considered today as state of the art, such as the "Private Eye" smart glass from Reflection Technology (1989) or the "VR Boy" from Nintendo (1995), both based on scanning displays rather than flat panel displays. Although such display technologies were well ahead of their time [5], [6], [7], the lack of consumer grade IMU sensors, low power 3D rendering GPUs and wireless data transfer technologies contributed to end this first boom. The other reason was the lack of digital content, or rather the lack of a clear vision of an adapted VR/AR content for enterprise or consumer [8], [9]. As a few decades earlier, the only sector that saw sustained efforts in AR/VR was the defense sector (flight training, helmet mounted displays for rotary wing aircraft and Head Up Displays –HUD- for fixed wing air crafts) [10].

Today's engineers, exposed at an early age to ever invading flat panel displays technologies, tend to act as creatures of habit much more than their peers 20 years ago who had to invent novel immersive display technologies from scratch. We therefore saw since 2014 the first generation implementations of immersive AR/VR HMDs based on readily available smart-phones and pico-projectors (HTPS-LCD, OLED, DLP, LCOS), and IMU/camera sensors. Currently, HMD display architectures are evolving slowly to more specific technologies, which might be a better fit for immersive requirements than flat panels were, and resembling the display technologies invented throughout the first AR/VR boom 2 decades earlier (iLEDs, 1D scanned arrays, laser/VCSEL MEMS scanners, ...).

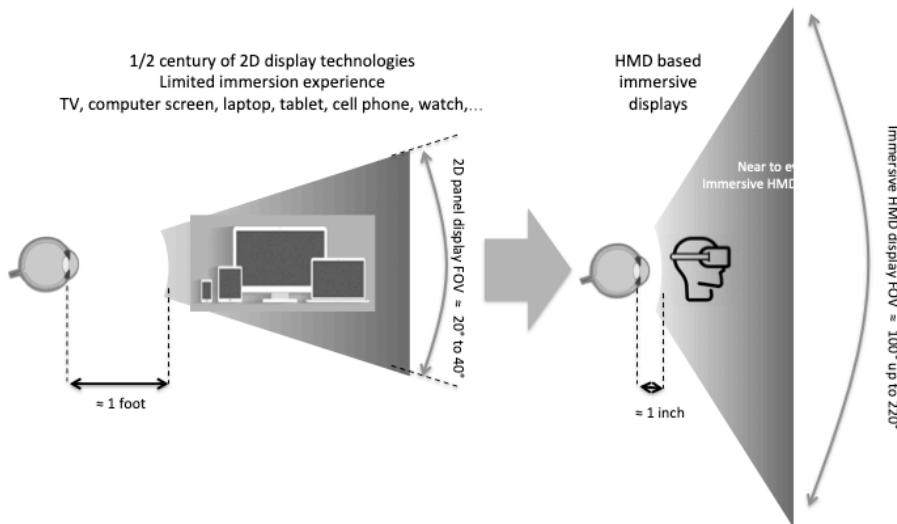


Figure 1: Immersive NTE displays: a paradigm shift for personal information display

The smart phone technology ecosystem, as in display, connectivity and sensors, shaped the emergence of the second VR/AR boom and have been certainly the lowest hanging fruits to be used as the first building blocks by product integrators. Such traditional display technologies might only serve as a initial catalyst for what is coming next. The immersive display experience in AR/VR is however a paradigm shift from traditional panel displays experiences, since more than half a century, starting from CRT TVs, to LCD computer monitors and laptop screens, OLED tablets and smart phones, LCOS or DLP digital projectors, to iLED smart watches (See Figure 1).

When flat panel display technologies and architectures (smart phone or micro-display panels) are used to implement immersive Near To Eye (NTE) display devices, étendue, fixed focus and low brightness become severe limitations. Addressing the needs for NTE immersive displays to match the specifics of the human visual system requires alternative display technologies.

The second emergence of VR/AR/smart-glasses boom in the early 2010s has also introduced new naming trends, more inclusive than AR or VR: Mixed - or Merged - Reality (MR), more generally referred today as "XR", a generic acronym as in "Extended Reality". Smart Eyewear (including both digital information display and prescription eyewear) tend also to replace the initial Smart Glasses naming convention.

2. AR/MR/VR markets.

Today, unlike in the previous AR/VR boom, investors, market analysts and AR/VR/MR system integrators as well as enterprise users are expecting to see a real Return On Investment (ROI) for these unique technologies in the next 5 years, as underlined by the 2017/18 Gartner Hype Cycles (see Figure 2).

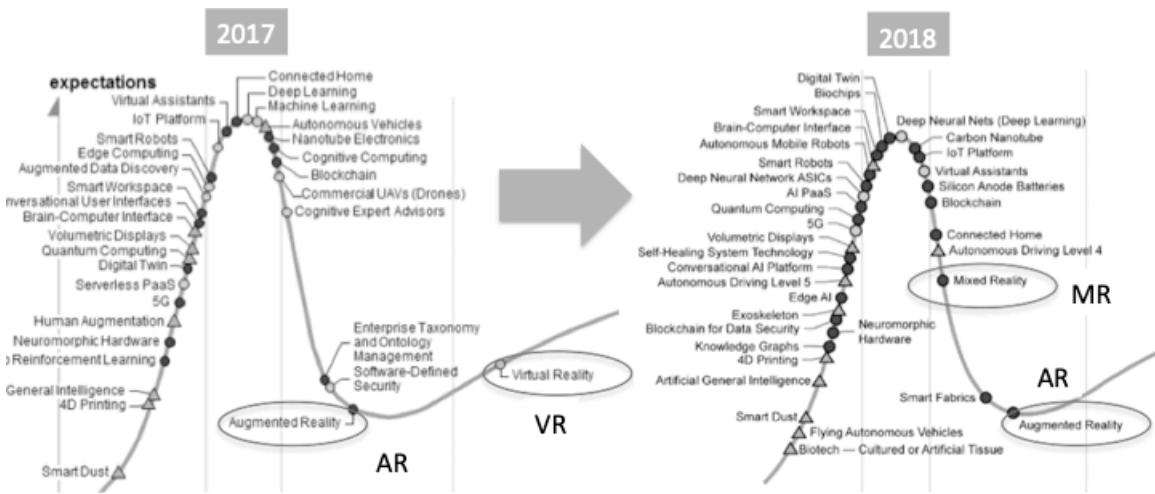


Figure 2: Gartner Hype Cycle showing VR, AR and MR technologies for 2017-2018.

The 2017 Gartner graph shows clearly AR and VR poised to reach the plateau of productivity in 2 to 10 years, with VR preceding AR by a few years. This is a notion shared by most AR/VR market analysts. It is interesting to note that the revised 2018 graph does not show VR anymore, instead introduced MR as departing from the peak of inflated expectations. VR approached in 2018 a more mature stage, even becoming a commodity, moving it thus off the emerging technology class of innovation profiles.

However, one has to take these expectations with a word of caution; the only market sector that has been proven to be sustainable today is MR for enterprise, where the ROI is mainly cost avoidance (*faster learning curves for employees, collaborative design, remote expert guidance, better servicing through overlaid manuals and instructions, enhanced monitoring, higher quality assurance in manufacturing, enhanced product display and demos and better end-user experiences*).

Enterprise sectors that have already shown tangible MR ROI are concentrated in the manufacturing (automotive, avionics, heavy industrial products), power, energy, mining and utilities, technology, media and telecom, as well as in healthcare and surgery, financial services, and retail / hospitality / leisure fields.

Proof of an existing consumer market for smart glasses/AR/MR is less obvious, hardware experiments have been tried out with mixed results for smart glasses (Google Glass, Snap Spectacles, Intel Vaunt). VR headsets developments have also slowed down recently (Oculus/Facebook VR, Google Daydream, Sony Playstation VR). Other VR efforts have been halted, such as the Video See through VR project Alloy from Intel and the StarVR wide FOV VR headset. 2018 also saw major restructuration and subsequent shut downs of medium sized AR headset companies such as MetaVision Corp. (Meta2 MR headset) and ODG Corp. (ODG R9 AR glasses), after both had a strong initial product introduction and VC capital support. On the other hand, the current VC investment hype fueling frenetically single start-ups such as Magic Leap Inc. (totaling \$2.4B VC investment pushing up a >\$6B valuation as of Q1 2018 before any revenues or product introduction has happened) is a harsh reminder of the ever present "Fear Of Missing Out" behavior from late stage

investors eager to jump on the bandwagon fuelled by the early investment decisions from the major VC firms.

No matter the investment hype, it might very well take a major consumer electronics company to create at the same time the ultimate consumer headset architecture (addressing visual / wearable comfort and immersion experience) and the subsequent consumer market. Unlike for the enterprise market where the content is provided by each individual enterprise through custom application developments for specific needs, the consumer market needs to rely solely on the entire MR ecosystem development, from generic hardware to generic content and applications.

Even though Q3 2018 saw for the first time a worldwide decline in both smart phone and tablet sales, hinting at Apple's stock fallout, it is unclear if such an MR consumer hardware has the potential (or even the will) to replace the existing smart phone/tablets, or alternatively be the ultimate companion to a smart phone, providing an immersive experience that is out of reach for any other traditional display screen concept.

Apart the consumer and enterprise markets discussed here, there is still today a considerable defense market for MR headsets. Microsoft has secured in Q4/18 a \$480M defense contract to develop and provide the US Army special versions of HoloLens, dubbed IVAS (Integrated Visual Augmentation System). Being the largest contract ever in AR/VR/MR, consumer, enterprise and defense combined, this will boost the entire MR ecosystem worldwide.

3. The emergence of MR as the next computing platform

Smart glasses (also commonly called Digital Eyewear) are mainly an extension of prescription eyewear, providing a digital contextual display to the prescription correction for visual impairment (see Figure 3). This is different from AR or MR functionality. Typical smart glass FOV remains small (less than 15 deg diagonal), often off-set from the line of sight. The lack of sensors (apart the IMU) allows for approximate 3DOF, and lack of binocular vision reduced the display to simple 2D text and images viewing. Monocular displays do not require rigidity in the frames as a binocular vision system would (to reduce horizontal and vertical retinal disparity variations producing eye strain).

The combination of a strong connectivity (4G, WiFi, Bluetooth) and the addition of a camera makes it a very convincing companion to a smart phone, for specific contextual display functionality such as virtual assistant, GPS and social network companion (thanks to the camera functionality). A smart glass does not aim at replacing a smart phone, but contributes to a good addition to it, as a smart watch would do.

VR headsets are an extension of gaming consoles, as shown by major gaming providers such as Sony, Oculus, HTC Vive and Microsoft Windows MR, with gaming content partners such as Steam VR. Such headsets are often also sold with gaming controllers (see Figure 3). Outside-in sensors (as the standalone with Oculus CV1 and HTC Vive 2016) lead the way to inside-out sensors in newer generation headsets, providing a more compact hardware. Although these high-end VR systems still require high end GPU in costly desktop or laptop gaming PCs, standalone VR headsets have recently been introduced (2018), such as the Oculus Go and the HTC Vive Focus, able to seduce a burgeoning VR consumer market base.

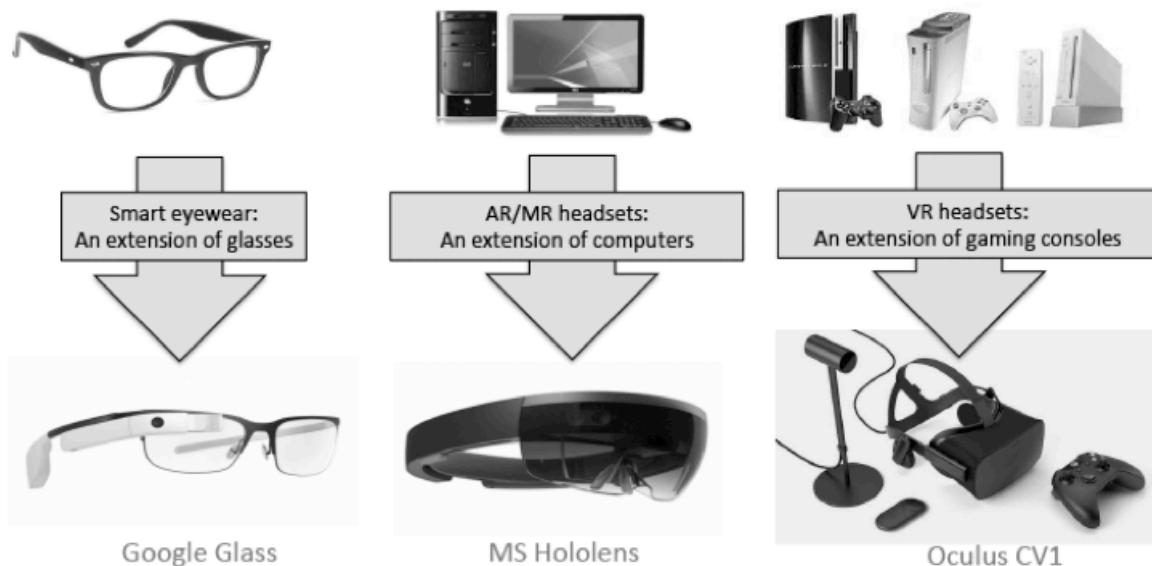


Figure 3: The emergence of Smart Glasses, AR/MR and VR headsets.

AR and especially MR systems are poised to become the next computing platform, replacing the ailing desktop and laptop hardware, even the now aging tablet computing hardware. Such systems are untethered for most of them (see HoloLens

in Figure 3), and require high-end optics for the display engine and the combiner optics as well as for the sensors (depth scanner camera, head tracking cameras to provide 6DOF, accurate eye trackers and gesture sensors). These are today the most demanding headsets in terms of hardware, especially optical hardware. These are the basis of this review paper.

Eventually, if technology permits, these three categories will merge into a single hardware concept. This will however requires improvements in connectivity (5G, WiGig), visual comfort (new display technologies) and wearable comfort (battery life, thermal management, weight/size).

The worldwide sales decline for smartphones and tablets is an acute signal for major consumer electronics corporations and VC firms to fund and develop the “next big thing”, whatever it may be. But MR is a good candidate!

4. The keys to the ultimate MR experience

The ultimate MR experience, for either consumer or enterprise, is defined along two main axes: **comfort and immersion**. Comfort comes in two declinations: wearable and visual. Immersion comes in various declinations, from display to audio, gestures, haptics, etc...

At the confluence of **comfort and immersion**, three main features are required for a compelling MR experience:

- *Motion to Photon latency below 10ms (through fast sensor fusion)*
- *Display locking in 3D world through continuous depth mapping and semantic recognition*
- *Fast eye and universal eye tracking is a required feature which will enable many of the features we list here*

Most of this can be achieved through a global sensor fusion process [5] integrated through dedicated silicon, as it has been implemented in HoloLens with the HPU (Holographic Processing unit) [11].

Comfort:

Comfort, both wearable and visual, is the key enabling a large acceptance base of any consumer MR headset candidate architecture.

Wearable comfort features include:

- *Untethered headset for best mobility (future wireless connectivity through 5G or WiGig will be of great help to reduce on-board compute and rendering)*
- *Small size and light weight*
- *Thermal management throughout the entire headset (passive or active).*
- *Skin contact management through pressure points*
- *Breathable fabrics to manage sweat and heat*
- *Center of Gravity (CG) closer to CG of human head*

Visual comfort features include:

- *Large eyebox allowing for wide IPD coverage. The optics might also come in different SKUs for consumer, Small, Medium and Large IPD, but for enterprise, as the headset is shared between employees, it needs to accommodate wide IPD range.*
- *Angular resolution close to 20/20 visual acuity (at least 45 Pixel Per Degree –PPD- in the central foveated region), and lowered to a few PPD in the peripheral visual region.*
- *No screen door effects (large pixel fill factor and high PPD), no Mira effects.*
- *HDR through High brightness and high contrast (emissive displays such as MEMS scanners and OLEDs/iLEDs vs. non emissive displays as LCOS and LCD)*
- *Ghost images minimized (<1%)*
- *Unconstrained 200+ deg see-through peripheral vision (especially useful in outdoor activities, defense and civil engineering).*
- *Active dimming on visor (uniform shutter or soft edge dimming)*

Visual comfort features based on accurate / universal eye tracking include:

- *Vergence/Accommodation Conflict (VAC) mitigation for close up objects located in foveated cone through vergence tracking from differential eye tracking data, (as vergence is the trigger to accommodation).*
- *Active pupil swim correction for large FOV optics.*
- *Active pixel occlusion (hard edge occlusion) to increase hologram opacity (more realistic).*

Additional visual comfort and visual augmentation features include:

- *Active vision impairment correction, with spherical and astigmatism diopters (can be implemented in part with hardware used for VAC mitigation) with display ON or OFF.*
- *If VAC mitigation architecture does not produce optical blur, render blur will add to the 3D cues and improve 3D visual experience (such as Chroma blur).*
- *Super vision features with display OFF, such as magnifier glass or binocular telescope vision.*

Immersion

Immersion is the other key to the ultimate MR experience, and is not only based on FOV: the FOV is a 2D angular concept, immersive FOV is a 3D concept, including the Z distance from the user's eyes, allowing for arm's length display interaction through VAC mitigation.

Immersion experience for the user comes in many flavors:

- *Wide angular Field of View (WFOV) including peripheral display regions with lower pixel count per degree (resolution) and lower color depth.*
- *Foveated display in either fixed version (foveated rendering) or dynamic version (through display steering, mechanically or optically).*
- *World locked holograms and hologram occlusion through the use of depth map scanners*
- *World locked spatial audio*
- *Accurate gesture sensing through dedicated sensors*
- *Haptics*

Figure 4 summarizes the main requirements to enable the analyst's optimist version of the VR/AR/MR/Smart glasses market, at the of confluence of immersion experience and wearable/visual comfort.

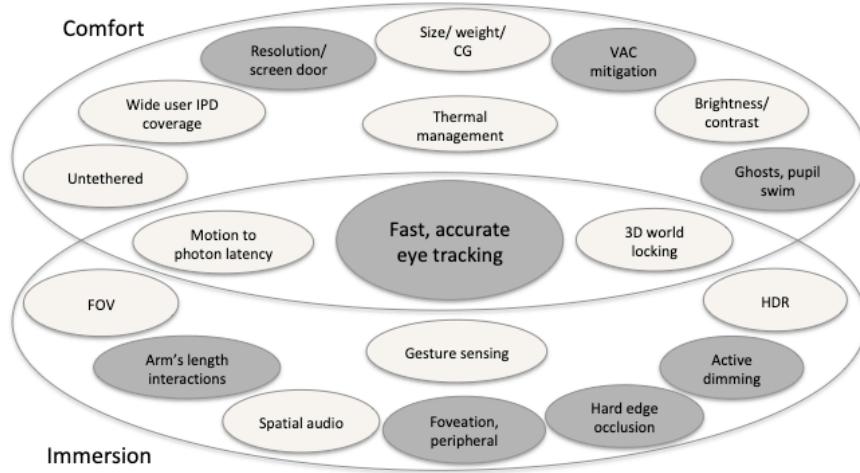


Figure 4: Main requirements addressing immersion and wearable/visual comfort for the ultimate MR experience.

The dark grey items in Figure 4 are based on a crucial enabling optical technology for next generation MR headsets: fast, accurate and universal eye/pupil/gaze trackers.

5. Human factors

In order to design a display architecture aiming at providing the ultimate MR comfort and immersion experience described in the previous section, one needs to consider the optical design task as a human centric task. This section analyzes some of the specifics of the human vision system [12], and how one can take advantage of such in order to reduce the complexity of the optical hardware as well as the software architecture, without degrading in any way the user's immersion and comfort experience [13].

The human Fovea, where resolution perception is at a maximum due to its high cone density, covers only 2-3 deg, and is set off-axis from the line of sight by about 5 deg. The human fovea is a result of early life visual experience, and grows from a small age to form a unique area set apart of the line of sight.

The human vision specifics are based on the cone and rod density over the retina, as described in Figure 5. The optical axis (or pupillary axis, normal to the vertex of the cornea) is slightly offset from the light of sight [14], (close to the visual axis) by about 5 degrees, and coincides with the location of the fovea on the retina. The blind spot, where the optic nerve is located, is offset by about 18 degrees off from the center of the Fovea.

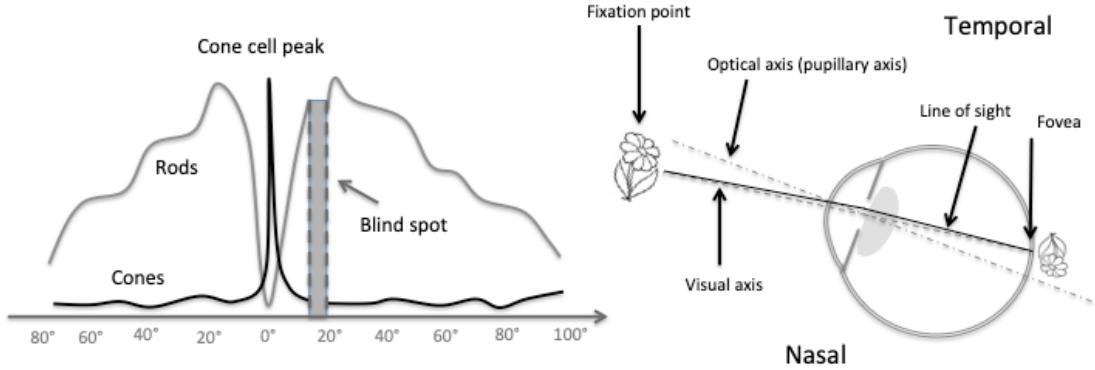


Figure 5: Rod and Cone cells density on the retina (left), optical axis and line of sight (right).

It is interesting to note that the human fovea grows slowly in early life based on specific human visual behavior, and is not a feature of our visual system given at birth. Therefore, the location of the fovea might drift to new positions on the retina with novel visual behaviors radically different from millennia of human evolution, such as the use of small digital displays held at close range by toddlers. Another severe change would be early childhood myopia due to the same cause [15] [16].

Figure 7 shows the horizontal extend of the different angular regions of the human binocular vision system. Although the entire FOV spans more than 220 deg horizontally, the binocular range spans only 120 deg in most cases (depending on the nose geometry). Stereopsis (the left and right monocular vision fusion [19] providing 3D depth cue) is however more limited (+/-40 deg) [17].

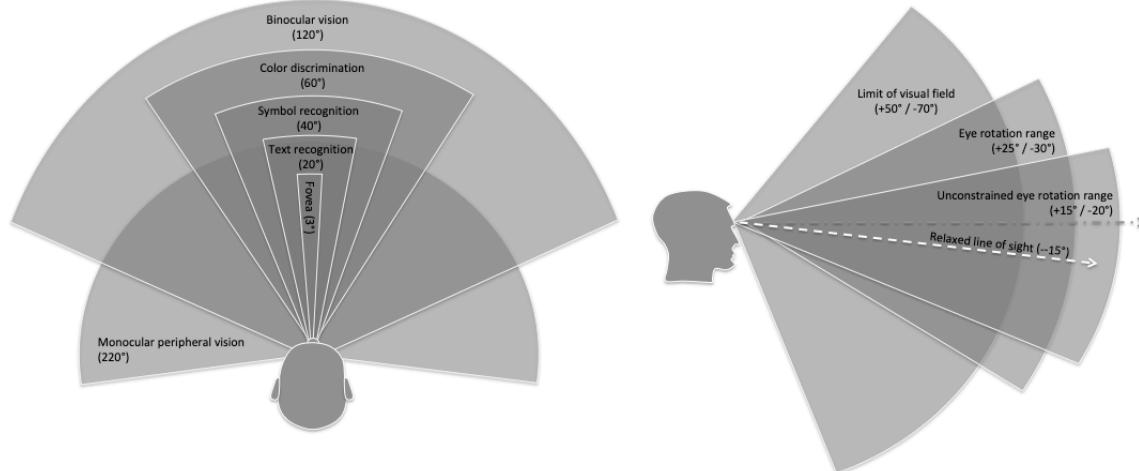


Figure 6: Human vision FOV (H and V)

The vertical FOV is similar in size to the horizontal FOV, and is set off-axis from the standard line of sight, by about 15 deg downwards (relaxed line of sight)

The human FOV is a dynamic concept, best described when considering the constrained and unconstrained eye motion ranges [14] (unconstrained: motions that do not produce eye strain and allow for steady gaze and subsequent accommodation reflex). While the mechanical eye motion range can be quite large (+/-40 deg H), the unconstrained eye motion over which gaze is possible without inducing the head turning reflex is much smaller, and covers roughly +/-20 deg FOV H. This in turn defines the static foveated region, ranging 40-45 deg FOV H. Fig. x shows the human binocular FOV, as the overlap of the left and right fields, as well as the parafovea and the center fovea region ranging 3 degrees full angle [18] (see Figure 7).

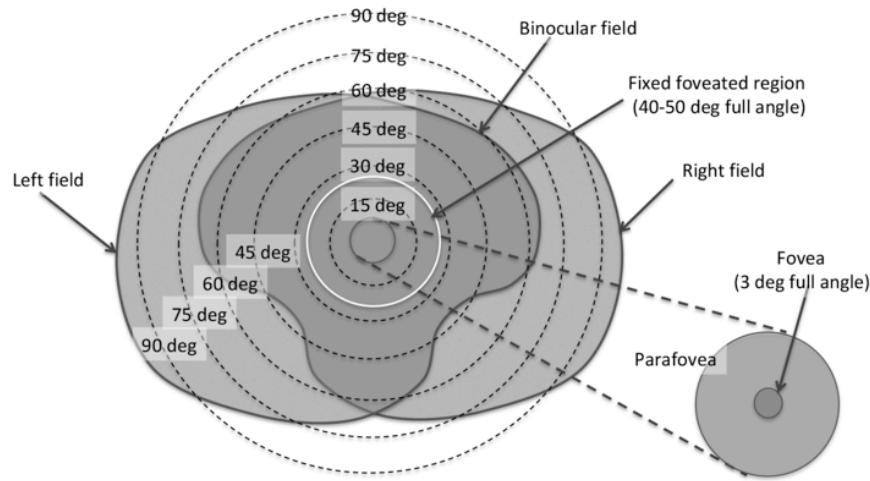


Figure 7: Human binocular field of view with fixed foveated region including unconstrained eye motion allowing for sustained gaze and accommodation.

The binocular FOV [20] is a quite large region, horizontally symmetric and vertically asymmetric, spanning +/-60 deg over a +55 deg upper and -30 deg lower, with a central lower region reaching also -60deg but over a smaller horizontal span of 35 deg full angle. The white circle showing the fixed foveated region over which sustained eyegaze is possible defines also the state of the art diagonal FOV for most high end AR/MR devices today, which also provide a 100% stereo overlap. Furthermore, for a given gaze angle, color recognition spans over a +/-60 deg FOV, shape recognition over a +/-30 deg FOV, and text recognition over a much smaller FOV of +/-10 deg.

Various FOVs from existing HMDs are shown in Figure 8. Standard VR headsets (Oculus CV1, HTC Vive, Sony PlayStation, Microsoft Windows MR) have all FOVs around 110 deg FOV diagonal, stretching towards 200 deg FOV for some others (PiMax and StarVR). Large AR FOV up to 90 deg can be produced by a large cell phone panel display combined with a large single curved free space combiner (Meta2, DreamGlass, Mira, NorthStar Leap Motion), to smaller FOV high end AR/MR systems with microdisplay panels such as Microsoft HoloLens V1 and Magic Leap One. Smaller smart glasses have typically FOVs ranging from 10-15 deg (Zeiss Smart Glasses, Google Glass) to 30-40 deg (Optinvent ORA, Lumus DK50, ODG R8).

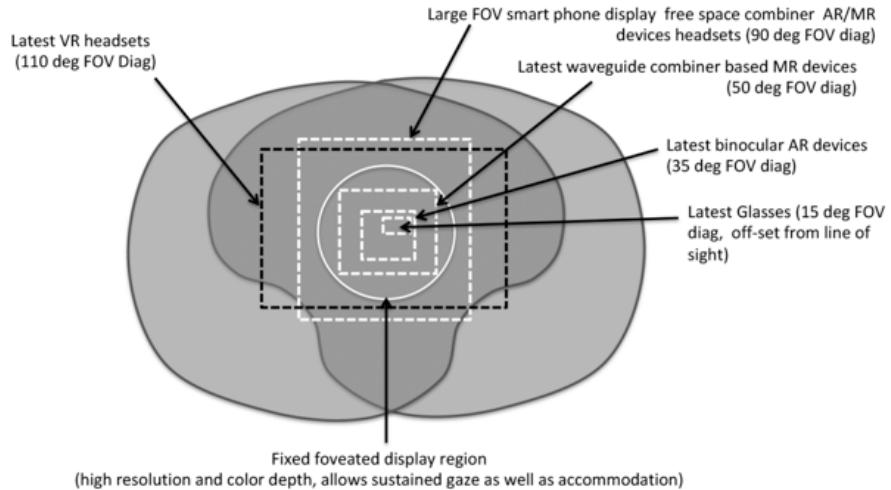


Figure 8: Typical FOVs for current state of the art Smart glasses, AR, MR and VR headsets, overlaid on the human binocular vision and the fixed foveated display region.

One other way to describe the FOV experience is to overlap the amount of see though FOV on the actual display FOV (see Figure 1). The foveated region is shown with dotted lines.

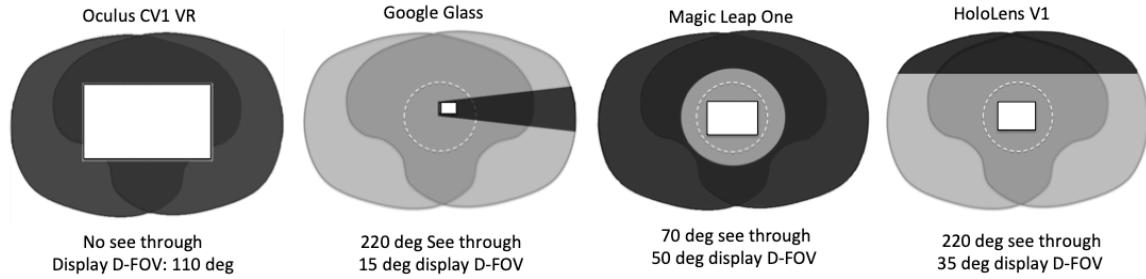


Figure 9: Display FOV and see through FOV for various VR, smart glasses and AR headsets.

For a VR system, there is no see through, and the display FOV can be quite large, 110 DFOV to 150 DFOV (left). For a smart glass (Google Glass, center left), the see through – or rather see around- experience is very wide, only hindered by the lateral display arm on one side, with an excentered display DFOV of 15 deg. For an the Magic Leap One MR headset (center right), the tunneling effect due to the circular mechanical enclosure of the glasses reduce the see through considerably, to a 70 deg circular cone, while the display has a DFOV of 50 deg. For Hololens (right), the lateral see-through (or see-around) FOV, is equal to the natural FOV of 220 deg, with a display diagonal DFOV of 35 deg.; only the top part of the FOV is capped by the mechanical enclosure holding the sensor bar, the display engines and the system board. There is no limitation to the bottom FOV.

In order to optimize an HMD optical architecture for large FOV, the various regions of the human FOV described in Figure 7 have to be considered in order not to overdesign the system. This allows, through a “human centric optimization” in the optical design process, to produce a system which is closely matched to the human vision system in terms of resolution, MTF, pixel density, color depth and contrast. Vergence Accommodation Conflict can also be considered as foveated

6. Optical specifications driving architecture and technology choices

Before discussing the factors influencing the choice of the various optical architectures and technologies available today, we define the main specifications which drive the optical design cost functions, such as: eyebox concept , eye relief, FOV, stereo overlap, brightness, angular resolution in foveated region, peripheral vision.

Eye box

In order to fit a device to a variety of users covering a large population, one critical feature is to cover a large IPD (Inter Pupillary Distance) population. Table 1 shows mean IPD values for various age groups for men and women.

Age group	Female IPD (mm)	Male IPD (mm)	Total IPD (mm)	Min/Max IPD (mm)
20-30	59.2	61.5	60.3	49-70
31-50	62.0	64.5	63.0	55-72
51-70	62.3	65.7	63.8	52-76
71-89	62.1	63.1	62.7	49-74

Table 1: Mean IPD values for various age groups for men and women.

To achieve this, a large eyebox is necessary. However, a static single exit pupil forming and eyebox is usually not the best solution, since there are varieties of techniques to increase the effective eyebox as perceived by the user (refer to section 8). There might be as many eyebox definitions as there are AR/MR devices out there. This critical and “universal” HMD specification seems to be actually the most volatile in the AR/MR field today. This is mainly due to the fact that there are a multitude of different optical combiner architectures and technologies, and thus a multitude of different prescriptions constraining the eyebox as experienced by the user.

The basic definition of the eyebox is the 3D region between the combiner and the human eye pupil over which the entire FOV is visible, for a typical pupil size [22].

The most straightforward criteria defining the eyebox is image vignetting [23]. To estimate the size of a vignetted eyebox, one can light up a vertical display sliver on the left side of the display (while the rest of the display is off), and plot the intensity of the resulting image (as an optical simulation or as an optical experiment) as a function of the position of the eye pupil toward the left, and start again with a vertical display sliver while moving the eye pupil to the right (see Figure 10).

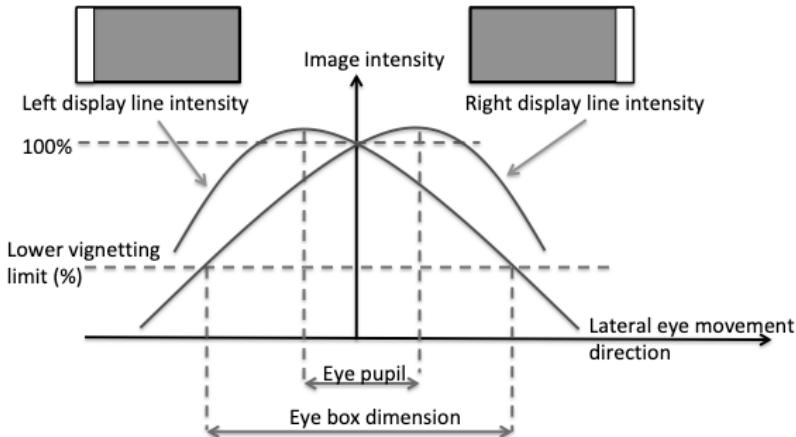


Figure 10: Definition of the eyebox through vignetting criteria.

The vertical eyebox can be measured or computed the same way, by projecting a horizontal light sliver (top and bottom) on the display and moving the eye pupil vertically.

There might be cases in which the eyebox might be more affected by distortion than vignetting (also called pupil swim, which is distortion variations as a function of the eye pupil position), or even LCA (Lateral Chromatic Aberrations). In these cases, the eyebox might be defined as a limit of distortion or LCA rather than vignetting.

