Digital Optical Elements and Technologies (EDO19): Applications to AR/VR/MR

Bernard C. Kress
Microsoft Corporation, Hololens – 1 Microsoft Way, Redmond WA 98052 USA

ABSTRACT

This paper serves as an introduction to the EDO19 conference "Digital Optical Technologies" dedicated specifically to digital optics for AR/VR and MR Head Mounted Display (HMD) systems, in all their incarnations: Display, Imaging and Sensors. We will introduce the field of AR/VR/MR optics historically, and review the most important criteria for AR/VR/MR optical design: matching the performances of the optics to the ones of the human visual system. This can be done for all Digital Optical Elements reviewed throughout this conference: free space optics, waveguide optics (waveguide combiners), diffractive, holographic and metasurface optical elements, tunable, switchable and reconfigurable optics, and novel computational techniques for display and imaging.

Keywords: Digital Optics, Virtual reality, Augmented reality, Mixed Reality, Smart Glasses, Waveguide Combiner, Diffractives Optics, Holographic Optics.

1. INTRODUCTION

Defense has been the first application sector for AR and VR, as far back as the 1950s [1]. Based on such early developments, the first consumer VR/AR boom has expanded in the early 90s and contracted considerably throughout the decade, as a poster child of a technology endeavor ahead of its time, and also ahead of its markets [2]. However, 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 VR boom in the 90s. 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].
This is very similar to the demise of the early iPod developments at General Magic Corp. in the late 90s, which provided an initial hardware concept however without wireless connectivity or on-line music store. Once all three conditions were met at Apple 10 years later (hardware, WIFI connectivity, on line music store), the iPod concept resonated with a strong consumer market, as a music experience ecosystem, not solely as a desirable music playing product, no matter how excellent it might have been designed and produced.
No matter how good a consumer AR/MR headsets can be, if it is nto shipped with its corresponding "iTune Store App", the consumer market will not take off.
The only AR/VR sector that saw sustained efforts and developments throughout the next decade was the defense industry (flight simulation and training, Helmet Mounted Displays -HMD- for rotary wing and Head Up Displays – HUD- for fixed wing aircrafts) [10]. The only effective consumer efforts during the 00s has been in the field of automotive HUDs and personal binocular headset video players.
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 2012 the initial implementations of immersive AR/VR HMDs based on readily available smart-phones display panels (LTPS-LCD, IPS-LCD, AMOLED) and pico-projectors microdisplay panels (HTPS-LCD, mu-OLED, DLP, LCoS), and IMUs, camera and depth map sensors (structured light and TOF). 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 (inorganic mu-LED panels, 1D scanned arrays, 2D laser/VCSEL MEMS scanners, …).

Digital Optical Technologies 2019, edited by Bernard C. Kress, Peter Schelkens, Proc. of SPIE Vol. 11062, 1106222 · © 2019 SPIE · CCC code: 0277-786X/19/$21
doi: 10.1117/12.2544404

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 early product integrators. Such traditional display technologies will serve as an 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, DLP and MEMS scanner digital projectors, to iLED smart watches (See **Figure 1**).

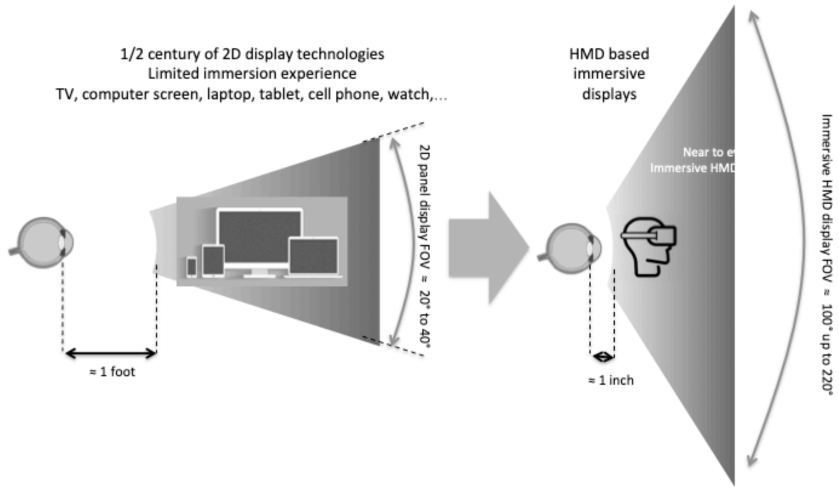

Figure 1: Immersive NTE displays: a paradigm shift in personal information display

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. SMART GLASSES/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**).

