University of Arizona

College of Optical Sciences

OPTI 909 – Master's Report

Professor Jim Schwiegerling

Academic Advisor: Dr. Hong Hua

Advancements in 3D Display Technologies for Single and Multi-User Applications

By: Alex Lyubarsky

Table of Contents

I. Preface/Introduction

The trend in 3D displays has grown exponentially in the past couple of decades due to the many advantages over a two-dimensional display. Since humans live in a three-dimensional world, it is only natural to want to create visual content that mimics our daily experiences. With recent advancements in display technologies, 3D displays have transitioned from being a fantasy to becoming more of a reality. As Allied Market Research predicts, the 3D display market is expected to reach $112.9 Billion globally by 2020 [1]. The most common displays that create a 3D visual experience can be found today in a local movie theater. Known as stereoscopic 3D, which although creates a fun visual experience, is also known to be uncomfortable especially over longer viewing times. The discomfort comes from what is known as the vergence-accommodation conflict where eyes are focused onto the screen while being crossed at a different point. While stereoscopic 3D has dominated in the 3D market, the uncomfortable viewing experience has driven researchers to find a better solution. Recently there has been much advancement made in the fields of autostereoscopic or automultiscopic displays from multi-view to super multiview. These technologies have also paved the way for what is known as a light field display that mimics that natural way the human visual system operates. Some other notable mentions of 3D displays include volumetric and holographic display but are still difficult to realize for the consumer market due to its complexity and size. While these technologies are gaining a lot of momentum for multi-user applications, we cannot forget about all the recent work and emerging technologies in head mounted single-user applications. Virtual reality has gained a lot of interest from the consumer industry ever since Facebook’s Oculus Rift or the Samsung Gear VR products were released. The issues with existing virtual reality products are similar to that of stereoscopic 3D displays. However, there have been some breakthroughs recently made in light field virtual reality displays. Similarly light field has been applied to augmented reality or what has recently been termed mixed reality.

This report will begin by introducing the functions of the human visual system. It is important to understand the basic operations of the human eye but more specifically how the human visual system is responsible for the perception of 3D. This will help us understand the issues with current stereoscopic 3D display and why the vergence-accommodation creates an uncomfortable viewing experience. The report will continue to delve deeper into 3D displays more specifically multi-view and super multi view autostereoscopic concepts which approach the human visual system functions to reduce or completely eliminate the vergence-accommodation conflict. The natural way humans view three-dimensional content is the ideal 3D display as is seen in the latest trend of light field displays. For completeness we will discuss some advancement made in holographic and volumetric displays as well and the challenges lying with such technologies. The report then will transition into single-user applications discussing technologies that create 3D in head mounted applications of virtual reality and augmented reality or mixed reality.

I would like to take this opportunity to thank the College of Optical Sciences at University of Arizona for the support of my distance learning master's studies in Optical Sciences from 2012-2017. I would like to thank Professor Jim Schwiegerling for guidance through the course of this master's report. I would also like to thank Dr. Hong Hua for being a great mentor and advisor throughout my years in the program.

II. Human Visual System

The human visual system which consists of the eyes and the brain is part of the central nervous system responsible for processing information viewed by the human eyes. The eyes are organs that convert viewed information into neural impulses sent to the brain. It is important to understand how the human visual system operates to determine the specifications required for a display more specifically a 3D display. Perception of color, spatial, depth, and temporal resolution are all functions of the human visual system.

The human eye consists of several components that are synonymous of a camera lens and sensor. With that being said, light entering the cornea gets constrained by the pupil or iris and is then focused by the crystalline lens onto the retina which acts like an imaging sensor. The retina is the main component which converts the focused image into an electric signal to be sent to the brain via the optic nerve. Photoreceptor cells on the retina, known as rods and cones, are responsible for low intensity light levels with low spatial acuity and high light levels with high spatial acuity and color perception, respectively. The structure of the human eye can be seen in Figure 1 along with other components not particularly applicable to this report. [2]

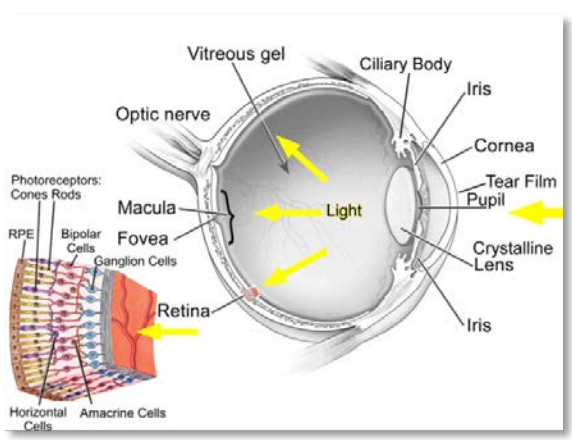


Figure 1. Structure of the Human Eye

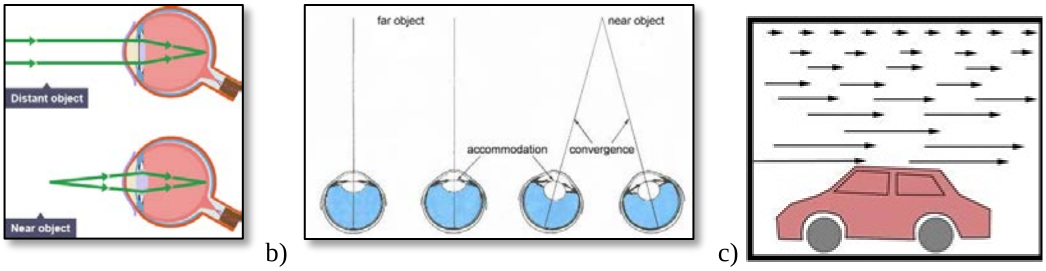
Additionally, some other important factors of the human eye for determining specifications of a display are spatial and depth resolution. As noted by D.G. Green et al, the human visual acuity is a function of the pupil size and the focal length of the crystalline eye lens. For a 4mm pupil, which is typical for daylight conditions, the human eye can resolve approximately 1 arc minute or $1/60^{th}$ of a degree. Similar calculations regarding depth of focus which are outlined in the same paper determine the human eye for 20/20 vision (1 arc minute) has a depth of focus of ±0.1 diopters. [3] However, the range of focus or accommodation range of the human eye meaning the eye can adjust focus on objects anywhere between 3 diopters and infinity.

It is also important to note the field of view capabilities of the human eye. When the eyes are in a steady fixation state, the monocular field of view extends 56° in the nasal direction (towards the nose) and 95° in the temporal direction (towards the temple). This gives a field of view of approximately 190° in the lateral direction but up to 220° when the eyes are allowed to move. The binocular field of view is the overlap regions of the monocular field of view and extends to approximately 114°. [4] The binocular field of view is more specific to the topic of this report since two eyes are required to view an object to get 3D perception.

a. Depth Cues: How We Perceive 3D

Now that we understand how the human eye accepts incoming light, it is important to understand how the brain interprets this information. The human visual system perceives the 3D information by various depth cues which can be characterized as physiological and psychological depth cues.

Oculomotor depth cues are a type of physiological depth cues that are based on the ability to determine the position of our eyes and sense the tension in the eye muscles. These depth cues are primarily due to convergence and accommodation of the human eye. Convergence is the angle created when the two eyes cross inwards to view an object up close. Similarly for far objects, our eyes tend to move outwards, which known as divergence. This is sometimes known as crossed or uncrossed viewing. Convergence is typically associated with accommodation. For example, when viewing a nearby object and the eyes are crossed inwards, the muscles which hold the crystalline lens tighten to change the focal length of the lens to focus the nearby object onto the retina. [5] Accommodation can also provide some depth information due to defocus where our brain can estimate distance of objects based on the blur. [6] It can be said that convergence and accommodation of the human visual system is always met. Note, convergence is a binocular depth cue, meaning requires both eyes to function, whereas accommodation is a monocular depth cue since each eye has the ability to focus on its own (i.e. if one eye is covered). Binocular disparity or stereopsis is another type of physiological depth cue which is a function of the horizontal separation of the human eyes known as the interpupillary distance. According to the Dictionary of Optometry, males have an average interpupillary distance of 64mm while females on average have eyes separated by 62mm [7]. This separation provides each eye with an image with slightly different perspectives. It is up to the brain to then fuse the images together to provide a sense of depth perception [8]. This binocular disparity or stereopsis happens to be the most important depth cue needed to create a 3D display. This leads us to motion parallax which is the last physiological depth cue. Motion parallax is a monocular depth cue but can also be a binocular depth cue in which the eyes distinguish between slow and fast moving objects in a scene. Far away objects within a scene will appear to move much slower than closer objects that would appear to move fast. Physiological depth cues are illustrated in Figure 2.

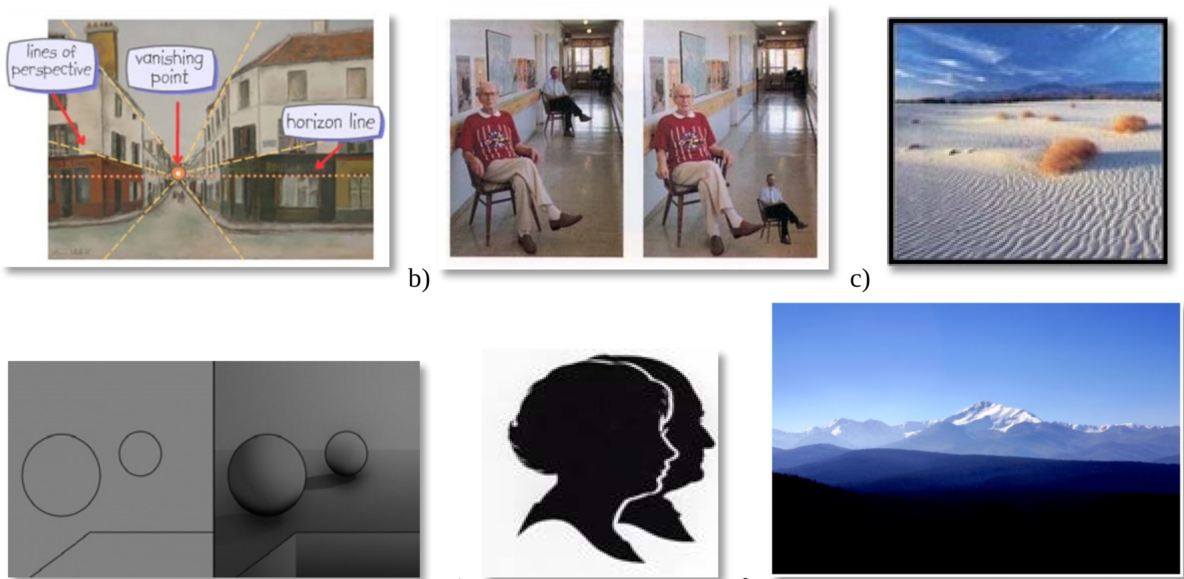


a) b) c)

Figure 2. Physiological Depth Cues
a) Accommodation[9], b) Convergence/Divergence[10], c) Motion Parallax[11]

.

There are quite a few psychological depth cues which all are considered monocular depth cues and are learned through everyday life [8]. Linear perspective is when parallel lines in a scene appear to converge to a point the further away they are from the viewer. Relative size is when our brain can estimate the distance of an object by comparing the perceived size to the already known size of the object. Texture and gradient can provide depth information by showing finer detail of a surface texture up close and will seem to fade at a farther distance on the surface. Similarly, shading and shadowing tells the brain objects are closer when they cast a shadow on other objects or appear closer if with shading, appear brighter than objects farther away. Occlusion or overlap provides the order of objects which can be interpreted as distance. By having an object in front of or overlapping another, the object will appear closer than the one behind it. Finally, aerial perspective provides distance information to the brain by showing distant scenes as hazy or faded. These psychological depth cues are more easily understandable in the illustrations shown in Figure 3.



a) b) c) d) e) f)

Figure 3. Psychological Depth Cues[11]
a) Linear Perspective, b) Relative Size, c) Texture & Gradient, d) Shading & Shadowing,
e) Occlusion/Overlap, f) Aerial Perspective

III. The Rise and Fall of Stereoscopic 3D Displays

Stereoscopy can be traced back to the 1830s when Sir Charles Wheatstone invented the stereoscope using a pair of mirrors to present a different image with slightly different perspectives to the left and right eye. He concluded that the brain would recombine the two images into a three dimensional image. It was not until the late 1840s, that the stereoscope became a portable device as seen in Figure 4 which was invented by David Brewster by improving on Wheatstone's invention with lenses versus mirrors.

Figure 4. David Brewster's portable stereoscope

This concept was later applied to larger audiences by Joseph d'Almeida in 1858. By using a red and green lantern projection slide, a red and green image would simultaneously be projected each with a slightly different perspective. The audience would wear glasses with red and green filters to allow the appropriate image into each eye. Many know this technique as the anaglyph. This was later adopted by Hollywood films in the 1920s. [12] A similar technique was invented by Edward H. Land in the 1930s using polarizing sheets to selectively distinguish different images to each eye which allowed for color stereoscopic films. 3D films were featured in American cinema throughout the 1950s but it was not until the late 1980s and early 1990s that IMAX and Walt Disney pictures really revived the 3D film market with high end cinema theaters and attractions. Most of the success of 3D films existed in the 2000s, with a major spike in interest and popularity in 2010 started by the film Avatar.