The eyebox is a three dimensional region located after the optical combiner. The perceived eyebox will thus vary when the eye relief changes (usually get reduced when the eyerelief increases). Eye relief can change when using prescription glasses between the combiner and the eye.

For a given optical system, the eyebox is inversely proportional to the field of view. The effective eyebox can be enlarged or reduced by simply lighting up a smaller or larger part of the display.

The perceived eyebox is also a proportional to the size of the human eye pupil, and can therefore be sensed as smaller in bright environment (bright sun light) or larger in darker environments (interior and/or with visor dimming).

Figure 11 summarizes the effects of eye relief, FOV and eye pupil size on the perceived eyebox as experienced by the user. The combination of all three can build up a more or less uniform eyebox size, no matter the size of the human eye pupil.

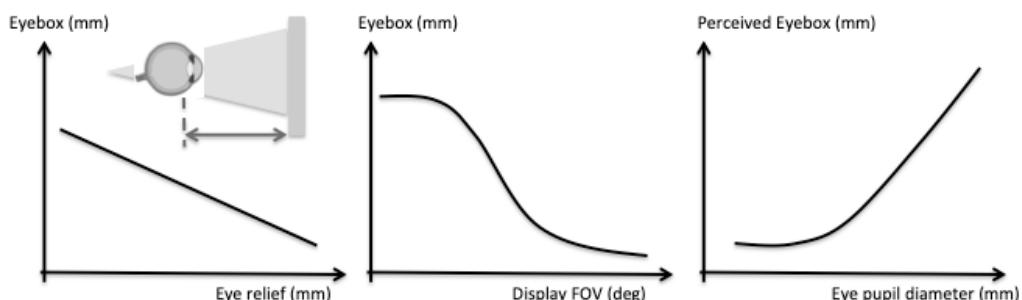


Figure 11.: FOV and eyebox are shared for a given optical system (center). Perceived eyebox is a function of the eye pupil (right).

Table 2. shows typical diameters of the human eye pupil as a function of the luminance.

Luminance (Cd/m ²)	10 ⁻⁶	10 ⁻⁴	10 ⁻²	1	10 ⁺²	10 ⁺⁴	10 ⁺⁶	10 ⁺⁸
Pupil size (mm)	7.9	7.5	6.1	3.9	2.5	2.1	2.0	2.0
Vision mode	Scotopic (starlight)		Mesopic (moonlight)			Photopic (office light to sunlight)		

Table 2: Eye pupil size as a function of luminance (and subsequent vision modes).

Eye relief and vertex distance

The eye relief is the distance between the vertex of the last surface of the optical combiner and the human eye (cornea). However, for most optical engineers, and in most optical models, the eye relief is usually the distance between the last

surface of the combiner and the human eye pupil, which increases the effective eye relief by about 2mm (3mm between the cornea to the human pupil in an aqueous humor media of index 1.33).

The eye vertex is more often used in the ophthalmic field as the distance between the base surface at its vertex (eye side lens surface) and the cornea.

Thus, in an AR headset in which the user is able to wear a prescription lens (such as the Microsoft Hololens for ex), both vertex distance and eye relief can be defined separately (see Figure 12).

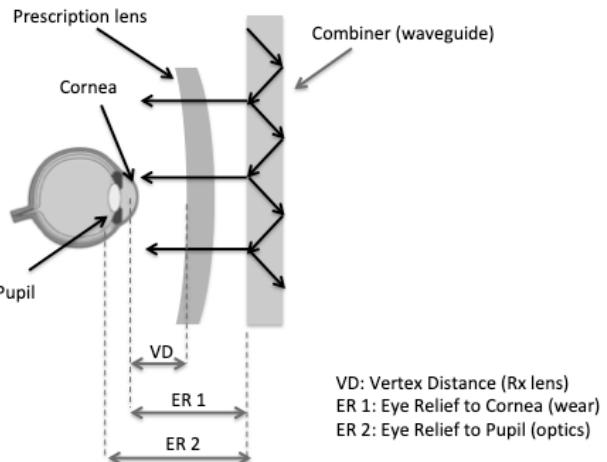


Figure 12: Vertex distance, eye relief for wear and optical design eye relief in AR headsets with waveguide combiner architecture.

Typical values of vertex distances in optometry range from 12mm to 17mm (depending on length of eye lashes and strength of the lens base curvature). The eye relief range over which the eyebox is defined, for an AR, VR or MR system should include the fact that the user might be wearing prescription lenses, although this is less critical in VR systems where the focus can be adjusted by moving the distance between the display and the lens.

Note that in a fixed focus stereo display in see through AR mode, with the focus typically set anywhere from 1.5m to 3m, only short sighted users (myopia) might be wearing their prescription glasses. Far sighted users (hyperopia) and users suffering from presbyopia might want to remove their glasses before using an AR/MR headset, since wearing them will not help much as for the hologram, especially if the display has a fixed focus (no VAC mitigation).

Typical values of eye reliefs in AR/MR headsets range from 13mm (user not wearing any prescription glasses) to 25mm (user wearing prescription glasses, and/or having long eyelashes). Prescribing a longer eye relief than 25mm reduces usually the eyebox to levels providing not much visual comfort.

Field Of View (FOV)

The FOV in an immersive display system is the angular range over which an image can be projected in the near or the far field [20]. It is measured in degrees and the resolution over the FOV is measured in pixels per degrees (PPD). Very often, the FOV is given as a diagonal measure of a rectangular aspect ratio image. For larger FOV values, aspect ratio can become square or even circular or elliptical. As pointed out in the previous section, the optical FOV can be larger than the experienced FOV if the eyebox is not wide enough.

The FOV is linearly proportional to the size of the microdisplay and inversely proportional to the focal length of the collimation lens or collimation lens stack. Keeping the same FOV while reducing the size of the microdisplay for industrial design reasons, requires the optical designer to increase the numerical aperture of the collimation optics, therefore increasing their size and weight and introducing also more aberrations especially at the edges of the FOV, and potentially introducing pupil swim (see below). A balance between microdisplay size and lens power in the optical engine is therefore needed to achieve the best MTF and best size/weight compromise.

The size of the FOV as measured by an optical metrology system may be different from the perceived FOV by the human eye. An AR system with a good MTF (crisper image) can be perceived as having a larger FOV than an MR system with a lower MTF and a similar (or even larger optical FOV). Similarly, a color uniformity or an LCA problem can result in a perceived FOV that is smaller than if the same image would have had a better color uniformity or a lower LCA.

Pupil swim

The collimation lens (or collimation lens stack) can introduce typical pillow distortions as well as lateral chromatic aberrations (LCA) to the immersive display. These aberrations can be compensated in software through pre-emphasis

over the original RGB image, by loading a pre-calculated distortion map for each color, such as a typical Barrel Distortion map, slightly different for R, G and B.

If the FOV gets larger, it may well be that this distortion will change as a function of the lateral position of the human eye pupil in the eyebox, as the eye gazes at the extremities of the FOV. This optical distortion variation is called Pupil Swim. One can compensate for pupil swim by using a pupil tracker and a library of R-G-B distortion maps stored in a look-up-table.

Note that pupil swim also occurs in prescription glasses for presbyopia patients, also called progressive glasses in the ophthalmic industry. In progressive lenses, pupil swim is caused by the variation of optical power as one looks around the field, horizontal and vertical. In this case, the patient's brain learns how to compensate for pupil swim, but this might take some time.

Display immersion

The immersion perception for the user is a much larger concept than only 2D FOV, and includes the third dimension, (the z reach) of the FOV, as shown in Figure 13. In most AR, MR and VR cases, the display is a stereo display which focus is set at a specific distance, say 1 or 2 meters in front of the user's eye (which is usually considered as to be part of the far field).

Increasing the Z reach of the FOV to enter the near field, and potentially be as close as a foot allows for the user to engage in arm's length display interactions, one of the many features increasing the immersion and the functionality of the MR system. However, using only stereo disparity to represent holograms at close range would introduce vision discomfort such as the VAC (Vergence Accommodation Conflict). Section 8 discusses such visual conflicts and lists some of the hardware and software solutions used in industry today to mitigate it.

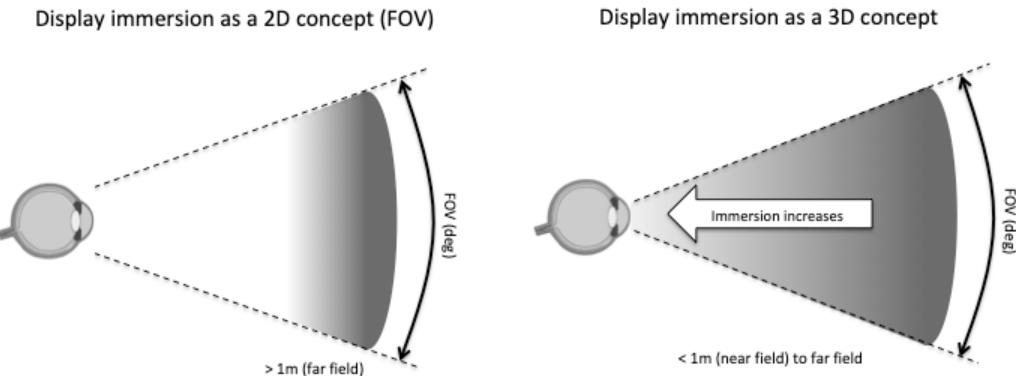


Figure 13. Increasing the z-extent of the FOV to increase the immersion experience and allow arm's length display interactions.

Stereo overlap

A large FOV is usually desired, especially in VR applications where the natural light field see through of 200+ degrees is totally obstructed, mainly to avoid the tunneling vision effect. In AR systems, the FOV might also be designed to be larger than the fixed foveated region.

By doing so, one might stretch the limits of the collimation optics by introducing parasitic distortion and LCA, and reducing the angular resolution, and/or stretching the size and resolution of the display panel. By considering the limited binocular overlap of the human visual system (see Figure 7), and the region over which stereopsis actually occurs, partial stereo overlap can be used to increase the binocular FOV without increasing the monocular FOV, and thus without stretching the display optics.

Partial overlap needs to be considered with a word of caution, since it should match the binocular overlap regions where the human visual system expects binocular vision [20]. Expecting binocular vision but seeing only with one eye might introduce vision discomfort, especially in VR systems.

Display brightness

In AR systems, the brightness measured at the display plane might be one or two degrees of magnitude higher than the brightness measured at the exit pupil. This is due to the efficiency of the collimation lens stack but also to the potential pupil increase or pupil replication scheme and the efficiency of the combiner [22] [23].

Therefore, two efficiency measures should be used. One based on nits gives the brightness efficiency of the system (as

experienced by the user's eye) and the other one given in lumens measures the efficiency of the optical system, and relates more to the luminous throughput of the AR system.

When a pupil replication scheme is used to increase the size of the eyebox, the Luminance value can be a degree of magnitude lower than the Illuminance value, for the exact same AR system.

A useful recap from photometry class:

- The amount of light in **Candela** per square meter is called **Luminance**, and measured in **nits**.
- The amount of light in **lumens** per square meter is called **Illuminance**, measured in **lux**.
- **FootCandle** is **lumens** per square **feet**.
- One **FootLambert** is 3.426 **nits**.

Angular resolution

In a decent AR/MR system, one would like to experience the same resolution over the digital hologram as over the see through reality, therefore expecting a 20/20 vision over the entire FOV, which turns out to be about 0.3 arcmin resolution (or 195 pixels per degree or PPD...). For most people, an angular resolution of less than 0.8 arcmin can be resolved, but this drops rapidly with age. A 1.3 arcmin angular resolution (relating to about 45 PPD) provides a decent MR experienced for most people in their 20s to 30s.

The human visual system is however only one part of the equation, capturing whatever has been resolved out of the display resolution by the HMD display optics MTF (Modulation Transfer Function).

The MTF of the optical system forming the image of the display in the far or near field has to match the resolution of the display over the FOV in PPD. A decent level for the MTF would be 30%.

However, for an immersive display, it is best to consider the through focus MTF to get a better appreciation of the MTF of the system over a large FOV. The human visual system can only resolve high resolution in a small angular cone, the Fovea, and constantly scans the FOV not only with lateral eye saccades, but also with small focus saccades ranging +/- 1/8th of a Diopter (D). These fast saccades over +/- 1/8th D are different in nature and much faster than the standard oculomotor accommodation reflex triggered by a 3D cues such as stereo disparity (inducing vergence) or Chromablu [24].

Figure 14 shows a standard MTF plot for 3 different fields in the display at a fixed focus distance. Such MTF can be low at field edges, but when computed over a +/- 1/8th D through focus, the maximum MTF over that region can be much higher, thus providing a much better experience for the user than a single focus MTF would predict.

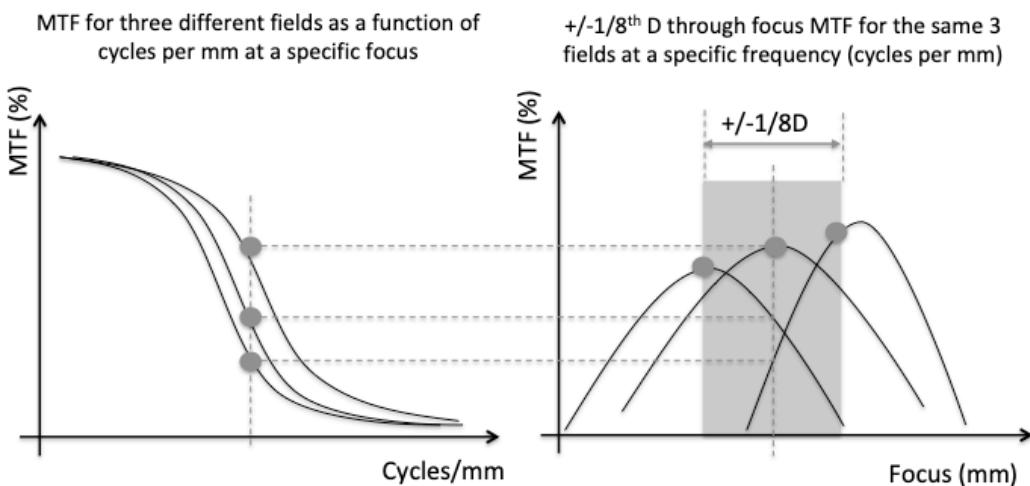


Figure 14: Standard MTF plot (left) and corresponding through focus MTF (right) addressing fast focus saccades of the eye over a +/- 1/8th D range.

This occurs as one compares the fixed focus MTF modeling (even when using Diffraction Based Polychromatic MTF), as compared to the user's own angular resolution experience when using the actual immersive display. This is a unusual positive effect (making things better in reality than they are in modeling – *usually things go the other way in optics...*).

Adverse display effects (screen door, aliasing, motion blur and Mura effects)

If the MTF of the display system is well resolving the pixels, especially in a panel based VR system, the user can see the pixel interspacing, which produces the parasitic and annoying "screen door effect". The screen door effect can be reducing the pixel interspacing region (OLEDs panels have smaller pixel gaps than LCD panels), or by reducing intentionally the MTF of the system so that the gaps are not resolved anymore (difficult).

Panel displays and microdisplay panels are usually made of pixels arranged in a grid. When it comes to displaying diagonal or curved lines, one is essentially forced to draw a curved line with square blocks placed along a grid, producing aliasing. Anything else but straight lines will naturally reveal the underlying shape of the pixels and the pixel grid. Increasing the pixel density can reduce aliasing. Anti-aliasing rendering can also reduce perceived aliasing by using different colored pixels along the edges of the line to create the appearance of a smoother line. Aliasing is particularly annoying in the peripheral region of an immersive display.

Motion blur is also detrimental to a high resolution virtual image perception: a 90 Hz refresh rate and very fast response times of between 3ms-6ms that will considerably reduce motion blur.

The Mura effect, or "clouding", is a term generally used to describe uneven displays, caused by the imperfect illumination of the screen or the unevenness of that screen. These effects can manifest themselves in areas or individual pixels that are darker or brighter, show poorer contrast, or simply deviate from the general image. As a rule, these effects are particularly noticeable in the reproduction of dark images. Generally speaking, the Mura Effect is a fundamental design feature of current LCD display panels. Mura Effects can also manifest themselves in displays based on OLED panels. An immersive display such as in VR increases the perception of the Mura effect. In AR headsets, the perceived Mura effect is much milder than in VR systems, as the see through background color and uniformity changes constantly as the head moves around the scene. Virtually no back-lit or edge lit displays are free of Mura Effects.

Figure 15 describes graphically these various effects.

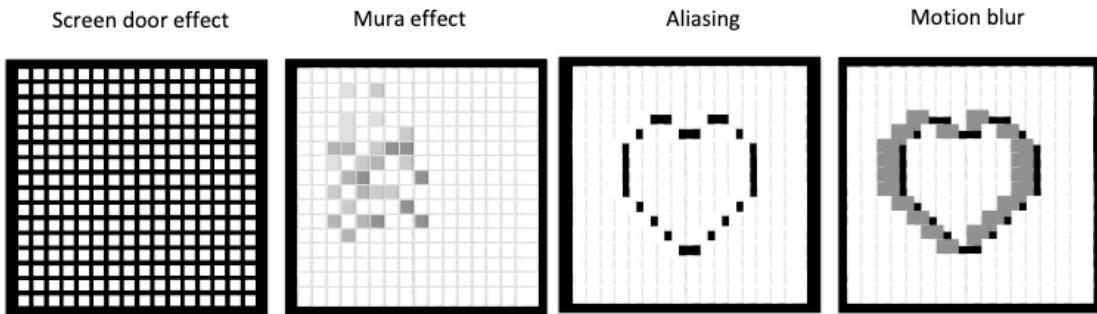


Figure 15: Screen door effect, Mura effect, motion blur and display aliasing.

Perceived resolution and FOV

Eventually, perceived FOV and resolution by the human visual system is the ultimate goal: we thus consider here again a human centric system design, in which resolution is rather a perceived spec (subjective) rather than a scientifically measured spec. For example, the second version of the Samsung Odyssey Windows MR headset includes a top cover pixel replication scheme (going from 616ppi to 1233ppi) providing a higher perceived resolution while the display pipeline remains at the unchanged lower resolution of the first version.

Perceived FOV can also be subjective, especially in AR systems. The quality of a display (high MTF, high resolution, absence of screen door and Mura effects, reduced aliasing and motion blur) contributes to a perceived FOV that is larger than that of the same immersive display architecture however with weaker imaging performances. The perception of the FOV by the user is a combination of the natural see-through FOV available, combined with the quality of the virtual image.

Foveation and Peripheral vision

Even a modest resolution of 45 PPD (1.3 arcmin) stretched over a 110 deg horizontal FOV would require a prohibitive display pixel amount of more than 4K in this single direction. Both foveation and peripheral display attempt to provide a large drop in the pixel count (and thus rendering requirements) while retaining a high angular resolution experience for the user.

Peripheral vision is a specific region of the human visual system in which flicker and aliasing effects are very critical, requiring high refresh rate, low latency and high persistence. The peripheral region is also very sensitive to clutter, unlike the foveated region.

The human acuity drops very fast when one departs from the Macula region, down from >100 PPD in the Fovea to less than 20 PPD at +/-25 deg FOV as shown in Figure 16 (left side) [25]. The right side shows the successive human vision field region related to visual acuity [26]. Peripheral region comprises all regions expect the macula region,

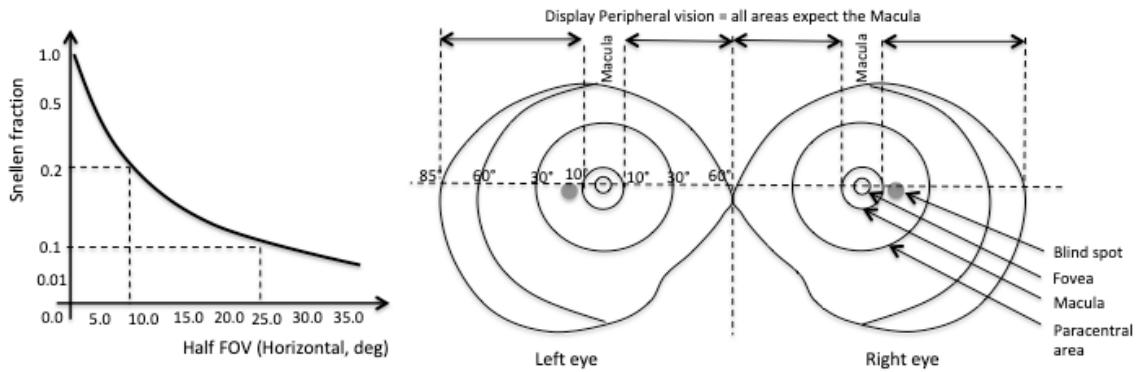


Figure 16: Visual acuity as a function of eccentricity (left), macula and peripheral vision regions (right)

The various human vision regions in Figure 16 are listed below, with their respective size and angular resolutions:

- **Fovea:** 3 degrees: Highest visual acuity (>150 ppp)
- **Macula:** Next highest area (+/-10 deg, down to 20ppd at the edges)
- **Paracentral area:** Visual acuity fair (+/-30 deg, down to 10 ppp at the edges)
- **Peripheral vision:** poor visual acuity but first alerting system for detecting movement, for orienting in space (balance) and moving around in the environment (below 10 ppp).

In order to provide a high resolution experience to the user while limiting the number of pixels in the display, one can use various foveation techniques [27], [28], [29], such as:

- Static digital foveation without gaze tracking (same static display, fixed foveation is rendered over central static 50-60 deg FOV cone).
- Gaze contingent dynamic digital foveation (same static display, but high resolution rendering is processed over dynamic foveated region over a moving 10-15 deg FOV cone).
- Gaze contingent dynamic optical foveation (this uses two different display systems, a static low resolution high FOV display over 60+ deg, combined with a dynamically steerable high resolution low FOV display over about 15-20 deg FOV cone – the steering direction is gaze contingent).

Figure 17 shows the drop in display resolution that can be implemented when using foveation techniques without any loss in high resolution perception by the user.

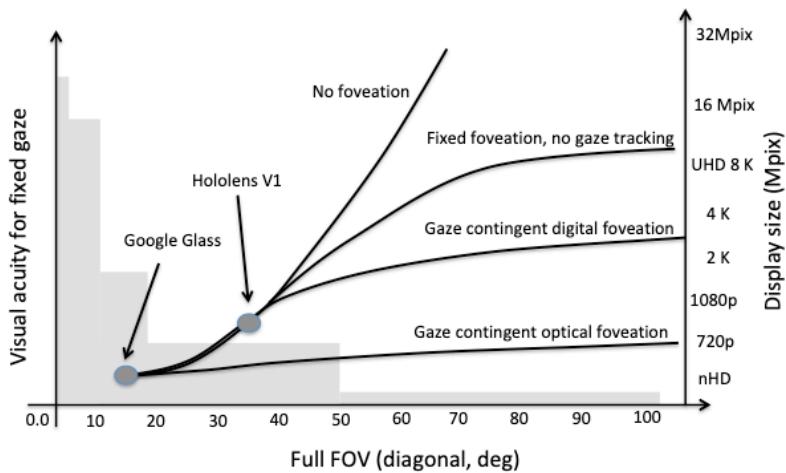


Figure 17: Decrease in pixel count in display for various foveation techniques

As seen in the previous section, partial stereo overlap can also reduce the pixel count without reducing the high resolution perception for the user, as long as the stereo overlap region is equal or larger than the stereopsis region.

Peripheral display can be added to the foveated area by using the same display engine or a different one. This will be reviewed later on in Section x on a few practical VR and AR architectures used today in industry.

7. Functional optical building blocks of an MR headset

Now that we have analyzed the specifics of the human visual system and defined the various optical specifications one needs to reach for a comfortable visual MR experience, we are ready to start design and optimize the display system and optical architecture. A typical functional optical building block suite of an MR system is shown in Figure 18.

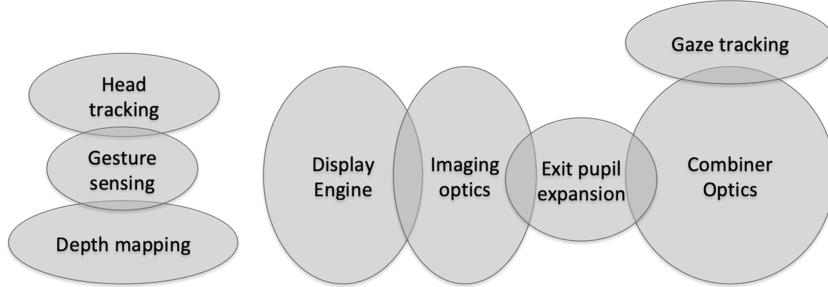


Figure 18: Functional optical building blocks of an MR system

The Display engine is where the image is formed, and then imaged onwards, forming a pupil (or not), and through an optical combiner which can include pupil replication scheme, to the eye pupil. Gaze tracking might or might not share optics with the display architecture (display architecture is usually an infinite conjugate system, and eye tracking is usually a finite conjugate system). Head tracking, gesture sensing and depth mapping are relying on external sensors (see last section).

Display engine

Early 90s VR and smart glass display engines were based of 1D LED scanners such as in the 1989 Private Eye smart glass by Reflection Technologies and VR Boy by Nintendo (1995). It is interesting to note that miniature Cathode Ray Tube (CRT) display units were used early on in VR as well as AR systems (Sword of Damocles, 1968), and are still in usage in some of the high end defense AR headsets today as in the current Apache Helicopter AR monocular headset.

Later, microdisplays using external illumination systems (and later backlights or frontlights) have been used in both smart glasses and AR headsets, such as HTPS LCD (Google Glass) [30], LCOS (Lumus, Hololens, Magic Leap) or DLP (Digilens, Aveyant) [31]. Emissive microdisplay panels have also been used extensively, such as OLED displays (ODG R9, Zeiss Smart Glasses). Higher brightness iLED microdisplays [32] are poised to revolutionize AR optical engines by providing the brightness (tens of thousands of Nits) and contrast required to compete with outdoor sun light without the use of a bulky illumination system.

More recently, 2D MEMS laser/VCSEL scanners, (Intel Vaunt, By North, QD laser) have proven to reduce dramatically the size of the optical engine. Redesigning the older 1D scanners with today's iLED and MEMS technology could prove to be an interesting solution to reduce the size of the optical engines and increase brightness/contrast. Similarly, digital 2D steering of smaller 2D panels is also an interesting option, which could account to "display wobulation". Other scanning technologies such as fiber scanners [33], integrated electro-optic scanners and Surface Acoustic Wave (SAW) scanners [34] have also been investigated.

It is noteworthy that laser based phase panels display engines (i.e. dynamic holographic projectors) which have entered recently the market through automotive Head Up Displays (HUDs) due to their high brightness, have been recently applied to the design of interesting display architectures that can provide a per pixel depth display, solving effectively therefore also the VAC conflict [35], [36]. Phase panels can come in many forms, from LCOS type platforms to MEMS pillar platforms.

Most of the scanner and phase panel based optical engines lack in exit pupil size (eye box), and need therefore fancy optical architectures to extend/replicate or steer such exit pupil to the user's eye.

Imaging optics

Once the image is formed over a plane, a surface, or through a scanner, there is a need to form an exit pupil, over which the image is either totally or partially collimated, and presented directly to the eye, or to an optical combiner. In some cases, an intermediate aerial image can also be formed to increase the étendue of the system.

Forming spatially demultiplexed exit pupils (either color or field separated) can also be an interesting option depending on the combiner technology used. Imaging optics are usually traditional free space optics, but in very compact form, including in many cases Polarization Beam Cubes (PBS) combined with Bird Bath architectures [37] to fold the optical path in various directions.

Combiner optics and Exit Pupil Expansion (EPE)

The optical combiner is very often the most complex and most costly optical element in the entire MR display architecture: it is the one seen directly by the user and the one seen directly by the world. It defines often also the size and aspect ratio of the entire headset. It is the critical optical element that reduces the quality of the see through and the one that defines the eyebox size, and in many cases also the FOV.

There are three main types of optical combiners used in most MR/AR/Smart glasses today:

- Free space optical combiners
- TIR prism optical combiners (and compensators)
- Waveguide based optical combiners.

These optical combiners are reviewed in details in the next sessions.

One important point to remember when optimizing an HMD display system is that the optical engine has to be optimized in concert with the combiner engine. Usually, having a team designing an optical engine without fully understanding the limitations and specifics of the combiner engine designed by another team, and vice versa, can result in a sub-optimal system or even a failed optical architecture, no matter how well the individual optical building blocks might be designed.

8. Circumventing the law of étendue in HMDs

We have reviewed in the previous section the various optical building blocks used in typical AR and MR headsets today. We review now the main challenges one has to overcome with such building blocks in order to provide at the same time:

- a large FOV and wide stereo overlap
- a large IPD coverage (large eyebox)
- a large eye relief allowing prescription lens wear
- a high angular resolution close to the 20/20 vision
- a small form factor, low weight and a CG close to the head

According to the law of étendue (Lagrange Invariant), when one attempts to expand the FOV by increasing the NA of the collimation lens, the eyebox get reduced (as well as the resolution), and the size of the optics increases. In an optimal system, it is of course interesting to have all four parameters maximized at the same time, calling for compromises as well as alternative architectures, carefully tuned to the specifics and limitations of the human visual system. We therefore come back to the concept of “human centric optical design” we introduced in previous sections of this paper.

When tasked to design an optical combiner, the optical designer has to check out various requirements, first with the User Experience (UX) team, which will indicate the IPD to cover (i.e. the target population for a single SKU), as well as with the Industrial Design (ID) team, which will indicate the minimum and maximum size of the display and combiner optics.

Figure 19 shows how a design space can be defined over a graph showing the combiner thickness as a function of target eye box size (IPD coverage). The Min and Max IPD values, as well as the Min (mechanical rigidity) and Max (aesthetics and wearable comfort) thickness of the combiner define a 2D window space over which the optical designer needs to specify the optical combiner. When contemplating the user of a bird-bath optical architecture (Google Glass, ODG R9, Lenovo AR...), simple in nature and relatively cheap to produce in volumes, the size of such optics is proportional to the eyebox (and also the FOV), and thus cannot usually satisfy the design window constrains. When contemplating the use of a waveguide combiner, one can notice that the waveguide thickness does not change when the eyebox increases (Figure 19). The lateral size of the waveguide combiner however increases with both FOV and eyebox. This is one reason why many AR/MR designers choose to use waveguide combiner architectures for AR/MR HMDs that need to accommodate a large population and at the same time produce a relatively large FOV.

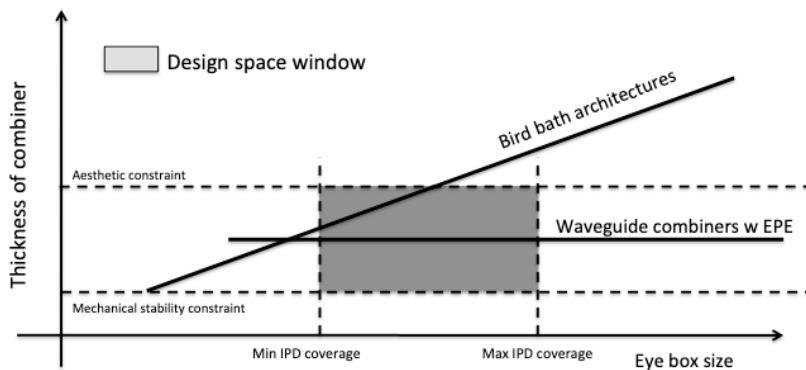


Figure 19: The design window addressing both IPD coverage and combiner thickness

As if this was not limiting enough, the law of étendue stipulates that the product of the microdisplay size by the numerical aperture of the display engine equals the product of the FOV by the perceived eyebbox (exit pupil).

$$(\text{Microdisplay Size}) \times (\text{Display Engine NA}) = (\text{EyeBox}) \times (\text{FOV in air})$$

As size matters, designing the smallest optical engine (small display aperture size and low NA lenses) while achieving a large FOV over a large eyebbox, would rather call for the following equation:

$$(\text{Microdisplay Size}) \times (\text{Display Engine NA}) < (\text{EyeBox}) \times (\text{FOV in air})$$

According to the law of étendue, this is not possible. However, as the final sensor is not a camera but the human visual system, various “tricks” can be played to circumvent this principle, in various dimensions (space, time, spectrum, polarization, etc...). This is in line with the already discussed principle of Human Centric Optical Design.

There are therefore various ways to circumvent the law of étendue. We list here 7 different architectural implementations which allow a larger perceived eyebbox by the user than what would be predicted by the strict law of étendue:

- 1) Mechanical IPD adjust
- 2) Pupil expansion
- 3) Pupil replication
- 4) Pupil steering
- 5) Pupil tiling
- 6) Pupil movement
- 7) Pupil switching

Mechanical IPD adjust

The majority of VR and smart glasses today incorporate a mechanical IPD adjust (Figure 20) in order to move the exit pupil of the imaging system to match the entrance pupil of the eye (Google Glass, Oculus VR, etc...). Although this is a very simple way to address a wide IPD range in monocular smart glasses and low resolution binocular VR headsets, it is a challenge for high resolution binocular VR/AR/MR system in which the vertical and horizontal binocular disparity mismatch needs to be controlled within milliradians.

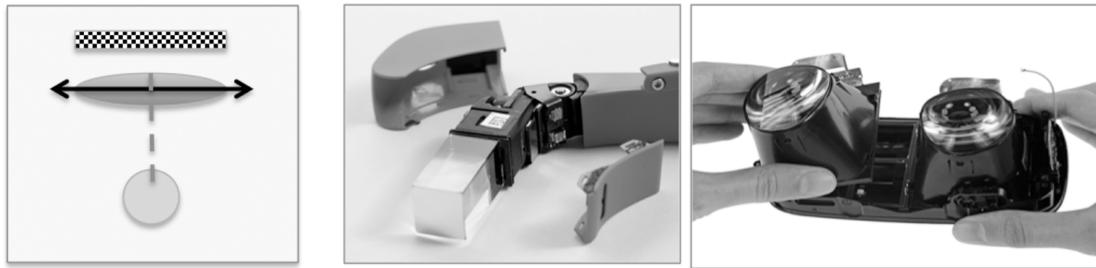


Figure 20: Mechanical IPD adjust (Google Glass, Oculus Rift).

Pupil expansion

When mechanical IPD adjust is ruled out for various reasons including binocular disparity mismatch, increasing the single exit pupil might be one solution. This is usually done in a “pupil forming” HMD architecture, in which an intermediate aerial image is created, in a plane or surface over which a diffuser might be located. This can be done through a conventional free space imaging system, a fiber bundle array or a waveguide system. The smaller or larger aerial image, diffused to a smaller or larger emission cone, can thus increase (or redirect through engineered diffusers) the field to a combiner which would produce an enlarged exit pupil (eye box). This can be implemented with a free space combiner, such as in Figure 21. This particular example (center) depicts a laser MEMS display engine forming an aerial image over a diffuser, but a microdisplay panel based optical engine can be used also. An SEM picture of a typical “engineered optical diffuser” which redirects the incoming light into a specific diffusion cone and direction is also shown. Switchable PDLC (Polymer Dispersed Liquid Crystal) diffusers can also be “engineered”, but are not as flexible as are wafer scale micro-optics based diffusers.