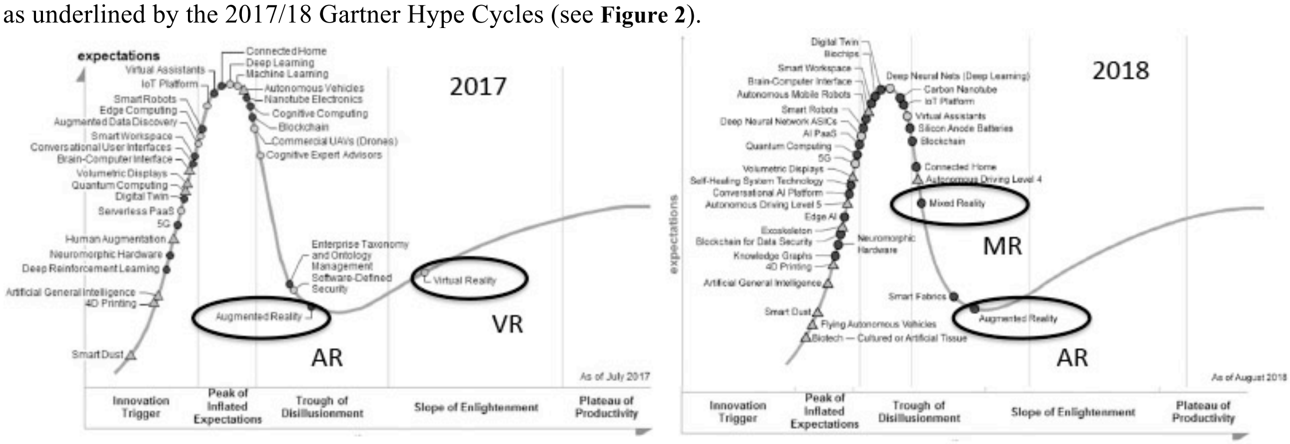

Figure 2: Gartner Hype Cycle showing VR, AR and MR technologies for 2017-2018.

The 2017 Gartner graph shows clearly AR and VR poised to reach the plateau of productivity within 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 new employees, lesser errors, higher yields*
- *collaborative design, remote expert guidance, better servicing, 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 or North “Focals”). 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 (Merged Reality) project Alloy from Intel and the ACER/StarVR wide FOV VR headset. However, the potential of video see through MR remains strong in the long term.
2018 also saw major shut downs of medium sized AR headset companies such as MetaVision Corp (Meta2 MR headset), CastAR Corp and ODG Corp (ODG R8 and R9 glasses), after all three had a strong initial product introduction and VC capital support. Others went through major re-structuration such as Avegant Corp. (Multifocal AR headset). MetaVision and CastAR saw also a recent rebirth in mid 2019, showing that the field is still uncertain and full of surprises. Others (Vuzix Corp) saw continuous growth as well as continuous VC backing (Digilens Corp) throughout 2019.
On the smart glass front, audio only smart eyewear has made a strong return. Audio smart glasses, providing audio immersion as well as world locked audio (based on IMU solely) are not new concepts, but have recently shown upgrades such as surround sound leaving the ear free (no bone conduction) and noise cancelling features towards the world (such as the Bose Frames). They can provide essential input and commands for both consumer and enterprise, and are an essential part of the augmented world experience. Major companies as Huawei have introduced their own version of audio augmented reality smart glasses recently. Camera glasses such as the Snap Spectacles (1^{st} and 2^{nd} generations) have had a tough time to show consumer acceptance, as did Google Glass Explorer version.
In addition to spatial world locked audio, if an IMU is present (as in the Bose AR Frames), a number of different head and body gestures that can also be detected including push-ups, squats, nod, shake, double tap, look up, look down, spin around, roll head around.
Table 1 summarizes the current adoption rate in consumer/enterprise/defense sectors for the different types of smart glasses/AR/VR/MR headsets available today.