As shown in Figure 5, while 3D film popularity peaked in 2010, as seen by the 3D box office sales compared to the total box office sales, the interest in 3D seemed to decline to the present day. [13]

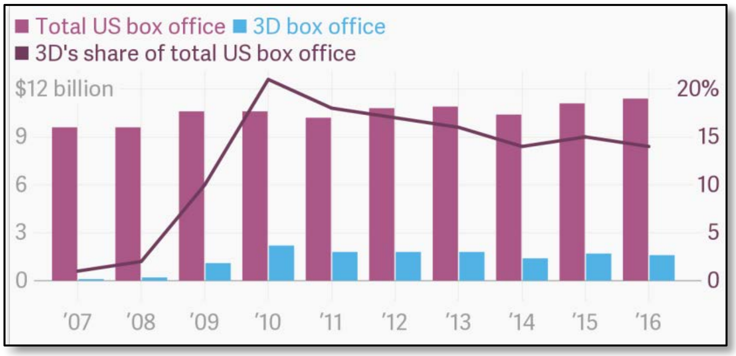
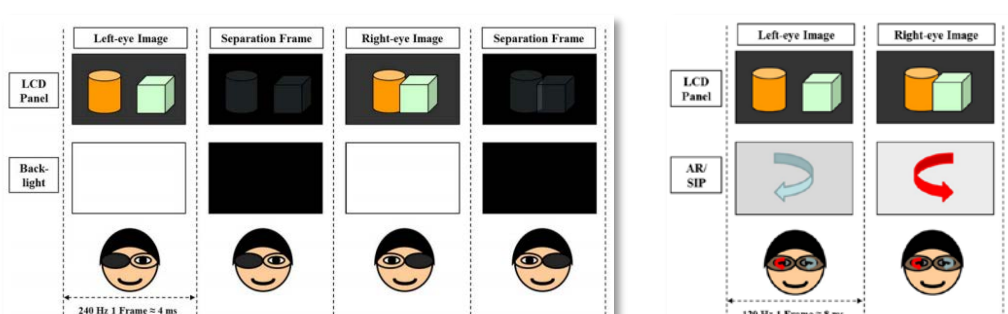


Figure 5. Trend in 3D films released vs. Box Office Sales

Most of the decline in popularity can be attributed to the visual discomfort experienced by people when viewing a 3D movie for an extended period of time. This visual discomfort has been known to cause altered eye strain, eye fatigue, headaches, altered vision, lightheadedness, confusion, nausea, and even convulsions to some [14]. To understand why stereoscopic 3D has such faults, it is important to discuss the technology a little more in depth.

Stereoscopic displays are mainly based on simulating binocular disparity depth cues by providing two separate images with slightly different perspectives to each eye. The two most common types of stereoscopic 3D displays are typically categorized by the technology of the glasses required, active liquid crystal shutters or passive polarization glasses. With active liquid crystal shutters, the glasses use two shutters that can open and close to pass or block light into the observer's eye. These are typically synchronized to the display device at a frame rate of at least 120Hz to provide a 3D image at 60Hz. It is important to have at least 60Hz otherwise image flicker can become noticeable to the human eye. During transition from left eye frame to right eye frame, there is also a dark frame used to enhance the quality or contrast of the 3D image but also reduce or eliminate any crosstalk between the left and right images due to the response of the liquid crystals. In any system if crosstalk is present, it can cause additional visual artifacts. Due to the time sequential nature of displaying the three different frames, the brightness is significantly lower of a 3D image than that of a 2D image therefore much brighter displays are necessary for 3D applications. Since active liquid crystal shutter stereoscopic 3D requires a lot of system integration and synchronization, the overall cost and size and weight of the glasses have prevented the adaptation of such a technology for the mass market. Passive polarization for stereoscopic 3D is a much simpler technology and much more commonly used. Similarly as Edward H. Land proposed in the 1930s, the display device outputs two different images one for each polarization for each eye at a fast frame rate. The glasses compose of a polarizer for each eye to only transmit the image corresponding to the particular polarization intended for the specific eye. Again, each image needs be displayed at 120Hz to provide a 60Hz 3D display. Both images have a slightly different perspective for the brain to fuse the images into 3D. Similarly to active shutter, brightness is decreased due to the time sequential nature and since a dark image may be necessary during the transition from left to right image. Due to the simplicity and lower cost as compared to active liquid crystal shutter, this technology has been commonly adopted by the mass market more specifically in 3D televisions and cinema applications. [15] A comparison of the two technologies can be seen in Figure 6.



a) b)

Figure 6. Operation of Stereoscopic 3D technologies
a) Active liquid crystal shutters b) Passive polarization

While both technologies are a unique implementation to provide a 3D experience using stereoscopy, they both suffer from the same technological limitation known as the vergence-accommodation conflict. This phenomenon is characteristic of stereoscopic displays because the content displayed is unnatural to how the human eye operates. As mentioned in section II, when viewing an object, the human eyes converge or cross to the point on the object. At the same time, the eyes focus on the object or accommodate to the object's focal distance. In stereoscopic displays, the image is always displayed at the same focal plane i.e. the screen or the display panel. However, the content displayed to create depth perception makes objects appear in front of the focal plane (or screen) or behind the focal plane causing a negative disparity or a positive disparity. This causes the eyes to converge to a different point than they are focused to. The human visual system then battles between the point of accommodation and the point of convergence which is known to cause eye fatigue. The vergence-accommodation conflict is illustrated in Figure 7.

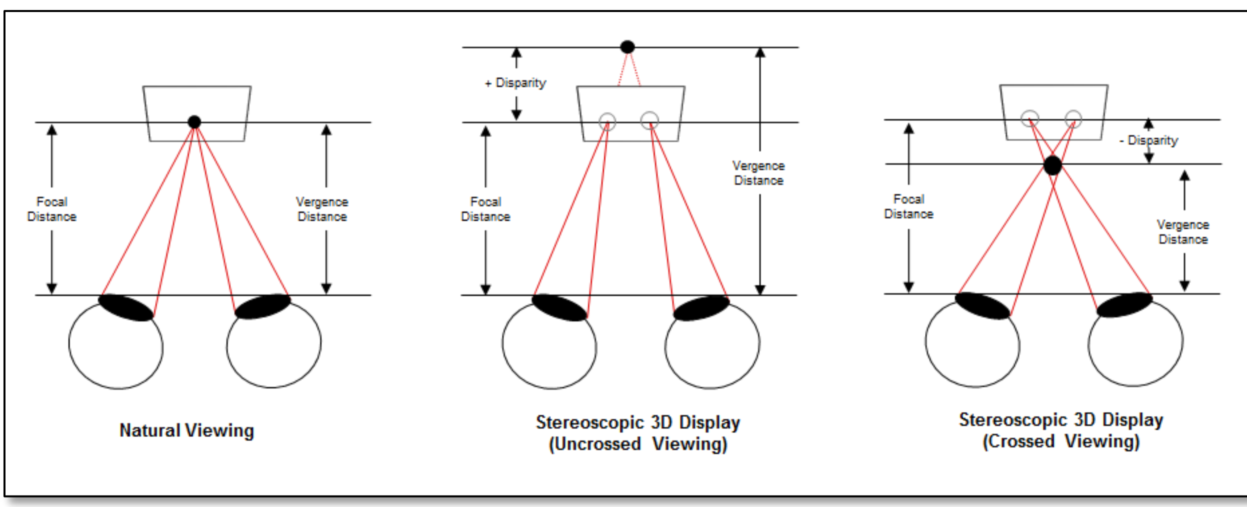


Figure 7. Natural viewing vs. Vergence –Accommodation conflicts in Stereoscopic 3D

Vergence-accommodation mismatch is not the only cause of visual fatigue in stereoscopic displays, however. The human visual system can actually tolerate some disparity as long as it is within the depth of focus. [16] However, this acceptable disparity is rather small compared to the typical stereoscopic depths to provide exciting 3D. In fact there exists a stereoscopic comfort zone where visual fatigue may not be experienced as seen in Figure 8. [14]

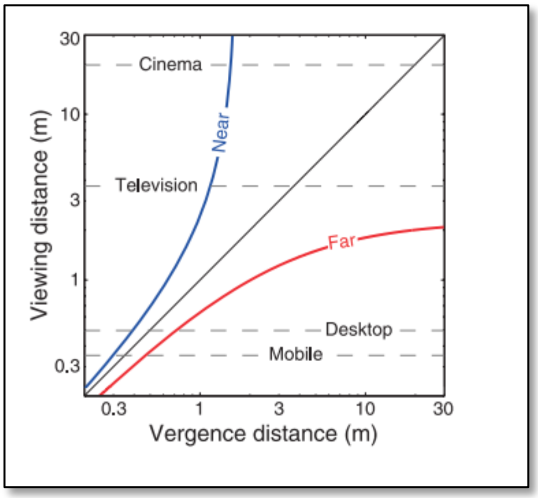


Figure 8. Stereoscopic Comfort Zone

Another cause of visual fatigue in stereoscopic displays in addition to vergence-accommodation mismatch includes distortions between left and right images or crosstalk which may lead to binocular rivalry. This occurs when each eye may see part of the image of the other eye and the human visual system battles between the two different images. [16]

With all of these aforementioned factors creating an unnatural blurring contributing to the issues experienced in stereoscopic 3D displays such as visual fatigue, it is clear as to why there has been a decline in 3D displays in the past few years. Furthermore, another main contributor to the decline in stereoscopic 3D displays is the need for some type of eyewear. With autostereoscopic displays, there no longer is a need for eyewear in addition to other added benefits depending on the type of technology used to create 3D. Other 3D display technologies such as holographic, volumetric, or light field displays all attempt to provide an alternate better 3D experience than stereoscopic displays by sampling the light field with as much angular and spatial resolution as possible to mimic the natural operation of the human visual system.

IV. What is the Light Field?

The light field is a vector representation of light rays travelling from any point into any direction in space. It was Arun Gershun in 1936 that was first to describe the vector "Light Field" as amount of light travelling in every direction through every point in space. [17] The light field was later defined in more detail by the 5D Plenoptic Function by Edward Adelson in 1991. [18] The 5D Plenoptic function is defined by:

$$L(x, y, z, \theta, \phi) \qquad \textbf{Equation 1}$$

where position is defined by *x*, *y*, and z coordinates and the direction of rays is defined by the angles θ and ϕ. Adelson even described this function as being expanded to a 7D function which includes time and wavelength parameters. Assuming light is infinite meaning travels in free space, the Plenoptic function can be simplified to a 4D light field as defined by Marc Levoy in 1996 for rendering light field content in computer graphics [19] which was also referred to by Steve Gortler as a Lumigraph [20]. The 4D light field becomes

$$L(x, y, \theta, \phi) \qquad \textbf{Equation 2}$$

defining a ray intersecting a point on plane with *x* and *y* plus the direction of the travelling ray with θ and ϕ.

To simply describe the light field, any object can be represented by "hogels" which in contrast to 2D pixels, define the position and direction and intensity of light rays. 3D displays attempt to recreate the light field by creating different views or perspectives of the 3D scene. While the light field can be considered as a continuously distributed number of views in all directions, in reality, displays can only have a finite number of views. Thus, it can be said the finite number of views of a 3D display subsample or approximate the light field as illustrated in

Figure 9. [21] It is important for the display to subsample the light field with enough views such that spatial and angular resolution matches the human visual system.

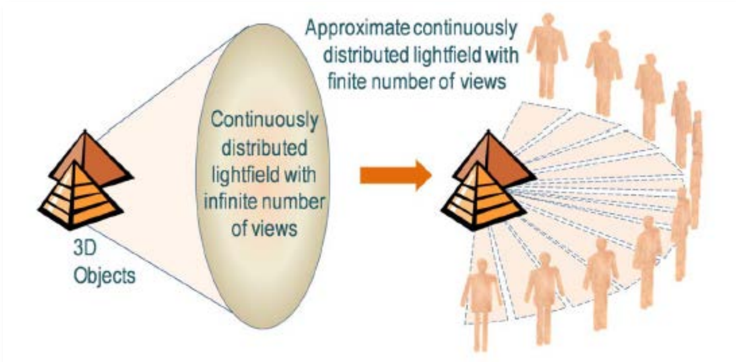


Figure 9. Finite number of views approximating the Light Field

V. <u>Various Autostereoscopic 3D Display Technologies</u>

The term autostereoscopic is typically reserved for displays that provide stereopsis based on directional-view images but can be applied to any display that does not require eyewear and can provide stereopsis. These displays are sometimes also referred to as glasses-free 3D displays. Additionally, they can be referred to as automultiscopic if producing more than two-views. An autostereoscopic display defines a display device which can reconstruct the 4D light field to reproduce the spatial radiance profile. Therefore, an ideal autostereoscopic 3D display is one that would represent the 4D light field space with infinite resolution. Compared to stereoscopic displays, where the viewer is focused on a 2D array of pixels, in the case of autostereoscopic displays, the display surface becomes a reference plane from which rays emitting from each pixel into some direction represent the angle space. [22] Therefore, the display provides both spatial and angular information.

Since infinite resolution is obviously impossible to create in an ideal 3D display, there are multiple implementations of autostereoscopic displays that have tradeoffs in spatial resolution or angular resolution yet provide a good enough quality 3D display. Some of those concepts include lenticular, parallax barrier, integral imaging, multi-view and super-multi-view autostereoscopic displays as well as multilayer displays all of which will be discussed in this section.

a. Parallax Barrier

The concept of a stereogram based on a parallax barrier was first introduced by Frederic E. Ives in 1902 but was more of a novelty gift of static images [23]. The technology became of interest almost a century later when it was implemented by Sharp and Hitachi for 3D laptop screen and 3D mobile phone screens [24]. Neither was a successful implementation of parallax barrier displays until Nintendo released the 3DS handheld gaming console which provided a 3D gaming experience in 2010 [25].

A parallax barrier is comprised of an array of slits or masks designed such that at a given viewing distance, the left eye would only see one set of alternating pixels while the right eye would see the adjacent set of pixels that would be blocked to the left eye. The display then produces one image for the pixels that are seen by the left eye and one image seen by the right eye in order to simulate stereopsis. This configuration would be the case for a two-view display which would have similar performance to that of a stereoscopic display only without the need for eyewear.