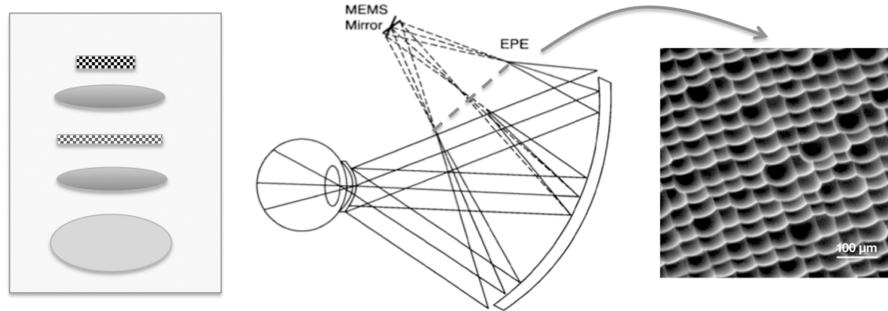


Figure 21: Single pupil expansion

Note that when using tunable focus lens in the display engine (or simply a laser retinal scanner), stacks of switchable diffusers (PDLC or other) can be used at various planes over which the aerial image might be formed, thus creating different image depths (one at a time). These can work either in transmission or reflection modes.

Exit Pupil replication (EPE)

Replicating the single exit pupil in a one or two dimensional array, where each image field appears at least once over the size of the human pupil can be a very effective way of enlarging the eyebox. The vast majority of 1D or 2D Exit Pupil Expanders (EPE) are waveguide based. The next section lists the various types of waveguide combiners, and the various type of waveguide couplers that can be used to perform the pupil replication. Examples of 1D EPE shown in Figure 22 are from Sony Ltd (Japan), Lumus Ltd (Israel), Optinvent SarL (France) and Dispelix Oy (Finland), and 2D EPE from BAE (UK), Digilens Corp. (USA), Vuzix Corp. (USA), Enhanced World (formerly WaveOptics) Ltd (UK), Nokia Oy (Finland), Magic Leap Corp. (USA) and Microsoft Corp. HoloLens (US).



Figure 22: Waveguide based Exit Pupil replication (EPE) in 1D (top) and 2D (center and bottom)

Gaze contingent exit pupil steering

When pupil expansion or pupil replication is not an option due to size and weight limitations, or even costs (waveguide gratings), one can implement a pupil steering scheme based on a gaze tracker (see Figure 23). Such steering matches the exit pupil of the imaging system to the human eye pupil at all times so that the user experiences a full FOV vision no matter where the users gazes at.

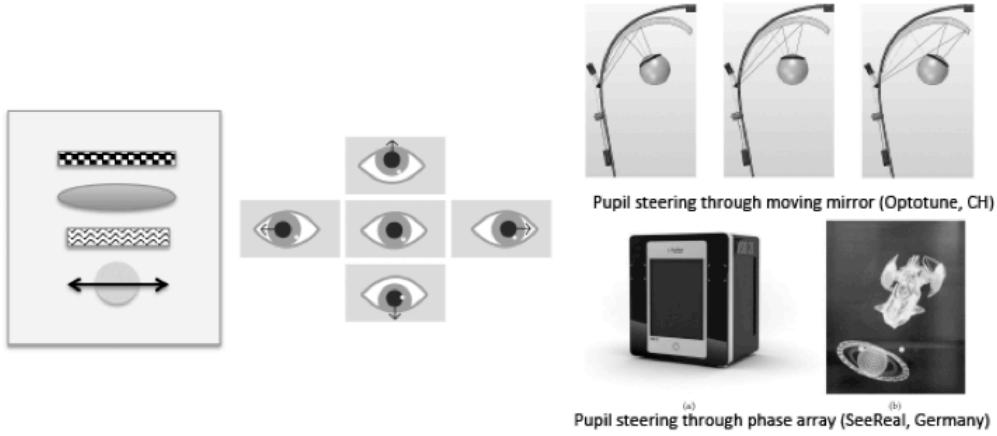


Figure 23: Gaze contingent pupil steering examples

Gaze contingent pupil steering relies however on a specific dynamic or tunable optical element, such as a slow movable mirror (such as a large MEMS mirror), or phase array steerers [38], or any other dynamic optical element such as switchable LC or PDLC based holograms. This can lead to a very compact form factor. SeeReal GmbH (Germany) [35] implemented such a gaze contingent exit pupil steerer, not on a wearable device, but rather on a desktop holographic 3D display device (Figure 23, right).

Display tiling

Yet another way to increase the eyebox without replicating the exit pupil is to simply replicate (or tile) the optical engines (display and lens). This would seem prohibitive if the display engine is large and bulky (such as an LCOS, LCD or scanner display), but makes sense if the resolution is kept low and the display optics are miniature (micro-optics). Implementation examples include “shell” type displays such as Lusovu (Portugal) shell display architecture based on transparent OLED curved panels with see through reflective MLAs (reflective, Fresnel, diffractive or holographic), and pin light displays (see Figure 24), either in transmission mode [39] or reflective mode [40] – see also section on VAC mitigation for more information on PinLight displays.

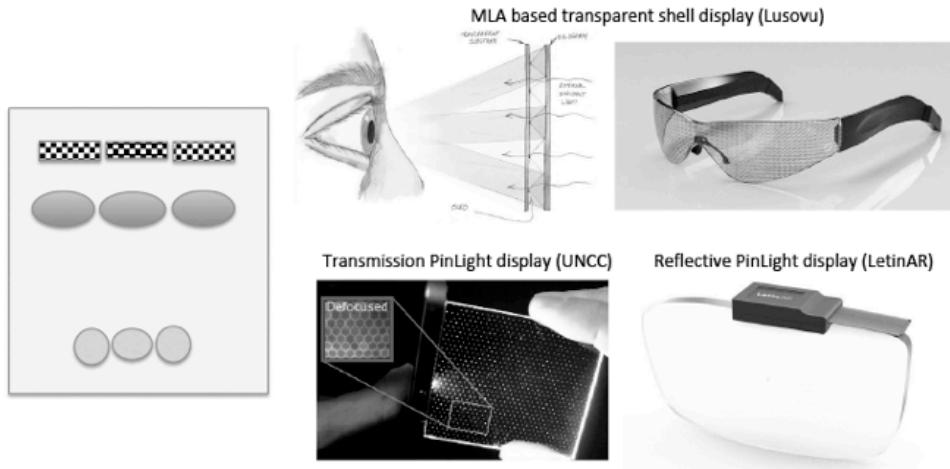


Figure 24: Display tiling examples

The display tiling (as in [41]) does not need to incorporate the entire scene under each microlens, as the scene FOV can be decomposed into various sub-scenes, each under an optical redirection element (such as an MLA), which would reconstruct the entire FOV. The various different sub-display/lens clusters would then build up the desired FOV over the desired eyebox.

Gaze contingent collimation lens movement

If there would be a world's eyebox size contest, the winning architecture would certainly go to one where the collimation lens would move physically as the eye moves around, following it closely. The simplest implementation of such a system

is to affix the lens directly to the cornea, such as with a contact lens. Such an architecture has been implemented by Innovega (now called Emacula). See Figure 25.

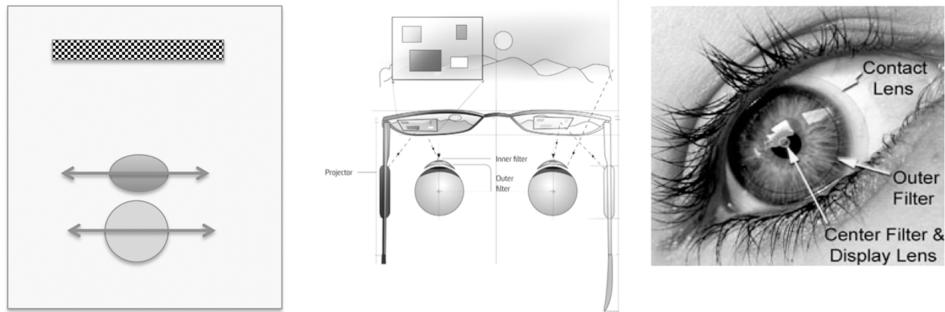


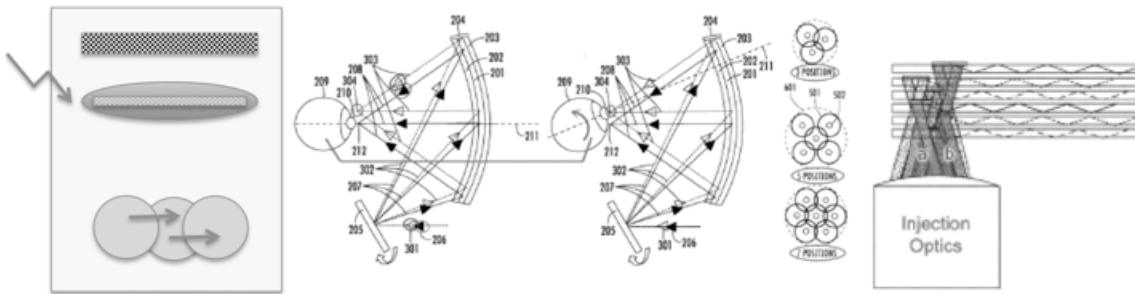
Figure 25: Collimation lens following gaze and pupil.

Here, the display (either an aerial image from a temple projector or an actual physical microdisplay panel) is located on a the back surface of the lens in pair of glasses. The display is polarized and the contact lens has a small collimation lens on its center. This mini lens is covered with a polarization film that matches the display polarization. The world is polarized the other way, and so is the peripheral portion of the contact lens. Thus, the see through field is not affected by the mini-collimation lens, only the display field is collimated and projected at near infinity. As the eye moves around, the mini lens on the contact lens moves accordingly, and thus moves around its exit pupil, its position matching at all times the position of the human eye pupil. However, the burden of wearing a pair of glasses in combination with a pair of contact lenses might be an issue for consumer adoption.

Exit pupil switching

Pupil switching is an interesting time domain eyebox expansion technique, when pupil steering might too complex to implement. As pupil steering, pupil switching is gaze contingent through an eye tracker (or rather a pupil tracker here). In many cases, pupil switching can be as simple as switching spatially de-multiplexed LED dies in the illumination part of an LCOS display engine, as it has been implemented in Magic Leap One [42] and Figure 26 (although the pupil switching was done there for other purposes – focus switching – rather than eyebox expansion). The pupil switching technique gets interesting only if the switching architecture remains simple and static, (illumination LED dies switch, phase multiplexed Bragg hologram couplers, angle selective metasurface couplers, etc...). The main difference between pupil switching and pupil steering is that in pupil switching there is no complex active steering mechanism (mirror, phase plate steerer, phase LCOS, ...). The VAC mitigation section provides more info on how the illumination path pupil switching works in Magic Leap One.

The Discontinued Intel Vaunt smart glasses ([43] and Figure 26 – center) is another example of pupil switching based on illumination switching. Vaunt operates through phase multiplexed reflective Bragg volume holograms and VCSEL wavelength switching. The imaging task is performed by a miniature MEMS scanner. This allows for multiple exit pupils to be formed by slightly different VCSEL wavelengths (as in 645nm, 650nm and 655nm), fooling the human eye without fooling the 3 phase multiplexed holograms inserted in the free space combiner lens (the Bragg selectivity of each hologram being smaller than the VCSEL spectral shift).



switching to provide to the user an expanded eye box.

9. Roadmap for VR headset optics

Before reviewing the combiner architectures for see through MR systems, we review here the various generations of collimation optics for non-see through MR (or VR) HMDs. Standard refractive lenses have strong limitations due mainly to angle of incidence, weight, size, which limits their optical power and thus the distance between the display and the optic (thus the size of the HMD and the weight and location of the headset's CG).

Hybrid Fresnel lenses have been used in most of the VR HMD in the past years, such as in the HTV Vive, Oculus CV1 and others (see Fig.X). Such lenses can either be refractive Fresnel or diffractive Fresnel. A hybrid diffractive Fresnel lens over a curved surface can provide an effective way to reduce Lateral Chromatic Spread (LCA) as an achromatic singlet. More recent developments attempt at designing lenses which center foveated region is pure refractive and becomes hybrid Fresnel towards the edges of the lens. This also reduces the total thickness without altering the central foveated high resolution area. Interesting Fresnel concepts such as the one from Wearality (far right on Fig. X) attempt to increase the FOV to very numbers without having to increase the display size, and reducing dramatically the weight and thickness of the lens. However, such lenses have also larger Fresnel zone ring artifacts in the image.

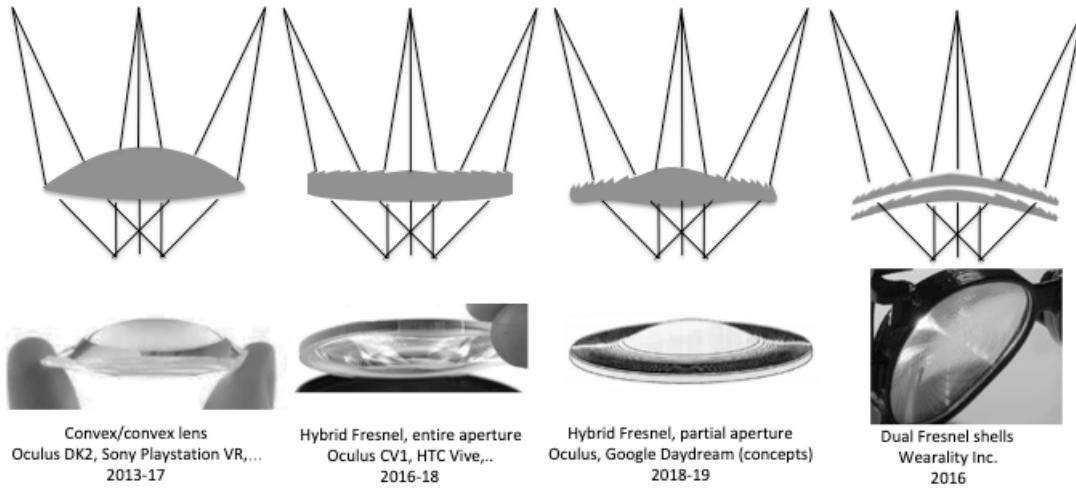


Figure 27: Successive VR lens configurations aiming at increasing the FOV while reducing the weight and size.

Reducing the weight and size of the lens is one aspect of comfort; reducing the distance between the lens and the display is also desirable since it improves the overall form factor and pushes the CG back to the head for improved wear comfort. Reducing the distance between the lens and the display requires to increase the power of the lens (reducing its focal length). Designing a stronger lens usually impacts the efficiency and MTF of the systems described in Figure 27. Other lens configurations, or rather compound lens configurations have to be investigated to do so. Figure 28 shows a few such compound lenses: polarization pancake lenses, multi-path compound lenses and MLAs, compared to traditional lenses.

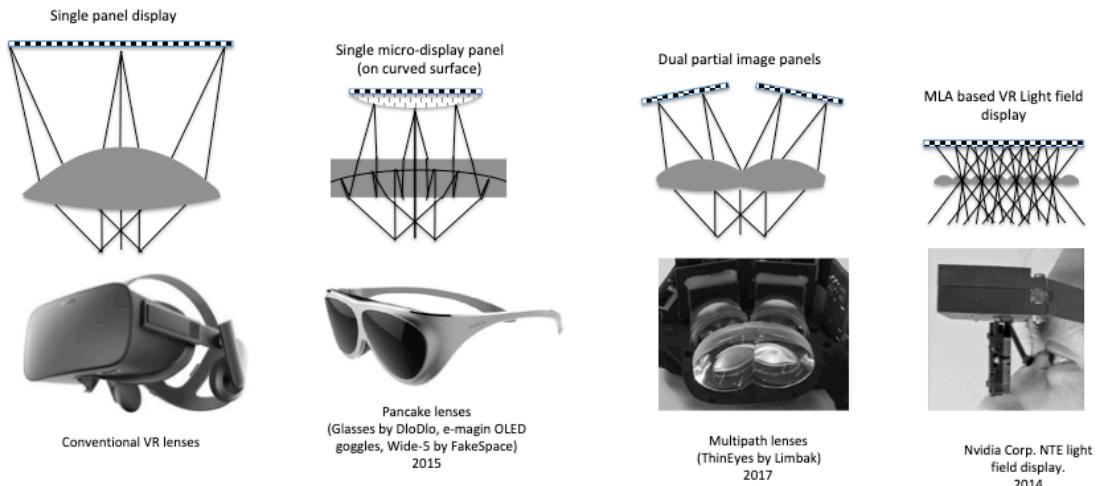


Figure 28: Compound VR lens configurations aiming at decreasing the distance between the display and the lens, reducing the form factor of the VR headset.

Pancake polarization lenses have been investigated since 1976 in order to increase the power of the lens without expanding its volume and weight [44]. However it provides new challenges such as polarization ghosts and low birefringence plastic is low cost are a requirement (see VR glasses by DioDlo, China [45]). A curved display plane can enhance the quality of the Pancake lens display (by using a fiber plate as in the e-magin OLED VR prototype).

Very interesting multipath lenses have also been investigated recently [46], [47], providing a smaller form factor while remaining high resolution optics. The concept is somehow similar to the MLA based light field display (introduced by Gabriel Lippmann, Nobel prize for integral imaging in 1908). However, in this particular case, the MLA array is reduced to 2 or 4 lenses. It uses multiple individual displays (2 in Figure 28), each depicting a partial image which is then fused together as the eye approaches the optimum eye relief. When this architecture is scaled towards using a large array of lenses such as in an MLA, the architecture comes close to the NVidia NTE Light Field VR display (far right on Figure 28).

We have reviewed in this section various novel VR lens configurations which can either reduce the size and weight of the lens, increase the FOV or reduce the overall size by reducing the distance between the lens and the display. Such lens configurations could also be used in AR and see through MR systems, with some modifications.

10. Free space optical combiners for AR

In section 8, we saw how both free space and waveguide based optical combiners can implement specific eyebox expansion schemes, either in VR or AR modes. We review in this section the various AR free space combiners, and in the next section, the waveguide combiner architectures used in industry today.

Free space combiners have been used extensively in defense applications, especially for HMDs in rotary wing aircrafts [1], [48], [49], from small FOV to mid FOV [50] to ultra wide FOV [51], [52].

The most straightforward free space combiner architecture would be a tilted flat half tone mirror [48], as used still today in defense AR systems such as the Apache Helicopter temple mounted monocular IHADSS (using a mini CRT imager), or in the consumer/enterprise market, the binocular ODG R7 top-down combiner plate from Osterhout Design Group, using a micro-OLED microdisplay (see Figure 29). Such architectures produces however a small eyebox, the collimation lens being located further away from the eye. These architectures are thus also limited to smaller FOV.

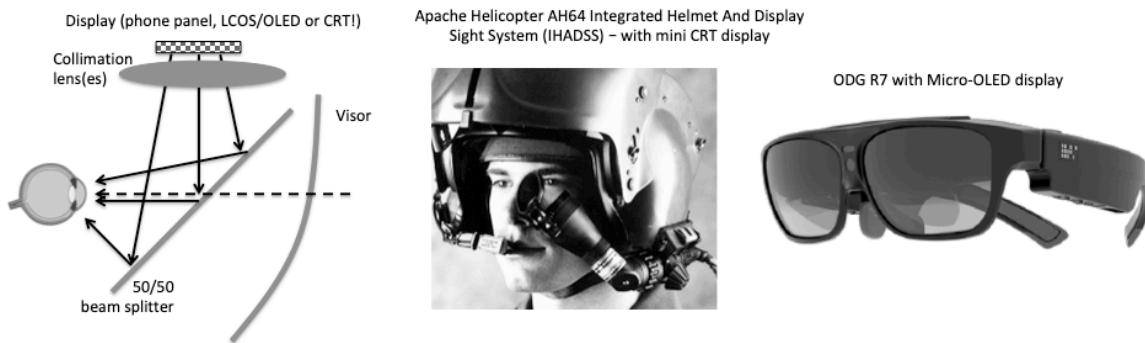


Figure 29: Tilted half tone mirror combiner architectures (IHADSS & ODG R7).

Free space combiners might also have optical power, working in reflection mode through a half tone coating (or polarization coating if the display is polarized). These yield good see through and reduced distortion as well as good color fidelity and low LCA (due to their operation in reflection mode). Also, there is no need for a compensator as in TIR prism combiners (see next section). However, they require a large tilted curved optical surface, and push the CG towards away from the head making the device bulky.

For the defense market, where a large FOV is usually desired and Industrial Design (ID) is of no concern, microdisplay based "bug eye" curved reflector combiner architectures are often used (see Figure 30 and [48], [49]). However, étendue and high resolution constrains call for often for very complex (and costly) off-axis optical relay systems as shown below (AHMD on the left with 65deg HFOV and 50 deg VFOV, BAE 40 deg DFOV fixed wing aircraft HMD in the center, and a 60 deg FOV design example on the right [7] showing the complexity of the off-axis temple mounted relay optics).



Figure 30: Single reflective combiners “bug eye” type, with temple side off-axis pupil forming imaging system (left: AHMD for rotary wing aircraft, center: BAE for fixed wing aircraft, right: relay optics design example).

More recently, similar architectures with much simpler relay optics have been developed for the consumer AR market, where ID concerns matter much more than for the defense market: these are based on large consumer smart phone display panels rather than complex off-axis imaging systems and a micro-display (Figure 31: top mount display for Meta 2 MR headset from MetaVision, Mira Prism and DreamGlass, and temple mounted display for the LeapMotion NorthStar AR reference design). The costs of such optical combiners remain low (\$49 for the Mira Prism, and \$99 for the Leap Motion reference design). When the display panel is provided along with a sensor bar (TOF depth map, gesture sensor and 6DOF head trackers), as in Dreamglass and Meta2, the costs are getting higher, from \$400 to \$1000.

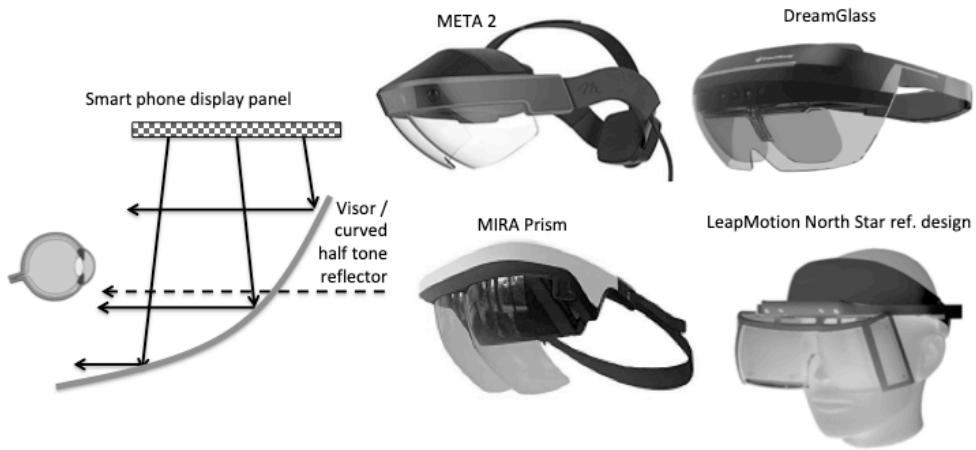


Figure 31: Curved halftone visor/reflectors used to generate large FOV from large panels.

These architectures are however prone to optical distortions [53], [54] due to the large FOV and the single optical surface for the entire imaging task. Distortion variations when the user moves his or her eye pupil around is referred to as “pupil swim” (the pupil swing can be as large as +/-5mm for a 90 deg FOV). Reducing the pupil swim leads to a better visual comfort, but is difficult to achieve in such architectures. We saw previously also that pupil swim can be actively compensated by using an eye tracker (or rather a pupil tracker in this case).

In order to reduce the protuberance of the optical combiner, additional optical elements are needed, especially when the display is a micro-display to reduce the size and weight of the headset. Such devices might be based on “air birefringent” architectures, where the birefringent is a reflective collimation lens working with an additional flat combiner, as depicted in Figure 32 (ODC R9, Lenovo AR “Star Wars”, and a variety of Chinese OEM display engines). The FOV generated from relatively small display panels can be quite large, reaching 50 or 60 deg diagonal, but not as large as the FOV generated by the previous architecture using large cell phone panels and single large curved combiner. However, the distortions and pupil swim effects can be better corrected here since there are more optical surfaces to work with.

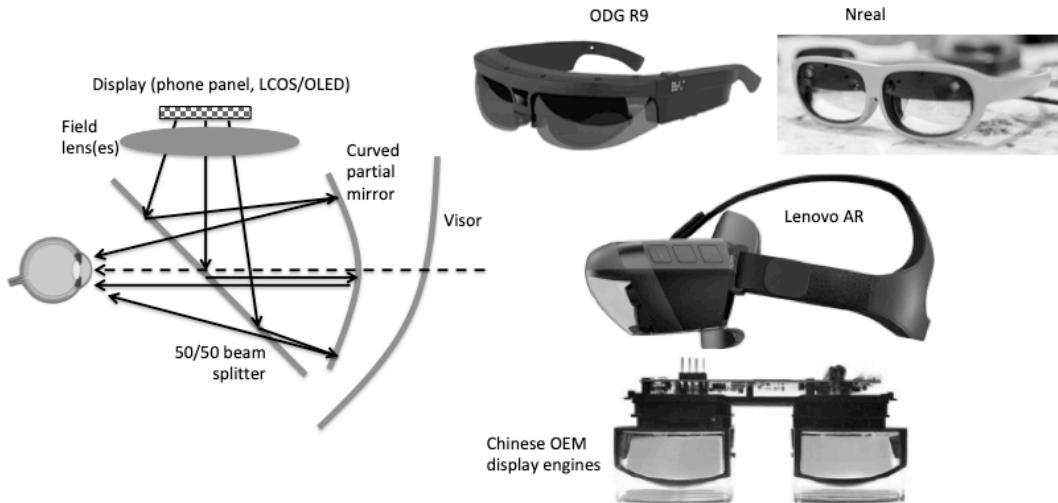


Figure 32: Typical air birdbath free space combiners (ODG R9, Nreal, Lenovo AR, Chinese OEM display engines).

Bird bath optical architectures may also be used in glass or plastic media rather than in air, and still be called free space architectures since there is no wave-guiding involved. Figure 33 shows such an example in Google Glass: the microdisplay (Himax LCOS with PBS cube backlight in V1 and Kopin LCD with backlight in V2) is temple mounted, and collimated by a 100% reflective metal-coated lens located on the nasal side. The collimated field is then redirected to the user's eye by a 50/50 beam splitter. The use of a PBS to redirect the field into the user's eye would have been much more effective, but the lack of available low birefringence plastic to make up the rod led to the optimal choice of a 50/50 beam splitter, as losing brightness is a better option than producing ghost images from unwanted polarization states.

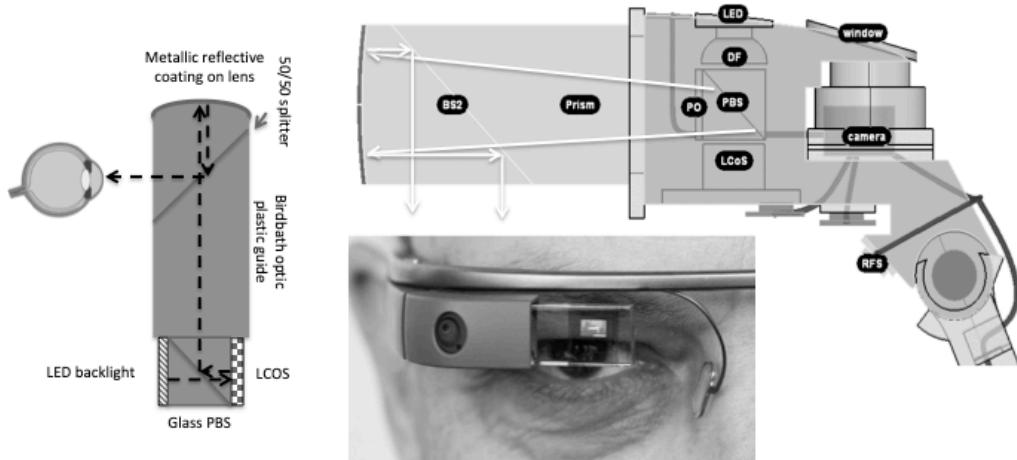


Figure 33: In media bird-bath combiners (Google Glass and copycats).

Other free space architectures might use more than one reflective curved mirror to build up a large FOV from a small microdisplay area. They usually have a top-down image injection as shown in Figure 34 – top (Immy Corp. AR headset based on a Micro-OLED display). Color fidelity and LCA are under control since only reflection optics are used. The concept of three off-axis mirrors is widely used in telescopes design with OAPs (Off-Axis Paraboloid mirrors). The FOV in IMMY is wide (60 deg).

The Raptor smart glass (Figure 34, bottom) for cycling enthusiasts from Eversight uses also a simple visor reflection, with an unusual position of the display engine on the nasal side. The Raptor display engine uses here a large transmission lens and an OLED microdisplay. The visor is coated with a partial mirror on the reflection area. Its FOV remains small, however well adapted for its particular use in cycling sports.

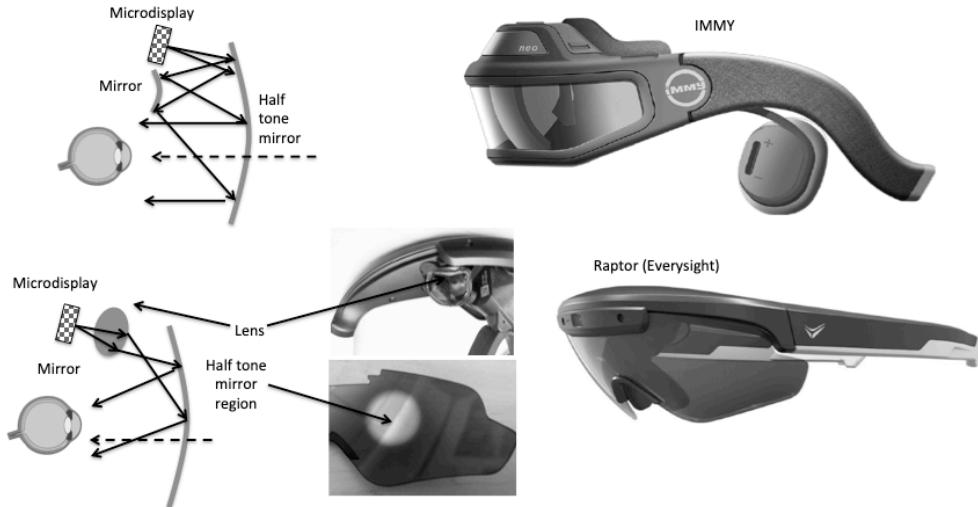


Figure 34: Off-axis visor reflection combiner with either mirrors (top) or refractive lens (down).

When the free space optical combiner size and curvature is too pronounced for a decent industrial design fit, especially when the combiner should look like standard prescription glasses, one can implement the optical power as a hybrid curved reflective/diffractive or curved reflective/Fresnel combiner. The compensation of the see-through is then implemented by index matching of microstructures coated with a partially reflective layer. The form factor can thus be close to a prescription lens. Figure 35 shows such examples (Toshiba Glass from Japan, upper, using a hybrid curved embedded Fresnel structure in a lens, and Glass-Up from Italy, lower, using an embedded surface relief diffractive structure in a lens).

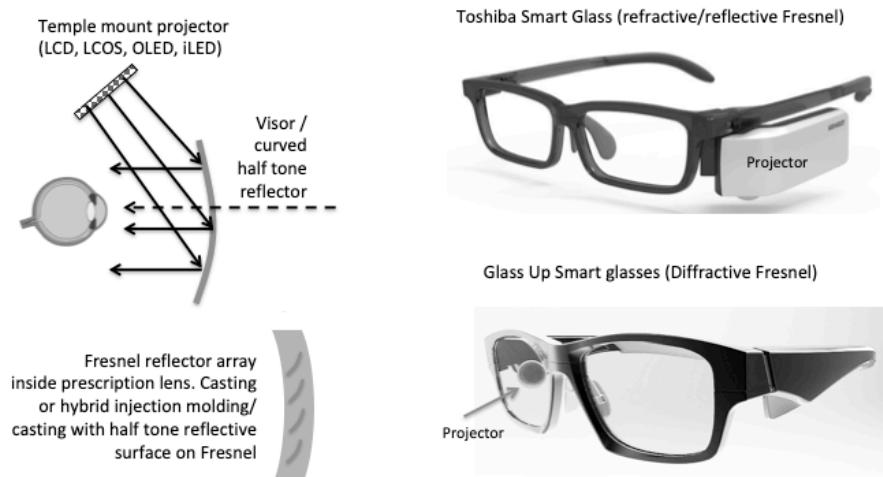


Figure 35: Hybrid refractive Fresnel combiner (Toshiba) and hybrid diffractive Fresnel (Glass-Up).

As Fresnel structures can produce LCA, one can use a diffractive structure on each Fresnel zone, reducing the LCA since the dispersion of diffractives is opposite of that of Fresnel zones (refractive). The diffractives and Fresnels can be injection molded in the same plastic, and have the same power sign, unlike with traditional hybrid achromatic lenses in which both optical power and index need to be different.

Another technique to reduce the curvature of the free space combiner is to use a holographic reflective layer(s) on the flat or curved combiner surface, either in air [55],[56] or in a lightguide media [57].

One of the problems of such architectures is the small FOV as well as the small eyebox produced by the lens combiner. In order to increase the FOV without increasing the size of the temple side optical engine (and microdisplay), one can use instead a laser or VCSEL MEMS scanner as in the early Brother AirScouter and QD Laser smart glasses, or in the more recent Intel Vaunt [58] or Thalmic Labs (now called By North, Canada) smart glasses, both of which are shown in Figure 36.

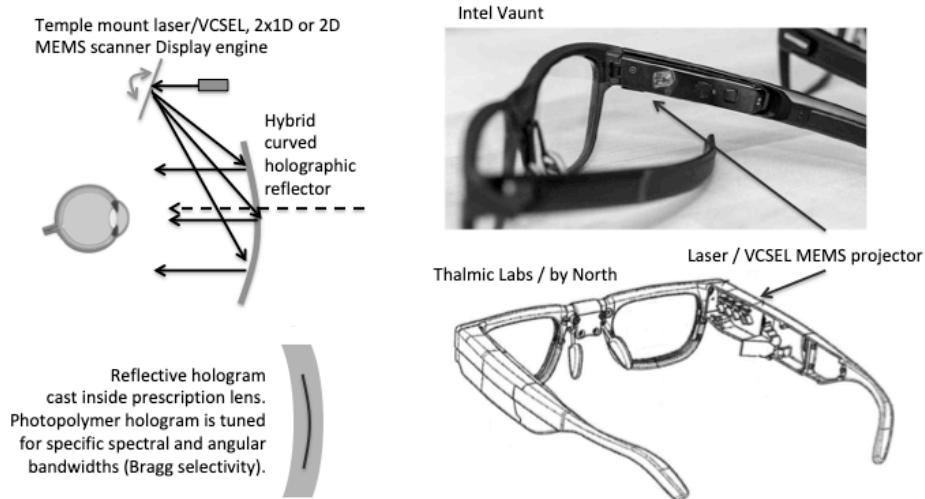


Figure 36: Temple mount laser/VCSEL MEMS scanner with hybrid volume holographic combiner.