			Consumer		Enterprise			Medical		Defense	
		Product examples	Daylong usage	Occasional indoor usage	Factory floor usage (shifts)	Heavy outdoor industry	R&D	Non surgical usage	Surgical usage	Training usage	Battlefield usage
Smart glasses	Audio only smart eyewear w prescription correction	Bose Frames Huawei Smart glasses	+++	+++	++	+++	+	++	++	+	+
	Rugged Smart Glasses, monocular, opaque	RealWear HMT-1 Vuzix m300	+	+	++	+++	++	+	---	---	++
	Smart Glasses, monocular, see-through	Vuzix Blade Digilens Mono HUD Optinvent ORA	+	++	++	++	++	+++	+	+	++
	Smart eyewear w display and prescription correction	Google Glass North Focals Lumus DK32	+++	+	+	---	---	+++	+	+	---
VR	Standalone VR without video see through (3DOF)	Oculus GO Google DayDream VR Samsung Gear VR	---	+++	---	---	+	---	---	+++	---
	Standalone VR with video see through	Oculus Quest NTC Vive Focus 2.0 Pico Neo	---	+++	+++	+	+++	+	+++	+++	---
	PC tethered VR with inside out sensors (6DOF)	Oculus Rift "s" HTC Vive Pro Windows MR 3rd party	---	+++	-	---		---	+	+++	---
	Large FOV PC tethered VR headsets	Varjo VR Foveated PiMax 8k Acer Star VR	---	+	-	---	+++	---	+	+++	---
AR	Tethered AR headsets to PC	Meta 2 DreamWorld Glasses	-	+	++	---	++	--	+++	++	---
	Standalone AR headsets	Epson Moverio Lumus DK50 / Vision Digilens Cristal	-	+++	++	++	++	++	+	+	++
	Standalone AR headsets w 6DOF and gesture sensing	ODG R9 nReal AR glasses Daqri, Atheer Labs,	--	++	+++	+++	++	+++	+	++	+++
MR	High end see through untethered MR	HoloLens V1 / V2	---	+	+++	++	+++	+	+++	+++	+++
	Pod-tethered high end see through MR	Magic Leap One Lenovo ThinkReality	---	++	++	+	+	+	++	+	+

Table 1: Current AR/VR/MR/Smart Glasses product offerings in consumer/enterprise/defense markets.

Small form factor Smart Glasses including minimal display (around 10 deg FOV monocular) and prescription correction saw a renewal in 2018 (after the 2014 Google Glass failure in the consumer market sector) with the hyped project Vaunt at Intel. However, this project has been halted later that year, as Intel invested rather in a very similar architecture the "Focals" developed by North Inc. in Kitchener, Canada. But after a price drop of nearly 50% early 2019, and a strong workforce layoff by North, the short term outlook for consumer smart glasses remains uncertain. Other smart glass concepts targeted solely to enterprise sectors have had a quiet but steady growth, such as the rugged RealWear headsets (Vancouver, Canada) and the more stylish Google Glass Enterprise V2 glasses.

On the other hand, the current VC investment hype fueling frenetically single start-ups such as Magic Leap Inc. (totaling >$3B VC investment pushing up a >$7B company valuation before seeing any revenues) is a harsh reminder of the ever present "Fear Of Missing Out" behavior from late stage investors (Alibaba, Singapore Temasek and Saudi funds) eager to jump on the bandwagon fueled by the early investment decisions from major tech VC firms (Google ventures, Amazon, Qualcomm). It is also interesting to note that the two last VC investments in Magic Leap (late 2018 and mid 2019) were by major communication companies (ATT/USA for an unknown amount and NTT/Docomo-Japan for $280M), pointing out that large bandwidth communication channels (5G, WiGig, etc…) will be fueled by future demanding AR markets, allowing also sustained ROI over MR cloud services (AR/MR hardware returns being razor thin).

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 Q4 2018 30% 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 Google Glass in **Figure 3**). This concept is very different from either 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 head tracking, and lack of binocular vision reduces the display to simple overlaid 2D text and images. Monocular displays do not require rigidity in the frames as a binocular vision system would (to reduce horizontal and vertical retinal disparity that can produce eye strain). Many smart glass developers provide also prescription correction as a standard feature ("Focal" by North or Google Glass V2).