The design of a parallax barrier display is rather simple since it only requires a design of the vertical mask array.

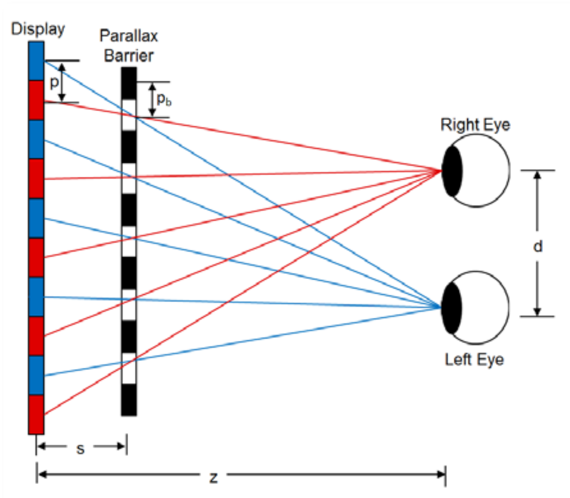


Figure 10. Basic Layout of Parallax Barrier

As illustrated in the basic layout of a parallax barrier in Figure 10, the pixel pitch, *p*, is typically known due to the display hardware used and the view width size, *d*, is based on the eye interpupillary distance, which as mentioned previously is on average 64mm. The view distance, *z*, is the distance from the display to the viewer and can depend on the 3D display application whether it is a handheld gaming console as the Nintendo 3DS or a standard TV monitor application. This leaves the only parameters to be solved which are p_b, the pitch of the parallax barrier vertical masks, and *s*, the distance of the parallax barrier mask to the display. [21] Using simple geometry, one can design a parallax barrier using the following:

$$s = \left(\frac{2p*z}{2p+d}\right) \quad \textbf{Equation 3}$$

$$p_b = 2p\left(\frac{z-s}{z}\right) \quad \textbf{Equation 4}$$

As an example, to view a 30-inch 1080p display at 1 meter distance, the pixel pitch of the display can be calculated to be 0.346mm which gives a distance of the parallax barrier mask from the display, $s=10.7mm$. Using this value, the pitch of the parallax barrier vertical masks becomes $p_b=0.685mm$. The concept of a parallax barrier display is simple when dealing with a two-view autostereoscopic display as seen above. However, they are not only limited to a two-view type display. In fact, parallax barrier can be designed for a multi-view application. In order to do so, the parallax barrier masks must now subtend multiple pixels rather than just a pixel for left and right eye. Therefore, a minor tweak to Equation 4 changes the factor of 2 to the number of views desired, *N:*

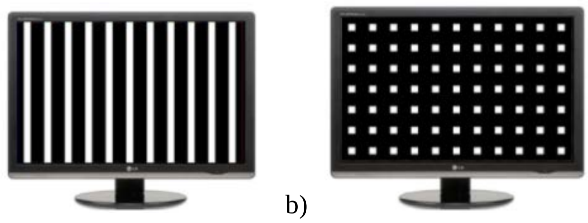
$$p_b = Np\left(\frac{z-s}{z}\right) \qquad \textbf{Equation 5}$$

From the equation, it can be seen that the area of the mask increases with an increase in number of views, thus more pixels are subtended by the parallax barrier slits.

Parallax barrier displays are commonly known to consist of vertical masks only. These types of displays are typically referred to as horizontal parallax only (HPO) displays. This means that the viewer can view a different perspective of an image as they move left and right only. Since there is no vertical parallax, meaning as the person moves up and down relative to the display a different perspective will not be seen. This is considered acceptable for many applications since rarely a user is moving up and down relative to the screen when in a seated position. Additionally, since the human eyes are horizontally separated, it is obvious to see that stereopsis in the horizontal direction or horizontal parallax is much more critical in a 3D display than in the vertical direction.

b. Pinhole Array

Parallax barrier displays are not limited to horizontal parallax only but can also be applied to a full parallax application. Full parallax provides both horizontal and vertical parallax in a display. This means movement in any direction will give a different perspective of the image to the viewer. Such displays which incorporate parallax barrier technology, are sometimes referred to as pinhole array displays. Similarly to a one-dimensional parallax barrier where a vertical slit subtends two or more pixels along the same axis, in a pinhole array display or two-dimensional parallax barrier, each pinhole subtends an array of M x N pixels. A comparison of the two types of parallax barrier displays can be seen in Figure 11. A pinhole array can be created using two rotated vertical slit masks. [22]



a) b)

Figure 11. a) 1-D parallax barrier, b) 2-D parallax barrier or pinhole array

While parallax barrier displays are a simple approach to providing glasses-free 3D, there are a number of drawbacks. Since parallax barriers use a mask array, there is a significant loss in brightness of the display since only a small amount of light passes through the slit or aperture. Resolution is also significantly reduced. This is primarily due to the number of views desired from the display. Essentially, the resolution is reduced by 1/N where N is the number of views desired. In a 1-D parallax barrier, resolution is only lost in one-axis whereas in a pinhole array, the resolution is lost in both directions. This limitation requires a significantly high pixel density from a display to achieve an acceptable 3D resolution. Additionally, these types of displays are limited to the number of views. It is merely a tradeoff between number of views (angular resolution) and spatial resolution. Some other limitations involve the viewing experience. For instance, parallax barriers due to their limited number of views may experience crosstalk or overlap where the right eye may see some of the view intended for the left eye or vice versa. This is known to create an image flipping phenomenon that confuses the viewer causing inaccurate depth perception. Additionally, dark vertical lines may appear between viewpoints due to the display pixel gaps which are referred to as the “picket fence effect.” [21]

An interesting implementation of parallax barrier has been demonstrated by New York University as described by Perlin et al. [26]. In their approach, they use a liquid crystal shutter to dynamically adjust the pitch and aperture of the parallax barrier to steer the light to the viewer. Using an eye-tracking system, they are able to determine the location of the left and right eye, and generate the views at those desired locations. The benefit is an increase in longitudinal movement and a potential increase in resolution since only two views are required in the viewing zone instead of multiple views. Nevertheless, many of the limitations of parallax barrier displays still apply.

c. Lenticular and Integral Imaging

Lenticular-based 3D displays are very similar to that of parallax barrier displays except that these displays use refraction instead of aperture masks. These displays use a technique known as integral imaging which was introduced by Gabriel Lippmann in 1908. What Lippmann referred to as "La Photographie Intégrale," the implementation required the use of an array of lenses similar to a fly's eyes in order to a photograph a scene with parallax in all directions. [27] Using this technique, a display can be created to create parallax in the horizontal, vertical or both directions by employing some type of lens array.

The concept of using a lenticular array in front of an image was first introduced by Walter Rudolph Hess in 1912 [28]. A lenticular or lenticular sheet is an array of cylindrical lenses. Similarly to a parallax barrier display, the lenticular array is arranged in such a manner that each lens subtends two or more pixels to distribute the pixels to appropriate viewpoints as seen in Figure 12. A lens element of the lenticular that subtends two pixels, one for each eye, only provides a two-view solution. However, if a lens element subtends multiple pixels, multiple viewpoints can be generated. When using a lenticular array, parallax is achieved in one direction, either vertical parallax or more commonly horizontal parallax only.

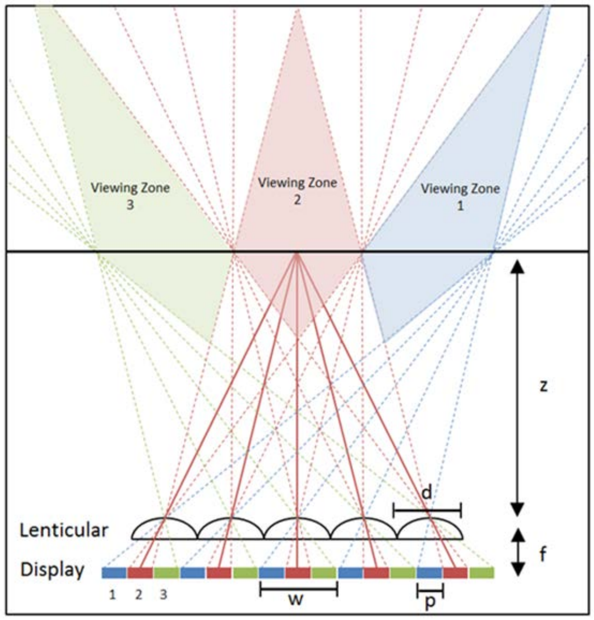


Figure 12. Geometry of a 3-view lenticular display

In order to design a lenticular-based 3D display, one must carefully design the parameters of the lenticular sheet. [22] Using the geometry of a multi-view lenticular display in Figure 12, the pitch, *d*, is selected such that each lenticule subtends a certain number of pixels, *w*, with a pixel pitch, *p*, that translates to the number of views, *n*. The focal length, *f*, is chosen such that the pixel structure is magnified and imaged at the viewing point which is at the center of each viewing zone at the optimal observer distance, *z*. It is important to note that since this is a lenticular-based display, the lenticule focal length is only in one axis. In other words, the lenticules are curved along one direction only. The lenticular-based display must satisfy the following condition:

$$\frac{d}{z} = \frac{n*p}{z+f}$$ **Equation 6**

From Equation 6, the optimum viewing distance can be determined to be the following:

$$z = f\,\frac{n*p}{n*p-d}$$ **Equation 7**

The center of each pixel defines viewing position within the viewing zone and the edges of each pixel define the boundaries of the viewing zone. This viewing zone defines the viewing angle, θ, of the display which from Equation 8, can be seen is proportional to the pitch of the lenticular and inversely proportional to the focal length of the lenticule.

$$\theta = 2\tan^{-1}\left(\frac{d}{2f}\right) \qquad \textbf{Equation 8}$$

As an example, if a 3D display application requires a user to be at a distance of 1 meter, a 30-inch 1080p panel providing 3-views with a 40° viewing angle, would give a pixel pitch of *p=0.346mm* which using equation 7 and 8, gives a focal length of *f=1.424mm* and a pitch of *d=1.036mm.*

While lenticular-based 3D displays only provide parallax in one direction, full parallax can be achieved using the integral imaging approach as used in photography. This technique is analogous to a pinhole array display in the sense that each lens subtends an M x N array of pixels to generate viewpoints in both the horizontal and vertical directions. The difference between integral imaging and lenticulars is the lens. Spherical lenses instead of cylindrical lenses in front of a display are used to provide parallax in both horizontal and vertical directions. Using Equations 7 and 8, a lens array can be designed to provide full parallax when the focal length of the lenslet is designed such that the element is curved in both axes.

While some of the drawbacks of parallax barriers may be applicable to lenticular or integral imaging displays, the biggest advantage is that since these displays are based on refraction, brightness is not sacrificed. Resolution reductions still apply depending on the number of viewpoints desired and the picket fence effect still exists. One method to eliminate the dark vertical lines is by arranging the lenticular lenses along a diagonal relative to the display. This slanted-lenticular display was introduced by Philips Research Laboratories in 1996 for a sevenview display [29]. In this implementation, a particular viewpoint is arranged by pixels subtended along the slated lenticular as seen by the pixels designated for viewpoint 3 between A-C lines in Figure 13. As the viewer moves left to right, a minor blending of views occurs from crosstalk since some pixels of the adjacent views are subtended by the lens. This implementation distributes the reduction of resolution in both the horizontal and vertical directions as compared to a standard lenticular display where reduction in resolution is only along the horizontal dimension which may create a non-standard image resolution.

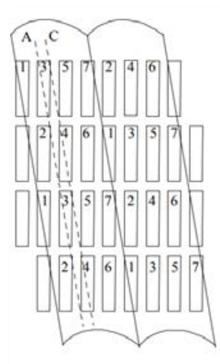


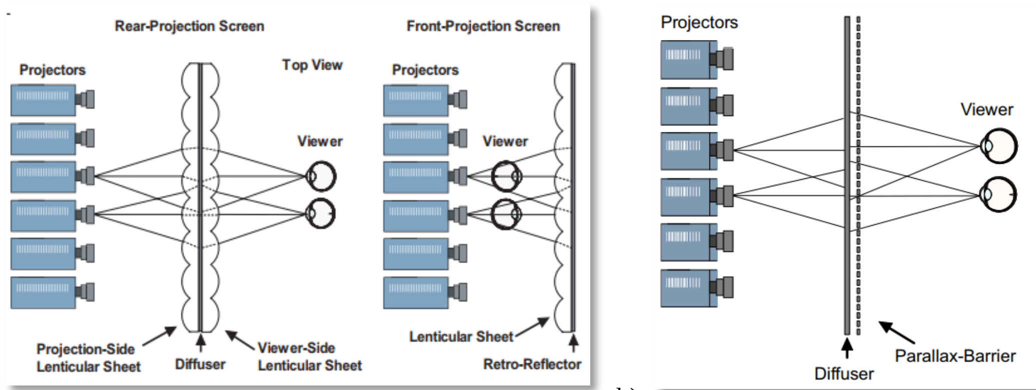
Figure 13. Slanted Lenticular by Philips

In order to address the loss in resolution in either direction, NLT Technologies in Japan had a different implementation. In their lenticular-based 3D display, they developed what they refer to as HxDP (Horizontally x-times Density Pixels) technology. The six-view display they developed is composed of horizontally striped RGB pixels each consisting of three horizontal sub-pixels and six vertical sub-pixels. This provides a resolution six times more than that of conventional square pixels. They use a special lenticular array designed to cover each of the sub-pixels so that each portion of the RGB pixel is directed toward a different view direction. Therefore, for a six-view display, each view sees the resolution of the original LCD panel. [30] Achieving such a higher pixel density in a LCD display to make up for the resolution loss in a multi-view display still becomes quite a challenge for LCD manufacturers. The tradeoff of view count to spatial resolution still exists.

d. Projection-Based Multi-View

The use of projectors as compared to an LCD panel in a 3D display provides several advantages. The main advantage is the ability to increase pixel density by adding more projectors. In an LCD panel, in order to increase pixel density, the pixel physical size must decrease to pack more pixels within a given area. The pixel size requirements exceed current manufacturing limits of LCD panels, thus projection based architectures become a more feasible solution. Multiple projectors are commonly used to overcome the spatial and angular resolution tradeoff in 3D displays. It is common to see two cases: 1) multiple projectors are used to reduce the magnification of the pixels such that more pixels are packed within the display area thus increasing the spatial resolution, and 2) multiple projectors are used to increase the number of views where each projector produces a single view in a multi-view system.