Intel Vaunt and By North Smart Glasses are using a hybrid optical combiner consisting of a reflective volume holographic off-axis lens embedded in an ophthalmic (or zero power) lens. The hologram is transparent to the natural field in see through mode but acts as an off-axis lens in reflection mode for the specific laser wavelengths. The embedding of such volume hologram in a lens can be done via casting, which is also the standard technique used to produce ophthalmic lens pucks. The photopolymer hologram on its underlying PET film is inserted in a standard ophthalmic glass cavity, and after casting from the bottom of the cast to avoid the formation of bubbles, UV curing is set, followed by a slow thermal annealing process to reduce stress and harden the cast polymer. This process also reduces dramatically the birefringence in the plastic when compared to injection molding.

In order to increase the eyebox size, especially when the FOV gets larger, one can use a pupil replication scheme by using several different laser or VCSEL emitters tuned to slightly different wavelengths, along with different phase multiplexed holograms in the photopolymer volume hologram, each sensitive to those specific wavelengths (as in the Intel Vaunt Smart Glass). The user does not perceive the wavelength changes but the hologram does thanks to the acute spectral Bragg selectivity of the embedded Photopolymer hologram.

Note that even with such exotic eyebox expansion schemes, the resulting eyebox can remain small. Transforming a limitation in a feature is definitely possible by explaining that if the user does not want to see the display, he or she has simply to move the line of sight back to front, as the eyebox location is set to the side. To see again the display, the user moves his or her gaze (and thus the pupil) back to the location of the small eyebox.

Free form TIR prism combiners.

Freeform prism based combiners have been extensively investigated [59], [60]. [61] since the emergence of free form diamond turning machines with 5 DOF axis a decade ago. These machines are now becoming standard fabrication tools in most optical manufacturing shops (for direct machining in plastic or for metal mold machining followed by pressure injection molding).

Typical freeform prism combiners include a semi-transparent coated surface through which the users looks and a TIR surface, with a top or temple image injection (see Figure 37). TIR prism combiners require a conjugate bounded prism to compensate for the see through distortion introduced by the combiner prism. Both can be injection molded and bonded together, but produce thick and heavy compounds, especially for large FOVs over 30-40 deg. Alternatives have been proposed with tiled prism combiners to combine various FOVs together [59]. Even more refined designs include glint based eye tracking through the same TIR prism combiner, providing an infinite conjugate to the display while providing a finite conjugate to the eye tracking architecture [59].

Examples of early prism combiner AR devices include Canon, Motorola AR headsets and Verizon Golden-i, more recent ones include Lenovo Daystar AR, ThirdEye (China), NED Glass X2 (China), and many others thanks to the commoditization of freeform prism design and fabrication through US (Rochester based optical fabs) and Chinese OEM optics vendors.



Figure 37: TIR combiner prism with compensator (left) and implementation examples (right).

These conventional TIR prism combiner have typically one TIR bounce and one bounce off a coated surface, which is fitted with a compensator. More complex curved longer TIR prisms can be designed, with the display engine on the temple side, interesting for thin smart glasses, with TIR bounces up to 3 or 4 (as depicted in Figure 38).

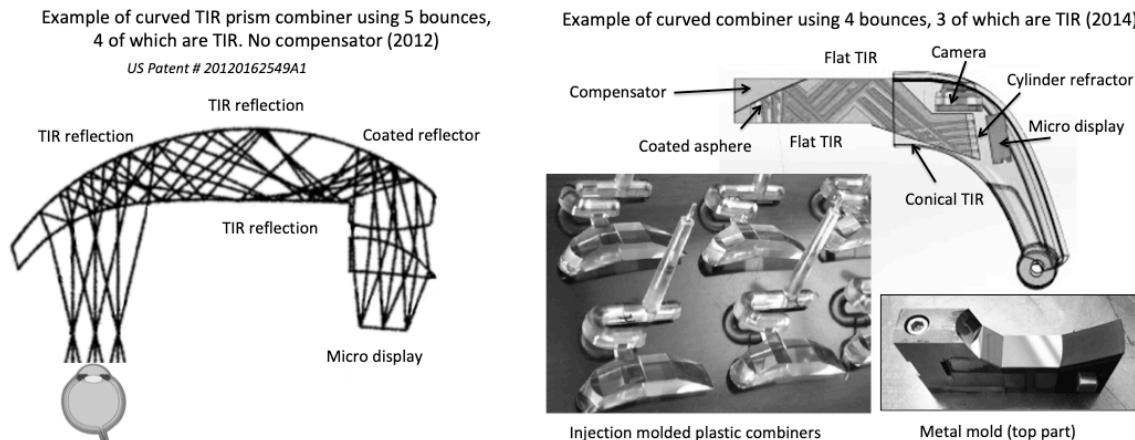


Figure 38: Example of multiple TIR bounces in temple mounted smart glass TIR prism combiners

The design on the left side shows a 5 bounces curved freeform TIR lightguide combiner without compensator, the first bounce being over a coated surface. The light engine also uses an additional off-axis lens (Augmented Vision Inc. 2012).

The design on the right side of Figure 38 shows a version of Google Glass with a TIR prism combiner that has an efficiency increase of at least 2x of that of the initial birdbath combiner design, as it does not use a double path through the half tone combiner (see Figure 20 and Figure 33). Its curved shape offers a generous space for the front facing camera without compromising the industrial design. Top part of the diamond turned metal mold assembly as well as a few injection molded parts are also shown, still with their gate and runner attached.

Freeform TIR prism combiners are at the interface between free space and waveguide combiners. When the number of TIR bounces increase, one might refer to them as waveguide combiners. Waveguide combiner architectures is the topic of the next section.

11. Waveguide optical combiners

Waveguide combiners are based on TIR propagation of the entire field in an optical guide, acting as a transparent periscope with a single entrance pupil and often many exit pupils.

The core of a waveguide combiner consists of the input and output couplers. These can be either simple prisms, micro-prism arrays, embedded mirror arrays, surface relief gratings, thin or thick analog holographic gratings, metasurfaces or

resonant waveguide gratings. All of the above have their specific advantages and limitations, all of which we discuss in this section. Waveguide combiners have been used historically or for tasks very different from AR combiner, such as planar optical interconnections [62] and LCD backlights [63], [64].

If the optical designer has chosen to use a waveguide combiner rather than a free space combiner for the various reasons listed in previous sections, the choice among existing waveguide techniques still remains vast, as much for expansion architecture as for actual coupler technologies. This section gives a tour of the current waveguide combiner zoo.

Waveguide combiners are an old concept, some of the earliest IP dates back to 1976 and applied to HUDs. Figure 39 (left) shows a patent by Juris Upatnieks dating back 1987, a Latvian/American scientist and one of the pioneers of modern holography [65], implemented in a Di-Chromated Gelatin (DCG). A few years later, 1D EPE architectures have been proposed as well as a variety of alternatives for in- and out-couplers technologies such as surface relief grating couplers by Thomson CSF (right).

Original patent on waveguide combiner with EPE (1987)

United States Patent [19]		[11] Patent Number: 4,711,512
		[45] Date of Patent: Dec. 8, 1987
[54]	COMPACT HEAD-UP DISPLAY	4,314,284 3/2/1982 Kramer ————— 350/182 R
[75]	Inventor: Juris Upatnieks, Ann Arbor, Mich.	Primary Examiner—Bruce Y. Arnold
[73]	Assignee: Environmental Research Institute of Michigan, Ann Arbor, Mich.	Attorney, Agent, or Firm—Kress & Young
[21]	Appl. No.: 754,406	[57]
[22]	Filed: Jul. 12, 1985	ABSTRACT
[51]	Int. Cl. ————— G02B 5/32; G02B 27/10	compact head-up display comprising a combiner which collects light energy from a remote viewer, reflects light energy through a beam splitter, and diffracts it out by means of an upper diffraction grating. The head-up display disclosed further comprises an instrument panel, a collimating lens, and a lower diffraction grating for diffracting the collimated light energy into the combiner.
[52]	U.S. Cl. ————— 350/3.7; 350/174	
[38]	Field of Search ————— 350/174, 3.7	
[56]	References Cited	
	U.S. PATENT DOCUMENTS	
	4,309,070 1/1982 St. Leger Seurle ————— 350/3.7	

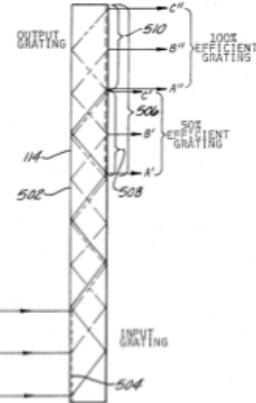
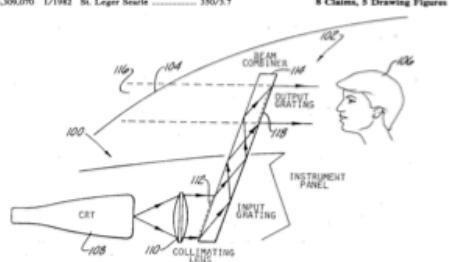


FIG. 5 is a schematic diagram of a combiner of the invention having an extended aperture.

Figure 39: Early patents showing the concept of holographic (left) and surface relief grating waveguide combiners for HUD and HMD applications.

Words of caution for the rigorous optical engineer reader: The term “waveguide” is here not the right term to refer to these combiners, there are rather thick “lightguides” with a high number of propagating modes, but due to the overwhelming use of the term “waveguide combiners”, this term is now fully accepted in industry. Similarly, when industry uses the term “hologram” to refer to fixed focus stereo imaging, or even the term “light field” to refer to any attempt, no matter how basic it might be, to solve the VAC conflict (see also section 8).

Single output pupil and curved waveguide

If the FOV is small (<20deg diagonal), such as in smart glasses, it might not be necessary to call for an exit pupil expansion scheme, making the waveguide design much simpler, and allowing for more degrees of freedom, such as curving the waveguide. Indeed, if there is a single output pupil, the waveguide can imprint optical power onto the TIR field, as it is done in the curved waveguide Smart Glass by Zeiss in Germany (developed now with Deutsche Telekom), see Figure 40.

Two bounded curved shells with coated Fresnel structures and freeform injector.

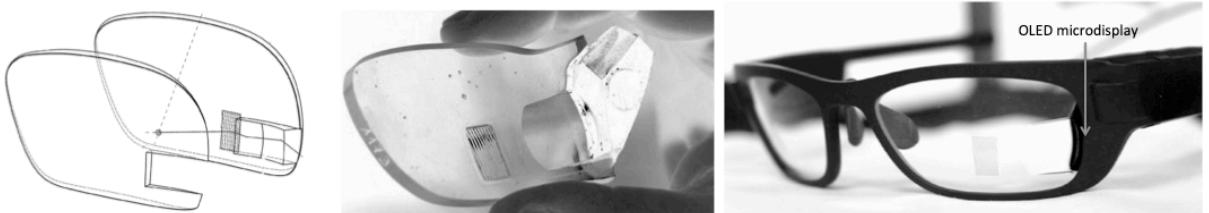


Figure 40: Zeiss monocular Smart Glass, single exit pupil allowing for curved waveguide.

In the Zeiss smart glass, the exit coupler is an off axis Fresnel reflector. The FOV as well as the outcoupler is ex-centered from the line of sight. The FOV remains small (11 deg), and the thickness of the guide relatively thin (3-4mm).

Single exit pupil have also been implemented in flat guides, as in the Epson Moverio BT100, BT200 and BT300 (temple mounted optical engine in 100m thick guide with curved half tone extractor in BT300) or in the Konica Minolta smart glasses, with top down display injection and flat RGB panchromatic volume holographic extractor (see Figure 41).

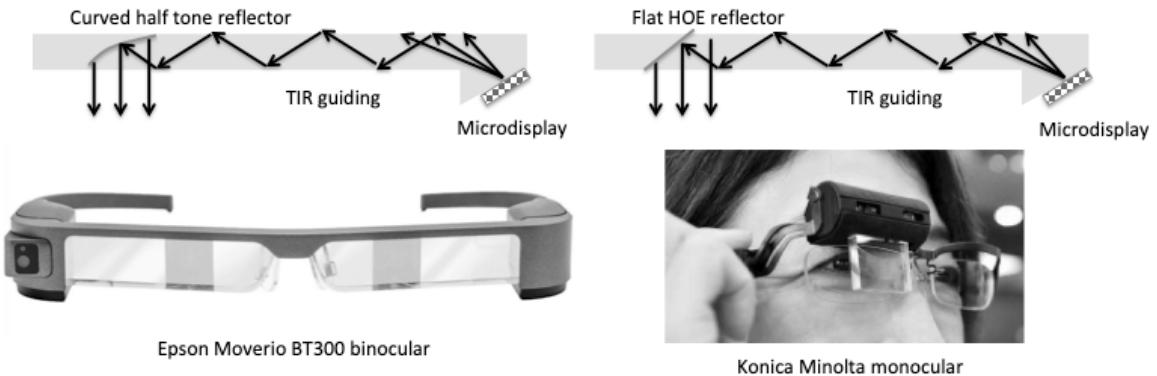


Figure 41: Single exit pupil flat waveguide combiners (with curved reflective or flat holographic outcouplers).

Single exit pupil (no EPE) are well adapted to small FOV smart glasses. If the FOV gets larger than 20deg, especially in a binocular design, 1D or 2D exit pupil replication is required. We discuss these in the following sections.

Covering a large IPD range (such as a 95 or 98 percentile of the target consumer population, including various facial types), requires a large horizontal eyebox, typically 10mm to 15mm. Also, due to fit issues and nose pad designs, a similar large vertical eyebox is also desirable, ranging from 8mm to 12mm.

1D eyebox expansion

As the horizontal eyebox is usually the most critical to cover large IPD ranges, a single dimensional exit pupil replication might be sufficient. The first attempts used holographic extractors (Sony Ltd) [66], [67], with efforts to record RGB holographic extractors as phase multiplexed volume holograms [68], and also cascaded half tone mirror extractors (LOE from Lumus, Israel) or arrays of micro-prisms (Optinvent, France) [69]. This also reduced the 2D footprint of the combiner, working only in one direction. However, in order to generate a sufficiently large vertical eyebox, the input pupil produced by the display engine need to be quite large, much larger than the horizontal size of the pupil since this one is replicated (see Figure 42).

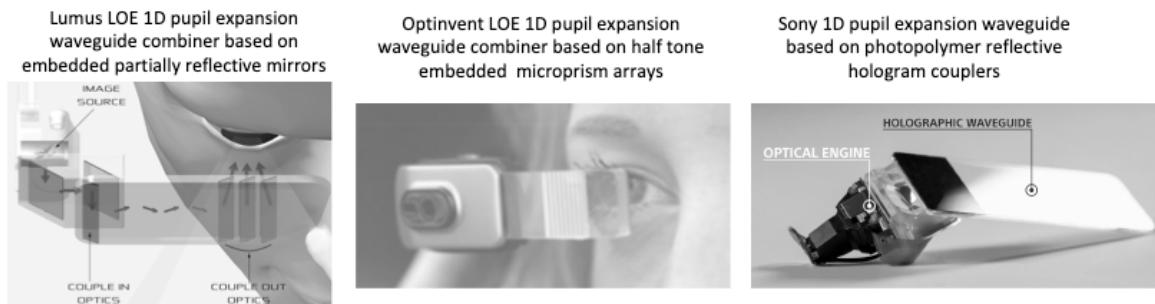


Figure 42: 1D Horizontal EPE with tall aspect ratio input pupil to form vertical eyebox.

In many cases, forming a tall aspect ratio input pupil can lead to larger optical engines. However, a single vertical pupil with natural expansion will provide the best imaging and color uniformity over the eyebox.

2D eyebox expansion

Two dimensional eyebox expansion is desired when the input pupil cannot be generated by the optical engine over a tall enough aspect ratio to form the 2D eyebox, because of FOV (é längde limitations) and related size/weight considerations. A two dimensional exit pupil expansion is therefore required. Various types of 2D EPE replication have been developed, from cascaded X/Y expansion as in the Digilens/Nokia/Vuzix/HoloLens/Magic Leap One combiner architectures [70], [71], [72], to combined 2D expansion [73], [74] as in the BAE Q-Sight combiner or the WaveOptics Ltd. Phlox 40deg combiner

(now Enhanced World Ltd) grating combiner architectures (see Figure 43).

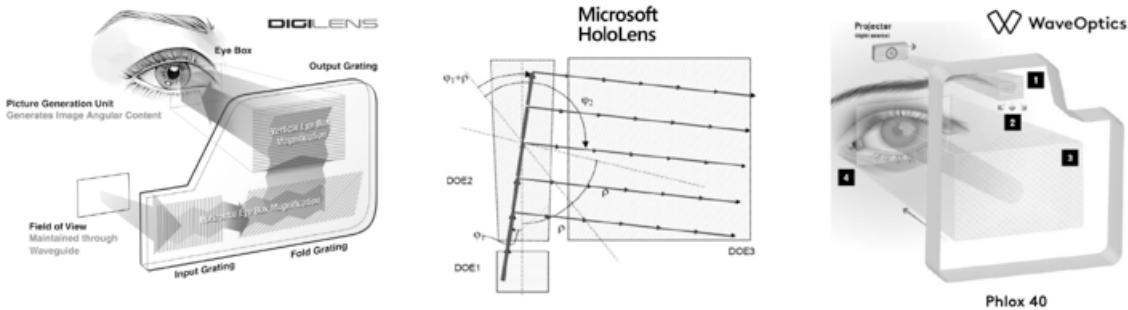


Figure 43: 2D pupil replication architectures in planar optical waveguide combiners.

Choosing the right waveguide coupler technology

The coupler element is the key feature of a waveguide combiner. The TIR angle is given by the refractive index of the waveguide, not the refractive index of the coupler nanostructures. Very often, the index of the coupler structure (grating or hologram) dictate the angular and spectral bandwidth over which this coupler can act, thus impacts the color uniformity over the FOV and the eyebox.

There are numerous of coupler technologies that have been used in industry and academia to implement the in-coupler and out-couplers, these include:

1) Macroscopic prism

A prism is the simplest TIR in-coupler one can think of and also the earliest one used. A prism can be bounded on top of the waveguide or the waveguide itself can be cut at an angle allowing normal incident light to enter the waveguide and be guided by TIR (depending on the incoming pupil size). Another way is to use a reflective prism on the bottom of the waveguide (metal coated).

2) Embedded cascaded mirrors

Cascaded embedded mirrors with partially reflective coatings are used as out-couplers in the Lumus Lightguide Optical Element (LOE) waveguide combiner. The input coupler remains a prism. As the LOE is composed of reflective surfaces, it yields a good color uniformity over the entire FOV. As with other coupler technologies, intrinsic constraints in the cascaded mirror design of the LOE might limit the FOV [75]. See-through is very important in AR systems: the Louver effects produced by the cascaded mirrors in earlier versions of LOEs have been reduced recently thanks to better cutting/polishing and coating and better design.

3) Micro-prisms array

Micro-prism arrays are used in the Optinvent (France) waveguide as out-couplers [69]. The in-coupler here is again a prism. Such microprism arrays can be surface relief or index matched to produce an unaltered see through experience. The micro-prisms can be coated with a half tone mirror layer, or can be alternatively totally reflective and transmissive to provide a 50/50 see through experience. The Optinvent waveguide is the only flat waveguide available today as a plastic guide, thus allowing for a consumer level cost for the optics. The micro-prism arrays are injection molded in plastic and bounded on top of the guide.

4) Thin reflective holographic couplers

Transparent volume holograms working in reflection mode, as in Di-Chromated Gelatin (DCG), bleached Silver Halides or more recently photopolymers such as Bayfol® photopolymer by Covestro/Bayer, (Germany) [76], photopolymers by DuPont (US), Polygramma (Brasil) or Dai Nippon (Japan), have been used to implement in- and out-couplers in waveguide combiners. Such photopolymers can be sensitized to work over a specific wavelength, or over the entire visible spectrum (panchromatic holograms).

Photopolymer holograms do not need to be developed as DCG, or do not need to be bleached as Silver Halides. Having a full color hologram based on 3 phase multiplexed single color holograms allows for a single plate waveguide architecture, simplifying the combiner and reducing weight, size and costs while increasing yield (no plate alignment required). However, the efficiency of such full RGB phase multiplexed holograms are still quite low when compared to single color photopolymer holograms.

Also, the limited index swing of photopolymer holograms allow them to work more efficiently in reflection mode than transmission mode (allowing for a better confinement of both the wavelength and angular spectrum bandwidths).

Examples of photopolymer couplers include Sony LMX-001 Waveguides for smart glasses, TrueLife Optics (UK) mastering the hologram in Silver Halide and replicating it in photopolymer.

Replication of the holographic function in photopolymer through a fixed master has been proven to be possible in a roll to roll operation by Bayer (Germany). Typical photopolymer holographic media thickness range from 16 microns to 70 microns depending on the required angular and spectral bandwidths.

5) Thin transmission holographic couplers

When the index swing of the volume hologram can get larger, the efficiency gets higher and the operation in transmission mode becomes possible. This is the case with Digilens's proprietary H-PDLC hologram material [77] (Holographic Polymer Dispersed Liquid Crystal). Transmission mode requires the hologram to be sandwiched in between two plates rather than laminating a layer on top or bottom of the waveguide as with photopolymers, DCG or silver halides. Digilens' H-PDLC has the largest index swing today, and can therefore produce strong coupling efficiency over a thin layer (typically 4 microns or less). H-PDLC material can be engineered and recorded to work over a wide range of wavelengths allowing therefore full color operation.

6) Thick holographic couplers

Increasing the index swing can optimize the efficiency and/or angular and spectral bandwidths of the hologram. However, this is very difficult to achieve with most available materials, and might also produce parasitic effects such as haze. Increasing the thickness of the hologram is another option, especially when sharp angular or spectral bandwidths are desired such as in telecom spectral and angular filters. This is however not the case for AR combiner, where both spectral and bandwidths need to be wide (to process a wide FOV over a wide spectral band such as LEDs). However, a thicker hologram layer also allows for phase multiplexing over many different holograms one on top of each other, allowing for multiple Bragg conditions to operate in concert to build up a wide synthetic spectral and/or angular bandwidth, as it can be modeled by the Kogelnik theory [79]. This is the technique used by Akonia Inc. (a US start up on Colorado, formerly InPhase Inc., which was originally funded and focused to produce high density holographic page data storage media, ruled by the same basic holographic phase multiplexing principles [78]).

Thick holographic layers, as thick as 500 microns, work well in transmission and/or reflection modes, but need to be sandwiched in between two glass plates. In some specific operation modes, the light can be guided inside the thick hologram medium, and therefore not limited anymore by the TIR angle dictated by the index of the glass plates. As the various holograms bandwidths are building up the final FOV, one needs to be cautious in developing such phase multiplexed holograms when using narrow illumination sources such as lasers.

Replication of such thick volume holograms are difficult in a roll-to-roll operation, as done with thinner single holograms (Covestro Photopolymers, H-PDLC), and require multiple successive exposures to build up the hundreds of phase multiplexed holograms constituting the final holographic structure. This can however be relatively easy with highly automated recording set-up as the ones developed by the now defunct holographic page data storage industry (In-Phase Corp, General Electric, etc...).

It is interesting to note that although the individual holograms acting in slivers of angular and spectral bandwidths spread indeed the incoming spectrum as any other hologram (especially when using LED illumination), but the spectral spread over the limited spectral range of the hologram is not wide enough to alter the MTF of the immersive image, and thus does not need to be compensated by a symmetric in- and out-coupler as with all other grating or holographic structures. This allows this waveguide architecture to be asymmetric, such as having a strong in-coupler as a simple prism: a strong in-coupler is always a challenge for any grating or holographic waveguide combiner architecture, and a macroscopic prism is the best coupler one can think of.

Figure 44 shows both thin and thick volume holograms operating in reflection and/or transmission modes. The top part of the figure shows a typical 1D EPE expander with a single transmission volume hologram sandwiched in between two plates. When the field traverses the hologram downwards, it is in off/Bragg condition, and when it traverses the volume hologram upwards after a TIR reflection, it is in an on/Bragg condition, or close to it, therefore creating a weak (or strong) diffracted beam which breaks the TIR condition.

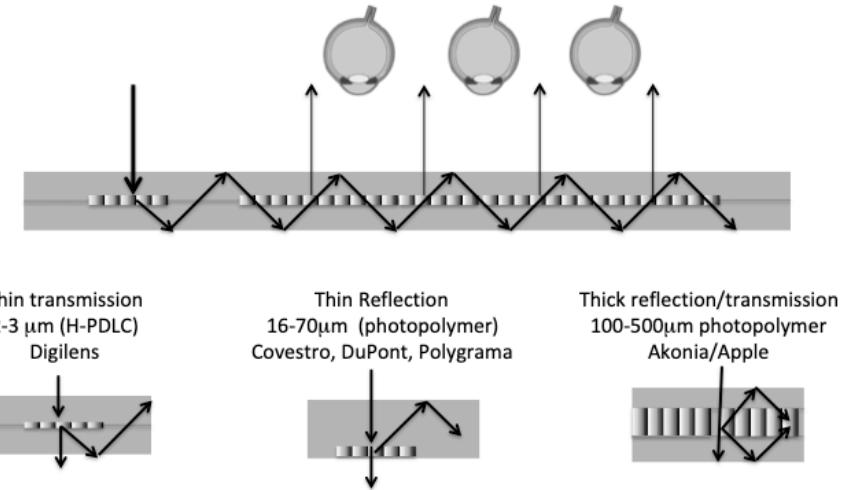


Figure 44: Different types of volume holograms acting as in and out couplers in waveguide combiners.

A hologram sandwiched in between plates might look more complex to produce than a reflective or transmission laminated version, but has the advantage that it can operate in both transmission and reflection modes at the same time, (for example to increase the pupil replication diversity).

7) Surface relief gratings couplers

Figure 45 reviews the various surface relief gratings used in industry today (blazed, slanted, binary, multilevel and analog), and how they can be integrated in waveguide combiners as in-coupling and out-coupling elements.

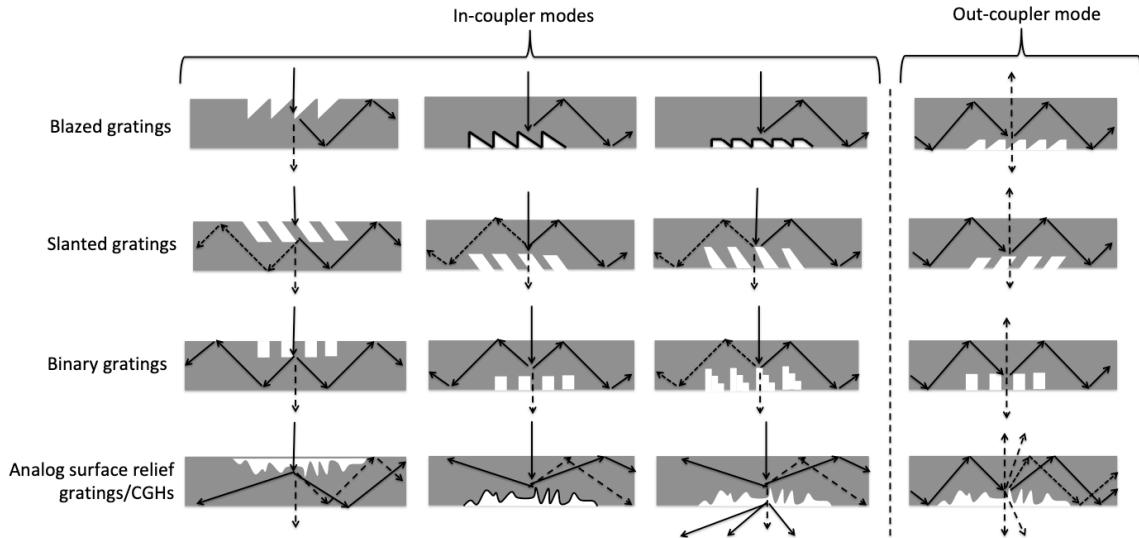


Figure 45: Surface relief grating types used as waveguide combiner in-couplers and out-coupler. Notes: a black surface line indicates a reflective coating on the grating surface. A long dashed line indicates the diffracted orders (that can be modulated in efficiency to produce a uniform eyebox). A short dashed line indicates parasitic orders difficult to suppress, creating either ghosts or stray light.

Covering a surface relief grating with a reflective metallic surface (see Figure 45) will increase dramatically its efficiency in reflection mode. A transparent grating (no coating) can also work both in transmission and reflection modes, especially as an out-coupler, in which the field has a strong incident angle.

Increasing the number of phase levels from binary to quarternary or even 8 or 16 levels increases its efficiency as predicted by the scalar diffraction theory, for normal incidence. However, for strong incidence angle and for small periods, this is not true anymore. A strong out-coupling can thus be produced, either in reflection or transmission modes.

Slanted gratings are very versatile elements, their spectral and angular bandwidths can be tuned by the slant angles. Front and back slant angles in a same period (or from period to period) can be carefully tuned to achieve the desired angular and spectral operation.

Surface relief gratings have been used as a commodity technology as soon as mastering and mass replication techniques technologies have been established and made available since the early 90s [88]. Typical periods for TIR grating couplers in the visible spectrum are below 500nm, yielding nanostructures of just a few tens of nanometers if multilevel structures are required. This can be achieved by either direct e-beam write, i-line (or DUV) lithography or even interference lithography (holographic resist exposure) [86]. Surface relief grating structures can be replicated in mass by nano-inprint, a micro-lithography wafer fabrication technology developed originally for the IC industry [89]. Going from wafer scale fabrication to panel scale fabrication will allow a reduction in costs allowing for consumer level grade AR and MR products.

Figure 46 and Figure 47 describe how the various surface relief gratings described in Figure 45 have been applied to the latest waveguide combiners such as Microsoft HoloLens V1 and Magic Leap One. Multilevel surface relief gratings have been used by other companies such as Dispelix Oy, and quasi analog surface relief CGHs (Computer Generated Holograms) by others such as EnhancedWorld Ltd. (formerly WaveOptics Ltd.).

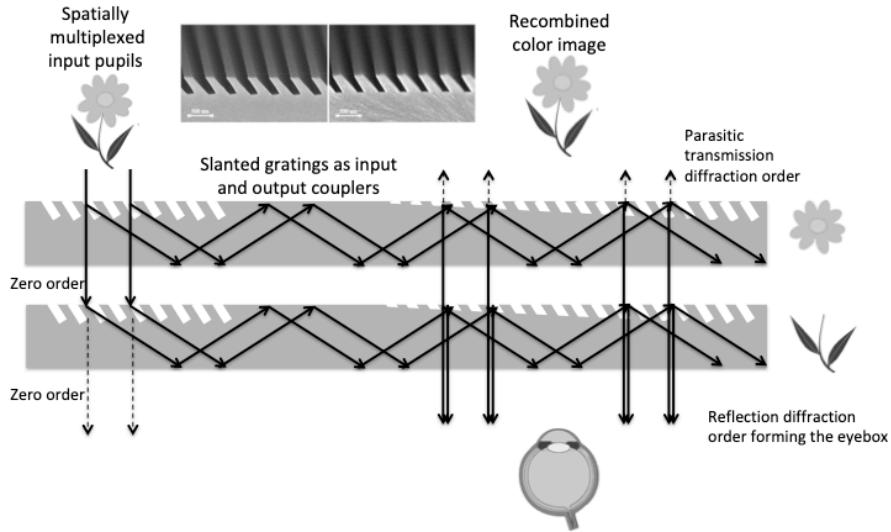


Figure 46: Spatially color-multiplexed input pupils with slanted gratings as in- and out-couplers working both in transmission and reflection modes (HoloLens V1 MR headset).

Figure 46 describes the waveguide combiner architecture used in Microsoft HoloLens V1 MR headset (2015). The display engine is located on the opposite side of the eyebox. The single input pupil carries the entire image over the various colors at infinity (here only two colors and the central field depicted for sake of clarity), as in a conventional digital projector architecture. The in-couplers have been chosen to be slanted gratings for their ability to act on a specific spectral range while letting the remaining spectrum unaffected in the zero order, to be processed by the next in-coupler area located on the guide below, and this for all three colors. Such uncoated slanted gratings work both in transmission and reflection modes, but can be optimized to work more efficiently in a specific mode. The out-couplers are here also slanted gratings, which can be tuned to effectively work over a specific incoming angular range (TIR range), and leave the see-through field quasi-unaffected. The part of the see-through field that is indeed diffracted by the out-couplers is trapped by TIR and does not make it to the eyebox. These gratings are also modulated in depth to provide a uniform eyebox to the user. Note the symmetric in- and out-coupler configuration compensating the spectral spread over the 3 LEDs bands.

The redirection gratings are not shown here. Input and output grating slants are set close to 45 degrees, and redirection grating slants at half this angle. The periods of the gratings are tuned in each guide to produce the right TIR angle for the entire FOV for that specific color (thus, same central diffraction angle in each guide for each RGB LED color band).

Figure 47 depicts the waveguide combiner architecture used in Magic Leap One MR headset (2018). The display engine is here located on the same side as the eyebox. The input pupils are here spatially color-demultiplexed, carrying the entire FOV at infinity (here again, only two colors and the central field depicted for sake of clarity).

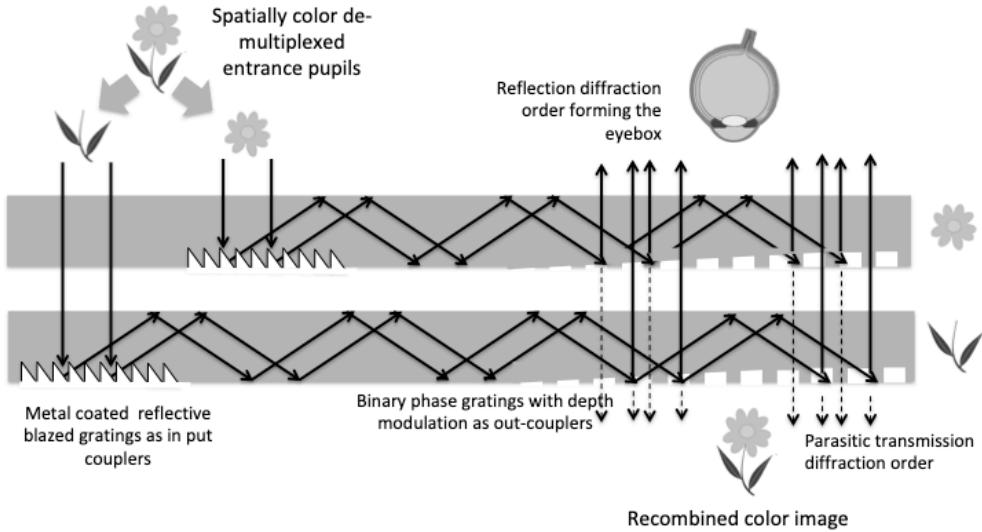


Figure 47: Spatially color-de-multiplexed input pupils with 100% reflective blazed gratings as in-couplers and binary phase gratings as out-couplers (Magic Leap One MR headset).