The combination of a strong connectivity (3G, 4G, WiFi, Bluetooth) and the addition of a camera makes it a convincing companion to a smart phone, for contextual display functionality or as a virtual assistant, GPS and social network companion (thanks to the camera). 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**). Early 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 (Windows MR headsets such as the Samsung Odyssey). 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 (3DOF-IMU) and the HTC Vive Focus, able to seduce a burgeoning VR consumer market base. More recently, further extensions of standalone VR headsets with inside-out sensors led to 2019 products such as the Oculus Quest (6DOF standalone VR headset).

However, tethered high end VR headsets with inside-out sensors have also been updated, by both Oculus and HTC providing in 2019 high end products such as the Oculus Rift "S" and the HTC Vive Pro. Other updates in Windows MR headsets have been done by Samsung "Odyssey 2019 with double the resolution of the first 2017 version).

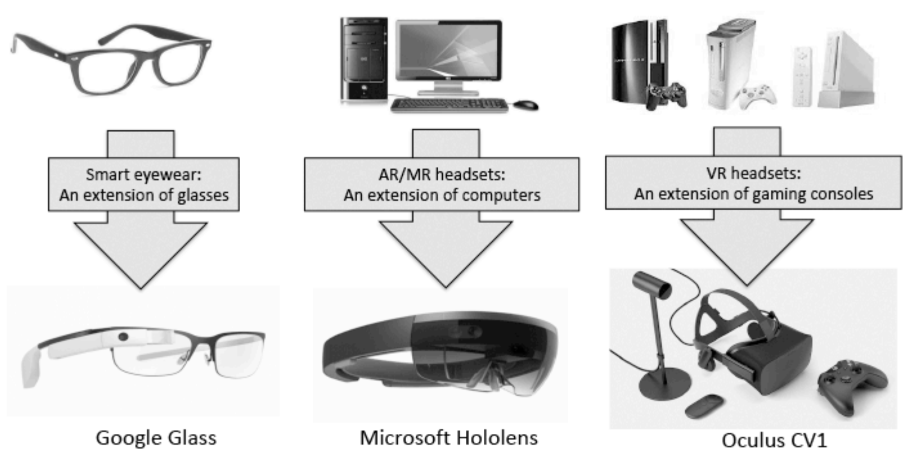

Figure 3: The emergence of Smart Glasses, AR/MR and VR headsets.

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 V1 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. Eventually, if technology permits, these three categories will merge into a single hardware concept. This will however require 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 in Q3 2018 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 optimized sensor fusion and low latency displays)*
- *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].

4.1 Wearable and visual 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 or magnifier glass or binocular telescope vision.*

4.2 Display 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.

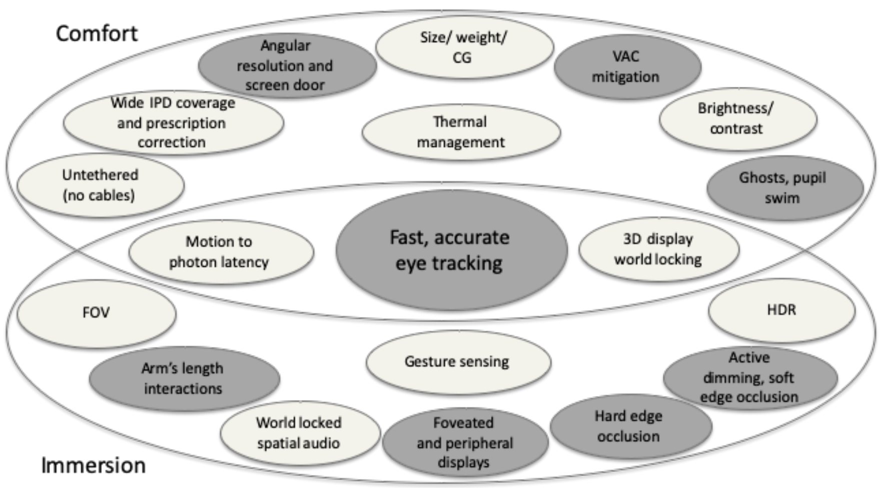

Figure 4: Comfort and immersion requirements for the ultimate MR experience.

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

4.3 Presence

We saw that immersion is a multisensory illusion for the HMD wearer (display, audio, touch, and more). Presence in MR is a state of consciousness, in which the HMD wearer truly believes he or her is in a different environment. Immersion produces the sensation of Presence.
However, in order for the Presence sensation to be vivid, various key factors need to be addressed and solved in the HMD, not only in the display (refresh rate, FOV, angular resolution, VAC mitigation, hard edge occlusion, optical foveation, HDR,…) but also in the sensor fusion process over the various sensors used in such devices (6DOF head trackers, depth map sensors, spatial mapping sensors, eye/pupil/vergence trackers, gesture sensors, etc…). The goal of MR is to create a high degree of Presence, to make the participants believe that they are really in another (virtual) environment.