Multiple projectors have been used in combination with a lenticular screen to achieve a multi-view display without the loss of resolution. As seen in Figure 14a, this has been demonstrated by Mitsubishi Research Laboratories where each projector forms an image on a lenticular-based transmissive or retro-reflective screen, which corresponds to each view. [31]



a) b)

Figure 14. a) Multi-projector lenticular-based display **b)** Multi-projector parallax-barrier display

Similarly multi-projectors can be used in combination with a parallax-barrier to provide more views in Figure 14b. As discussed previously, in a lenticular-based or parallax-barrier display, the spatial resolution of the panel gets reduced by the number of views desired. Using the multi-projector method, since each projector corresponds to a different view, the spatial resolution is preserved because within each view, the observer sees the entire resolution of the corresponding projector.

Other techniques using multiple projectors in a 3D display have also been implemented. 3rd Dimension Technologies [32] and Holografika [33] both use similar concepts of multiple projectors illuminating a 'holographic screen’ to generate a sufficient number of views to produce a 3D display. Within each view, each corresponding projector displays a 2D image of slightly different perspective.

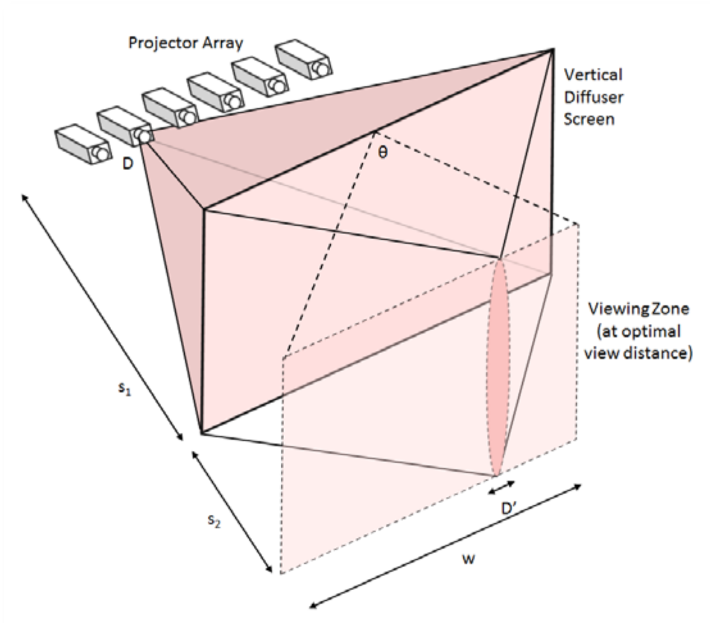


Figure 15. Geometry of a Multi-Projector Multi-View Display

Looking at the geometry in Figure 15, the pupil diameter of the projector, *D*, gets magnified horizontally to the viewing zone at the optimal viewing distance, s_2, based the following:

$$D' = m * D = \frac{s_2}{s_1} * D \qquad \textbf{Equation 9}$$

where s_1 is the throw distance of the projector. The holographic screen uses 1-D diffusion in the vertical without altering the horizontal information to expand the pupil thus creating a thin vertical slit as the view window. The viewing zone width, *w*, defines the full viewing angle, θ, of the 3D display. By defining the required angular resolution, $\Delta\theta$, of the display, the number of projectors or number of views, *n*, can be calculated using either of the following:

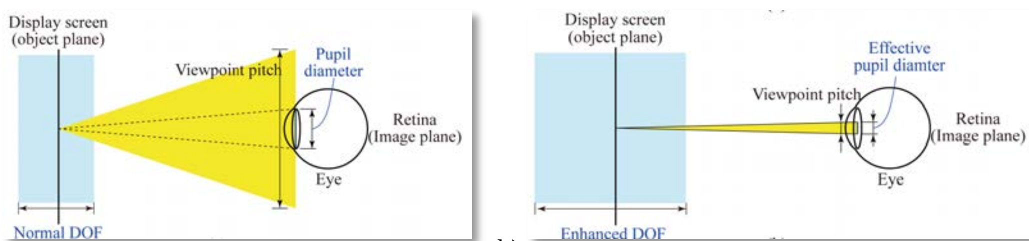
$$n = \frac{\theta}{\Delta\theta} \quad \text{or} \quad n = \frac{w}{D'} \qquad \textbf{Equations 10 and 11}$$

The spacing of the projectors can be set to *D*, the width of the projector pupil to fill the view spacing within the viewing zone. Furthermore, the angular resolution desired sets the requirements back onto the projector. Using elementary optics, the projection pupil can be calculated as a function of the imager size, F/# and focal length. It then becomes obvious that such a multi-projection system becomes feasible for longer throw projectors since the pupil is larger or at very long observer distances or by using a projector with a larger imager size. A method to slightly reduce the number of projectors requires a blending of the views to provide a continuous horizontal parallax that can be achieved with a very slight diffusion in the horizontal direction. Currently, Holografika's HoloVizio systems employ as many as 80 projectors each with 720p resolution, thus creating ~73 million voxels. Within a 70° viewing angle, this translates to approximately 0.88° angular resolution. It is said that this angular resolution may not be sufficient enough to sample the light field for a large depth of focus of the 3D image to reduce or eliminate the vergence-accommodation conflict of 3D displays. However, the 3D experience is much improved over stereoscopic two-view displays.

e. Super-Multi-View and Light Field Displays

Most of the previously mentioned implementations of a 3D display approach a physical limit on how many views a multiview display can create. This creates a motion parallax that is either provided in a step-wise fashion or view blending is required to smoothen out the jaggedness between views. In either case, this reduces the effectiveness and performance of a 3D display. Since the light field is not sampled with enough views, the vergence accommodation conflict may still exist. However, it may be significantly reduced as compared to a stereoscopic display since the depth of focus of the display is increased. In order to eliminate the vergence accommodation conflict, the view window size generated by a display must be close to that of a pinhole. In this case, the depth of focus approaches infinity.

Super multi-view displays sample the light field with a large number of views to enable a more natural 3D viewing of a display. To do so, the super multi-view (SMV) displays reduce the pupil size of the view windows to less than the diameter of the human eye pupil. Since the pupil size is much smaller than that of a standard multiview display, the depth of focus is significantly increased. The difference in depth of focus between multiview and super multiview displays can be seen in Figure 16. [34]



a) b)

Figure 16. Depth of focus comparison of **a)** Multiview display **b)** Super Multiview Display

Multiple views with slightly different perspective images enter the eye simultaneously. The super multi-view (SMV) condition is met when more than two views enter the eye. According to Lee et. al [35], the depth of focus can be extended to near 2 diopters when four simultaneous view images enter the eye. While the depth of focus of the human eye is 3 diopters, a super multiview can provide nearly correct focus cues. This is primarily due to the fact that the point of vergence and point of accommodation may lie within the same depth of focus thus practically eliminating the vergence accommodation conflict. Additionally, these types of displays provide very smooth motion parallax. Realizing a super multi-view display can be rather challenging and extremely expensive due to the very high pixel density needed which may not be possible unless using multiple projectors which can also be costly.

Takaki [36] calculated the number of ray directions necessary or the required angle pitch to achieve correct accommodation responses in a super multiview display or as he sometimes referred to as high density directional display. The angle pitch is given by:

$$\delta = \tan^{-1}\left(\frac{d}{z}\right) \qquad \textbf{Equation 12}$$

where *d* is the eye pupil diameter which ranges from 2-8mm and *z* is the distance from the 3D image to the eye. The required angle pitch is such that the view window pupil size is at most the size of the eye pupil diameter. The required angle pitch for a given eye pupil diameter and view distance can be seen in Figure 17.

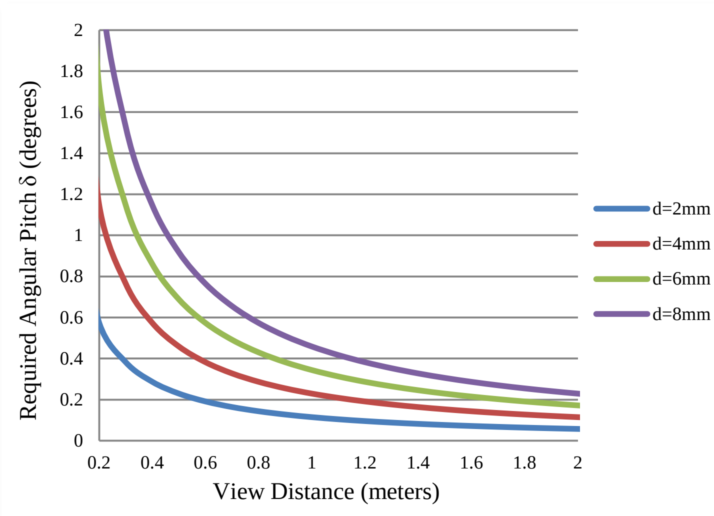
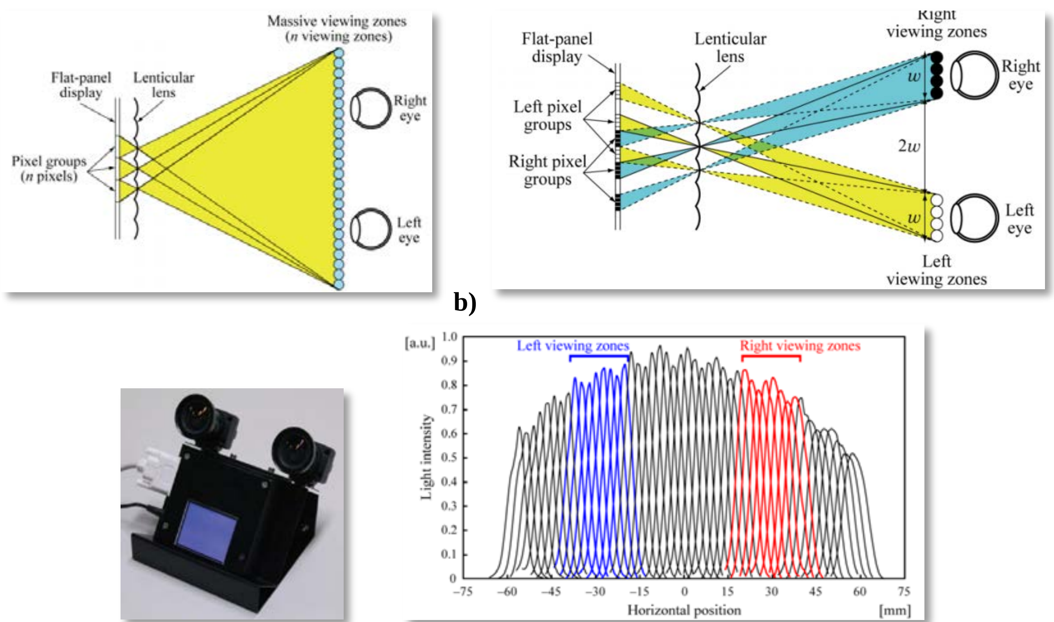


Figure 17. Required angular pitch as a function of pupil size and viewing distance in a SMV display

Takaki tried several implementations of a super multiview display. His first prototype was composed of a stack of projectors followed by a micro lens array, an aperture array, and a common lens. Each projector projected images in different horizontal directions very similar to that of multi-projection multiview displays mentioned previously. His unique arrangement allowed 128 projectors to be stacked in a compact space. This system created a 12.8” display with 128 viewpoints each with 800 x 600 pixels. The horizontal ray angle pitch achieved was 0.28°. [34]

Another implementation of a super multiview display by Takaki involved the use of a high resolution flat panel display using a lenticular lens. As mentioned previously the resolution requirement for a flat panel display must be very high in order to achieve a high number of viewpoints. Multiple systems were prototyped that achieved up to 72 viewpoints with as low as

0.38° horizontal ray angle pitch. However, each system was limited to a 3D resolution of 320 x 400 pixels or 640 x 400 pixels on a 22.2 inch display. [34] The spatial resolution versus angular resolution tradeoff is quite apparent when using a lenticular-based display as shown by these systems. In order to overcome this issue, Takaki et. al. [37] developed the reduced-view technique so that the high resolution required of the display can be reduced. In this prototype, the views are partitioned for the left and right eyes only. By carefully designing the lenticular lens array, a group of pixels subtended by the lenticule are directed to the left eye to generate those views and the next adjacent group of subtended pixels is directed to the right eye with its corresponding number of views. In a lenticular-based display, the viewpoints are replicated creating multiple viewing zones of the same views. This occurs when oblique light rays from a pixel enter adjacent lenticules. Taking advantage of this phenomenon, the display does not need to generate a large number of viewpoints to fill the viewing zone as in a conventional super multi-view display. Figures 18a-b show the advantage of a “reduced-view” super multiview display as compared to a conventional super multiview display. Additionally, Takaki expanded on this prototype further by adding a stereo camera for eye-tracking to take advantage of the replicated viewpoints such that content can be changed based on the viewers eye positions (Figure 18c). Eight viewpoints around each eye were demonstrated with a pitch of 2.6mm which is much smaller than the average pupil diameter of 5mm. The viewpoints for each eye can be seen in Figure 18d which also shows the replicated viewpoints that are used in the eye-tracked method.



a) b)

c) d)

Figure 18. a) Conventional SMV display technique **b)** “Reduced-view” SMV display **c)** Eye-tracked “Reduced view” SMV display **d)** Generated viewpoints

Using the slanted lenticular approach, the 3D resolution was still reduced from 1,024 x 768 pixels down to 256 x 192 pixels. However, they were able to achieve the super multiview condition by only generating a small number of views compared to the previous 72 viewpoints achieved with a flat panel display and without needing a panel with a much higher pixel density. Although this concept was developed with the intent of a mobile application thus the display size being rather small, the concept can be scaled to larger display applications.