Spatial color de-multiplexing can be done conveniently with a color sequence LCOS display mode for which the illumination LEDs are also spatially de-multiplexed. In this configuration, the input grating couplers are strong blazed gratings, coated with a reflective metal (such as Al). They do not need to work over a specific single color spectral width since the colors are already de-multiplexed. The out-couplers are here simple top-down binary gratings, which are also depth modulated in order to produce a uniform eyebox for the user. These binary gratings are shallow, acting therefore very little on the see through, but have much stronger efficiency when working in internal reflection diffraction mode, since the optical path length in this case is longer by a factor of $2n\cos(\alpha)$ than from the transmission mode, (where n is the index of the guide and α the angle of incidence in the guide). As in HoloLens V1, most of the see through field diffracted by the out-couplers is trapped by TIR.

The redirection gratings (not shown here) are also composed of binary top-down structures. The periods of the gratings are tuned in each guide to produce the right TIR angle for the entire FOV for that specific color (same central diffraction angles for each RGB LED color band).

Other companies such as Enhanced-World Ltd in the UK (Formerly Wave-Optics Ltd) uses multilevel and/or quasi analog surface relief diffractive structures to implement in- and out-couplers (see also Figure 43). This choice is mainly driven by the complexity of the extraction gratings, acting both as redirection gratings and out-coupler gratings, making them therefore more complex than linear or slightly curved (powered) gratings, and more similar to iteratively optimized Computer Generated Holograms (CGHs) [89]. Allowing multilevel or quasi analog surface relief diffractive structures increases the Space Bandwidth Product of the element to allow more complex optical functionalities to be encoded with relatively high efficiency.

8) Resonant waveguide grating couplers

Resonant Waveguide Gratings (RWGs), also known as Guided Mode Resonant (GMR) gratings or waveguide-mode resonant gratings [90], are dielectric structures where these resonant diffractive elements benefit from lateral leaky guided modes. A broad range of optical effects are obtained using RWGs such as waveguide coupling, filtering, focusing, field enhancement and nonlinear effects, magneto-optical Kerr effect, or electromagnetically induced transparency. Thanks to their high degree of optical tuning (wavelength, phase, polarization, intensity) and the variety of fabrication processes and materials available, RWGs have been implemented in a broad scope of applications in research and industry. RWGs can therefore also be applied as in- and out-couplers for waveguide gratings [91].

Figure 48 shows a RWG on top of a lightguide (referred often incorrectly through the popular AR lingo as a “waveguide”), acting as the in- and out-couplers.

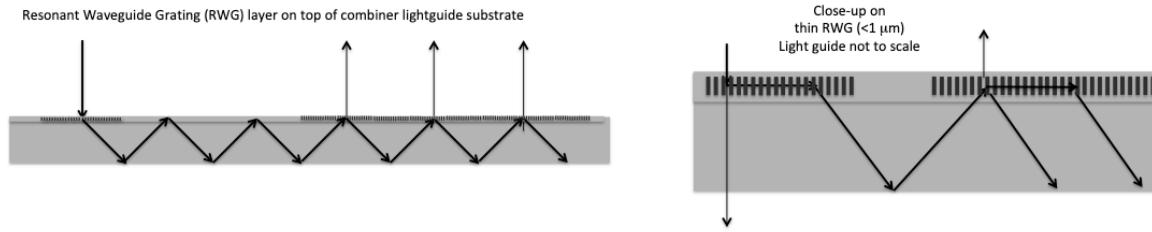


Figure 48: Resonant Waveguide Gratings as in- and out-couplers on a waveguide combiner

Roll to roll replication of such gratings structures can help bring down overall waveguide combiner costs. The CSEM research center in Switzerland developed the RWG concept back in the 80s, companies are now actively developing such technologies [91].

9) Metasurface couplers

Couplers based on metasurfaces have been recently investigated to implement new types of flat optical elements, operating in either free space or waveguide space [92] and well as in industry (Metalenz Corp). Replicating however the functionality of a flat diffractive structure with novel metasurface nanostructures have very limited benefits. Implementing novel optical functionality based on complex diffractive or refractive effects, which cannot be implemented with their traditional counterparts diffractive or refractive elements [93], [94], are more interesting prospects. Such prospects include phase multiplexing of different functionalities and/or angular selective coupling functionality.

Metasurface nanostructures can be produced the same way as traditional nanostructures, via lithographic mastering and nano-imprint in high index organic or inorganic materials, with or without additional layers. The critical dimensions of such structures are however close to an order of magnitude smaller than their counterparts diffractives discussed in previous sections. A fabrication advantage of meta-surfaces is that they can implement quasi-analog or multilevel diffractive elements effects with a simple (although smaller) binary structures, requiring only one lithography process and no field alignment steps, rather than multiples lithography and field alignment steps in the case of conventional flat diffractive optics [89], where each lithography step might introduce systematic and/or random fabrication errors.

Table 3 summarizes the various waveguide coupler technologies reviewed here, along with their specifics and limitations.

Waveguide coupler technology	Operation	Reflective coupling	Transmission coupling	Efficiency modulation	Lensed out-coupler option	Spectral dispersion, LCA	Color uniformity over FOV eybox	Dynamically tunable	Polarization maintaining	Mass production	Company/Product
Embedded mirrors	Reflective	Yes	No	Complex coatings	No	Minimal	Good	No	Yes	Slicing, coating, polishing,	Lumus Ltd. DK SD
Micro-prisms	Reflective	Yes	No	Coatings	No	Minimal	Good	No	Yes	Injection molding	Optinvent Sarl. ORA
Surface relief slanted grating	Diffractive	Yes	Yes	Depth, Duty cycle, slant	Yes	Strong	Needs compensation	Possible with LC	No	NIL (wafer, plate)	Microsoft Hololens, Vuzix Inc, Nokia Oy
Surface relief blazed grating	Diffractive	Yes	No	Depth	No	Strong	Needs compensation	Possible with LC	No	NIL (wafer, plate)	Magic Leap One
Surface relief binary grating	Diffractive	Yes	Yes	Depth, Duty cycle	Yes	Strong	Needs compensation	Possible with LC	No	NIL (wafer, plate)	Magic Leap One
Multilevel/analog surface relief grating	Diffractive	Yes	Yes	Depth, Duty cycle	Yes	Strong	Needs compensation	Possible with LC	Possible, but difficult	NIL (wafer, plate)	WaveOptics Ltd (Phox), BAE, Dispela Oy
Thin volume photopolymer hologram	Diffractive	Yes	Yes	Index swing	Yes, but difficult	Strong	Needs compensation	Possible with shear	No	NIL (wafer, plate)	Sony Ltd, TruelifeOptics Ltd,
H-PDL volume holographic	Diffractive	No	Yes	Index swing	Yes, but difficult	Strong	OK	Yes [electrical][]	No	Exposure	Digilens Corp. (Monosh HUD, AeroHUD)
Thick photopolymer hologram	Diffractive	Yes	Yes	Index swing	Yes, but difficult	Minimal	OK	No	No	Multiple exposure	Akonia Corp (now Apple Inc.)
Resonant Waveguide Grating (RWG)	Diffractive	Yes	Yes	Depth, Duty cycle	Yes	Can be mitigated	NA	Possible with LC	Possible	Roll to roll NIL	CSEM / Resonant Screens
Metasurface coupler	Diffractive	Yes	Yes	Various	Yes	Can be mitigated	Needs compensation	Possible with LC	Possible	NIL (wafer, plate)	Metalenz Corp.

Table 3: Benchmark of various waveguide coupler technologies.

Although Table 3 shows a zoo of optical couplers candidate, today most of the AR/MR/smart glass products are based on only a handful of traditional technologies as thin volume holograms, slanted surface relief gratings and embedded half tone mirrors.

Modeling optical waveguide combiners

Modeling of the angular and spectral Bragg selectivity of volume holograms, thin or thick, in reflection and transmission modes, can be performed with the 1969 Kogelnik theory [80], [81].

Similarly, modeling of the efficiency of surface relief gratings can be performed accurately with RCWA (Rigorous Coupled Wave Analysis) [82], [83], especially the FMM (Fourier Modal Method). The FDTD (Finite Difference Time Domain), also a rigorous electromagnetic (EM) nanostructure modeling method, can in many cases be a more accurate modeling technique but is also much more heavy and CPU time consuming. However, FDTD will show all the diffracted fields, the polarization conversions and the entire complex field, while the Kogelnik model and the RCWA will only give efficiency values in particular diffraction orders.

FDTD can model non periodic nanostructures, while RCWA can accurately model quasi periodic structures. Thus, FDTD might help in modeling k-vector variations (rolled k-vector) along the grating, slant, depths and duty cycle variations, as well as random and systematic fabrication errors in the mastering and replication steps.

Free versions of the RCWA-FMM [84] and FDTD [85] codes can be found on the net. Kogelnik theory can be easily implemented as a straightforward equation set for transmission and reflection modes. Commercial software suites include R-Soft from Synopsys and Lumerical.

These models predict the efficiency in each order for a single interaction of the light with the coupler element. In order to model the entire waveguide combiner, especially when a pupil replication scheme is used, RCWA or Kogelnik need to be used as DLLs in conventional ray tracing optical design software such as Zemax or more specific light propagation software such as the ones provided by LightTrans, Germany, [86] (see also Figure 49).

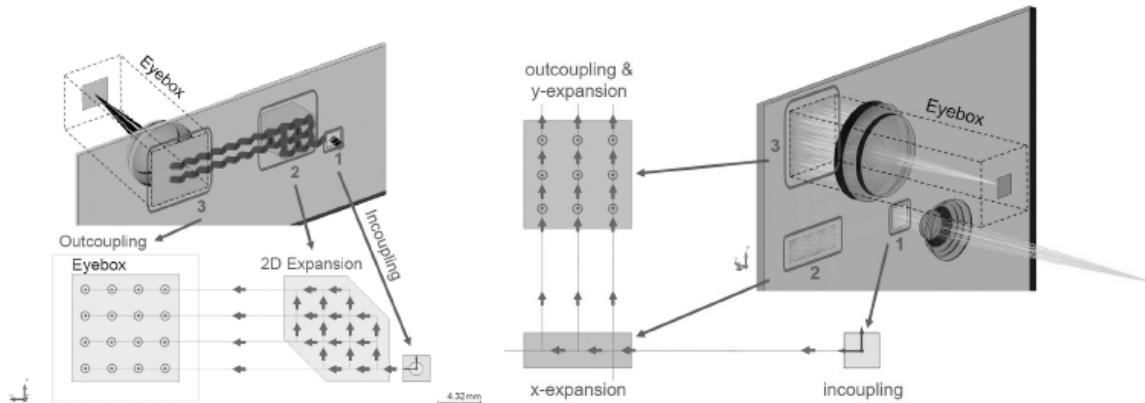


Figure 49: Waveguide grating combiner modeling by LightTrans (Germany), as 1D EPE (left) or 2D EPE (right).

The interaction of the EM field with the coupler regions (surface relief structures or index modulations) modeled through RCWA or Kogelnik can be implemented via a DLL (Dynamically Linked Library) in a conventional optical design software based on ray tracing (as C or Matlab code). As the FDTD numerical algorithm propagates the entire complex field rather than predicting only efficiency values (as in RCWA or Kogelnik model), it is therefore more difficult to implement as a DLL.

Ray trace based optimization of the waveguide combiner architecture with accurate EM-light/coupler interactions models are critical to design a combiner with good color uniformity over the FOV, uniform eyebox and high efficiency. Inverse propagation from the eyebox to the optical engine exit pupil is a good way to simplify the optimization process. The design process can also use an iterative algorithm to optimize color over FOV/eyebox or efficiency, or even reducing the real estate of the grating areas by making sure that no light gets lost outside the effective eyebox (see also Figure 60).

Waveguide couplers have their specific angular and spectral bandwidths, which are affecting both the FOV and the eyebox uniformity. A typical breakdown of the effects of a 2D EPE waveguide architecture is shown in Figure 50:

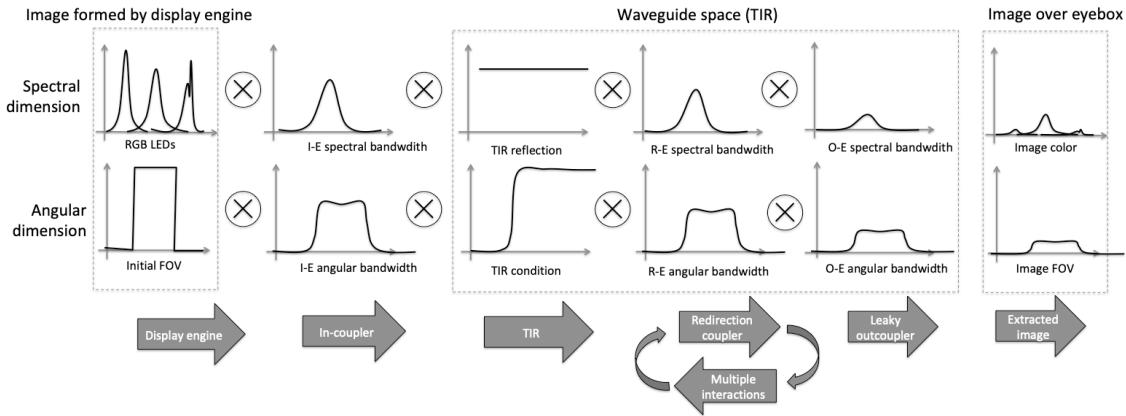


Figure 50: Cascaded effects of the field/coupler interactions on the FOV uniformity.

As one can see from Figure 50, the coupler's spectral and angular bandwidths are critical to the FOV uniformity, especially color uniformity. While embedded mirrors and micro-prisms have a quasi uniform effect on color and FOV, others do not, such as gratings and holograms. It is therefore interesting to have the flattest and widest bandwidths possible. For volume holograms, this means operating in reflection mode and having a strong index swing (Kogelnik), and for surface gratings, this means a high index (RCWA/FDTD). The position of the angular bandwidth can be tuned by the slant angle in both holograms and surface gratings. Multiplexing bandwidths can help to build a larger overall bandwidth, both spectral and angular, and is used in various implementations today. Such multiplexing can be done in phase, in space or in time, or a combination of all the above. Finally, as spectral and angular bandwidths are closely linked, impacting on spectral behavior has an impact on FOV and vice versa.

Polarization and degree of coherence are two other dimensions one should need to investigate especially when lasers or VCSELs are used in the optical engine or if polarization maintaining (or rather polarization conversion) is required. The multiple interactions in the R-E region can produce multiple miniature Mach-Zehnder interferometers, which might modulate the intensity of the particular fields.

Choosing the waveguide coupler layout architecture

We have seen that couplers can work in either transmission or reflection modes to create a more diverse exit pupil replication scheme (producing a more uniform eyebox), or to increase the compactness of the waveguide by using both surfaces. The various couplers might direct the field in a single direction or in two directions, or even more, increasing potentially the FOV that can propagate in the waveguide without increasing necessarily its index.

Figure 51 shows how the optical designer can expand the functionality of in- or out-couplers, with architectures ranging from bi-dimensional coupling to dual reflective/transmission operation in the same guide with sandwiched volume holograms or top/down grating couplers.

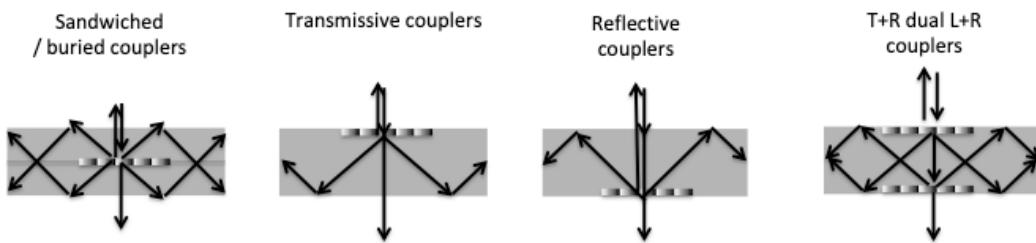


Figure 51: More functional coupler architectures yielding compact and efficient waveguide combiners.

When using more complex and more functional coupler architectures, the optical designer has however to keep in mind that each have their specific effects on MTF, efficiency, color uniformity and FOV.

For example, while the index of the guide allows for a larger FOV to propagate, the index of the grating structures in air would increase the spectral and angular bandwidths in order to process a larger FOV without compromising color uniformity or efficiency. The waviness of the waveguide itself will impact the MTF as random cylindrical powers added to the field. Multiple stacked waveguides might be efficient as processing single colors, but their misalignment will impact the MTF as misaligned color frames. Similarly, hybrid top / bottom couplers will affect the MTF if they are not perfectly aligned (angular alignment within a few arc seconds).

Building a uniform eyebox.

As the TIR field gets depleted when the image gets extracted along the out-coupler region, to produce a uniform eyebox, the extraction efficiency of the out-coupler needs then to gradually increase in the propagation direction. This of course complicates the fabrication process of the couplers, especially when the gradual increase in efficiency needs to happen in both pupil replication directions.

For volume holograms, the efficiency can be increased by a stronger index swing in the photopolymer or PDLC (through a longer exposure or a thickness modulation). For surface relief gratings, the optical designer has a few options to do so, as shown in Figure 52. This is true for the redirection grating (R-E) as well as the outcoupler (O-E).

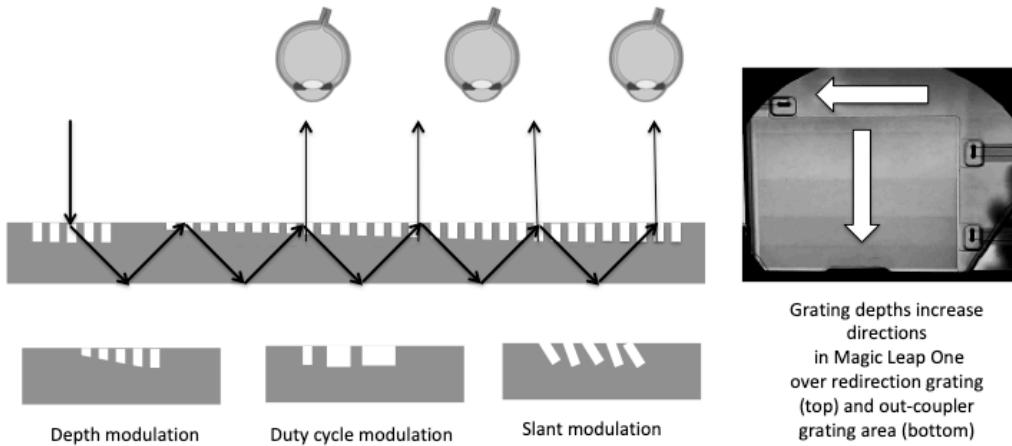


Figure 52: Modulation of the extraction efficiency along the grating out-couplers to build a uniform eyebox

Groove depth and duty cycle modulation can be performed on all type of gratings, binary, multilevel, blazed and slanted (see also Figure 45). Duty cycle modulation has the advantage of modulating only the lateral structures, not the depth, making it an easier mastering process. Modulating the depth of the gratings can be done in binary steps (as in Figure 52 – Magic Leao One), or in a continuous way (Digilens waveguide combiners).

Slanted grating angle modulation (both in a single grating period – such as back and front slants – or over a larger grating interaction length can change the angular and spectral bandwidths to modulate efficiency and other aspects of the coupling (angular, spectral, polarization). Periodical modulation of the slant angles is sometimes also called the “Rolling K vector” technique, and can allow for larger FOV processing due to specific angular bandwidth management over the grating area. Once the master has been fabricated with the correct nanostructure modulation, the Nano-Imprint Lithography (NIL) replication process of the gratings is the same no matter how complicated the nanostructures might be, with one word of caution for slanted gratings where the NIL process needs to resolve the undercut structures. The slanted grating NIL process is mastered today by many foundries [86].

Spectral spread compensation in diffractive waveguide combiners.

Spectral spread comes to mind as soon as one speaks about gratings or holographic elements. This is especially critical when the display illumination is broadband such as with LEDs (as in most of the waveguide grating combiner devices today – Digilens, Nokia, HoloLens, Vuzix, Magic Leap,...). The straightforward technique to compensate for the inevitable spectral spread is to use a symmetric in-coupler / out-coupler configuration, in which the gratings or holograms work in opposite direction, thus compensating in the out-coupler any spectral spread impacted in the in-coupler (Figure 53).

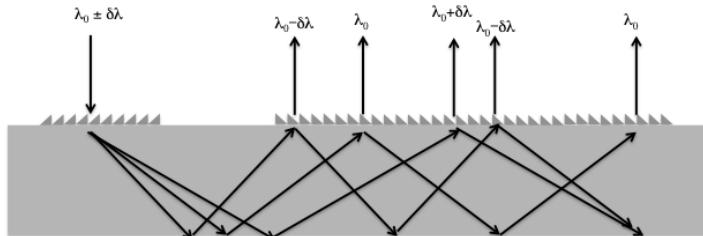


Figure 53: Spectral spread compensation in symmetric in-coupler / out-coupler waveguide combiner.

Although the spectral spread might be compensated, one can notice on Figure 53 that the individual spectral bands are

spatially de-multiplexed at the exit ports while multiplexed at the entry port. A strong exit pupil replication diversity is thus required to smooth out any color non uniformities generated over the eyebox.

This symmetric technique might not be used to compensate for spectral spread across different colors (R,G,B LEDs), but rather for the spread around a single LED color. Spread across colors might stretch the RGB exit pupils too far apart, and also reduce the FOV over which all RGB colors can propagate by TIR.

Field spread in waveguide combiners.

The different fields propagating by TIR down the guide are also spread out, no matter what the coupler technology is (mirrors, prisms, gratings, holograms,...) – see Figure 54.

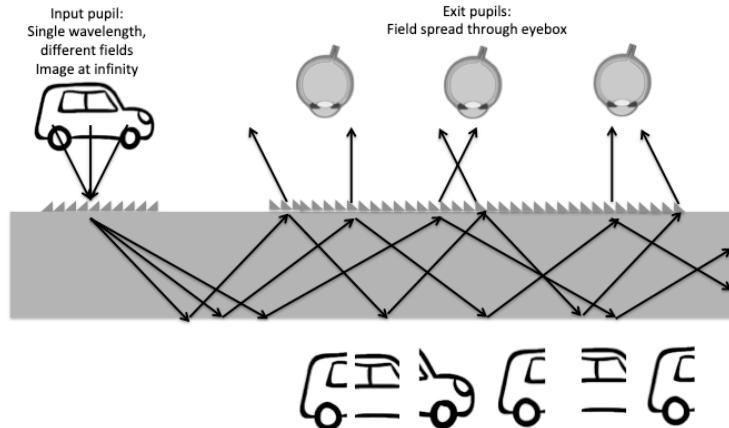


Figure 54: Field spread in a waveguide combiner.

One can however form again a uniform FOV (i.e. all fields appearing) over the eyebox with a strong exit pupil diversity scheme. This is a concept often misunderstood as in many cases only one field is represented when schematizing a waveguide combiner. Figure 54 shows the field spread occurring in a diffractive waveguide combiner. The number of replicated fields is also contingent on the size of the human eye pupil. If the ambient light gets bright, thus the human eye pupil getting small, only part of the FOV might appear to the user, missing a few fields. This is very similar to the eyebox reduction effect discussed in section 8.

Focus spread in waveguide combiners.

When a pupil replication scheme is used in a waveguide combiner, no matter what the coupler is, the input pupil needs to be formed over a collimated field (image at infinity / far field). If the focus is set to the near field instead of far field in the display engine, each waveguide exit pupil will produce an image at a slightly different distance, producing therefore a very mixed visual experience, overlapping the same image with different focus depths. It is quasi impossible to compensate for such focus shift over the exit pupils because of both spectral spread and field spread over the exit pupils, as discussed in the previous paragraphs. Figure 55 shows such a focus spread over the eyebox from an input pupil over which the image is formed in the near field.

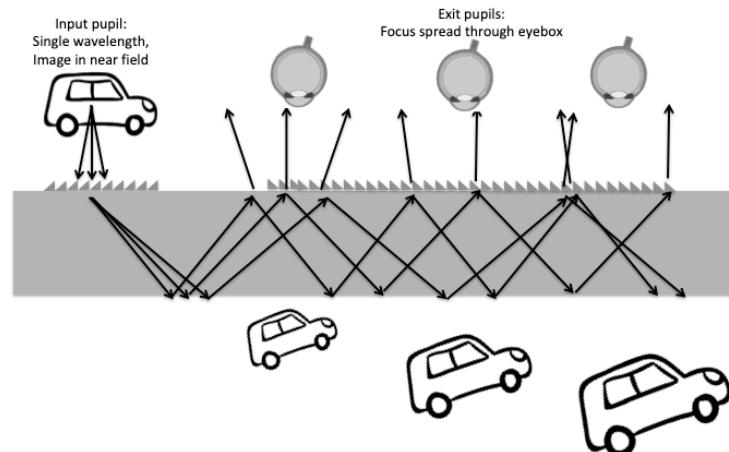


Figure 55: Focus spread in a waveguide combiner with a non collimated input image pupil.

The image over the input pupil can however be located in the near field when there no pupil replication scheme is performed in the guide, such as in the Epson Moverio BT300 or in the Zeiss Smart Glass (yielding small FOV and small eyebox).

When pupil replication is used in the guide, the virtual image can also be set at a closer distance for better visual comfort by using a static (or even tunable) negative lens acting over the entire eyebox. For an unperturbed see through experience, such lens needs to be compensated by its conjugate placed on the world side of the combiner waveguide. This is the architecture used in Microsoft HoloLens V1 (2015) [37].

Another way, more compact, would be to introduce a slight optical power in the output coupler (O-E), so that this coupler takes the functionality of an off-axis lens (or an off-axis diffractive lens) rather than that of a simple linear grating extractor or linear mirror/prism array. Although this is difficult to implement with a mirror array (as in an LOE), it is fairly easy to implement with a grating or holographic coupler. The grating lens power does not affect the zero diffraction order which travels by TIR down the guide, but affects only the out-coupled (or diffracted) field. The see-through field is also not affected by such a lensed out-coupler since the see-through field diffracted by such element would be trapped by TIR and will thus not enter the eye pupil of the user.

All three configurations (no lens for image at infinity, static lens with its compensator, and powered O-E grating) are shown in Figure 56. The left part of the eyebox shows an extracted field with image at infinity (as in Lumus DK40 - 2016), the center part shows an extracted field with image at infinity which passes through a negative lens to form a virtual image closer to the user and its counterpart positive lens to compensate for see through (as in Microsoft HoloLens V1, 2015), and the right part of the eyebox shows an extracted field with the image directly located in the near field through a powered grating extractor (as with an off-axis diffractive lens as in Magic Leap One, 2018).

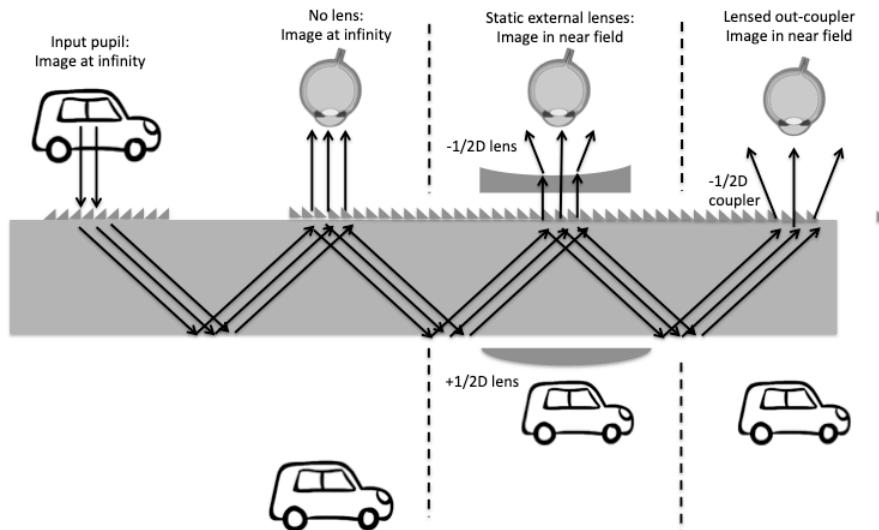


Figure 56. Three different out-coupler architectures positioning the virtual image in the same near field over all exit pupils.

For example, a half diopter negative lens power would position the far field image to a more comfortable 2m distance, uniformly over the entire eyebox.

Having a powered out-coupler grating might however reduce the MTF of image especially in the direction of the lens offset (direction of TIR propagation), since the input (I-E) and output couplers (O-E) are no more perfectly symmetric (the input coupler being a linear grating in both cases and the outcoupler an off axis diffractive lens). Thus, the spectral spread of the image in each color band cannot be compensated perfectly anymore and will produce LCA in the direction of the lens offset. This can be critical when using an LED as a illumination source, but would affect the MTF much less when using a narrower spectral sources such as lasers or VCSELs.

One of the main problems with such a lensed outcoupler grating configuration is when one attempts to propagate two colors in the same guide (for example 2-guides RGB waveguide architecture as in Figure 63), generating longitudinal chromatic aberrations (focus changing with color since the lens is diffractive). Using a single color per guide and a laser source can greatly simplify the design task.

Polarization conversion in diffractive waveguide combiners.

Polarization conversion can be a problem when using diffractive or holographic couplers, since these are often optimized to work best for a single polarization, usually "s" (orthogonal to the grating lines). Polarization conversion might occur in

the guide through diffraction, and reduce the overall efficiency by producing more light in weaker "p" polarization, which would interact less with the gratings or holograms. The light engine can easily produce polarized fields, such as with an LCOS or a laser scanner. Another downside of having a mixed polarization over the eyebox is that the use of polarization optics is not possible anymore (such as tunable liquid crystal lenses for VAC mitigation, see also section 12). Mirrors or micro-prism based couplers maintain the polarization state better than grating or holographic based couplers.

Note that polarization conversion can also have a benefit, allowing the in-coupled field to interact again with the in-coupler grating or holographic region without getting strongly out-coupled by the same in-coupler due to the time reversal principle: this allows for thinner waveguides while maintaining a large input pupil size (see for example Figure 61).

Propagating full color images in the waveguide combiner over a maximum FOV.

We have seen in the previous paragraphs that the spectral spread of grating and holographic couplers can be perfectly compensated with a symmetric in- / out-coupler configuration. This is possible over a single color band but will reduce considerably the FOV if used over the various color bands (assuming that the couplers will work over these various spectral bands).

In order to maximize the RGB FOV in a waveguide combiner, one solution is to use stacked guides optimized each for a single color band, each coupling a maximum FOV by tuning the diffraction angle of the in- and out-couplers accordingly. This is the architecture used in both HoloLens V1 and Magic Leap One (see Figure 57), although the position of the input pupil (light engine) is opposite in both devices.

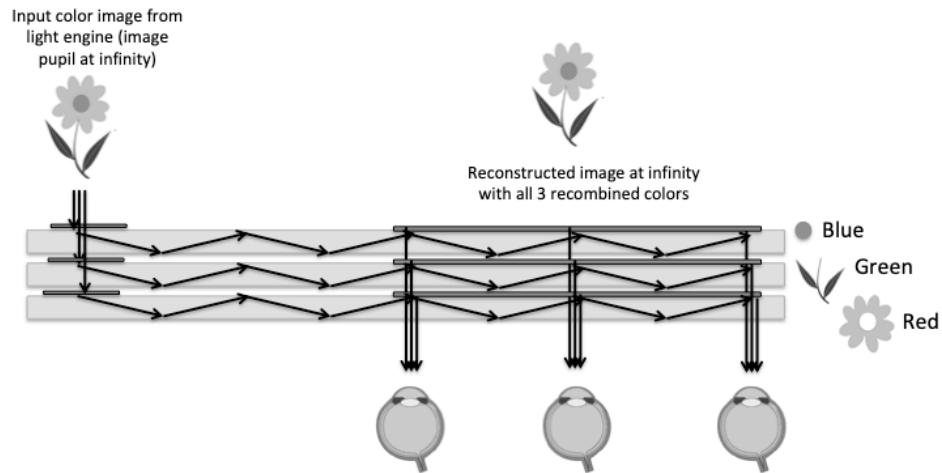


Figure 57: Stacked waveguides combiners providing largest FOV TIR propagation over 3 colors.

Air gaps in between all plates are required to produce the TIR condition. Such gaps also allow for additional potential filtering in between plates for enhanced performance (such as spectral and polarization filtering).

Figure 58 Shows the functional diagram of such a single color plate as a top view as well as its k-vector space depiction [63], [64]. Here again, I-E (as in Incoupling Element) refers to the in-coupler, R-E (as in Redirection Element) refers to the leaky 90 degree redirection element and O-E (as in Outcoupler Element) refers to the leaky out-coupler forming the final eyebox (for 2D pupil replications).

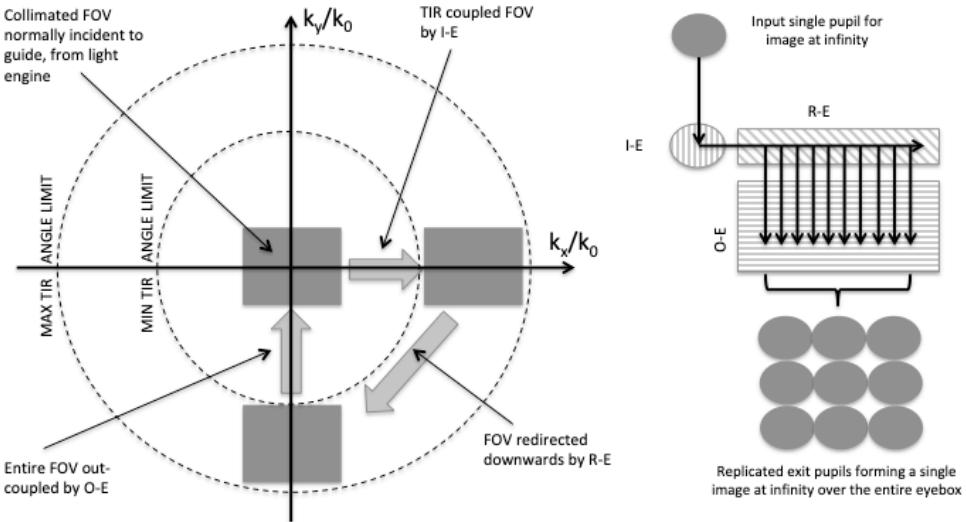


Figure 58: k-vector diagram and lateral pupil replication layout for single guide and single color

Note that the entire FOV is shown on the K vector diagram (left, Figure 58), but only a single field (central pixel in the FOV, with entry normal to the guide) is shown on the eyebox expansion schematic (right, Figure 58). Refer also to previous paragraphs showing how spectral spread within a color and field spread complicates such a diagram.