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].

5.1 The human visual system

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.

Line of sight and optical axis

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.

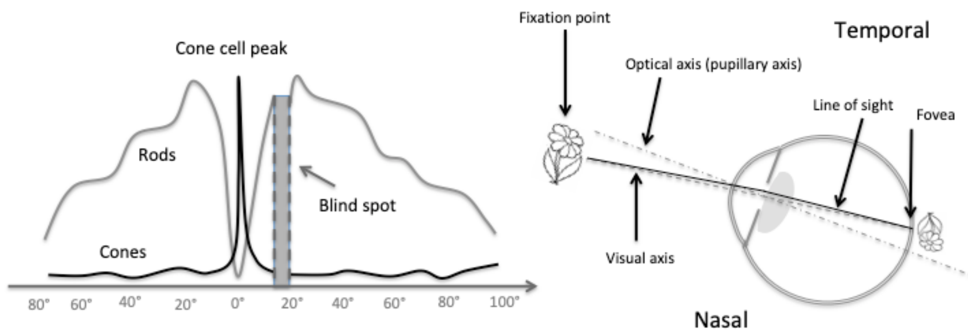

Figure 5: Rod and Cone cells density on the retina (left), optical axis and line of sight (right).

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].

Lateral Chromatic Aberrations or LCA

Chromatic aberrations result in the separation of colors through Fresnel rings, gratings or traditional refractive lenses, all of which are dispersive. The "L" in LCA reflects both "lateral" and "longitudinal" chromatic spread, where different colors are focusing at different depths, depending on the Abbe V number of the lens (refractive elements having the opposite dispersion of diffractive elements). Reflective optics do not produce LCA, and are therefore used extensively in AR display applications.

Correcting for LCA is usually done in software, by pre-compensating each color frame in a field sequential mode or pre-compensating the entire color image in an RGB display panel (equivalent of having three distortion map compensations, one for each color). However, this can lead to display artifacts such as color aliasing, and needs high angular resolution to achieve a good effect. Optical dispersion compensation is a better way, but requires also more complex optics (hybrid refractive diffractive) or symmetric coupling architectures such as in waveguide combiners using grating or holographic couplers, or replacing refractive optics by reflective optics.

As the main detector is here the human eye, it is also interesting to analyze the natural LCA spread of the eye, which is surprisingly strong. **Figure 6** shows the measured LCA of the human eye (left, as an aggregate of measurements from the past 50 years), yielding a 2 diopters spread over the visible spectrum, and how color images appear on the retina, on axis and off-axis (center), and how they appear as "seen" or "experienced" by the viewer after being processed by the visual cortex (right).

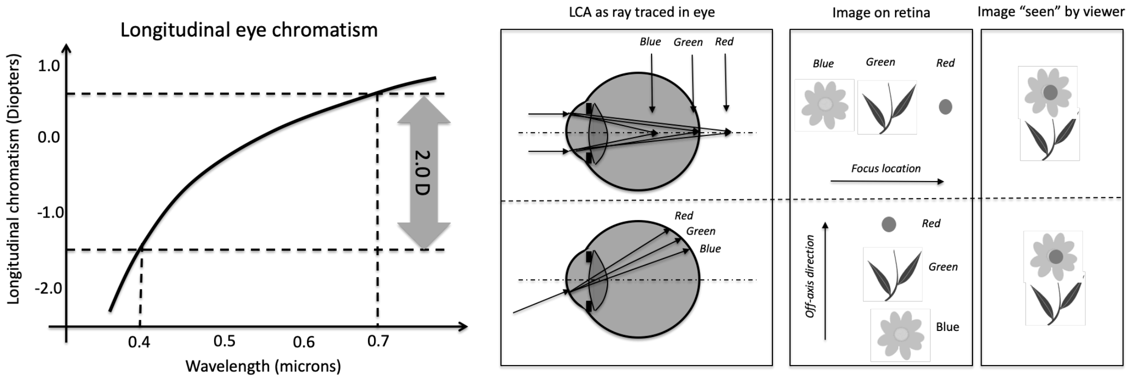

Figure 6: Natural LCA of the human eye

The natural LCA of the human eye is also the basis of an interesting digital rendering 3D depth cue, called "Chromablur". The blue and red blur (optical or rendered), depending on which side of a white object they appear, can tell a far defocus from a near defocus, and thus inform the eye's oculomotor muscles on the sign of the accommodation power change to get the image back in focus (that is, the green part of the image focused on the retina).