Another two techniques to construct a super multiview view display both require the use of projection. In one configuration, a hybrid system of a flat panel array and an array of projection lenses are used. Each flat panel uses a lenticular array to generate 16 views. The flat panel then uses a projection lens to image onto a screen. This configuration is repeated 16 more times to create an array of flat panels plus projection lenses that image onto a common screen to generate 256 views. Although the image size is small at 10.3 inches and the 3D resolution is rather low at 256 x 192 pixels, the horizontal interval of viewing zones achieved was 1.3mm.

Thus a super multiview with high angular resolution was achieved. Again, the spatial versus angular resolution tradeoff is apparent. [38] A time-multiplexed method of an array of projectors was also implemented by Kanebako and Takaki [39]. In this prototype, a digital micromirror device (DMD) is illuminated at different angles with 15 individual LEDs. The LEDs emit light one after another in synchronization with the DMD to display a different 2D perspective image at high frame rates. This projector then produces 15 directional images. Using an array of four time-multiplexed projectors, 60 directional images along the horizontal direction were generated. This enabled a 0.31° horizontal angle pitch within an 18.3° viewing angle all while maintaining the XGA 2D resolution of the projector for 3D imagery.

Most of the super multiview displays mentioned are designed for fairly close viewing distances at approximately 1 meter. This is because the required angular pitch as shown in the graph of Figure 17 translates to 0.2°-0.4°. Higher density views or finer angular pitch requires an ultra-high pixel density for a flat panel or requires more tiling of displays. To achieve a finer angular pitch, Lee et. al. [40] prototyped a 100-inch 3D display using 300 Megapixel projectors. Using two side mirrors and an optimal configuration for stacking 300 projectors, they were able to achieve 600 viewpoints within a 40° viewing angle. The stacking of 300 projectors was critical in their system since each projector needed to have the projection pupil adjacent to the next. However, projection lenses are typically larger in diameter than the pupil. Therefore, the projectors were arranged in a diagonal stacking such that the pupils lined up adjacent to each other along the horizontal axis. This system provided 0.06° angular resolution for horizontal parallax only and at a viewing distance of 3 meters, each view window becomes ~3mm wide which is smaller than a 5mm eye pupil in bright conditions. No spatial resolution was lost since the number of views equaled the number of projectors.

While super-multiview displays are the closest approach to a light field display, most implementations only provide horizontal parallax. Even doing so, it is clear to see the difficulty in realizing such a display whether it is due to the high pixel density required or the high number of displays required. A full parallax ‘light field display’ has been demonstrated by Zebra Imaging with their ZScape Motion Display (ZMD). [41] In this implementation, they try to achieve a very high pixel density by tiling 216 transmissive LCD panels with 1080p resolution to fill a 12” x 12” display. The system generates an angular pitch of 1.18° in horizontal and vertical directions with 252 x 216 hogels of 1.6mm in diameter which determines the smallest resolvable spot size of the 3D image. The company evolved into FOVI3D which has since improved on the angular and spatial resolution of their full parallax light field displays. They define the system performance based on the total number of pixels or rays. In other words, they specify the total number of hogels and the number of views per hogel. In a 1000mm x 750mm display, the ZMD architecture generated 278,852 hogels with 5,776 views per hogel which equals 1,610,648,424 pixels. Their latest Gen 3 prototype of the same display size generates 3,000,000 hogels with 14,400 views per hogel which produces 43,200,000,000 pixels. This equates to a 0.75° angular view pitch in horizontal and vertical directions. [42] Although neither example comes close to the angular pitch of a super multiview display, these companies were the first to truly demonstrate a full parallax display over horizontal parallax only by generating views which opens up many new applications for a 3D display.

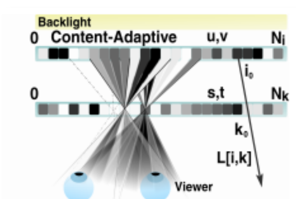
f. Cascaded/Stacked Displays

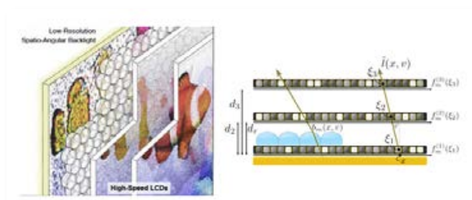
Cascaded or stacked displays are a multilayer display concept where by time multiplexing attenuation, light rays from a backlighting unit can be selectively directed to represent light directions of a sampled light field. In many cases, these types of displays employ a concept known as “directional backlighting.” The MIT Media Lab has explored several implementations of multilayer displays. Lanman et. al [43] implemented dual-stacked LCD panels in front of a backlight such that the second panel served as a dynamic parallax barrier. As discussed previously, a parallax barrier creates multiple views by the pixel subtended behind the masks. However, in conventional parallax barrier displays, the masks are fixed which means only a fixed number of views are generated. Having a temporally multiplexed parallax barrier can increase the number of views or tradeoff angular and spatial resolution based on the content being displayed. Therefore, they propose a content-adaptive automultiscopic dual-stacked LCD concept. The disadvantage of this solution is common with parallax barrier displays with reduced brightness, possible crosstalk, and even moiré artifacts created by the stacking of the layers.

In a similar study by the same research group, tensor displays were introduced as a family of compressive light field displays. [44] These tensor displays are comprised of stacks of light-attenuating layers such as multilayer transmissive LCD panels that can rapidly modulate the backlight to selectively address directional components representing a light field. In addition to introducing and characterizing tensor displays, they introduce a tensor representation for light field that requires multi-frame decompositions of images to increase image performance. These algorithms developed are outside the scope of this report. In their prototype, they were able to generate 9 x 3 views with a field of view of 50° x 20° and resolution 840 x 525 pixels (half of the resolution of the LCD panel used) using either five layers and uniform backlighting or a single layer and directional backlighting for each of the 12 decomposed frames to sample the light field.

A projection-based compressive light field display was also developed by the same research group. [45] This implementation requires the use of a light field projector and a passive novel angle expanding screen. Within the light field projector, two spatial light modulators are rapidly modulated in sequence such that a light ray angle from the first SLM maps to the position of the second SLM thus providing images of the 2nd SLM as if they are emitted from the projector at different angles. Using a two-plane parametrization with the two spatial light modulators, a 4D light field can be generated. This light is then incident on a screen which must redirect the light to a viewing zone. The screen is comprised of two lenticular sheets to form a Keplerian telescope to expand the angle of the projector to a wider pupil or viewing zone. Typically this is achieved using a diffuser screen. However, in order to maintain the angular information, a diffuser cannot be used as it would scatter the angles in the expansion. Therefore, they develop an angle expanding screen. The limitation however, is that the angle expanding screen seems to only work for low magnification. As the angle is increased, the apparent pixel size is decreased causing a poor fill factor of the 3D image. Although they were able to achieve 25 views, they were only limited to a FOV of 5°.

In any of the three prototypes by MIT Media Lab as shown in Figure 19, multiple layers of displays were used to demonstrate a directional backlight scheme to generate a light field.

a) 

b) 

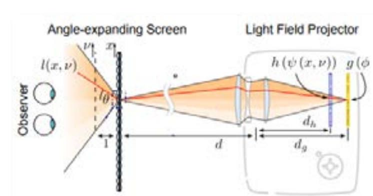
c) 

Figure 19. a) Content Adaptive Parallax Barrier **b)** Tensor Display **c)** Compressive Light Field Projector

Angular and spatial resolution tradeoff is still apparent even when compression algorithms are applied to increase the performance. While, this technology is another method in creating a light field display, performance is reduced as compared to multiview or super multiview displays discussed in previous sections since the light field is not sampled enough.

VI. <u>Holographic Displays</u>

The concept of holography was first developed by Dennis Gabor in 1947 where light wavefronts scattered from an object were recorded and recreated or displayed when the object was no longer present. By recording and recreating the light waves, all information including phase, amplitude, and wavelength is maintained. The term holography itself comes from the Greek words “holos” meaning “whole” and “grafe” meaning “drawing.” Thus, a hologram is an image that represents a “whole drawing” by recreating all of the light wave information from an object. In an ideal case, the natural object or scene should have no difference from a hologram.

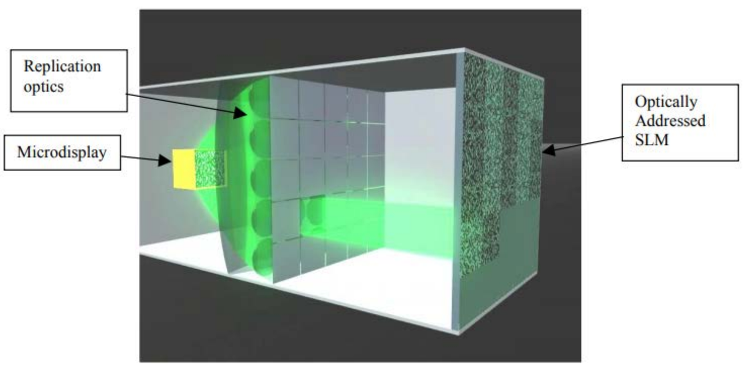
In a traditional hologram, using coherent light such as a laser, an interference pattern is recorded and reconstructed to display an object. Recent studies have explored the use of electronic or digital holography in which the recording process and reconstruction process can be separated. In this implementation, a camera is used to record the interference pattern and a spatial light modulator is used to display the recorded pattern. Thus, the hologram is created using computational calculations without the need of the real object. This method, in which a hologram is generated by computations, is known as computer generated holography (CGH).

3D displays using holography are thought of as the best method for displaying 3D content since all characteristics of 3D objects can be recreated. However, there are multiple challenges that prevent such displays from practical implementation. The toughest hurdle for holographic 3D displays is the pixel density required. According to Stanley et. al. [46]:

$$\Delta h * \Delta\theta \approx \frac{n*\lambda}{2} \qquad \textbf{Equation 13}$$

where Δh = size of image volume, λ = wavelength of laser source, $\Delta\theta$ = field of view, and n = number of pixels across 1 side of CGH. Thus, using Equation 13, for a 30 inch diagonal image with a 20 degree field of view, the hologram pixel count is on the order of ~10^9 pixels. This translates to pixel sizes smaller than 1µm for gigapixel or even terapixels requirements for a reasonable sized display. While the pixel sizes of current spatial light modulator technologies are nowhere near the required sizes, one common technique is thru the use of tiling subholograms to achieve higher pixel densities required for a holographic display. If the pixel count is solved, there still is a challenge of the tremendous amount of content information and data processing required to address the high pixel count.

Qinetiq managed to overcome the issue of low pixel count by using electronically and optically addressed spatial light modulators combined with replication optics to create “active-tiling” of a subholograms. The subholograms are loaded onto the electronically addressed SLM which using replication optics creates multiple images onto the optically addressed SLM to generate a large hologram. This configuration can be seen in Figure 20a. A pixel count of 100 Megapixels was achieved that was able to produce a 300mm wide color image as seen in Figure 20b. [42]

a) 

b) 

Figure 20. a) Qinetiq’s “Active Tiling” Subhologram modulator **b)** 300mm color holographic image

SeeReal has taken a different approach to resolve the high pixel count requirements for a holographic display. Using eye-tracking, SeeReal is able to track the user’s position and generate a 3D scene point by properly illuminating and encoding only a small subhologram as opposed to the entire hologram display. As seen in Figure 21, by using these “view windows,” the diffracted angle from the holographic display is significantly reduced and thus the pixel size can be significantly larger. Compared to 0.5µm-1.0µm pixel reqirements, SeeReal is able to achieve similar performance with pixel sizes from 20µm-70µm. This implementation makes a holographic display feasbile using existing technologies. [47]

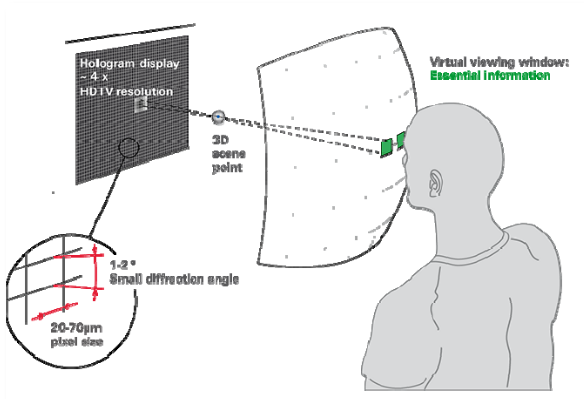


Figure 21. SeeReal's view window eyetracking for holographic display

University of Arizona had a unique approach to holographic displays that is rather different than computer generated holography. They developed a photorefractive polymer material that takes in an interference pattern of two fast pulsed coherent laser beams to create and store a 3D image much like in a hologram recording process. Each laser pulse records a hogel in the polymer which then fades away by natural decay in minutes or even seconds or by recording a new 3D image with a new diffraction structure. A 4 inch x 4 inch display was demonstrated to show horizontal parallax with 1mm hogel resolution which can be said to be fairly low resolution imagery. [48] Although this implementations of holographic displays, along with SeeReal or Qinetiq are quite novel, the technologies still struggle with sufficient pixel count but more importantly fast refresh rates to generate real-time 3D holographic imagery.