The FOV in the direction of the in-coupling can be increased by a factor of two when using a symmetric in-coupling configuration [94], in which the input grating or hologram (or even prism(s)) would attempt to couple the entire FOV to both sides, with one of the input configurations described in Figure 59.

As the TIR angular range does not support such an enlarged FOV, part of the FOV is coupled to the right and part of the FOV is coupled to the left. Due to the opposite directions, opposite sides of the FOV travel in each direction. If such TIR fields are then joined back with a single out-coupler one can reconstruct the original FOV by overlapping both partial FOVs, as in Figure 59, left and center.

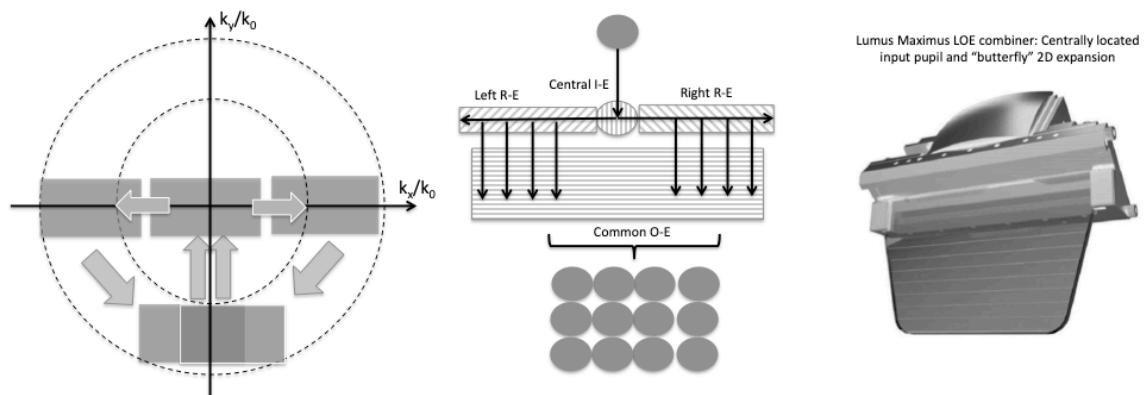


Figure 59: Symmetric in-coupling allowing for a FOV increase in the direction of in-coupling and LOE example.

In the orthogonal direction, the FOV that can be coupled by TIR remains unchanged. This concept can be taken to more than one dimension, but the coupler real estate on the waveguide can become quickly prohibitive.

Waveguide coupler geometries

We have reviewed the various coupler technologies that can be used in waveguide combiners, as well as the two dimensional exit pupil expansion that can be carried out in waveguide combiners. Waveguide combiners are desirable since their thickness is not impacted by the FOV (see section 7), unlike with other combiner architectures such as free space or TIR prisms (see section 9). However, the lateral dimensions of the waveguide (especially the redirection coupler and out-coupler areas over the waveguide) are closely linked to size of the in-coupled FOV, as shown in Figure 60. For example, the R-E region geometry is dictated by the FOV in the waveguide medium: it expands in the direction orthogonal to the TIR propagation, forming a cone geometry.

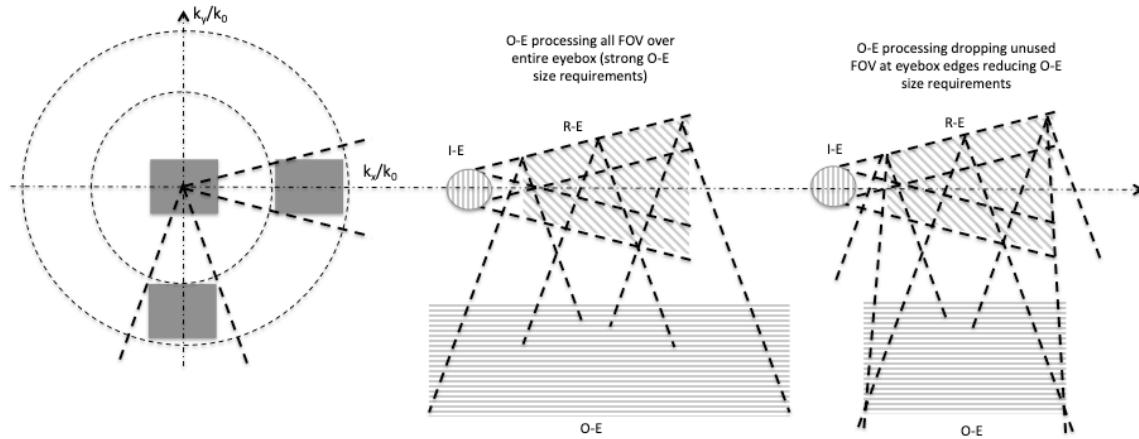


Figure 60: Redirection and out-coupler area requirements as dictated by the incoupled FOV.

The largest coupler area requirement is usually the outcoupler element (center), aiming at processing all FOVs and building up the entire eyebox. Eye relief is impacting this strongly also. However, its size can be reduced, in a “human centric optical design” approach, as we did for other aspects of the combiner design (see sections 6 and 7): the right part of the FOV at the left edge of the eyebox as well as the left part of the FOV at the right edge of the eyebox can be discarded thus reducing considerably the size of the O-E without compromising the image over the eyebox. Note that in Figure 60 the K-vector diagram (left) shows the FOV whereas the lateral schematics of the waveguide (center of right) show the actual size of the coupler regions.

Reducing the input pupil can help to reduce the overall size and thickness of the combiner. However, the thickness of the guide has to be large enough not to allow for second I-E interaction with the incoming pupil after the first TIR bounce. If there is a second interaction, by principle of time reversal, part of the light will be then out-coupled and form a partial pupil (partial moon if the input pupil is circular) propagating down the guide instead of the full one. This is more pronounced for the smallest field angle as depicted in Figure 61.

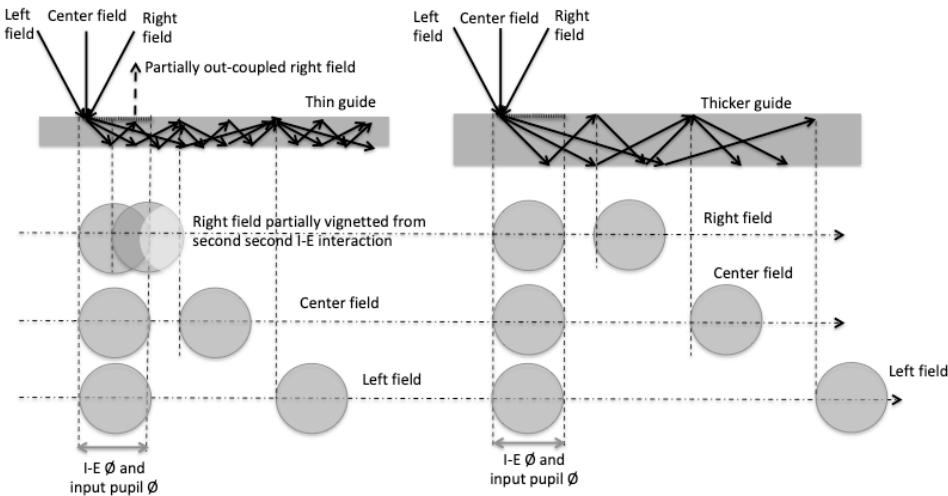


Figure 61: Effects of input pupil size (and size of I-E) or thickness of guide on TIR pupil.

However, if the polarization of the field is altered after the first TIR reflection at the bottom of the guide, the parasitic outcoupling can be reduced, if the I-E is made to be highly polarization sensitive.

Reducing the waveguide thickness is also interesting in order to produce a stronger pupil diversity over the eyebox, and thus a better eyebox uniformity. If reducing the guide is not an option (for parasitic out-coupling of the input pupil and also for étendue limitations in the display engine), one can use a semi transparent Fresnel surface inside the guide (as in two guides bounded together), which would reflect only part of the field, and leave the other part unperturbed, effectively increasing the exit pupil diversity.

Reducing the number of plates for full color display over the maximum allowed FOV.

Reducing the number of plates without altering the color of the image while propagating the maximum FOV allowed by the index of the guide is a desirable feature since it reduces the weight, size and complexity of the combiner, and make it also less prone to MTF reductions due to guide misalignments. Both lateral and longitudinal angular waveguide misalignments will contribute to a reduction the MTF built up by the display engine. Waveguide surface flatness issues are yet more causes for MTF reduction.

Due to the strong spectral spread of the in coupler elements (grating, holograms, RWGs or meta-surfaces), the individual colors fields are coupled at higher angles as the wavelength increases, thus reducing the overall FOV that can propagate in the guide within the TIR conditions (smallest angle dictated by the TIR condition and largest angle dictated by pupil replication requirements for uniform eyebox). This issue is best shown in the K vector diagram (see Figure 62). Note that this is not the case for embedded single prism in-couplers which yield a much lower spectral spread, therefore allowing single guide RGB waveguide combiner over near maximum FOV such as an LOE (embedded partial mirrors outcoupplers) from Lumus or the micro prism array from Optinvent.

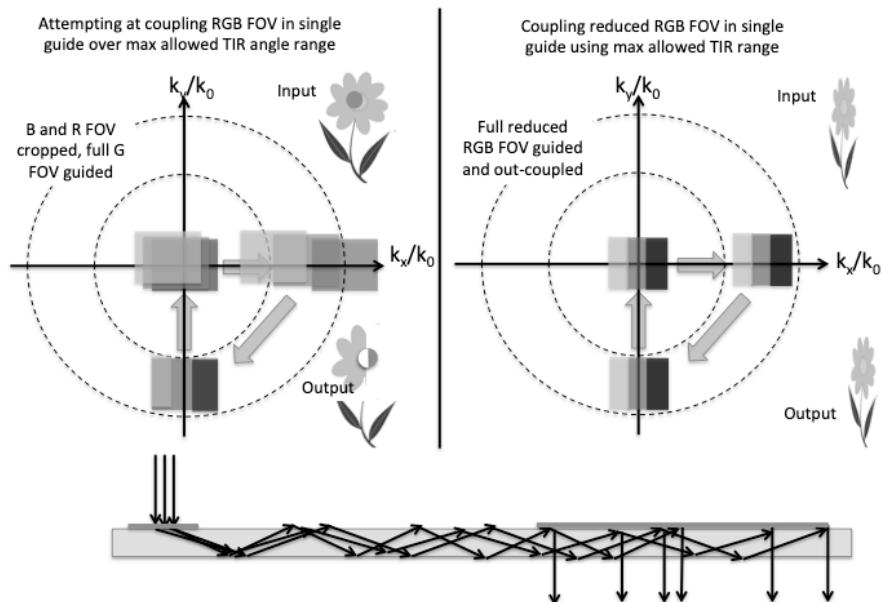


Figure 62: k-vector diagram of a single plate waveguide combiner using (left) RGB FOV coupling over a single color TIR angular range condition, and (right) RGB reduced FOV sharing the same TIR range.

The left configuration in Figure 62 acts as a hybrid spatial/spectral filter, filtering the left part of the Blue FOV, allowing the entire Green FOV to be propagated (if the grating coupler periods have been tuned to match the green wavelength), and filtering the right part of the Red FOV. The right configuration in Figure 62 propagates the entire RGB FOV (assuming the couplers can diffract uniformly over the entire spectrum), at the costs of the FOV extend in the direction of the propagation (for ex: Dispelix Oy).

Recently, two guide RGB combiner architectures have been investigated, reducing by 30% the weight and size of traditional 3 guide combiners, where the green FOV is shared between the top and bottom layer. Figure 63 shows such a waveguide combiner configuration. Various companies are using this 2 plate RGB waveguide combiner architecture today (Vuzix, EnhancedWorld, Digilens).

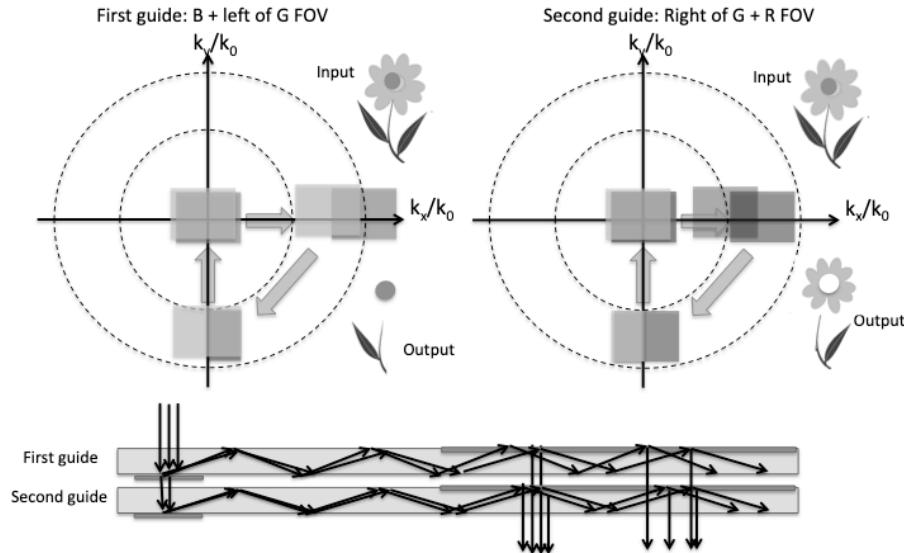


Figure 63: Two guides RGB waveguide combiner configuration.

This however requires the grating (or holograms, RWGs or metasurfaces) to be efficient over a larger spectral band. This implies for surface relief gratings to be replicated in a higher refractive index, widening their spectral (and angular) bandwidths. High index grating replication by NIL stretches the traditional wafer scale NIL resin material science (inclusion of TiO₂ or ZrO₂ nanofiller particles). Nano-imprint at Gen2 panel size of higher index inorganic spin-on glass material might be the best fit, solving also the resin or photopolymer reliability issues over various environmental conditions (temperature, pressure, shear, UV exposure and humidity).

This two guide RGB configuration splits the green FOV in two at the in-coupler region, and merges them again together over the out-coupler region. For a good color uniformity over the FOV and the eyebox, especially in the green field, this technique requires a perfect control of the two guides efficiency balance. Pre-emphasis compensation of the guide mismatch is possible using the display dynamic range, but this requires precise calibration, reduces the final color depth and does not solve the stitching region issue where the two fields overlap.

An alternative to the architecture is to use the first guide to propagate Green and Blue FOVs, and the second guide to propagate only the Red FOV, as Green and Blue are closer spectrally to each other than Red. This however reduced slightly the allowed FOV travelling without vignetting but solves the Green FOV stitching problem.

The holly grail would of course to design a single guide waveguide combiner with highly dispersive coupler elements such as gratings or holograms mass replicated by NIL. As there would be no multiple guides to align, by using accurate lithographic grating alignment on the same guide, this would yield the best possible MTF and the lowest costs.

One solution is to phase multiplex three different color couplers with three different periods into a single layer, and tune it so that there is no spectral overlap (no color ghost images over the eyebox). Such phase multiplexing is theoretically possible in volume holograms. This might be achieved in the Akonia (now Apple) thick holographic material (500 microns). If a thinner photopolymer (less than 20 microns) is desired for better reliability and easier mass production, a large holographic index swing is required. Standard photopolymers can be panchromatic, and can also be phase multiplexed, but the resulting efficiency still remains low, and color cross contamination between holograms are an additional issue. This is also theoretically possible with surface relief gratings, but having at the same time high efficiency and high extinction ratio over the 3 color bands is very difficult to achieve. Meta-surfaces and RWGs can theoretically also produce such phase multiplexed layers, however with the same limitations.

Another solution is to spatially interleave various grating configurations, by varying periods, depths and slant angles. This is however also difficult to achieve practically. Yet another solution to solve the single RGB guide problem would be to time multiplex RGB gratings through a switchable hologram such as the ones produced by Digilens Corp. This switching technique could also produce much larger FOVs multiplexed in the time domain and fused in the integration time of the human eye.

12. Vergence-Accommodation Conflict (VAC) Mitigation

Three dimensional display is a key feature and a perfect fit for immersive displays, especially for MR where the natural depth cues compete directly with the digital depth cues formed by the NTE display architecture. These two cues need to

agree with each other in order to provide a comfortable visual experience no matter where the hologram is placed over the reality, and no matter where the user decides to focus his visual field.

Visual depth cues are numerous, the most obvious ones are motion based parallax and dynamic occlusion [95], both of which can be easily implemented with a 6DOF head tracker (IMU + lateral cameras). Occlusion of the hologram by real objects requires a fast and accurate depth map scanner (see next section).

Binocular disparity presented to the eye by a stereo display is the other main visual depth cue. Stereo disparity can be rendered by most binocular NTE headsets today. Stereo disparity photographs in 3D stereoscopes was actually a worldwide consumer hit for the burgeoning photography technology towards the end of the 19th century (see Figure 64).



Figure 64: Late 19th century stereoscope (left) and stereoscopic photographic plates (right).

Stereo disparity induces an occulo-motor depth cue: the vergence of the eyes, measured in prism diopters, which is in turn a trigger to the accommodation of the eyes, measured in spherical diopters. In a HMD binocular stereo display, vergence of the eyes is triggered by a stereo disparity rendering. Occulo-motor vergence sets in at about 200 ms, and subsequent accommodation takes slightly longer, around 300ms. The unique relation (depending on IPD) between vergence and accommodation is summarized in Figure 65.

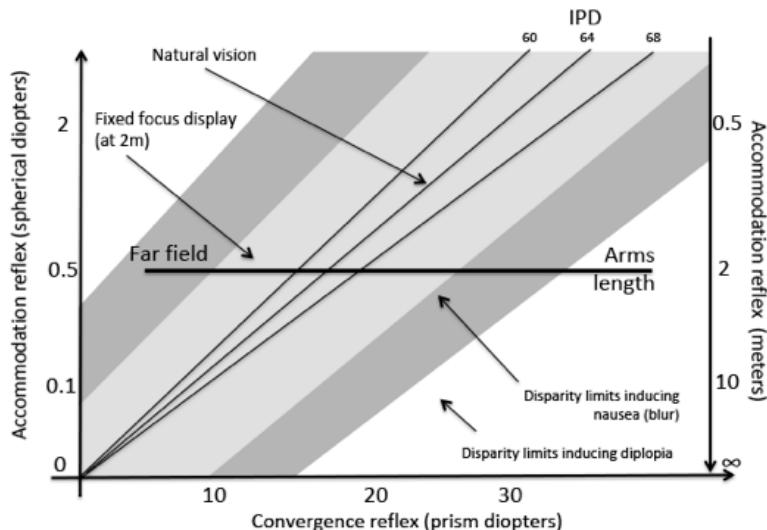


Figure 65: Vergence-accommodation diagram showing VAC mismatch with fixed focus stereo display

Both vergence and accommodation are closely linked together within the human visual system [96]. Unfortunately, most of the binocular HMDs presenting stereo disparity to the users rely on a fixed focus display, introducing the Vergence Accommodation Conflict (VAC) [97]. VAC can yield visual discomfort and reduce the quality of the 3D immersive visual experience. This is why in most headsets the stereo disparity is intentionally limited to present “holograms” over a set depth range, which would keep the VAC discomfort within acceptable limits (see the disparity limits inducing nausea, and the disparity limits inducing diplopia in the graph on Figure 65).

Although limiting stereo vision at a minimum distance to the eyes might be acceptable for VR and AR applications, it is not acceptable for MR applications, where the experience relies on the interaction between the digital hologram and reality, very often at arm's length. The VAC conflict produces the most visual discomfort (and eventually also nausea) when the vergence is set at arm's length while the display is still set at near infinity (far field starting at 1.5m). Arm's length display interaction is a key feature for MR headsets: VAC mitigation technologies and algorithms are thus starting to be investigated in various headsets, VR and AR/MR.

Based on numerous human visual perception and psychophysics studies [98], and also on smart marketing claims [99], it is interesting to note that solving the VAC conflict is perceived by the tech investment community as a crucial feature for next generation MR headsets, to the point of investing >\$2.4B in single start-ups differentiating themselves from others by attempting to solve that specific visual comfort issue.

We reviewed in the previous section that the display location can be set directly in the display engine for free space combiners, or by an additional lens covering the entire eyebox in waveguide combiner using EPE. This lens can also be integrated directly in the out-coupler grating (as in Magic Leap One), or as an external refractive negative lens with its positive compensator on the world side (as in HoloLens V1).

Focus tuning through display to lens movement

The most straightforward way to change the location of the virtual image in front of the user, especially in a VR system, is to change mechanically the distance between the display panel and the collimation lens. This has been investigated by the computational display group at Stanford University [100] and in the Oculus Half Dome prototype unveiled in 2018 by Facebook Reality Labs (see Fig. Figure 66).



Stanford (G. Wetzstein) varifocal prototype
with external motor based on Gear VR (2015)

Oculus Half Dome varifocal prototype
(D.Lanman) with integrated motors (2018)

Figure 66: Focus tuning of entire scheme by moving the display panel to lens, in VR systems.

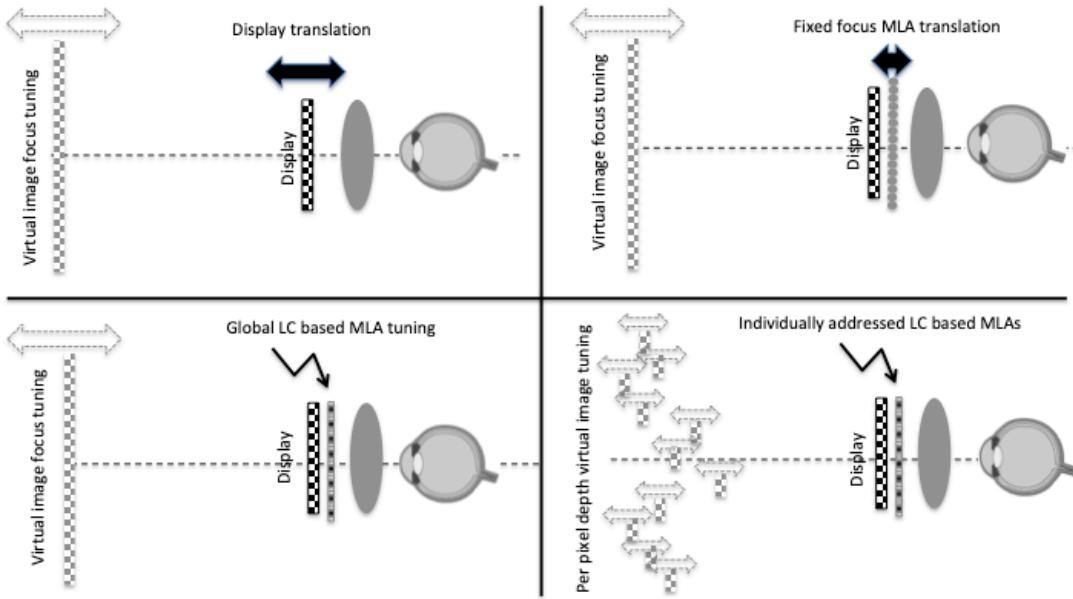
Focus tuning through the use of MicroLens Arrays (MLAs)

Another way to change the position of the virtual image in a VR system is to use an MLA in front of the display, and translate such MLA rather than translating a larger single lens or display. If the MLA is carefully aligned to the display, a very small movement of the MLA combined with a single fixed lens can produce a large focus position movement for the virtual image. This allows the use of fast and small actuators working over a limited motion range, pushing or pulling a thin MLA plate rather than moving a thicker display panel or bulkier lens over larger ranges to achieve the same focus change.

In other implementations, such an MLA can be dynamic, created by an array of Liquid Crystal flat micro-lenses, that can take on various power, therefore pushing or pulling the virtual image without any mechanical motion.

Yet in another implementation, the MLA can be addressed independently, producing a specific focus for each lens in the array, and thus providing the potential of developing a true “per pixel depth” display (or “per pixel-cluster depth”), which would produce true optical blur and be very close to a light field display experience. However, such tunable MLA arrays are yet to be developed by industry.

Figure 67 summarizes various focus tuning techniques based on display movement, MLA tuning or electronically addressable MLA arrays, to dynamically change the position of the entire virtual image, or only parts of the image, or even acting on individual pixel focus depths.

**Figure 67: Display translation and display MLA focus tuning**

These techniques are best suited for VR headset displays as well as for free-space based optical combiners in AR or MR headsets. They are however not suited for waveguide combiners using exit pupil replication for the reasons addressed in section 11.

Binary focus switch

Switching in a binary mode between a near field focus plane and a far field focus plane can help mitigate the VAC conflict and allow the user to interact with the hologram at arm's length for long periods of time. One of the first architectures to implement such was based on polarization switch: the display engine switches the image polarization between 's' and 'p' polarizations and a set of polarization sensitive reflectors change the distance between the display panel and the collimation lens, therefore changing the virtual image focus as per its polarization state. Thin LC polarization rotators are best to use in this case (over large angular and spectral operation).

A similar focus switch could be achieved in the spectral field, where each plane could be affected to a specific color, and tuned for a specific combiner lens power, such as in the Intel Vaunt example (see also Figure 36). In this particular example however, the spectral switch was used to induce an exit pupil move to enlarge the perceived eyebox. Spectral color switch can be done in the illumination engine within same color bands (5nm or 10nm color shifts).

A different focus switch architecture has been implemented recently in the Magic Leap One MR headset. The architecture uses two sets of three colors waveguides (6 high index waveguides, 325 microns thick each, see also Figure 47), with two sets of extraction grating types, each having a different diopter power (see section 11 and Figure 56 for details on the powered out-couplers). One set of 3 guides positions the virtual image in the far field at 1.5m (-0.67 diopter powered outcoupler grating), and the other set of 3 guides positions the image in the near field at 40cm from the device (-2.5D powered outcoupler grating). Figure 68 shows the Magic Leap One illumination configuration LED producing the LCOS display engine exit pupil switch in order to coupler in the near field or the far field guide sets.

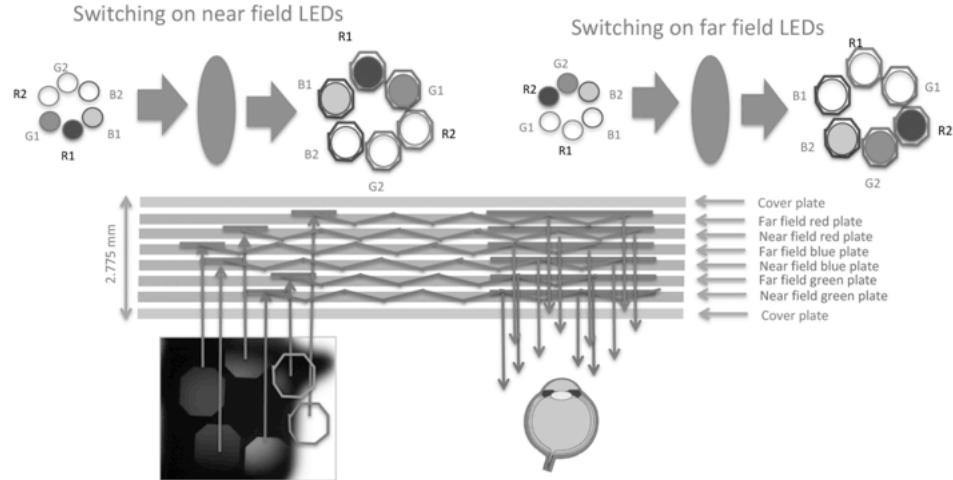


Figure 68: Binary focus switch through display engine exit pupil switch in waveguide combiners

Although such dual focus switch is interesting and relatively easy to implement, both the reduction in MTF due to the LED spectral spread caused by the in- and out-coupler grating period mismatch (since the exit coupler is powered), and the fact that the focus change is not smooth limits its visual comfort experience.

Since vergence is the reflex to accommodation in human vision, the trigger to the focus switch is here implemented with a vergence sensor based a differential left/right eye gaze tracker (here glint based eye trackers). The focus switch occurs a few hundred of milliseconds after the eye's vergence have changed from far field to near field, or vice versa.

Varifocal display architectures

We have seen previously that the MR immersion experience can be increased by allowing the FOV content to get closer to the user to allow arm's length display interaction: this is a key feature for any MR experience.

By using a tunable lens in the light engine, one can change slightly the vergence of the field, and therefore changing the location of the virtual image [101]. Tunable lenses can be implemented in various ways [102], liquid oil push/pull [103], liquid crystal (LC) [104], reflective MEMS, deformable membranes, Alvarez lenses, Multiorder DOE (MDOEs), et...). Often they are best used in conjunction with fixed refractive or reflective lenses, the compound lens system providing the mean focus as well as the slight change of focus to move the virtual image from infinity to the near field of the user (from -0.5D to -3.0D, to move the virtual image from near infinity to slightly more than a foot away from the user).

Accommodation is the reflex to eye vergence, and eye vergence is a result of stereo disparity. Similarly to binary focus switch (previous paragraph), active continuous focus tuning can be vergence contingent and therefore relying on a vergence tracker (ET based). However, for a VR varifocal system, a simpler gaze tracker might be sufficient, as the digital depth scene is known and thus a specific gaze direction can be linked to a specific depth of the digital scene. This is not the case in a see through system since the user might want to focus on a specific close up real object around which angular cone might be located one or more digital hologram(s). The hologram(s) should also look real, therefore being rendered out of focus, either behind and/or in front of the object over which the user wishes to focus.

Lens tuning in the optical engine can be done in various ways, by moving a lens, using a tunable lens in either reflective or transmissive modes, and in compact forms as in PBS based birdbath architectures. Figure 69 reviews implementations of focus tuning with free space combiners. Left, the Avegant AR example which switches the focus over multiple different planes (multifocal rather than varifocal), and right, an example of a variable power visor reflector based on membrane deformation (also based on liquid pressure).

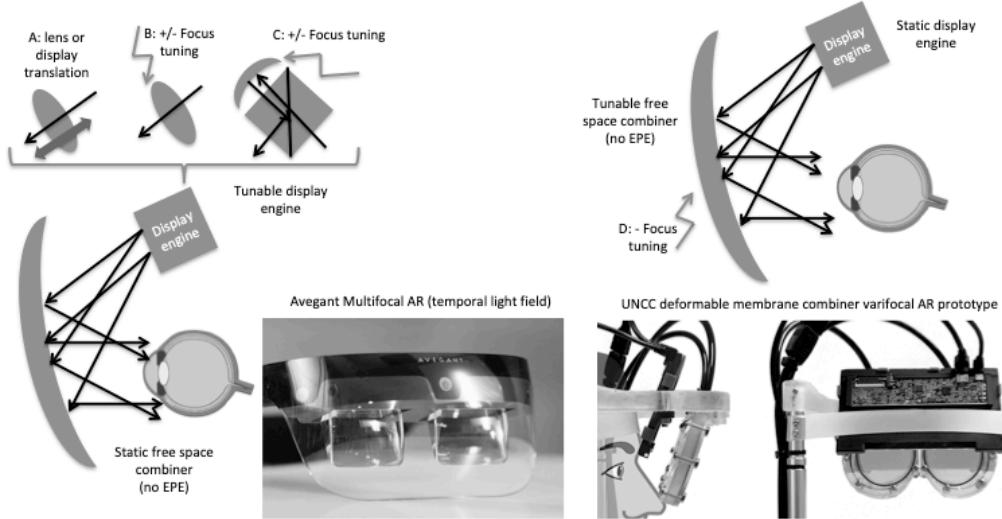


Figure 69: Free space Varifocal optical architecture implementations

When a waveguide combiner is used, if there is no exit pupil replication, the focus tuning can be performed in the light engine (left on Figure 70). This is not the case anymore when EPE techniques are used (1D or 2D).

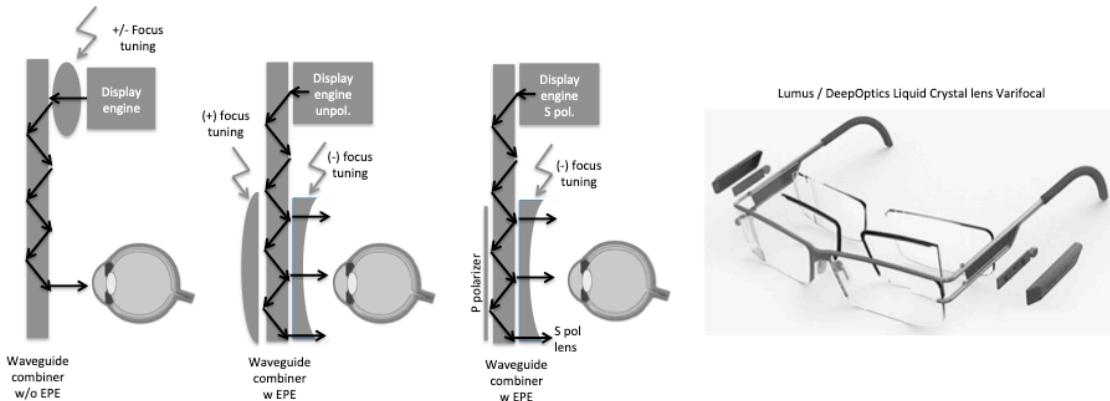


Figure 70: Waveguide combiner Varifocal optical architecture implementations

If the waveguide combiner uses a pupil replication scheme such as in a conventional waveguide combiner (1D or 2D EPE), the in-coupled field in the waveguide needs to have its image located at infinity so that all replicated pupils remain in the same depth plane over the entire eyebox. In this case, the tunable lenses might be used over the entire eyebox, only in transmission mode, with a tunable compensation lens to compensate the see-through and provide an unaltered visual experience (center in Figure 70).

The varifocal system for a pupil replication scheme can be further simplified if the see-through is polarized in one direction and the digital image in the orthogonal direction (right on Figure 70). By using polarization sensitive tunable lenses (such as LC lenses), there is no need for a compensation lens in the see-through mode anymore since the see-through field is not affected by the polarization selective tunable lens, working on the opposite polarization state. Similar configurations can be implemented by using circularly polarized states and Geometric-Phase (GP) holograms or metasurface diffractive elements as the technology base to build up the tunable lenses.

Note that there is a significant difference between the underlying technology in transmission tunable lenses (such as fluid pressure or fluid injection lenses, micro-electro-fluidic lenses, liquid crystal lenses, or even mode complex phase arrays, acousto optic refractive index tuning or other electro-optic refractive index tuning techniques) and much faster and compact reflective lens technologies such as MEMS reflective and/or Fresnel structures or reflective membrane techniques. Reflective lenses are smaller and faster than transmission lenses (best used in display engine where the single pupil remains small also), which tend to be also larger and heavier (esp. liquid filled lenses). Gravity sag and liquid flow management is a concern for all liquid lenses, and see-through artifacts a concern in liquid crystal lenses.