The human eye LCA varies slightly from one individual to the other: therefore, changing slightly the natural LCA by increasing or decreasing the spectral spread with external optics will not affect dramatically the visual acuity. However, if part of the field has a specific LCA (see through in AR), and the other part has a different LCA (digital image in AR), this could lead to some visual discomfort.

Visual acuity over fovea and peripheral region

The measured polychromatic MTF of the human eye is plotted in **Figure 7**. This MTF represents mainly photopic vision over the on-axis field (optical axis of the eye or pupillary axis), which is close to the line of sight field centered on the foveated area on the retina (around 3-5 degrees offset).

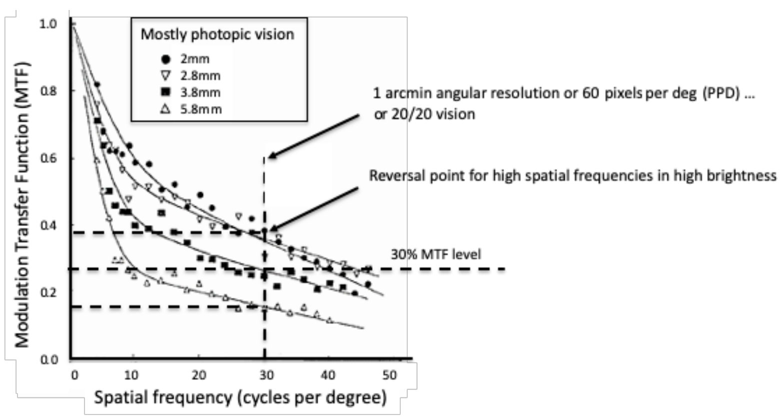

Figure 7: Polychromatic Modulation Transfer Function (MTF) of the human eye for various pupil diameters

The ability of the eye to resolve small features is the "visual acuity." The teenage and early adult human eye can distinguish patterns of alternating black and white lines as small as one arcmin (or 30 cycles per degree or 60 PPD). That is also the definition of 20/20 vision. A few of us might distinguish smaller patterns (having a higher MTF at these cycles per degree), but most of us will see these patterns as grey shades (having and low MTF, below 0.1).

For all pupil sizes, the MTF at 20/20 for photopic vision is higher than 30%, only for scotopic vision for pupils over 5mm would the MTF drop to lower levels.

Note that MTF50 (50% of the MTF for low frequency) or MTF50P (50% of MTF from the peak value) are good indicators, but rather used for cameras. The human eye especially as it moves constantly, can still distinguish features

well in the 30% MTF levels. Unlike cameras, the eye's MTF also drops for very low frequencies (mainly due to lack of movement at such low frequencies).
Due to high aberrations in the human eye, the higher the pupil size the lower then resulting MTF, which is the opposite to diffraction limited high end camera objectives where the MTF increases with the aperture. It is interesting to note that higher spatial frequencies are actually distinguishable when the pupil increases in photopic vision, allowing over 20/20 vision over lesser bright fields.
This said, the human vision system cannot be limited to its pure optical properties as a simple camera would, but as a computational imaging system (where the CPU is the visual cortex). The impressive LCA discussed in the previous section is a testimony to this effect.

Specifics of the human vision FOV

Figure 9 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].