VII. Volumetric Displays

Volumetric displays produce 3D imagery in actual 3D space as compared to multiview or holographic displays that generate ray directions to represent a 3D object. These displays typically require a medium in which 3D objects are displayed through means of emission, scattering, or illumination of individually addressable voxels. The term voxel originates from combining the words "volume", "pixel", and "element" but can be simply considered as a 3D pixel. It is common to volumetric displays to describe their resolution in voxels, i.e. 256 x 256 x 256 voxels. Volumetric displays can be considered autostereoscopic or automultiscopic since they create 3D imagery to an unaided eye. Most achieve near 360-degree spherical viewing angle in which the image changes as the viewer moves around, thus providing parallax in one or both directions. Additionally, since volumetric displays rely on a true 3D space to create 3D pixels, both physiological and psychological depth cues can be met. Volumetric displays can be traced back to the 1960s which used rapid varifocal mirror oscillations to generate a 3D image [49]. More recently with the advancements in display technologies, volumetric displays have provided better performance and have become a more commercially viable technology. This section will provide an overview of true volumetric displays which are typically categorized as static volume displays or swept-volume displays.

a. Static Volume Displays

Volumetric displays using static volume techniques provide 3D imagery by generating voxels in a three dimensional image space. Recent advancements in static volume displays mainly focus on the medium material and up-conversion techniques. In these types of systems, the display volume is filled with special material that can be selectively excited by two independently controlled radiation beams, which activate a voxel when they intersect. The volumetric display materials are typically either gas or a solid state material such as crystal or glass. As for the radiation beams, electron beams cannot be used for this purpose, however, laser beams can and are more commonly used in these types of systems.

One common technique known as two-photon up-conversion uses two infrared photons to pump a material into an excited level, from which it can make a transition to a lower level with visible fluorescence. A voxel is formed at the intersection of two independently scanned laser sources that exhibit two photon absorptions from two different wavelengths [21]. This process can be seen in Figure 13.

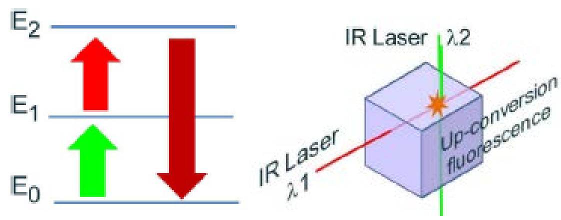


Figure 22. a) Energy level diagram of an active ion
b) Two intersecting laser beams addressing a voxel in a medium with such an ion

The use of up-conversion in a solid state material dates back to Lewis et al. in 1971 [50], although complex pictures or moving points was not achievebale due to the lack of appropriate excitation sources or medium materials. Elizabeth Downing in 1994, however, improved the technique using a ZBLAN cube which consists of a doped medum material with three dopants combined with three infrared lasers to achieve a full color 3D image. [51] It is apparent that the material selection is quite critical in the performance of such static volumetric displays. This can be seen in the studies by Koudsi et. al. in their CScape display, as they increased the number of dopants in the medium material to improve the up-conversion of two intersecting laser beams of wavelengths 1532nm and 850nm, scanned using digital micromirror devices from Texas Instruments DLP. [52] Similar up-conversion techniques have been done in a gas medium as compared to a solid state material for a volumetric display. This technology relies on the intersection of two laser beams on the same gas or vapor atoms for an excitation process resulting in fluorescene at that point. By scanning the intersection points fast enough, a 3D image can be drawn in the vapor. This implementation suffers from technical difficulties such as vacuum chamber requirements and temperature stability in addition to the number of voxels limited by the speed of the scanners and the safety of using laser beams in general. The latter two however are exhibited by both volumetric displays using gas or solid state mediums.

b. Swept Volume Displays

Volumetric displays categorized as swept volume displays typically require a time sequential 2D image on a display synched with mechanical motions to sweep out a display volume at a frequency higher than the human eye can resolve. Essentially a 3D object is displayed by rapidly slicing 2D images. The human visual system will then integrate the light and perceive the rapidly synched images as a 3D object. Swept volume displays have been a common form of 3D displays dating back to the 1960s and 1970s. In 1963, Robert Schipper developed a swept volume display that used an electroluminescent panel with a high speed light emitter array that rotated within a plastic ball. [53] Similarly in 1979, Edwin Berlin improved on the design by using a 2D LED array covering half of a cylindrical display that rotated to sweep out a 3D image. [54] The difficulty with these displays was achieving high resolution since it is a function of the LED density, and the speed of the rotating mechanism and the pulse of the LEDs. Much advancement has been made in swept volume displays that improve on the works of Schipper and Berlin. One example is the swept volume display from University of Oklahoma. They use an active LED array of 16 back to back pairs of image panels, each consisting of 16

FPGA display controllers of which each is connected to a 4x4 pixel array. By displaying 1,024 2D cross section images per revolution and at speeds of 1800 revolutions per minute, they achieve a 3D resolution of 64 x 64 voxels at a 30Hz refresh rate. The Oklahoma University swept volume display (OU-SVD) system can be seen in Figure 23. [55]

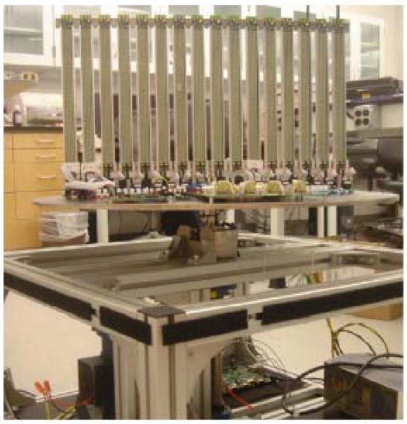
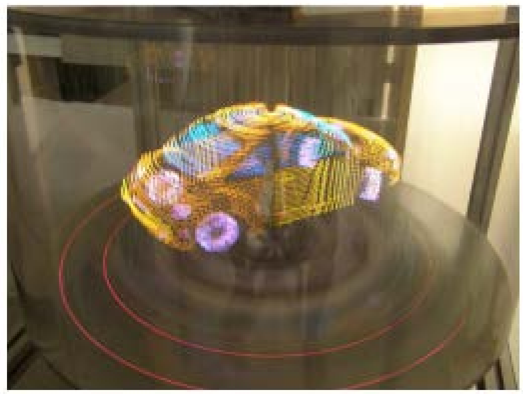
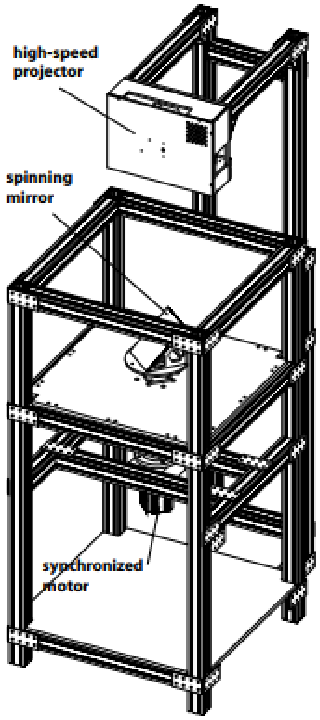
a)  b) 

Figure 23. Oklahoma University Swept Volume Display **a)** off mode **b)** in operation

Other sweeping volumetric displays commonly use a rotating mirror instead of an LED array. In this kind of a display, an image is typically projected onto the rotating mirror and synched with the movement to sweep out a 3D volume. As demonstrated by researchers at University of Southern California, they use a high speed projector to create an image onto a diffuser bonded to a mirror that rotates 20 times per second. They are able to achieve 1.25° angular resolution across a full 360° horizontal field of view at 768 x 768 pixels resolution per view. [56] This system can be seen in Figure 24. Similarly the Perspecta 3D display by Actuality uses rotating mirrors and a rotating double helix screen on which projected images are modulated to create a stack of spatial image layers that can be perceived as 3D volumetric images. [57]



a)

b)

Figure 24. USC swept volumetric display **a)** System layout **b)** in operation

Multi-planar imaging- or layered imaging-based volumetric displays also can be characterized as a swept volume display. LightSpace Technologies currently has a line of commercially available products that use this technology. [58] In their systems, they use a high speed projector and image onto a stack of up to 20 switchable liquid crystal scattering shutters or multi-planar optical elements (MOEs) that when voltage is applied, change the transmission of the component. The projector is then synched to the on and off states of each liquid crystal to display a 2D image at each plane. At a 20Hz refresh rate, the human eye will perceive the sliced 2D images into a 3D image, however it noticeable image flicker.

VIII. 3D Displays for Near Eye applications

Prior to this section, 3D displays mainly for multi-viewer applications were discussed. In order to have multiple viewers, the displays must have a wide viewing angle or a large viewing zone. Within this wide viewing angle, the 3D display must have a fine angular view pitch in order to properly sample the light field to achieve nearly correct depth cues. As described, this requires a tremendous amount of 2D pixels for high pixel density or requires the use of multiple display technologies or an array of displays to increase the spatial and angular resolution. Near eye 3D displays reduce the requirements since these types of displays are limited to a single user and the display viewing angle is limited to the observer's eye only, thus not needing such high view count.

A near eye display (NED), also referred to as head mounted display (HMD) projects images into a viewer's eyes, creating a virtual image within the field of view. The virtual image appears at some distance and appears larger due to the magnification of the optics used to create the image. These types of displays have several advantages of traditional displays, such as compact size, lightweight, or require low power. Thus, a virtual image that looks like a big TV screen can be created within a small form factor. Near eye displays fall into two main categories: immersive and see-through. Immersive near eye displays create a large field of view and block the user's view of the real world. Such displays are termed Virtual Reality (VR) glasses since the user can be immersed in a virtual world. See-through near eye displays allow the user to see the real world but can overlay virtual images onto it. These displays are commonly referred to as Augmented Reality (AR) glasses since the real world is augmented with virtual objects. Additionally, another term used to characterize near eye displays is Mixed Reality, which encompasses both augmented reality and augmented virtuality (AV) where real objects are overlayed into a virtual world. [59] It is common to see existing technologies or products being referred to by incorrect terminology, although all are still near eye displays.

Although near eye displays have a long history dating back to the 1980s, they have gained a lot of interest and momentum in the recent years due to its vast range of applications and much advancement in technologies that has enabled products to enter commercial markets. In addition to gaming and entertainment, near eye displays enable applications such as training and simulation, education, teleconferencing, and even medical applications. What restarted the virtual reality craze can be contributed to the release of the Oculus Rift in 2010. Since then companies like Samsung, HTC, Playstation, Microsoft and many others followed with competing products that either exist as a standalone product or required a mobile phone to be used as the display. Although augmented reality glasses have been used for quite some time in military applications, the recent augmented reality consumer craze can be owed to the growth of VR headsets with users seeing the cool features of VR and wanting to overlay such content onto the real world. According to Digi-Capital, shown in Figure 25, the craze of AR and VR could enable such markets to reach $120 billion by 2020 which is obvious to see why many consumer electronic companies working on their own headset products. [60] However, there are several drawbacks to the current AR and VR technologies such as size, mobility, battery power, and most importantly visual performance especially when it comes to 3D that need much improvement before reaching such market milestones.

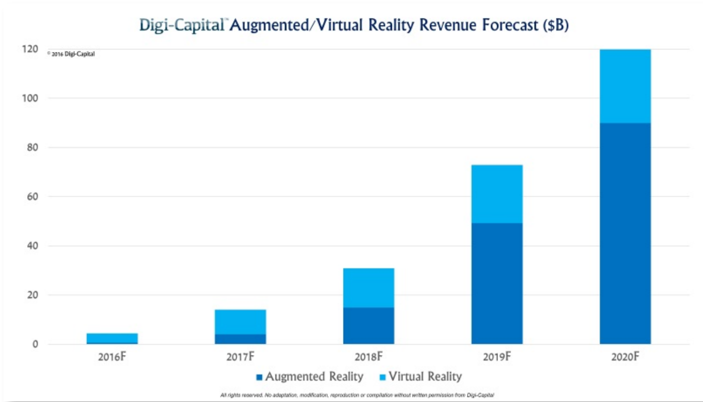


Figure 25. AR/VR market forecast

a. Virtual Reality

Personal headsets using virtual reality provide an immersive environment to a user by blocking out the real world and creating a virtual world with a display of high field of view. As MIT Technology Review shows in Figure 26, these headsets commonly consist of a small screen placed directly in front of a user's eyes and magnified with a set of optics to provide fields of view of up to 120°.

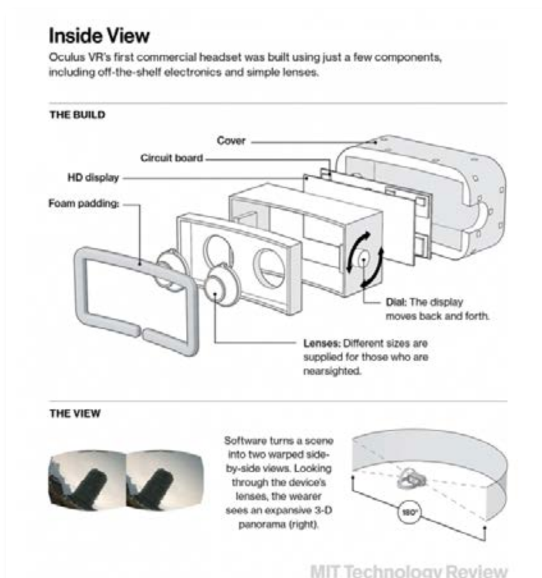


Figure 26. Basic architecture of VR headset (Oculus Rift example)

Separate images are provided for each eye thus capable of providing stereoscopic 3D imagery. Additionally, these headsets may have stereo sound, head tracking, or even eye tracking systems.