Developers of such tunable lenses include Adlens (UK) and Optotune (CH) for liquid pressure or liquid injection lenses, DeepOptics (IS) and Liqxtal (Taiwan) for LC lenses and SD Optics (Korea) for reflective MEMS tunable lenses.

As the focus tuning acts on the entire scene at the same time, digital render blur might be implemented in order to provide a more comfortable 3D visual cue for the user, at least in the foveated region. Specific digital blur rendering techniques providing better focus cues have been proposed such as Chromablur [105].

Multifocal display architectures

Unlike with Varifocal, the Multifocal procedure [106] renders and produces “at the same time” multiple depth scenes at a few predetermined positions (from 2 and up). If a fast display and a fast tunable lens are used in the optical engine, within the integration time of the eye, the user will see all rendered focus planes at the same time, and thus see an true optical blur [107], [108] (no need for render blur as with Varifocal). This is why multifocal display techniques are sometimes referred to as “light fields”, or even temporal light fields.

Display technologies such as a DLP would be fast enough to display up to 4 focus planes within a 90Hz frame rate (thus using a 360Hz display refresh rate, and more for RGB color sequence operation). Reflective MEMS tunable lenses are also fast enough to provide a focus shift of more than 3D at 360Hz. Such refresh rates cannot be achieved with liquid filled or liquid crystal lenses. SD Optics develop such fast tunable MEMS lenses.

A variety of multifocal display architectures have been proposed, such as by Avegant’s “temporal light field” AR headset and Oculus’ Focal Surfaces VR headset. In Avegant, the scene is split over 2 or 4 different planes, each having a different depth, in time sequence. In the Oculus example, the multiple focii planes are modulated by a phase panel in order to provide “focal surfaces” rather than focal planes, which can enhance the 3D cues over specific scenes.

Although a gaze/vergence tracker might not be necessary in this case since the entire scene is rendered and projected over various physical depths (unlike with Varifocal), a gaze tracker might still be needed to avoid parasitic plane to plane occlusions or dead-spaces due to lateral pupil movements over the FOV changing the parallax.

Light field displays

Light field capture (i.e. integral imaging) and light field display is not a new concept. Gabriel Lippmann, a Franco-Luxembourgish physicist and inventor [110], received the Nobel Prize for its invention in 1908. He also invented holography (which he called natural color photography) well before Denis Gabor, a Hungarian-British electrical engineer who received the Nobel prize for its invention 63 years after Lippmann’s first “impromptu” holography discovery. There are many ways to implement a light field display, Lipmann used the first ever Micro Lens Array (MLA) to provide a true light field display from an integral imaging capture. Such architecture has been used again recently in light field displays for VR (and potentially AR) applications [111], [112]. Other implementations of light field displays include multiple scenes projection (Holografika), and directional backlights (Leia), and tensor based displays as investigated by Prof. Gordon Wetzstein at Stanford U. Such “spatial” light fields are very different from “temporal light fields” as discussed previously.

The main drawback when attempting to implement light field displays in a product is either a resolution loss (spatial light fields) or a refresh rate increase (temporal light fields). These two limitations tempered their introduction in products, although Leia’s directional backlight display technique has been implemented recently in the new Hydrogen Smartphone from Red Corp.

Pin light displays for NTE Display

Pin light displays are an interesting concept and have been implemented both in a transmission mode [113] and more recently in reflective mode (LetinAR PinMR™ lens module, see Figure 71). In conventional pin light displays, a virtual aperture encoded on the display allows virtual projectors to be tiled, creating an arbitrary wide FOV. The image projected is rearranged into tiled sub-images on the display, which appear as the desired image when observed out of the viewer’s accommodation range. Eye tracking can enhance the resolution of such a display. FOV stretching over 110deg in a small form factor have been reported. As with spatial light field displays, the resolution hit has limited its introduction in products.

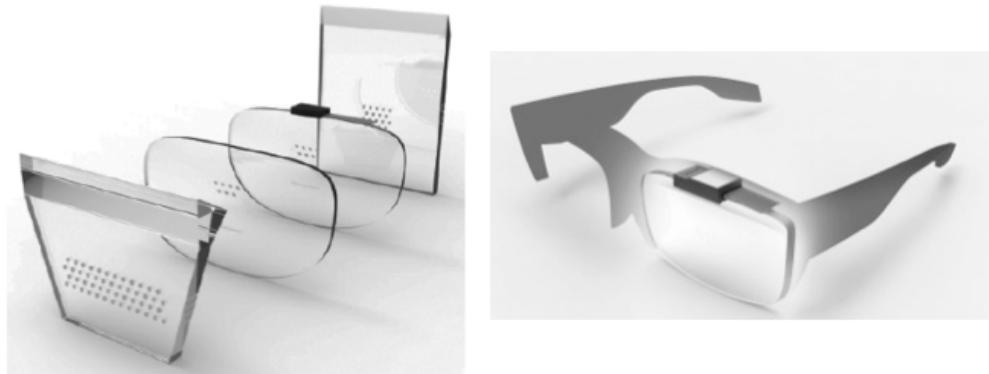


Figure 71: LetinAR PinMR™ miniature mirror array waveguide combiner

The PinMR™ combiner developed by LetinAR is more a hybrid mirror array extractor rather than a traditional pinlight display. The image is coupled in the same way as with other reflective combiners (Lumus LOE, Epson Moverio, ...). As the mirrors are smaller than the user's eye pupil, the user can "see around" the mirrors and have a decent unaltered see through experience. Due to the size of the mirrors, this architecture also produces an extended depth of focus, similar to traditional transmission pin light displays, but with a higher resolution potential. The resulting FOV can be wide, without requiring a thick combiner, as it is the case with traditional reflective mirror waveguide combiners (Zeiss Smart Glasses, Epson Moverio BT300). However, as the entire mirror array build up both the FOV and the eyebox, a single fused image is possible only at a specific eye relief distance.

Retinal Scan Displays for NTE Display

Virtual retinal Displays (VRD), or Retinal Scan Displays (RSD) draw directly an image onto the retina in a raster scan form. VRDs have been used for decades (e.g. Kazuo Yoshinaka at Nippon Electric Co. in 1986) in enterprise and defense as miniature NTE displays. They are compact, efficient if using lasers and especially VCSELs (lower threshold level than edge emission lasers), and can produce a highly contrasted virtual image. As these are scanned images, if linked to fast tunable lenses these can produce effective volumetric displays [114]. However, various basic problems have hindered their introduction into mainstream products such as a small eyebox limiting its effective FOV, speckle and phase artifacts in the image linked to the high coherence level of the source passing through random phase objects in the eye's aqueous humor. Nevertheless, several VRD NTE products have been developed (QD laser, Brother Air Scouter, Intel Vaunt and most recently By North – formerly Thalmic Labs – Focals).

Due to the small diameter of the laser beams entering the eye, such display architecture produce a virtual image that appears to be in focus no matter where the user's accommodation is set at. To note that this does not solve the VAC conflict since it does not produce realistic 3D virtual objects with true optical blur as light fields would, but rather produces a 2D virtual image which appear to be always in focus. This can be interesting for monocular smart glasses where 2D text display is prevalent. Text will never be experienced by the user as a potential real object, and thus can be presented in focus anywhere in the field without compromising the user's visual comfort. The aim of digital text superimposed on a virtual scene has to be always in focus (so it can be read by the user independently of the accommodation state), which is quite different from a 3D object which has to compete with reality (optical blur, parallax, etc.).

This extended depth of focus effect is lost when an exit pupil expansion/replication is used. In this case the eyebox might be increased to comfortable levels at the expense of the extended image depth of focus.

Finally, laser scanners producing an aerial image on a reflective diffuser (before being reflected by a half tone curved combiner) have also been used extensively in automotive HUDs due to the high brightness being able to compete with direct sun light (>10,000 nits). See for example Pioneer, Panasonic, Microvision, Mirrocle, Navdi, and more recently Wayray.

Digital holographic displays for NTE Display

Similar to laser scanners, digital holographic displays have been used in automotive HUDs for some time, due to their particularly high contrast and their relative low price (Daqri / Two Trees Photonics Ltd. for Jaguar Ltd. cars). Transmission (HTPS LCD) or reflective (LCOS based) phase panels are used to implement holographic displays. These have been produced for more than a decade (Aurora, HoloEye, Jasper, Himax), in either ferroelectric or nematic LCs, with either analog or digital drives. Other phase panel technologies such as MEMS pillars (Texas Instruments Inc) implementing reflective phase panels are subject to current R&D efforts. The pattern to be injected in the phase panel can be either a Fresnel transform (3D object in near field) or a Fourier transform (2D far field image) of the desired image. They are often referred to as CGHs (Computer Generated Holograms).

Digital holographic displays can produce "per pixel depth" scenes, in which each pixel can be located physically in a different depth plane [115], thus producing a infinite true light field experience [116]. However, occlusion has to be taken into account, as well as other parasitic aspects such as speckle and other interference issues.

Large FOV holographic display can be produced by either reducing the pixel size and at the same time reducing the pixel interspacing (challenging for most phase panel technologies) or by using non collimated illumination (diverging waves) [36]. Real time 60Hz to 90Hz hologram calculation (either direct or with an iterative IFTA algorithm) is requiring a very strong CPU/GPU support [117], [118], and custom IC development might be needed (such as hard wired FFTs).

Unlike with amplitude LCOS modulation, phase panel modulation for digital holography is very sensitive to flicker from digital drive and phase inaccuracies from analog drive over large panels.

Complex amplitude/phase encoding over a single panel with accurate phase and amplitude levels would increase the contrast by reducing quantization noise [116]. Speckle can be reduced by classical hardware methods (phase, amplitude, polarization or wavelength diversity [119]), or by software methods with higher refresh rates. Color display can be produced by either lateral RGB panel split especially for Fourier type CGHs (4K or larger panels) or conventional color sequence, but the latter puts more pressure on the panel refresh rates.

Diffraction efficiency remains low with binary phase states (1 bit, on/off pixels), and can be increased by going to 2,3 or 4 bits phase encoding. Only a few bits of phase and/or amplitude levels (however very accurately achieved in the phase panel) can produce much larger dynamic range in the resulting image (as the image is not produced by classical imaging, but rather by diffraction). This could allow for a potential 256 phase levels image (8bits color depth) generated by a single bit depth phase panel.

A typical holographic display can yield a large FOV (80deg demonstrated [36]), but the eyebox remains small. We saw

previously that typical eyebox expansion techniques (such as waveguide combiners with EPE) cannot be used anymore with an image field not located at infinity.

Figure 72 demonstrates an example of true full color holographic display in a pair of glasses [36], where each pixel is located at a different depth. It also shows a typical phase pattern to be injected in the panel such that the diffraction pattern in the near field produces the desired 3D object. Non iterative algorithms will soon allow the computation of such holograms at 90Hz in real time over a 1080p phase panel arrays.

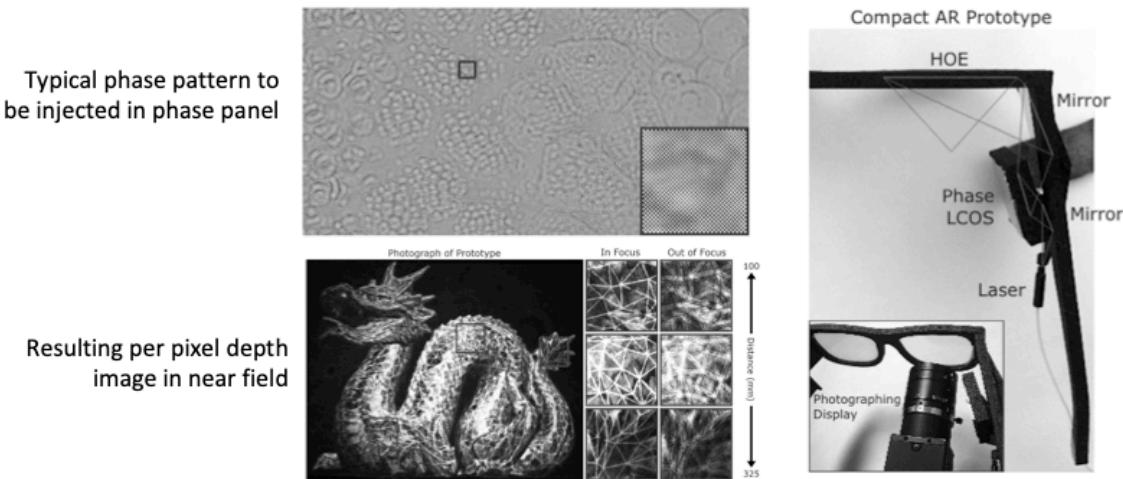


Figure 72: Per pixel depth synthetic holographic display example in smart glasses.

Currently the NTE holographic display hardware ecosystem is getting fragmented, various companies focusing on each individual building block rather than on the entire system, such as:

- Phase panel development (based on LCOS or MEMS)
- Illumination systems based on lasers, VCSEL or reduced coherence laser sources
- Custom IC (hard wired FFT, etc...)
- Custom algorithms for real time hologram calculation
- Specific combiner technologies for holographic fields

Such a fragmentation of the hardware/software ecosystem might be the solution to come up with a low cost consumer level hardware in the next years, solving effectively the VAC conflict and providing at the same time a small and compact light engine providing a large FOV with high efficiency.

Figure 73 and Table 4 summarize the various VAC conflict mitigation techniques we reviewed in this section, from binary focus switch to Varifocal, Multifocal, Retinal Scan Display, Light field display and true dynamic holographic display.

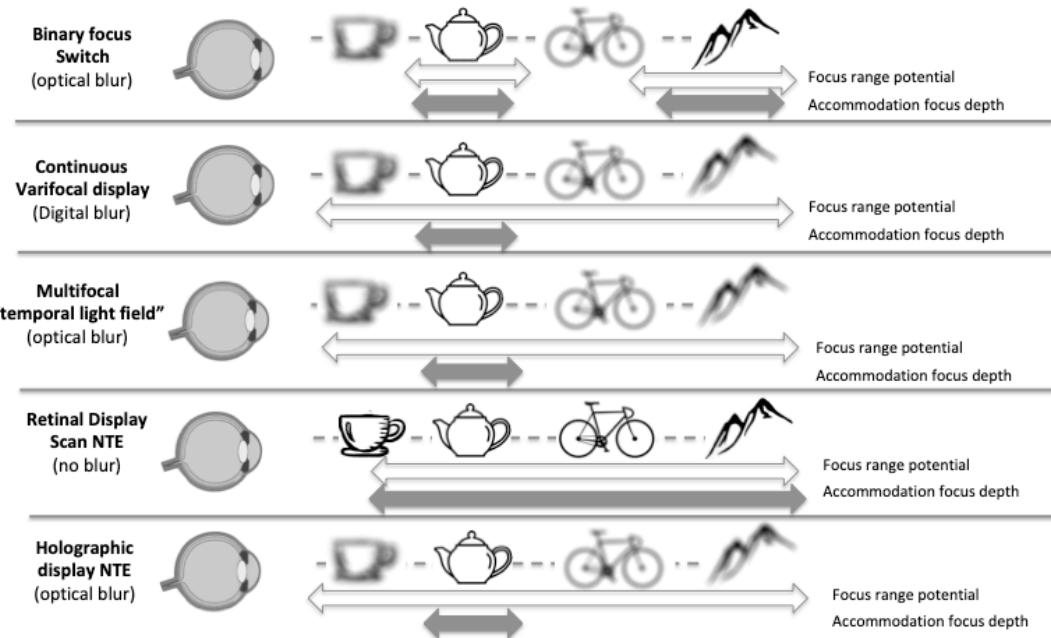


Figure 73. Summary of the various VAC conflict mitigation techniques (in this figure, the user's accommodation is set on the tea pot at arm's length distance for all 5 VAC mitigation architectures)

Figure 76 shows also the focus range potential and the accommodation range for the various VAC mitigation techniques. VRD/RSD scan techniques might provide the widest focus depth but not the best VAC mitigation.

VAC mitigation	Display engine	Illumination engine	Free space combiner	EPE Waveguide combiner	VAC	Depth planes	Blur	Eye tracking	Accommodation experience	Limitations	Company/product/prototypes
Binary focus plane	Orthogonal pol. states display engine	Orthogonally polarized	Yes	No	Partially solved	2	Optical	Not required	Realistic	Requires dual pol. display or fast pol. switch	Various prototypes and demonstrators
Dual focus switch	Any	Spatially multiplexed	Yes	Yes	Partially solved	1	Digital	Vergence tracker	Realistic	Requires multiple waveguide stacks and multiple illumination	Magic Leap One
Continuous varifocal	Any	Any	Yes, w internal Tr./Refl. lenses	Yes w large external Tr./lenses	Solved	1	Digital	Vergence tracker	Realistic (digital)	Requires large Tr. tunable lenses with waveguide EPE	Lumus/DeepOptics, Lemnis, Oculus Half Dome
Multifocal	DLP, laser scanner, fast LCOS, ...	Any	Yes	No	Solved	2<N<8	Optical	Pupil tracker for late stage occlusion rendering	Realistic	Refresh rate hit. Requires fast tunable lens and display	Avegant Corp. AR
Focal surfaces	Display with additional phase panel	Any	Yes	No	Solved	2<N<8	Optical	Pupil tracker for late stage occlusion rendering	Realistic	Refresh rate hit. Requires phase panel	Oculus prototype
Retinal Scan Display	Single or dual mirror laser scanner	Laser/VCSEL	Yes	No (OK only if no EPE)	Not solved	analog	No blur	No required	Very large (unrealistic)	Small eyebox. Parasitic phase objects. Best w text in smart glasses	QD laser, Intel Vaunt, By North, Brother airscouter, ...
Light Fields display	High res. 2D display w MLA	Any	Yes	No	Partially solved	-	Optical	Not required (can increase resolution)	Realistic	Spatial resolution hit	Nvidia, Leia, Stanford tensor display prototype
Holographic display	Phase panel (LC or MEMS)	Laser/VCSEL	Yes	No	Solved	N>=2	Optical	Pupil tracker for late stage occlusion rendering	Realistic	Heavy real time calculation requirements	Holoeye, Eyeway, Vivid-Q, Microsoft Research, prototypes

Table 4. Specifications of various VAC mitigation techniques.

Considering the amount of VC investment and technology excitement around VAC mitigation today, it seems that VAC mitigation solutions will be implemented in most of next generation MR devices (AR requiring it more than VR). The questions remains which technique or technology might be best suited. It might well be that different VAC technologies

and architectures might suit best different specific hardware and experience requirements (defense, enterprise, consumer,...)

13. Soft edge and hard edge occlusion

The MR experience aims at merging seamlessly 3D digital content over a 3D scanned reality to provide a realistic 3D visual experience, and eventually merge both in a single visual experience.

We have seen previously that occlusion is a very powerful 3D cue: therefore hologram occlusion by reality is crucial, and can be done through a real time depth scan generating an occlusion map over the holograms.

Pixel occlusion (or sometimes called hard edge occlusion) is different from hologram occlusion: a realistic hologram requires the virtual images to be realistic not only with true 3D cues, resolution and high dynamic range, but also in opacity [120], [121]. Increasing the opacity of the hologram can be done by increasing the brightness over the hologram while reducing the brightness of the see through (through a static or tunable dimming visor for example). This is however not the best solution or even the easiest solution since display brightness is an expensive spec.

A video see through experience can effectively provide a perfect pixel occlusion over the reality and overlay the digital information with the same dynamic range as reality. In doing so, one trades a true infinite resolution over 220+ FOV deg FOV light field experience (e.g. the natural see through) for a limited FOV single focus display with lower resolution. The Intel Alloy Video See Through project was based on this concept (Figure 74), however, this project was cancelled recently.

An alternative solution not requiring video see through is to use an SLM (transmission or reflective) which can reduce the reflectivity – or totally absorb– selectively the see through images over a 2D array of pixels and then use a single traditional optical combiner to merge the digital image over the occluded see through field (this is done with a TIR prism combiner in reflection mode in Figure 74). This architecture does not sample the see through, only the occlusion over the see through, and does not alter the real light field. The occlusion can also happen in a single depth plane, for which the image is focused on the SLM. All the other depth planes can only be dimmed partially. This in turn has severe limitations in parasitically occluding important out of focus fields. It also reduces the FOV down to the maximum FOV the combiner can provide. Finally, such a system can also become quite large and heavy.

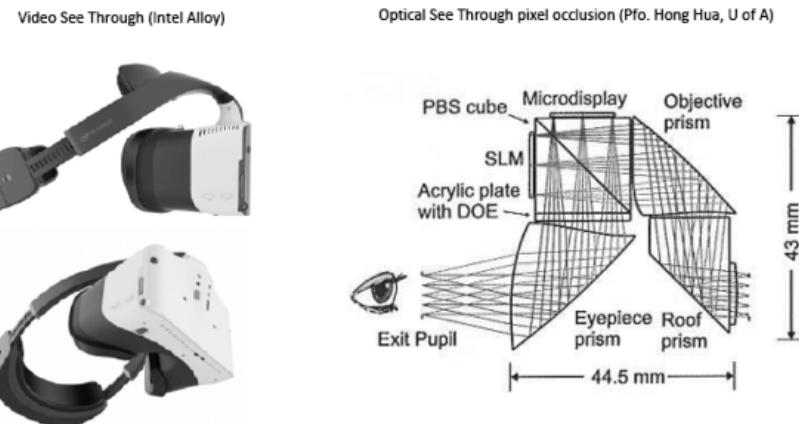


Figure 74: Pixel occlusion architectures (hard edge occlusion).

While hard edge pixel occlusion needs be processed over a focused aerial image, soft edge occlusion can be done over a defocussed image, for example through a pixelated dimming panel on a visor [122]. Such pixelated dimmers can be integrated as LC layers, both as polarization dimmers (only acting on one polarization, from 45% down to 0%), or as an amplitude LC dimmer, based on dyed LCs layers (from 50% down to 5% diming typically).

14. Peripheral display architectures

We have seen previously that the performance of the human peripheral vision system is quite different from the central foveated region. The peripheral region might lack dramatically in angular resolution and color depth, but is more sensible to clutter, jitter, aliasing and other display phenomenon than the foveated region. These peripheral display effects can dramatically reduce the visual experience for a wide FOV HMD user.

As there are two different physiological visual systems in the human eye, simply increasing the FOV reach of the foveated region display architecture to cover the peripheral regions might not make sense for most of the architectures we have discussed in this paper (both free space and waveguide based). For very large FOV values, the optimal solution might

actually be a tiled display architecture, where the two display architectures might be, or not, based on the same technology.

Recently, several optical architectures based on two different display engines have been proposed, one for the fixed foveated region and the other for the peripheral region. Examples of architectures stretching a single display system to larger FOVs (200deg +) include the Wide 5 from FakeSpace (150 deg FOV with a pancake lens), the dual panel Pimax 8K VR headset, dual panel StarVR/Acer headset and XTAL 180 deg FOV H with single non-Fresnel lens from VRgineers (left, Figure 75).

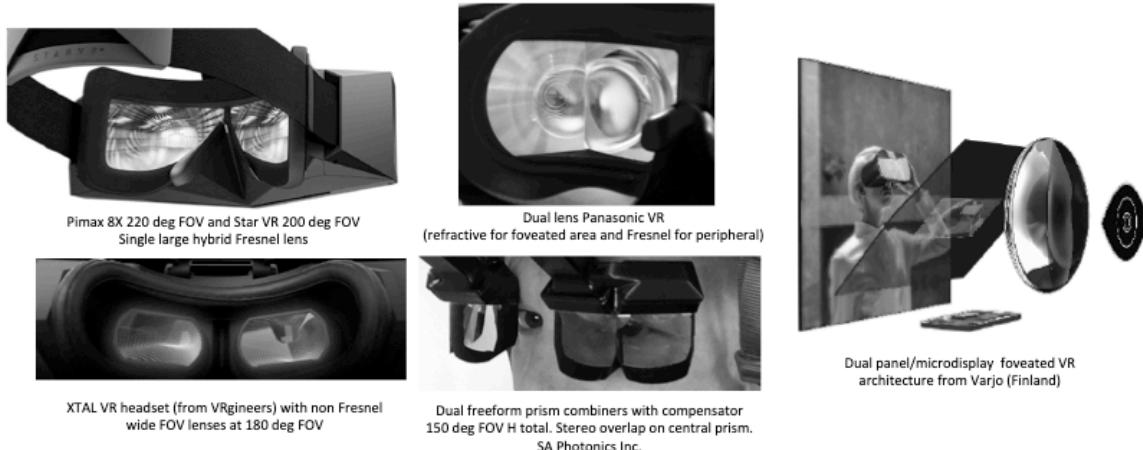


Figure 75: Peripheral VR display architectures: stretched single display (left) and dual display architectures.

Dual display architectures to provide at the same time foveated display and peripheral display include the Panasonic dual lens (space multiplexed refractive/Fresnel) VR system and the foveated VR display from Varjo (Finland). The latter is using a low res large panel display for the peripheral region and a high res microdisplay for the foveated region, which are both combined together with a gaze contingent steerable half tone combiner plate (see center and right in Figure 75). Dual freeform prism combiners per eye have also been investigated to provide a centrally foveated region with 100% stereo overlap, as well as a peripheral display region with no overlap (Prof Hong Hua of U of A). FOV up to 150 deg H have been reported (bottom center on Figure 75 - SA Photonics Inc.).

15. Optical sensor fusion in MR headsets

An MR headset experience is only as good as its combined display and sensor systems [123]. Thus, Motion to Photon (display) latency is a critical spec related to the quality of the visual and global sensory experience. Motion to Photon latency is also instrumental in reducing the well documented VR/AR motion sickness [124].

A low motion to photon latency (< 20ms, targeting 10ms and below) is necessary to convince the user's mind that he or she is in another place [125]: this is also called presence. Presence is key to the MR experience.

Display refresh rate is of course one aspect of latency, as is persistence: persistence is linked to the display technology itself. Laser scanners and DLPs have very low persistence, while LC, LCOS and OLED displays have higher persistence.

Sensor fusion, aiming at reducing latency, is enabling presence. To point out the importance of sensor fusion in MR, some companies go to the extend of developing their own specific GPU silicon chips which includes custom sensor fusion to allow the lowest latency. This provides the best experience with a minimum of discomfort for the user (see for example Microsoft's HoloLens HPU - Holographic Processing Unit). Figure 76 depicts a typical sensor fusion architecture for an MR system, along with the sensors arrays discussed below.

Sensors arrays include optical Head Tracking (HeT) to lock the hologram in place while the head (and/or the body) moves around. IMUs linked to a dual camera HeT can provide a 6 Degree of Freedom (6DOF) sensing, required for decent head tracking [126]. Depth mapping sensors allow the hologram to be world locked over the 3D scanned reality allowing also for hologram occlusion by reality to enhanced realism.

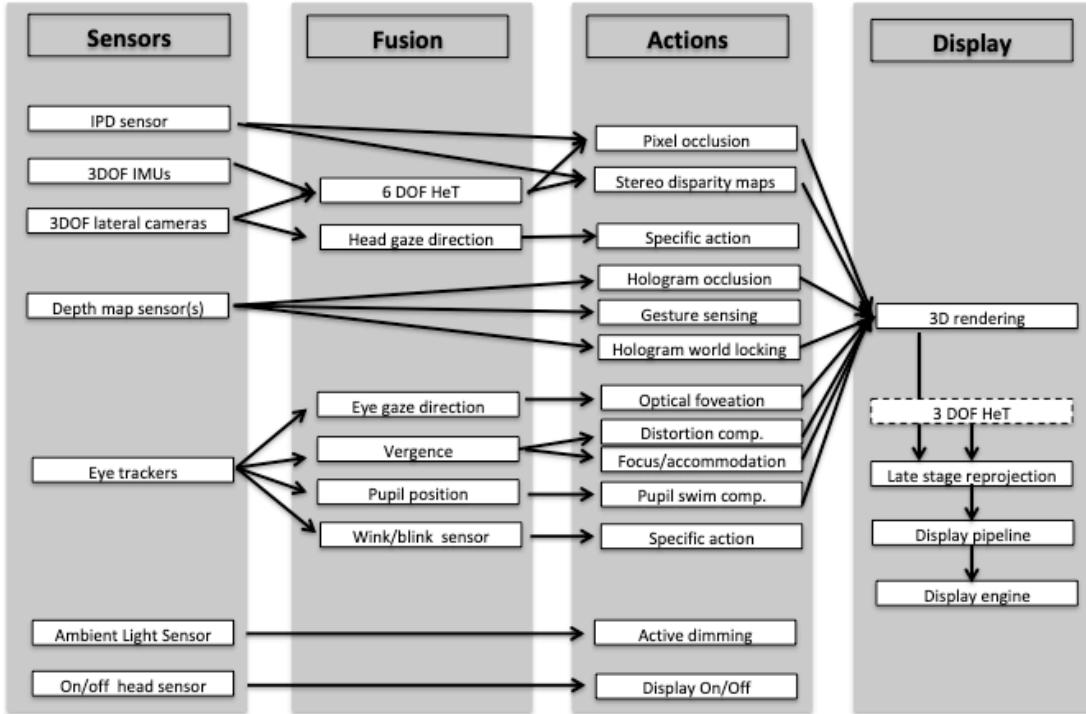


Figure 76: Sensor fusion flow in typical MR systems

Depth map sensors

Depth mapping can be implemented in a wide variety of ways [127], through stereo cameras, structured illumination or time of flight (TOF) (see Figure 77). In some cases, two operation modes might be required for the depth map sensors: near field mode (accurate gesture sensing) and far field mode (accurate reality scanning).

Semantic depth scanning is becoming also a standard in MR, recognizing the 3D structures beyond the 3D scanning, so that they can be used as intended in an MR environment (chair, table, floor, wall, person, animal, computer, toy...). Artificial Intelligence through Deep Neural Networks (DNN) can help recognize on the fly the 3D scanned objects: DNN are thus starting to be integrated in next generation custom sensor fusion IC chips.

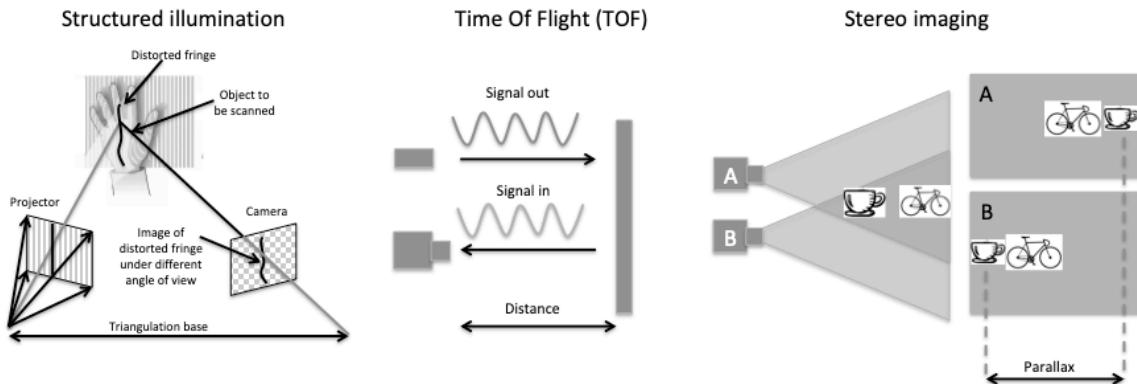


Figure 77: 3D depth mapping sensors used in industry

Stereo Cameras

Stereo cameras simulate human binocular vision by measuring the displacement in pixels between the two cameras placed a fixed distance apart and use that to triangulate distances to points in the scene. Conventional sensor arrays (CMOS) can be used. Parallax (thus depth resolution) in stereo cameras is a function of the camera separation, thus

increasing potentially their required sensor bar footprint.

Structured Light Sensor

Structured Light sensing works by projecting an IR light pattern (grid, fringes, spot patterns, etc...) onto a 3D surface and using the distortions to reconstruct surface contours [127]. Ideal projectors are Far field pattern projectors such as Fourier CGHs (Computer Generated Holograms). CGHs work well with IR lasers or VCSELs around 850nm-900nm. The sensor does not need to be custom (CMOS). FOV (both in projection and sensing) and also lateral resolution limit their reach. Popular structured light depth map sensors are the Kinect 360 (Xbox 360) from Microsoft Corp., the Structure sensor bar from Occipital Inc., the RealSense sensor bar from Intel Corp, and the depth sensors on Magic Leap One MR headset.

Time-of-Flight Sensor

TOF sensors work by emitting rapid pulses of IR light that are reflected by objects in its field of view [128]. The delay of the reflected light coming back is used to calculate the depth location at each pixel in the angular space. Such sensor architectures can be implemented with a 2D scanner and a single detector, a 1D source array scanned in the orthogonal direction and sensed back onto a linear detector array, or a single pulse light sensed by a 2D detector array. More sophisticated TOF sensors encode the phase rather than the amplitude. Such sensors chip layouts can be highly custom [130]. Double reflections are one of the limits to overcome by TOF sensors. A popular TOF sensor is the Kinect One (Xbox One) and its modified version on Hololens V1, from Microsoft Corp.

All these sensors (stereo cameras, structured light and TOF) have their specific features and limitations. Most of them are based on IR illumination and have a hard time functioning outdoors as bright sunlight can wash out or add noise to the measurements. B&W stereo cameras have no problems working outdoors and consume less power, but they work best in well-lit areas with lots of edge features and high contrast.

Head trackers and 6DOF

DOF (Degrees Of Freedom), are the number of different “directions” that an object can move in the 3D space. 3DOF headsets can track the head orientation, (where the user is looking). The 3 axis are roll, yaw and pitch. 6DOF headsets will track orientation and position (the headset knows where the user is looking and also where the user is located in space). This is sometimes referred to as *roomscale* or *positional* tracking. 6DOF can be accomplished by dual front or lateral facing B&W environmental understanding cameras compiled with the data from dual IMUs.

In previous generation VR and MR configurations, outside-in (e.g. sensors not located on the headset, but rather scattered throughout the room) Head Trackers (HeT) and gesture sensors have been used (Oculus DK1, DK2, CV1, Sony Playstation VR, HTC Vive). Outside/in sensors are being replaced in current hardware generations by inside/out sensors (all sensors located on the headset) for a more comfortable MR experience (Oculus Quest, HTC Vive Pro).

SLAM

SLAM (or Simultaneous Localization And Mapping) is a critical technology for all AR applications, no matter if HMD based or simply smart-phone based. SLAM allows the device to understand its environment and recognize it through visual input [129]. SLAM can be based solely on cameras (as with HeT) or with depth scanners (structured illumination, TOF, stereo vision, etc...). Google’s ARCore and Apple’s ARKit make heavy use of SLAM in smart phone implementations with standard camera or more complex sensors. A typical SLAM data cloud is shown in Figure 78.