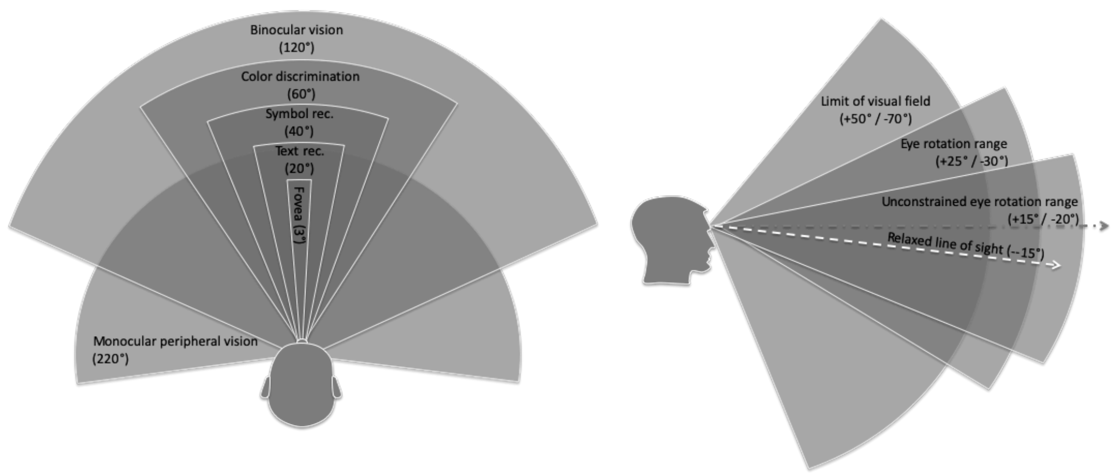

Figure 8: Human vision FOV (H and V)

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 vision 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 9**).

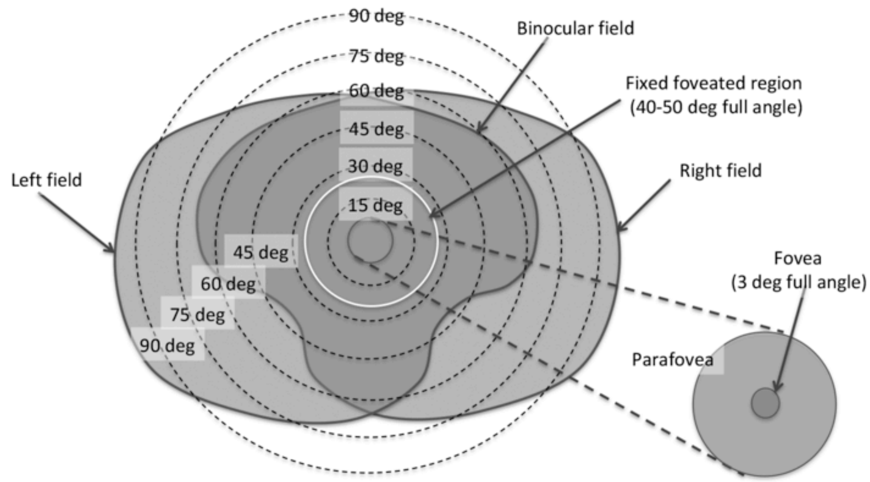

Figure 9: Human binocular field of view with fixed foveated region including unconstrained eye motion allowing for sustained gaze and accommodation.

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.

5.2 Adapting display hardware to the human visual system

Various FOVs from existing HMDs are shown in **Figure 10**. 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 AR, NorthStar Leap Motion AR), to smaller FOV high end AR/MR systems with microdisplay panels such as Microsoft HoloLens V1 and Magic Leap One. Smart glasses have typically D-FOVs starting from 10-15 deg (Zeiss "Tooz" Smart Glasses, Google Glass, North Focals) to larger FOV starting at 25 deg, up to 50 deg D-FOV (Vuzix Blade, Digilens, Optinvent ORA, Lumus DK50, ODG R9).

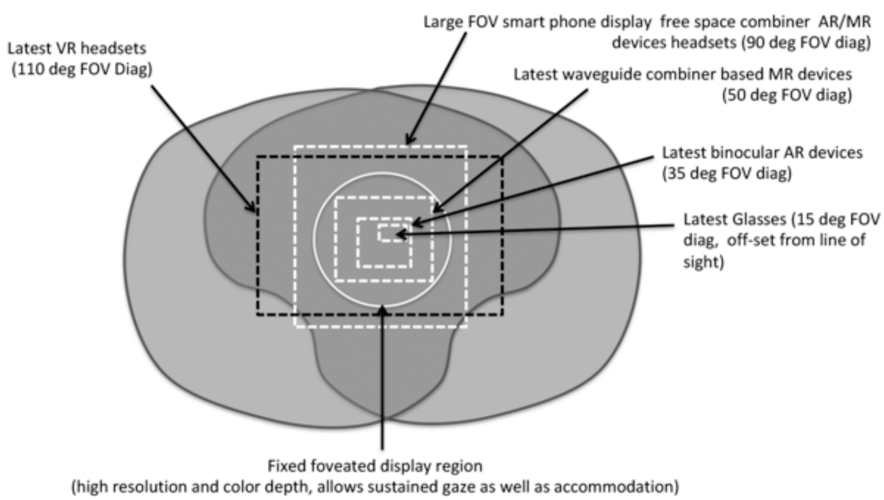

Figure 10: Typical FOVs for current state of the art Smart glasses, AR, MR and VR headsets, overlaid on the human binocular vision and the fixed foveated display region.