In any case, since only a single display or screen is used and 3D imagery is provided, it is clear as previously mentioned in this report, that such a display would create a vergence accommodation conflict for the user. Other image quality issues exist such as apparent pixel gaps that are magnified with the optics for larger fields of view causing a “screen-door” effect or the use of simple lenses which provide distortion and other aberrations that need to be corrected by software [61]. Since the human eye can resolve 1 arc minute, in order for the eye to not see such gaps, it is important for a display to have high pixel density if used in a VR headset as seen in Figure 27.

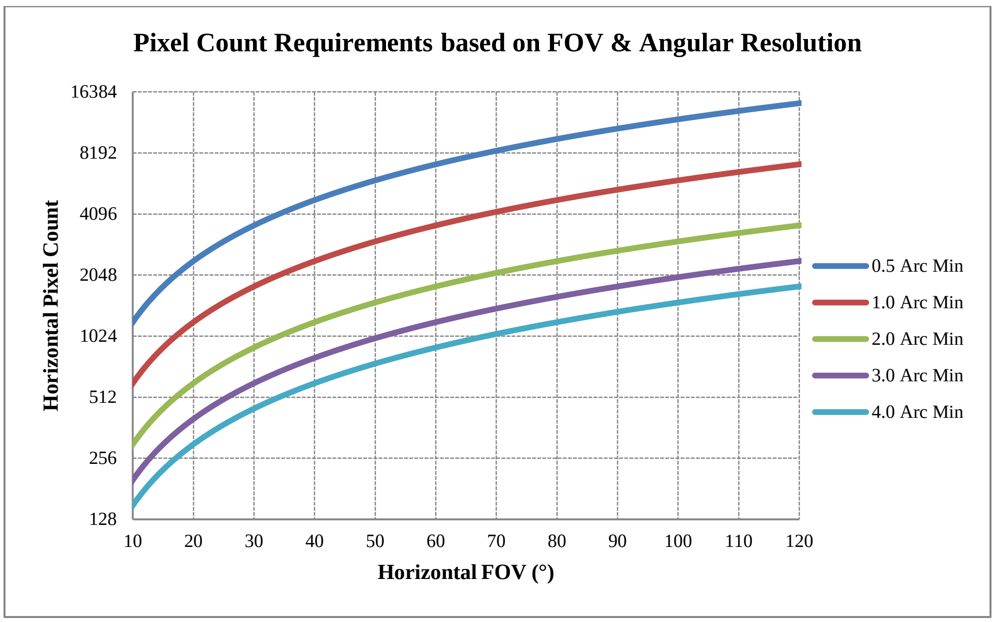
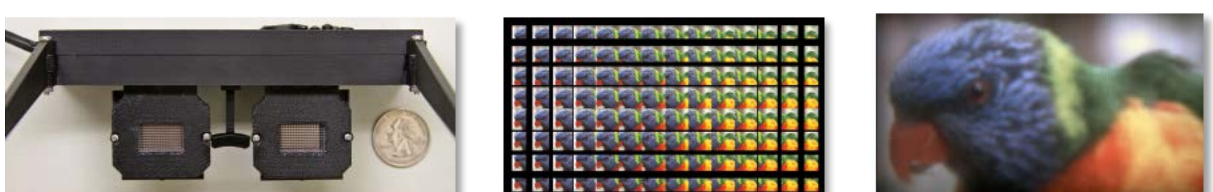


Figure 27. Display pixel count requirements based on FOV and angular resolution

Due to these image quality issues, virtual reality headsets have been known to cause eye strain and fatigue in addition to motion sickness. It can be safe to assume that most of these issues can be contributed to the vergence-accommodation conflict due to the stereoscopic nature of VR headsets. [59]

There have been only a few attempts to improve the 3D imagery by reducing the vergence-accommodation conflict in a virtual reality headset. One attempt involved the tiling of sixteen SVGA microdisplays to achieve a 150° x 100° field of view and a 3 arc minute resolution. The microdisplays were tiled in a spherical array and viewed through a curved, multi-faceted lens array. [63] This approach uses the integral imaging method as mentioned previously in this report. Similarly Lanman and Luebke from the Nvidia Research group [64] created a "near-eye light field display" using a microlens array covering a pair of OLED panels. The prototype, shown in Figure 28, achieved a resolution of 534 x 534 pixels with a 7.6mm x 7.6mm eyebox and a ~67° field of view. Although this implementation does not allow viewing the real world, the prototype specifications are low to be labeled as a virtual reality headset primarily due to the small field of view which would not provide an immersive experience. As they claim, spatial resolution is significantly reduced which is proportional to the ratio of the microlens focal length to the distance of the display from the eye. In order to create a thin form factor, they must sacrifice some resolution. The elemental images used to display a reduced resolution 3D image can be seen in Figures 28b-c. Larger microdisplays with smaller pixel pitches would enable a wider field of view with higher resolution.



a) b) c)

Figure 28. a) Nvidia's Near Eye Light Field Display **b)** Elemental images **c)** Perceived 3D image

One of the most common implementations of a light field virtual reality headset is the "Light Field Stereoscope" by Huang et al. from Stanford University [65]. In this implementation, they use a directional backlighting method with two transmissive liquid crystal panels to create a 3D image with appropriate focus cues and binocular disparity. A schematic of their prototype is shown in Figure 29a. The two liquid crystal panels function as spatial light modulators and are critical in providing appropriate focus cues. The panels are stacked such that the backlight is modulated in a multiplicative manner. With the optics placed in front of the panels, a virtual and magnified image of each panel is created, one at infinity and one close to the observer to provide a near and far plane. Using the two spaced virtual images and a two-plane parametrization method [19], the light field can be generated. The technical layout can be seen in Figure 29b.

a) b)

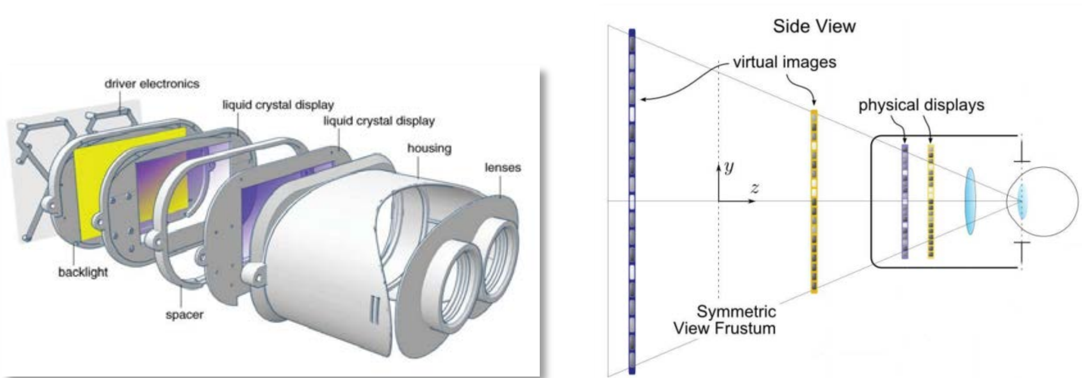
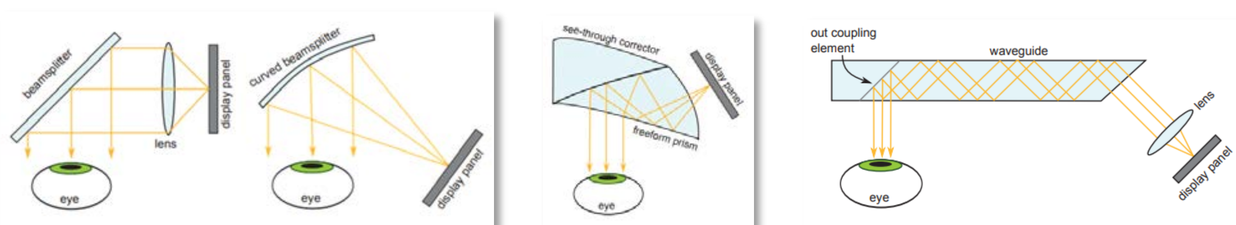


Figure 29. a) The Light Field Stereoscope by Stanford University **b)** Technical layout

Having two layers does limit the angular resolution sampling of the light field which has been argued by Ryana et. al. [66], that at least five display layers are needed to achieve natural accommodation of observers. The use of multi-focal plane for near eye displays will be discussed in more detail in section VIIIb. Nonetheless, the Light Field Stereoscope was able to achieve an accommodation range of 0.2m to 1.2m. Additionally, this display was able to provide approximately 90° field of view. Although the light field stereoscope seems to be a promising direction for reducing vergence accommodation conflicts in VR displays, the system still suffers from slow latency (65ms) which is important in VR gaming systems, light field rendering, and the largest limitation which is diffraction caused by the fill factor of the cascaded LCD panels. In future work, they hope to address these issues in addition to increasing field of view in order to reach a competitive commercial product.

b. Augmented and Mixed Reality

Augmented reality (AR) headsets allow virtual images or video content to be overlaid onto a user's field of view of the real world. These headsets are sometimes also referred to as smartglasses or wearable computer glasses since information mirrored from a mobile device can be displayed within the headset. Architectures for augmented reality headsets are much more complex than that of virtual reality headsets. Much like VR displays, size and weight are still strong considerations during the design process of these displays. However what differentiates these displays is an optically transparent component that must be used such that the user can see-through the device and view the virtual image overlaid onto the real world. The term "see-through head mounted display (HMD)" is commonly used as a result of using such optical components. As shown in Figure 30, these see-through HMDs typically fall into three categories based on the optical component used: combiner or curved mirror based, freeform prism based and waveguide or light-guide based. [67]



a) b) c)

Figure 30. See through HMD architectures **a)** Combiner-based **b)** Freeform prism **c)** Waveguide based

Using optical combiners such as a flat partially reflective, partially transmissive mirror or curved combiners are of the most common optical architectures used in HMDs. Large fields of view can be achieved at the cost of much larger optical components especially when a large eyebox size is desired. The reflectors themselves become rather large which causes the entire assembly to be bulky and heavy. These combiner based architectures are sometimes referred to as "birdbath optics." Freeform prisms are used as an alternative to birdbath optics in order to reduce the bulkiness. These components replace all of the airspace with fast diverging rays with glass or plastic such that the ray divergence is reduced thus reducing overall size. The freeform surface is used as a replacement for the optics to create a pupil for the eye. A see through corrector is commonly used at the output to correct for any aberrations of viewing the real world. Waveguide based HMDs provide a more compact solution by propagating light through a thin optical slab using total internal reflection. Waveguides allow the display device to be tucked away from the user's eyes typically at the temples or ears providing a more traditional comfortable eyewear solution. Waveguides have been increasingly popular due to the compactness advantages ever since the release of the Google Glass headset which used a simple beamsplitter cube as a reflective waveguide which is similarly used in Epson Moverio products. Other waveguide architectures include using polarized reflectors as demonstrated by Lumus, using surface reflectors and conventional coatings from Optinvent, diffractive waveguide structures from Nokia or Vuzix, stacked holographic optical elements from Sony and Konica Minolta or TruLife, or electrically switchable waveguides from DigiLens. [68] The use of waveguides, however, does limit the field of view capabilities of these displays as well as efficiency. On the other hand, some waveguides are capable of pupil expansion in one or even two dimensions to expand the eyebox while reducing the size of the optics. Furthermore, these head mounted displays typically require a microdisplay such as OLED or microLED or some sort of spatial light modulator such as LCoS or DLP in order to generate the virtual images. Since OLED and microLED technologies are self-emissive, they provide a compact solution with only a few lenses required to form a pupil for the virtual image. However, currently such technologies are limited to brightness output as well as speed especially in applications where 3D is desired. To overcome such issues, projection based architectures using LCoS and DLP panels have been adapted at the expense of slightly larger assemblies.

Most augmented reality headsets commercially available today only provide 2D information such as the Vuzix Blade 3000 [69]. However, the desire from companies to look into new technologies to release 3D capable headsets has become increasingly so with the release of Microsoft's HoloLens headset. HoloLens is a mixed reality smartglasses headset which is a term lately used to describe augmented reality headsets that overlay 3D imagery as opposed to 2D information onto the real world. According to Kress and Cummings [70], Hololens uses RGB LEDs to illuminate a 0.57-inch diagonal LCoS panel. Using a birdbath beam cube and some optics, they form an exit pupil as the input to a diffractive waveguide using three stacked surface relief gratings. The waveguide expands the pupil in two dimensions to provide a larger eyebox and providing 34° field of view. 3D imagery is achieved using traditional stereoscopic content meaning the system still suffers from the vergence-accommodation conflict. Overall, the performance of the HoloLens has been well received without any complaints of visual fatigue which may let us assume the stereoscopic 3D content may be within the zone of comfort for the human eyes.