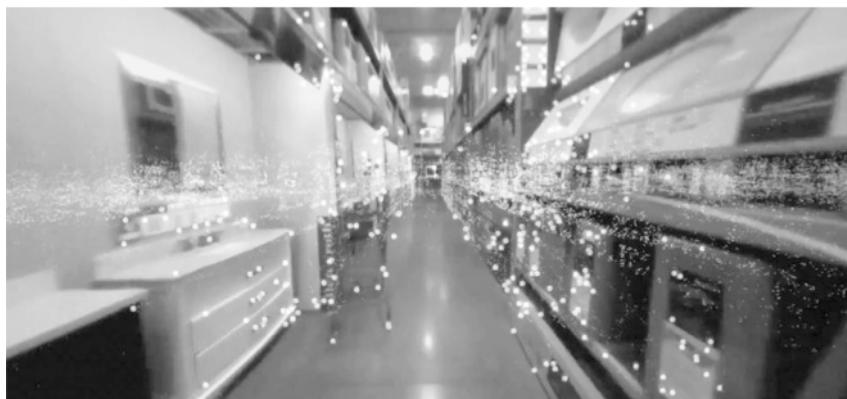


Figure 78:Typical SLAM data cloud from ARCore.

Using these points, the AR device can understand its surroundings, and thus can create more interactive and realistic AR and MR experiences. The algorithm has two tasks:

- build a map of the environment through scanning
- locate the device within this scanned environment.

SLAM data can be shared between interconnected devices, scanning different parts of the same environment and building a single environment database. Early implementations of SLAM used graphical markers or beacons. More recent implementations do not need markers anymore (Metaio/Apple, Wikitude, Google Tango, etc...). More advanced SLAM technologies might also use semantic recognition of the scanned environment, based on AI (DNN).

ET, Gaze, pupil and vergence trackers

Eye Tracking (ET), gaze tracking, pupil position tracking and eventually also vergence tracking (differential left/right eye ET) rely on a sensor fusion architecture which has not only high angular resolution (sub-degree), but is also universal (each eyeball being slightly different). Allows the user to wear prescription glasses is a challenge for most eye trackers based on traditional glint imaging architectures (SMI/Apple, Tobii). Other eye tracking architectures include non-glint based ET, scanner ET (Digilens, AdHawk) and lens-less sensors ET (Rambus).

Eye tracking can be used to interact with the display, but is often also an enabling technology for active optical foveation or display compensation such as pupil swim correction, FOV color uniformity, eyebox uniformity, exit pupil switching or steering, etc... Vergence tracking is key to VAC mitigation, and might be used to set the focus of the display to adapt to the current accommodation state (see section on VAC compensation).

Hand and gesture sensors

Gesture sensing is a critical feature for any MR device, allowing for arm's length display interactions. There are various type of optical gesture sensing techniques used in industry today. Most of them rely on depth map sensors as described previously (single camera, stereo cameras, structured illumination or TOF). However, these have to be working rather in the close near field, which is usually a different setting than the far field scanning setting for most sensing technologies.

Popular hand gesture sensors include the Leap Motion gesture sensor, which is based on IR flood illumination and stereo camera sensing the hands motions, with heavy lifting algorithms (this is not a depth map sensor).

Yet other gesture sensors might use radar technology in a miniature package, such as in the Google/ATAP team "Solis" sensor. Leap motion and Solis gesture recognition sensors are shown in Figure 79.

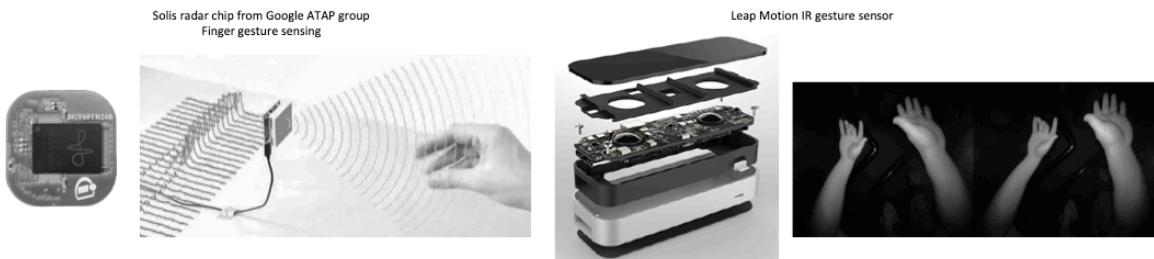


Figure 79: Google ATAP Solis radar chip (left) and Leap Motion optical sensor (right) for gesture sensing.

Other critical hardware requirements

Throughout this review paper, we have focused our attention on optical hardware and optical architecture requirements for next generation MR headsets. Various other critical hardware components are required to get there, but are out of the scope of this review paper. Among those are passive thermal management, novel battery technology, wireless links such as BT, WiFi and 3G/4G, [131] and eventually the long awaited 5G and WiGig networks which will enable remote rendering, reduce on-board compute requirements and thus allow smaller headset form factors and cooler operation.

16. Conclusion

The aim of this review paper was to capture the state of the art in optics and optical technologies for VR, AR, MR and smart glasses, in their display engines, optical combiners as well as sensors. We demonstrated that the key to choosing

the right optical building blocks and the right display/sensors architecture is to match closely their performances to the specifics and limitations of the human visual and sensory system, which we dubbed “human centric optical design”.

Both free space and waveguide combiner architectures have been reviewed, as well as techniques to mitigate the limitations of étendue, so that a large FOV can be produced over a generous eyebox, without affecting the high angular resolution of the immersive display or the display uniformity perception, thus providing a new level of visual comfort to the user.

Special emphasis has been put on waveguide combiner architectures and subsequent optical in- and out-couplers. These tend to become the “de facto” optical building blocks for tomorrow’s light-weight MR headsets addressing at the same time immersion and wearable comfort.

Optical foveation, pixel occlusion, peripheral displays and Vergence Accommodation Conflict (VAC) have also been addressed as keys issues for visual comfort. These provide the immersive display and sensory experience that would allow not only enterprise and defense markets to flourish, as they do already today, but also the burgeoning consumer market to eventually reach the expectations of the market analysts.

17. REFERENCES

- [1] W.S. Colburn B.J. Chang , **Holographic combiners for head up displays**, Tech Report No. AFAL-TR-77-110 , 1977
- [2] Jason Jerald, **“The VR book: human centered design for virtual reality”**. ISBN 978-1-97000-112-9 (2016).
- [3] Woodrow Barfield, **“Fundamentals of wearable computers and augmented reality”**, second edition, ISBN 978-1-482243595 (2015).
- [4] Laura Inzerillo, **“Augmented reality: Past, Present and Future”**. The Engineering Reality of Virtual Reality 2013, edited by Margaret Dolinsky, Ian E. McDowall, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8649,
- [5] R. T. Azuma, **“A survey of augmented reality: Presence, Teleoperators and Virtual Environments”** 6, 355- 385 (1997).
- [6] O. Cakmakci, and J. Rolland, **“Head-worn displays: a review,”** J. of Disp. Tech. 2, 199-216 (2006).
- [7] J. Rolland, and O. Cakmakci, **“Head-worn displays: The future through new eyes,”** Opt. and Phot. News 20, 20-27 (2009).
- [8] D. W. F.Van Krevelen, and R. Poelman, **“A survey of augmented reality technologies, applications and limitations,”** The International Journal of Virtual Reality 9, 1-20 (2010).
- [9] Kok-Lim Low, Adrian Ilie, Greg Welch, Anselmo Lastra, **“Combining Head-Mounted and Projector-Based Displays for Surgical Training”** , IEEE Virtual reality 2003, Proceedings 10.1109/VR.2003.
- [10] Y. Amitai, A. Friesem, and V. Weiss, **“Holographic elements with high efficiency and low aberrations for helmet displays,”** Appl. Opt. 28, 3405-3416 (1989).
- [11] Nick Baker **“Mixed Reality”** Keynote at Hot Chips HC28 – Symposium for High Performance Chips, Aug. 21-23 2016, (www.hotchips.org).
- [12] Michael F. Deering, **“The Limits of Human Vision”, Sun Microsystems, 2nd International Immersive Projection Technology Workshop**, 1998.
- [13] Antonio Guirao Conception Gonzalez, Manuel Redondo, Edward Geraghty Sverker Norrby and Pablo Artal: **“Average Optical Performance of the Human Eye as a Function of Age in a Normal Population”** Investigative Ophthalmology & Visual Science: January2014. Vol. 40, No. I.
- [14] Darryl Meister , **“Memorandum to Vision Council lens Technical committee”**, Carl Zeiss Vision GmbH (2013)
- [15] Eli Peli **“ The visual effects of head-mounted displays are not distinguishable from those of desktop computer display”**, Vision Research 38 (1998) pp. 2053-2066
- [16] Philip R. K. Turnbull, John R. Phillips , **“Ocular effects of virtual reality headset wear in young adults”** , Nature, Scientic Reports 7: 16172 (2017)
- [17] Martin S. Banks, David M. Hoffman, Joohwan Kim, and Gordon Wetzstein , **“3D Displays”**, Annu. Rev. Vis. Sci. 2016. 2:19.1–19.39
- [18] Di Zhang, Pascaline Neveu, Yulia Fattakhova, Stéphanie Ferragut, Mathieu Lamard, Béatrice Cochener & Jean-Louis de Bougrenet de la Tocnaye, **“Target Properties Effects on Central versus Peripheral Vertical Fusion Interaction Tested on a 3D Platform”** Current Eye Research, Taylor and Francis, ISSN: 0271-3683 (Print) 1460-2202.
- [19] Robert Patterson, **“Human Factors of Stereoscopic Displays”**, SID 09 DIGEST • 805 (2009)
- [20] Brian Wheelwright, Yusufu Sulai, Ying Geng, Selso Luanava, Stephen Choi, Weichuan Gao, Jacques Gollier, **“Field of view: not just a number”**, Proc. SPIE 10676, Digital Optics for Immersive Displays, 1067604 (21 May 2018)
- [21] Kosei Oshima, Kazuhiko Naruse, Katsutoshi Tsurutani, Junichi Iwai Takahiro Totani, Shinichi Uehara, Satoshi Ouchi, Yoshihiko Shibahara, Hiromitsu Takenaka, Yoshitaka Sato, Takashi Kozakai, Makio Kurashige, and Hirofumi Wakemoto, **“Eyewear Display Measurement Method: Entrance Pupil Size Dependence in Measurement Equipment”** , Book 2: Session 79: Measurements for AR/VR Displays, Volume 47, Issue 1, San Francisco, CA, May 22–May 27, 2016.

- [22] Thomas H. Harding Clarence E. Rash “**Daylight luminance requirements for full-color, see-through, helmet-mounted display systems**”, Opt. Eng. 56(5), 051404 (2017)
- [23] Peter G.J. Barten , “**Formula for the contrast sensitivity of the human eye, Image Quality and System Performance**”, edited by Yoichi Miyake, 231 D. René Rasmussen, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 5294 (2004).
- [24] Steven A. Cholewiak, Gordon D. Love, Pratul P. Srinivasan, Ren Ng, Martin Banks, ” **Chromablur: rendering chromatic eye aberration improves accommodation and realism**”, Journal of the ACM Transactions on Graphics (TOG), Vol. 36 Issue 6, Nov. 2017. Article No.210.
- [25] Westheimer G. “**Visual Acuity**”, chapter 17. In Moses, R.A. and Hart W.M. (ed), Adler’s Physiology of the eye, Clinical Application. St. Louis, The C.V Mosby Company, 1987
- [26] Josephine C. Moore “**Divisions of the visual field related to visual acuity**”, PhD, OTR. Pedretti and Early, 2001
- [27] Hong Hua and Sheng Liu , “**Dual-sensor foveated imaging system**”, Applied Optics, Vol. 47, No. 3, 20 January 2008
- [28] T.Ienaga, K.Matsunaga, K.Shidoji, K.Goshi, Y.Matsuki and H. Nagata, “**Stereoscopic video system with embedded high spatial resolution images using two channels for transmission**” in *Proceedings of ACM Symposium on Virtual Reality Software and Technology* (2001), pp. 111–118
- [29] Anjul Patney, Marco Salvi, Joohwan, Kim Anton Kaplanyan, Chris Wyman, Nir Bentv, David Luebke and Aaron Lefohn , “**Towards Foveated Rendering for Gaze-Tracked Virtual Reality**”, ACM Trans. Graph., Vol. 35, No. 6, Article 179, 2016
- [30] Linghui Rao, Sihui He, and Shin-Tson Wu, “**Blue-Phase Liquid Crystals for Reflective Projection Displays**”, JOURNAL OF DISPLAY TECHNOLOGY, VOL. 8, NO. 10, OCTOBER 2012
- [31] “**Understanding Trade-offs in Microdisplay and Direct- view VR headsets designs**”, Insight Media Display Intelligence report, 2017
- [32] Zhaojun Liu, Wing Cheung Chong, Ka Ming Wong and Kei May Lau, “**GaN-based LED micro-displays for wearable applications**”, Microelectronic Engineering 148 (2015) 98–103
- [33] Brian T. Schowengerdt, Mrinal Murari and Eric J. Seibel, “**Volumetric Display using Scanned Fiber Array**”, SID 10 DIGEST 653 , 2010.
- [34] Sundeep Jolly, Nickolaos Savidis, Bianca Datta, Daniel Smalley and V. Michael Bove, Jr, “**Near-to-eye electroholography via guided-wave acousto-optics for augmented reality**”, Practical Holography XXXI: Materials and Applications, edited by Hans I. Bjelkhagen, V. Michael Bove, Proc. of SPIE Vol. 10127 (2017).
- [35] Enrico Zschau, Robert Missbach, Alexander Schwerdtner and Hagen Stolle, “**Generation, encoding and presentation of content on holographic displays in real time.**” Three-Dimensional Imaging, Visualization, and Display 2010, Bahram Javidi and Jung-Young Son, Editors, Proc. SPIE, Vol. 7690, 76900E (2010)
- [36] Joel Kollin, Andreas Georgiu, Andrew Maimone, “**Holographic near-eye displays for virtual and augmented reality**”, in ACM Transactions on Graphic 36(4):1-16 · July 2017.
- [37] Bernard Kress and William Cummins “**Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices**”, SID 2017 Book 1: Session 11: AR/VR Invited Session II.
- [38] Paul F. McManamon, Philip J. Bos, Michael J. Escuti, Jason Heikenfeld, Steve Serati, Huikai Xie, and Edward A. Watson, “**A Review of Phased Array Steering for Narrow-Band Electrooptical Systems**”, Proceedings of the IEEE , Vol. 97, No. 6, June 2009
- [39] Maimone A and Fuchs H. “**Computational Augmented Reality Eyeglasses**”. IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2013.
- [40] <http://www.letinar.com>
- [41] <http://www.lusovu.com/>
- [42] <https://www.ifixit.com/Teardown/Magic+Leap+One+Teardown/112245>
- [43] <https://www.youtube.com/watch?v=bnfwClgheF0>
- [44] <http://www.kessleroptics.com/wp-content/pdfs/Optics-of-Near-to-Eye-Displays.pdf> (slide 15)
- [45] <https://www.dlodlo.com/en/v1-summary>
- [46] Narasimhan, Bharathwaj “**Ultra-Compact pancake optics based on ThinEyes super-resolution technology for virtual reality headsets**”. Proc. SPIE 10676, Digital Optics for Immersive Displays, 106761G (21 May 2018).
- [47] Grabovičkić, Dejan & Benítez, Pablo & Miñano, Juan & Zamora, Pablo & Buljan, Marina & Narasimhan, Bharathwaj & I. Nikolic, Milena & lopez risco, Jesus & Gorospe, Jorge & Sánchez, Eduardo & Lastres, Carmen & Mohedano, Rubén. (2017).”**Super-resolution optics for virtual reality**”. Proc. SPIE 10335, Digital Optical Technologies 2017, 103350G.
- [48] Clarence E. Rash “**Helmet-mounted Displays: Design Issues for Rotary-wing Aircraft**” (Press Monograph) Hardcover 31 Jan 2001, ISBN-10: 0819439169, SPIE; illustrated edition.
- [49] Melzer, J.E., “**Head-Mounted Displays**,” [The Avionics Handbook] 2nd Edition, C.R. Spitzer, ed, CRC Press, Boca Raton, FL (2006).
- [50] Jonathan P. Freeman, Timothy D. Wilkinson, Paul Wisely , “**Visor Projected HMD for fast jets using a holographic**

- video projector**" 3D Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV, edited by Bahram Javidi, Proc. of SPIE Vol. 7690
- [51] H. Nagahara, Y. Yagi, and M. Yachida, "A wide-field-of-view catadioptrical head-mounted display," Elec. and Com. in Japan 89, 33-43 (2006).
- [52] Kocian, D. F., "Design considerations for virtual panoramic display (VPD) helmet systems," AGARD Conference Proceedings No. 425, The man-machine interface in tactical aircraft design and combat automation, 22-1 (1987).
- [53] Matt Novak , "Meet Your Augmented and Virtual Reality Challenges Head-On: Design Your Next System with 2D-Q Freeforms in CODE V" , white paper, Synopsys Corp. 2018
- [54] R. Martins, V. Shaoulou, Y. Ha, and J. Rolland, "A mobile head-worn projection display," Opt. Express 15, 14530 (2007).
- [55] Takahisa Ando, Koji Yamasaki, Masaki Okamoto, "Head Mounted Display using holographic optical element", SPIEVOL.3293.
- [56] Y. Ha, V. Smirnov, L. Glebov, and J.P. Rolland, "Optical Modeling of a Holographic Single Element Head-mounted Display" in Helmet- and Head-Mounted Displays IX: Technologies and Applications, edited by Clarence E. Rash, Colin E. Reese, Proceedings of SPIE Vol. 5442.
- [57]. Kasai, Y. Tanijiri, T. Endo, and H. Ueda, "A practical see-through head mounted display using a holographic optical element" Opt. Review 8, 241-244 (2001).
- [58] Mickaël Guillaumée, Seyed Payam Vahdati, Eric Tremblay, Arnaud Mader, Victor J. Cadarso, Jonas Grossenbacher, Jürgen Brugger, Randall Sprague and Christophe Moser, "Curved Transflective Holographic Screens for Head-Mounted Display", in Advances in Display Technologies III, edited by Liang-Chy Chien, Sin-Doo Lee, Ming Hsien Wu, Proc. of SPIE Vol. 8643 (2013).
- [59] Hong Hua, Dewen Cheng, Yongtian and Wang Sheng Liu "Near-eye displays: State-of-the-art and emerging technologies", in 3D Imaging, Visualization, and Display for Defense, Security, and Avionics IV, edited by Bahram Javidi, Jung-Young Son, John Tudor Thomas, Daniel D. Desjardins, Proc. of SPIE Vol. 7690 (2010).
- [60] Hong Hua "Past and future of wearable augmented reality displays and their applications, Fifty Years of Optical Sciences at The University of Arizona", edited by Harrison H. Barrett, John E. Greivenkamp, Eustace L. Dereniak, Proc. of SPIE Vol. 9186, (2001)
- [61] Ichiro Kasi, Hiroaki Udea, "A forgettable Near To eye Display", Digest of papers, 4th International Symposium on Wearable Computers, 10.1109/ISWC (2000).
- [62] J. Michael Miller, Nicole de Beaucoudrey, Pierre Chavel, Jari Turunen, and Edmond Cambril, "Design and fabrication of binary slanted surface-relief gratings for a planar optical interconnection", 10 August 1997, Vol. 36, No. 23, APPLIED OPTICS
- [63] Jyrki Kimmel, Tapani Levola, Pasi Saarikko and Johan Bergquist, "A novel diffractive backlight concept for mobile displays", Journal of the SID 16/2, 2008
- [64] Jyrki Kimmel, Tapani Levola, "Diffractive Backlight Light Guide Plates in Mobile Electrowetting Display Applications", SID 09 Paper 471 Page 2
- [65] Juan Liu, Nannan Zhang, Jian Han, Xin Li, Fei Yang, Xugang Wang, Bin Hu, and Yongtian Wang, "An improved holographic waveguide display system", Applied Optics Vol. 54, Issue 12, pp. 3645-3649 (2015)
- [66] Takuji Yoshida, Kazutatsu Tokuyama, Yuichi Takai, Daisuke Tsukuda, Tsuyoshi Kaneko, Nobuhiro Suzuki, Akira Yoshikiae, Katsuyuki Akutsu and Akio Machida "A plastic holographic waveguide combiner for light-weight and highly-transparent augmented reality glasses", Journal of the SID 26/5, 2018
- [67] Hiroshi Mukawa, Katsuyuki Akutsu, Ikuo Matsumura, Satoshi Nakano, Takuji Yoshida, Mieko Kuwahara, Kazuma Aiki, Masataka Ogawa, "A Full Color Eyewear Display using Holographic Planar Waveguides" 8.4, SID 08 DIGEST 2008.
- [68] Takashi Oku, Katsuyuki Akutsu, Mieko Kuwahara, Takuji Yoshida, Eisaku Kato, Kazuma Aiki, Ikuo Matsumura, Satoshi Nakano, Akio Machida, and Hiroshi Mukawa , "High-luminance See-through Eyewear Display with Novel Volume Hologram Waveguide Technology" , 15.2, 192 • SID DIGEST 2015
- [69] Khaled Sarayeddine, Pascale Benoit, Gilhem Dubroca and Xavier Hugel "Monolithic Low Cost plastic light guide for full colour see through personal video glasses", in ISSN-L 1883-2490/17/1433 ITE and SID (IDW 10), pp1433-1435. (2010)
- [70] Tapani Levola "Exit pupil expander with a large field of view based on diffractive optics", Journal of the Society of Information Display SID 17: 659-664 (2009).
- [71] Tapani Levola "Diffractive optics for virtual reality displays", Journal of the SID 14/5, 2006
- [72] Bernard Kress "Diffractive and holographic optics as optical combiners in head mounted displays," Proceedings of the 2013 ACM conference on Pervasive and Ubiquitous computing - Ubicomp'13, pp1479-1482, (2013).
- [73] Alex Cameron "Optical Waveguide Technology & Its Application In Head Mounted Displays" in Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI, edited by Peter L. Marasco, Paul R. Havig II, Daniel D. Desjardins, Kalluri R. Sarma, Proc. of SPIE Vol. 8383, 83830E
- [74] Malcolm Homan, "The use of optical waveguides in Head Up Display (HUD) applications" in Display Technologies and Applications for Defense, Security, and Avionics VII, edited by Daniel D. Desjardins, Kalluri R. Sarma,

Proc. of SPIE Vol. 8736,

[75] Dewen Cheng, Yongtian Wang, Chen Xu, Weitao Song, and Guofan Jin, “**Design of an ultra-thin near-eye display with geometrical waveguide and freeform optics**” Optics Express, Vol 22, Issue 17 pp.20705-20719 (2014).

[76] David Jurbergs, Friedrich-Karl Bruder, Francois Deuber, Thomas Fäcke, Rainer Hagee, Dennis Hönel, Thomas Rölle, Marc-Stephan Weiser, Andy Volkov, “**New recording materials for the holographic industry**”, Practical Holography XXIII: Materials and Applications, edited by Hans I. Bjelkhagen, Raymond K. Kostuk, Proc. of SPIE Vol. 7233

[77] www.digilens.com

[78] Kevin Curtis and Demetri Psaltis, “**Cross talk in phase coded holographic memories**”, Vol. 10, No. 12, December 1993 . J. Opt. Soc. Am. A 2547.

[79] Herwig Kogelnik, “**Coupled Wave theory for Thick Hologram Gratings**”, The Bell System Technical Journal, Vol 48, No 9 (1969).

[81] M.A. Golub, A.A. Friesem and L. Eisen, “**Bragg properties of efficient surface relief gratings in the resonance domain**”, Optics Communications 235 (2004) 261–267

[82] M.G. Moharam “**Stable implementation of the rigorous coupled wave analysis for surface relief gratings: enhanced transmittance matrix approach**”, J. Opt. Soc. Am. A, Vol 12, Issue 5 pp. 1077-1086 (1995)

[83] L. Alberto Estepa, Cristian Neipp, Jorge Francés, Andrés Márquez, Sergio Bleda, Manuel Pérez-Molina, Manuel Ortúñoz and Sergi Gallego “**Corrected coupled-wave theory for non-slanted reflection gratings**”, Physical Optics, edited by Daniel G. Smith, Frank Wyrowski, Andreas Erdmann, Proc. of SPIE Vol. 8171

[84] <http://www.kjinnovation.com/>

[85] <https://meep.readthedocs.io/en/latest/>

[86] Pasi Laakkonen and Tapani Levola, “**A new replication technology for mass manufacture of slanted gratings**”, Europhotonics, pp 28-29, June July 2007.

[87] <https://www.lighttrans.com/applications/virtual-mixed-reality/waveguide-huds.html>

[87] Michael W. Farn, “**Binary gratings with increased efficiency**”, Applied Optics, Vol 31, Issue 22, pp. 4453-4458 (1992).

[88] Bernard Kress and Patrick Meyrueis “**Applied Digital Optics: From Micro-optics to Nanophotonics**” 1st Edition, John Wiley and Sons Publisher, ISBN-10: 0470022639 (2007)

[89] Giorgio Quaranta, Guillaume Basset, Olivier J. F. Martin and Benjamin Gallinet, “**Steering and filtering white light with resonant waveguide gratings**”, Proc. SPIE 10354, Nanoengineering: Fabrication, Properties, Optics, and Devices XIV, 1035408 (2017)

[90] Guillaume Basset, “**Resonant screens focus on the optics of AR**”, Proc. SPIE 10676, Digital Optics for Immersive Displays, 106760I (2018).

[91] Federico Capasso “**The future and promise of flat optics: a personal perspective**”, Nanophotonics Vol. 7 Issue 6. (2018).

[92] Wei Ting Chen, Alexander Y. Zhu, Jared Sisler, Yao-Wei Huang, Kerolos M. A. Yousef, Eric Lee, Cheng-Wei Qiu and Federico Capasso, “**Broadband Achromatic Metasurface-Refractive Optics**” *Nano Lett.*, 2018, 18 (12), pp 7801–7808

[93] Shuyan Zhang, Alexander Soibel, Sam A. Keo, Daniel Wilson, Sir. B. Rafol, David Z. Ting, Alan She, Sarath D.Gunapala and Federico Capasso, “**Solid-immersion metaleenses for infrared focal plane arrays**” *Appl. Phys. Lett.* 111, 111104 (2018)

[94] <https://hololens.reality.news/news/microsoft-has-figured-out-double-field-view-hololens-0180659/>

[95] Kellnhofer, P., Didyk, P., Ritschel, T., Masia, B., Myszkowski and K., Seidel, H. “**Motion Parallax in Stereo 3D: Model and Applications**”. ACM Trans. Graph. 35, 6, Article 176 (November 2016).

[96] Jason Geng, “**Three-dimensional display technologies**”, in Advances in Optics and Photonics Vol. 5, [Issue 4](#), pp. 456–535 (2013)

[97] Martin S. Banks, David M. Hoffman, Joohwan Kim, and Gordon Wetzstein , “**3D Displays**”, Annu. Rev. Vis. Sci. 2016. 2:19.1–19.39 (2016)

[98] <https://variety.com/2017/gaming/news/magic-leap-impressions-interview-1202870280/>

[99] Bereby-Meyer, Yoella & Leiser, David & Meyer, Joachim. “**Perception of artificial stereoscopic stimuli from an incorrect viewing point**”. Perception & psychophysics. 61. 1555-63 (1989).

[100] Nitish Padmanaban, Robert Konrad, Tal Stramer, Emily A. Cooper, and Gordon Wetzstein, “**Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays**”, PNAS 114 (9) 2183-2188 (2017).

[101] Praneeth Chakravarthula, David Dunn, Kaan Aksjt and Henry Fuchs, FocusAR: Auto-focus Augmented Reality Eyeglasses for both Real and Virtual, [IEEE Trans Vis Comput Graph.](#) 2018 Nov;24(11):2906-2916.

[102] Robert E. Stevens*, Thomas N. L. Jacoby, Ilinca Š. Aricescu, Daniel P. Rhodes , A Review of Adjustable lenses for Head Mounted Displays, [Proceedings Volume 10335, Digital Optical Technologies 2017](#)

[103] Jin Su Lee, Yoo Kwang Kim, and Yong Hyub Won, “**Time multiplexing technique of holographic view and Maxwellian view using a liquid lens in the optical see-through head mounted display**”, Optics Express Vol. 26, Issue 2, pp. 2149-2159 (2018)

- [104] Su Xu , Yan Li , Yifan Liu , Jie Sun , Hongwen Ren and Shin-Tson Wu , “**Fast-Response Liquid Crystal Microlens**”, *Micromachines* 2014, 5, 300-324;
- [105] Steven A. Cholewiak Gordon D. Love, Pratul Srinivasan, Ren Ng, Martin S. Banks, “**ChromaBlur: Rendering Chromatic Eye Aberration Improves Accommodation and Realism**” *ACM Transactions on Graphics*, 2017.
- [106] Gordon D. Love, David M. Hoffman, Philip J.W. Hands James Gao Andrew K. Kirby and Martin S. Banks, “**High-speed switchable lens enables the development of a volumetric stereoscopic display**”, 31 August 2009 / Vol. 17, No. 18 / *Optics Express* 15716
- [107] Rahul Narain, Rachel A. Albert, Abdullah Bulbul, Gregory J. Ward, Martin S. Banks and James F. O'Brien “**Optimal Presentation of Imagery with Focus Cues on Multi-Plane Displays**” *ACM Transactions on Graphics*, Vol. 34, No. 4, Article 59, 2015.
- [108] Hong Hua “**Optical methods for enabling focus cues in head-mounted displays for virtual and augmented reality**”, Proc. SPIE 10219, Three- Dimensional Imaging, Visualization, and Display 2017,
- [109] Hong Hua “**Head Worn Displays for Augmented Reality applications**” SID Display Week Seminar – short course, June 1st, 2015.
- [110] Joseph W. Goodman, “**Holography Viewed from the Perspective of the Light Field Camera**”, W. Osten (ed.), *Fringe 2013*, DOI: 10.1007/978-3-642-36359-7_1, Springer-Verlag Berlin Heidelberg 2014
- [111] Huan Deng, Han-Le Zhang, Min-Yang He and Qiong-Hua Wang, “**Augmented reality 3D display based on integral imaging**” , *Advances in Display Technologies VII*, edited by Liang-Chy Chien, Tae-Hoon Yoon, Sin-Doo Lee, Proc. of SPIE Vol. 10126 (2017)
- [112] Byoungho Lee, Jong-Young Hong, Changwon Jang, Jinsoo Jeong, and Chang-Kun Lee, “**Holographic and light-field imaging for augmented reality**”, *Emerging Liquid Crystal Technologies XII*, edited by Liang-Chy Chien, Proc. of SPIE Vol. 10125 (2017)
- [113] Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke and Henry Fuchs, “**Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses using Defocused Point Light Sources**” , *ACM Transactions on Graphics*, Vol. 33, No. 4, Article 89, 2014
- [114] Brian T. Schowengerdt and Eric J. Seibel , “**True 3-D scanned voxel displays using single or multiple light sources**” *Journal of the SID* 14/2, 2006
- [115] Hagen Stolle, Jean-Christophe Olaya, Steffen Buschbeck, Hagen Sahm, and Armin Schwerdtner, “**Technical solutions for a full-resolution auto-stereoscopic 2D/3D display technology**”, *SPIE Proceeding Stereoscopic Displays and Applications XIX*, Volume 6803, 2008.
- [116] Gang Li, Dukho Lee, Jeong Youngmo, and Byoungho Lee , “**Fourier holographic display for augmented reality using holographic optical element**”, *Advances in Display Technologies VI*, edited by Liang-Chy Chien, Sin-Doo Lee, Ming Hsien Wu, Proc. of SPIE Vol. 9770, (2016)
- [117] Deniz Mengü, Erdem Ulusoy, and Hakan Urey, “**Non-iterative phase hologram computation for low speckle holographic image projection**”, *Optics Express* 4462, 7 Mar 2016, Vol. 24, No. 5.
- [118] J.-S. Chen and D. P. Chu, “**Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications**”, *Optics Express*, 13 Jul 2015 | Vol. 23, No. 14
- [119] Joseph Goodman, “**Speckle Phenomena in Optics: Theory and Applications**”, Roberts and Company Publishers, ISBN-13: 978-1936221141 (2007)
- [120] Ozan Cakmakci, Yonggang Ha and Jannick P. Rolland, “**A Compact Optical See-through Head-Worn Display with Occlusion Support**”, *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004)*
- [121] J. P. Rolland, R. L. Holloway, and H. Fuchs, “**A comparison of optical and video see-through head- mounted displays**,” Proc. SPIE 2351, 293–307 (1994).
- [122] Ariela Donval, Noam Gross, Eran Partouche, Ido Dotan, Ofir Lipman and Moshe Oron, “**Smart filters - operational HMD even at bright sunlight conditions**”, *Proceedings Volume 9086, Display Technologies and Applications for Defense, Security, and Avionics VIII; and Head- and Helmet-Mounted Displays XIX* (2014)
- [123] Kay M. Stanney, Ronald R. Mourant and Robert S. Kennedy, “**Human Factors Issues in Virtual Environments: A Review of the Literature**”, *Presence*, Vol. 7, No. 4, August 1998, 327–351.
- [124] Alex D. Hwang “**Instability of the perceived world while watching 3D stereoscopic imagery: a likely source of motion sickness symptoms**”, *i-Perception* (2014) Vol. 5, pp 515-535.
- [125] Rachel Albert, Anjul Patney, David Luebke, and Joohwan Kim “**Latency Requirements for Foveated Rendering in Virtual Reality**” *ACM Trans. Appl. Percept.* 14, 4, Article 25 (2017).
- [126] Michael J. Gourlay et al. “**Head Mounted Display Tracking for Augmented and Virtual reality**”, *SID Magazine*, January-February 2017, pp 6-10.
- [127] P. Zanuttigh et al. , “**Time-of-Flight and Structured Light Depth Cameras, - Chapter 2: operating principles of structured light depth cameras**” Springer International Publishing 2016, 43., DOI 10.1007/978-3-319-30973-6_2.
- [128] Matti Laukkanen, “**Performance Evaluation of Time-of-Flight Depth Cameras**”, Aalto University, Finland, Master Thesis, 2015.

[129] F. Lu and E. Milios, “**Globally Consistent Range Scan Alignment for Environment Mapping**” *E. Autonomous Robots* (1997) 4: 333.

[130] C. S. Bamji *et al.*, “**Mpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3µm global shutter pixels and analog binning,**” *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, San Francisco, CA, 2018, pp. 94-96. doi: 10.1109/ISSCC.2018.8310200

[131] Aykut Cihangir, Will G. Whittow, Chinthana J. Panagamuwa, Fabien Ferrero, Gilles Jacquemod, Frédéric Ganesello, and Cyril Luxey, “**Feasibility Study of 4G Cellular Antennas for Eyewear Communicating Devices**”, *IEEE Antennas and wireless propagation letters*, Vol. 12, 2013 .