One other way to describe the FOV experience is to overlap the amount of unobstructed see though FOV over the actual display FOV (see **Figure 1**). The fixed foveated region (scanned through eye movements) is shown in dotted lines.

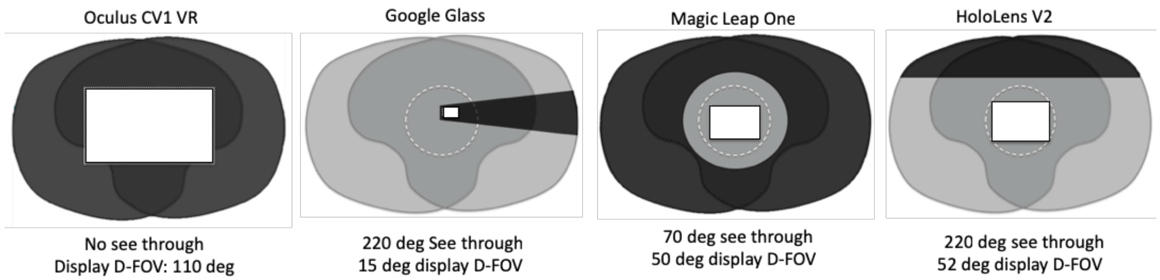

Figure 11: Display FOV and see through FOV for various VR, smart glasses and AR headsets.

For a VR system, there is no see through, and thus the display FOV can be quite large, 110 D-FOV to 150 D-FOV (left), up to 200+ FOV with products from PiMax or StarVR. For a smart glass (Google Glass, center left), the see through – or rather see around- experience is wide, only hindered by the lateral display arm, with an ex-centered display D-FOV of 15 deg. For 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 about a 70 deg circular cone, while the display has a D-FOV of 50 deg. For Hololens V2 (right), the lateral see-through (or see-around) FOV, is equal to the natural juman FOV of 220 deg, with a display diagonal D-FOV of 52 deg., covering most of the foveated region; only the top part of the FOV is capped by the mechanical enclosure holding the sensor bar, the laser/MEMS display engines and the system board. There is no limitation to the bottom FOV. One other particularity of HoloLens V2 is the ability to move the entire display visor up to reveal a totally unobstructed view.

In order to optimize an HMD optical architecture for large FOV, the various regions of the human FOV described in **Figure 9** 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

5.3 Perceived resolution and FOV

The bottom line for an AR/VR system is the perceived FOV and perceived resolution by the human visual system: considering a human centric system design, in which resolution is rather a perceived spec (subjective) than a scientifically measured spec.

For example, one way to increase the perceived resolution in an immersive display without increasing the GPU rendering burden is simply to duplicate the pixels on the physical panel side. This has been done in the latest version of the Samsung Odyssey Windows MR headset (2018 version), in which the display pipeline renders and drives the display at 616ppi whereas the resulting physical display shows 1233ppi. This has been demonstrated to reduce the screen door effect and increase the user's perceived resolution.

Perceived FOV span 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 a similar immersive display architecture which would however boast 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.

6. CONCLUSION

This paper served as an introduction to the EDO19 conference "Digital Optical Technologies" dedicated specifically to digital optics for AR/VR and MR systems, in all their incarnations: Display, Imaging and Sensors. We introduced the field of AR/VR/MR optics historically, and reviewed the most important criteria for AR/VR/MR optical design: matching the performances of the optics to the ones of the human visual system. This can be done for all Digital Optical

Elements reviewed throughout this conference: free space optics, waveguide optics (waveguide combiners), diffractive, holographic and metasurface optical elements, tunable, switchable and reconfigurable optics, and novel computational techniques for display and imaging.

7. 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, Scientiic 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)