There have been some interesting advancements in the academia realm in regards to near eye displays with appropriate focus cues for 3D imagery. The use of a lens array was demonstrated by Hua and Javidi [71] in their integral imaging optical see through head mounted display (InI-OST-HMD). A lens array of 18 x 11 elements is placed in front of a 0.8-inch microdisplay having 1920 x 1080 resolution. After using a freeform prism, they create a 6.5mm pupil, with approximately 34° field of view. The limitation of this approach is as seen with all integral imaging 3D displays, is the tradeoff between angular resolution and spatial resolution. The system can be seen in Figure 31.

a) b) c)

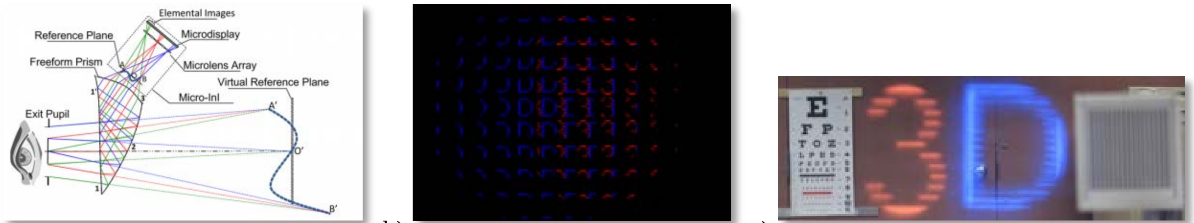
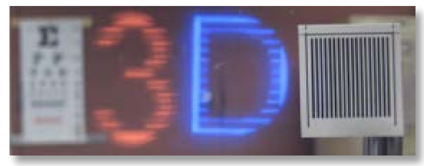


Figure 31. a) Integral Imaging HMD layout **b)** Elemental Images **c)** Focus at 4m **d)** Focus at 30cm



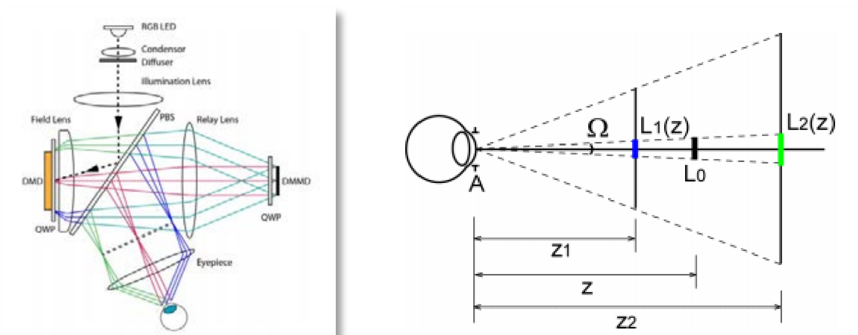
d)

By having 18 elements across a 34° field of view, the angular resolution can be calculated to be approximately 1.9°. Within each view, a very low spatial resolution of 102 x 102 pixels was achieved. While the angular resolution may seem low, in a near eye display, it all depends on the virtual image distance and the depth range of which the 3D content is placed at. Additionally, as seen in the graph in Figure 27, with their 34° field of view, in order to achieve human eye resolution of 1 arc minute, more than 2048 pixels are needed. Thus, integral imaging still proves to be a challenging approach to 3D as a very high pixel density display is required.

Multi-focal plane displays have become an interesting implementation for augmented reality. Compared to Lightspace's volumetric multi-focal plane display, this technology is easier to implement in a near eye display since virtual image distances do not require a physical screen. There have been a few implementations of a multi-focal plane distance for augmented reality. Hu and Hua [72] demonstrated a multi focal plane near eye display by varying the virtual image distance thru the use of a deformable membrane mirror device (DMMD). Liu and Hua [73,74] similarly created a multi focal plane system using a liquid lens. Both components are electro-optical devices that change in curvature as voltage is applied. They are carefully placed within the optical path of the near eye display in order to change the focal length of the eyepiece to produce an image at varying distances. By displaying a new image at different virtual image distances at a fast rate, the human eye integrates the rapidly switching 2D images to perceive a 3D image. These displays are designed such that the accommodation range matches the depth of focus of the human eye. The accommodation range or depth range of the 3D image that is being generated is provided by:

$$\Delta D_{depth} = \frac{\Phi_{eye}^2}{\Phi_{relay}^2} * \Delta\Phi_{VE} \qquad \textbf{Equation 14}$$

where Φ_{eye} is the power of the eyepiece, Φ_{relay} is the power of the relay optics, and $\Delta\Phi_{VE}$is the optical power range of the varifocal elements whether it is a deformable membrane mirror or liquid lens. A schematic optical layout using a DMMD can be seen in Figure 32a.



a) b)

Figure 32. a) Schematic optical layout of multi-focal plane HMD **b)** Depth-blending architecture

These systems work by time multiplexing methods of the spatial light modulator such that a 2D image is generated at a single focal plane for one frame and a new 2D image is generated at the next focal plane for the next frame. According to Rolland et. al [75], in order to cover the range of accommodation from 0.5m to infinity for the human visual system with 1 arc minute resolution, 14 focal planes are required. However, generating this large amount of focal planes with current technologies may be impractical due to the speeds necessary to generate 14 images with at least 60Hz frame rate per plane to provide flicker-free 3D images. This would require a microdisplay of at least 840 Hz and a varifocal element with matching speeds. To realize such a display, methods of depth blending or depth fusing were introduced by Hu and

Hua and Liu and Hua. The concept of depth-blending or depth-fusing is when a pixel that represents an object at the midpoint between two focal planes is simulated by illuminating each plane with a proper intensity such that the sum of the intensities of the two focal planes approximates the intensity of the object as if a real object was viewed. Following the layout in Figure 22b, the perceived luminance of the simulated pixel is represented by:

$$L_0 = L_1(z) + L_2(z) = w_1(z)L_0 + w_2(z)L_0 \qquad \textbf{Equation 15}$$

where $w_1$and $w_2$are the depth-weighted fusing functions. Additionally, the perceived depth may be considered as the weighted sum of the depth of the two focal planes as:

$$\hat{z} = w_1(z)z_1 + w_2(z)z_2 \qquad \textbf{Equation 16}$$

Hu and Hua determined an optimal 0.6 diopter spacing of focal planes is necessary based on the point spread function and MTF of the retinal image. Thus, to cover a human eye depth of focus of 3 diopters to infinity with 0.6 diopter spacing, 6 focal planes was determined to be the minimum number of required focal planes. According to Ravikumar et. al [76], a linear weighting function maximized the retinal image contrast of the perceived pixel as well as provided an appropriate contrast gradient to drive the eye's accommodative response. Additionally, the perceived images appeared smoother than a non-linear or boxcar weighting. Furthermore, it is important to note that multi-focal plane displays are independent of near eye display system requirements such as field of view or eyebox size which means such technology should be compatible with any type of near eye display architecture with the exception of waveguides since most are only designed for infinite virtual image distances. Compelling 3D imagery using multi focal plane technology for a near eye display have been demonstrated by these mentioned university groups but also by a corporate company, Ricoh Innovations research group [77]. While these implementations do seem promising to produce 3D with appropriate focus cues, in either case, the systems are rather large and difficult to fit within a compact head mounted display. Perhaps further development of compact microdisplays and varifocal elements as well as waveguides that can support varying virtual image distances may be required to reduce the size to make this a commercially viable solution.

A more compact solution to a near eye display with appropriate focus cues has been proposed by Maimone and Fuchs of University of North Carolina at Chapel Hill [78]. This approach uses the cascaded display concept as mentioned in previous sections where multiple displays are stacked and modulated to direct light rays that correspond to some physical virtual image distance. The system is composed of two or more thin high speed transmissive spatial light modulators such as transparent LCD panels that control the intensity of the transmitted light. A thin, transparent backlight unit such as a transparent OLED or an edge-lit waveguide distributes light uniformly over the panels and is rapidly modulated. Additionally, a thin high speed liquid crystal shutter is used to provide occlusion of the real world such that the virtual objects can be placed over them. The architecture and operation of the system can be seen in Figure 33.

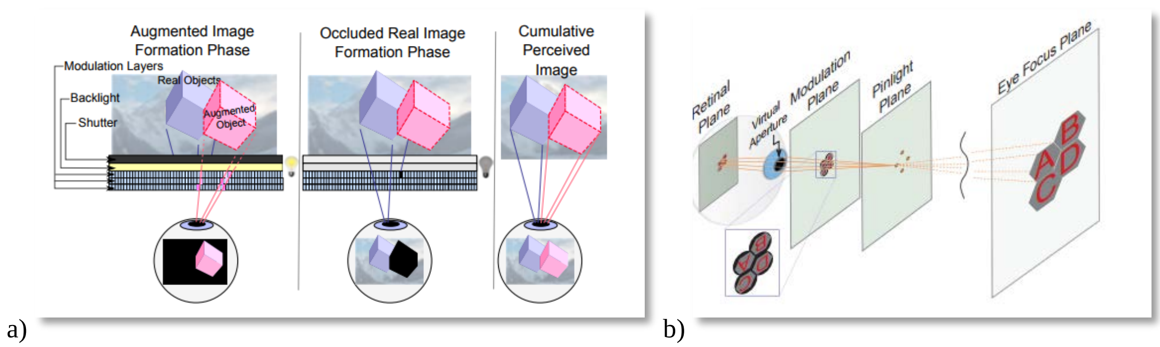


Figure 33. a) Modulated layers in a near-eye display **b)** Pinlight display configuration

A similar technology was demonstrated by members of the same research group that is referred to as a "pinlight display" [79]. In this implementation, a microlens array or an etched waveguide is edge-lit to produce defocused "pinlights". The spatial light modulator is coded such that the defocused pinlights from each microlens or etched divots produce an image on the retina with a narrow field of view. The spatial light modulator is composed of one or more transparent LCD panel to direct the light very similarly to the stacked modulator approach. In either concept, although high fields of views are achievable with larger panels, diffraction effects caused by the pinlight or the direct viewing of the pixels reduce image quality which does not rival performance in current head mounted displays.

There are several companies currently in a race to achieve a mixed reality headset to compete with Microsoft's HoloLens but also be the first 'Light Field' near eye display. These include Avegant which is a successful startup after the release of their personal video headset, the Glyph, and Magic Leap which is a highly secretive startup with over $1.9 billion in funding. Avegant has demonstrated a "Light Field" headset but they claim it is not production ready. Although, how they generate 3D content with appropriate focus cues is not mentioned, it is rumored that they may be using a fixed multi focal plane approach. [80] Magic Leap on the other hand has not released a single product nor has demonstrated anything publicly. However, with an extensive patent portfolio, it seems they have tried technologies anywhere from "wiggling fibers," to stacked waveguides, multi focal planes using adaptive optics, zone plate diffraction displays, combining display technologies, metamaterials and nanoparticles, and many more of which some are too farfetched or commercially feasible. [81] In any case, it will be exciting to see what kind of performance these products will have and what technologies enable it to provide a light field near eye display to the consumer market.

IX. Conclusion

3D displays enable many applications such as education, military, medical, gaming and entertainment, and more. It is obvious that as we live in a 3D world it is only natural that we want a display to produce 3D content to provide a better visual representation. These 3D displays have evolved in the past century but all can be tied back to stereoscopic 3D imagery. To display a 3D object, a view for the left and right eye must be created to provide parallax for the human visual system to perceive as 3D. While that has been accomplished with stereoscopic 3D using eyewear as seen in today's movie theaters or home entertainment systems, the visual discomfort caused by a vergence accommodation mismatch can be contributed to the recent decline in such technology.

Light field is used to represent the ray positions, directions, and intensities of a 3D object and can be used to mimic the operation of the human visual system. With advancements in current display components, the decline in stereoscopic 3D has forced researchers and companies to revive traditional concepts to achieve a glasses free 3D display but more specifically attempt to subsample a light field to provide appropriate focus cues to eliminate visual discomfort. These autostereoscopic displays employ concepts of parallax barrier masks and integral imaging displays using microlens arrays to produce a 3D image with appropriate focus cues. However, as discussed in this report, such displays suffer from a spatial versus angular resolution tradeoff. To have a better angular resolution, more views or perspectives of the 3D scene are required. This places tougher requirements on the display technology to achieve finer pixel sizes for higher pixel density panels. With the current manufacturing limits of producing such small pixels on an LCD display, projection based applications becomes more attractive since multiple projectors can be used to increase the pixel count or the view count. Multiple projectors have been used in concepts such as multiview or even super multiview that subsample the light field with higher angular resolution. These concepts however require a great number of projectors which increases the cost of these solutions preventing such technologies from reaching an affordable consumer market. Other methods to achieve a 3D display with appropriate focus cues include stacked display panels which although may provide a compact flat panel solution, suffer from many image quality issues that yet to be resolved. Volumetric displays are able to achieve 3D with appropriate focus cues however such technologies require a physical volume. With fast mechanical movements as in a volume swept display, such displays may also be considered rather dangerous. Holographic displays are considered the holy grail of 3D displays as they can provide all the necessary information to represent a 3D scene or object. These displays, however, require sub-micron sized pixels equating to gigapixel or terapixel counts for a reasonably sized display. Not only is this currently difficult with today's display technologies but costs of such displays can be astronomical. While these technologies can enable 3D viewing for single or multi-user applications, the increase in performance and reduction in cost of these systems will need to be further developed before such displays reach our homes.

Near eye displays provide a single-user viewing experience as seen in today's virtual reality and augmented reality glasses or head mounted displays. Since the viewing angle is limited to the user's eye only, requirements on view count are greatly reduced yet still subsampling the light field with high angular resolution. However, requirements on weight and size or form factor pose challenges on creating 3D near eye displays. In virtual reality headset, a stacking of display panels is commonly used and seems to be a promising technology once visual artifacts caused by fill factor and diffraction are resolved. In augmented reality glasses, multi-focal plane technologies seem to be a more feasible solution with faster display technologies and vari-focal elements developed that enable such concepts. However, system size is still a challenge and may be reduced with a waveguide based architecture which can expand the eyebox while reducing the size of the optics. The issue lies with waveguide technologies today that are limited to only a single plane at infinity which prevents a compact multi-focal plane solution. Nevertheless, companies such as Avegant have demonstrated a light field near eye display with the assumption that it uses multi focal plane technology.

In either a single or multi user 3D display application, there has been a great deal of development both by university research groups and even corporate research and development teams. While no truly light field display is currently available to the consumer market, the trend is apparent with multiview technology already being incorporated into laptop and mobile phones. It will be interesting to see the advancements in technologies in the next coming years and the corporate race of producing the first light field display.

X. References

1. "3D Display Market Is Expected to Reach $112.9 Billion, Global, by 2020 - Allied Market Research." *PR Newswire*. Allied Market Research, 26 Feb. 2015. Web. 07 Dec. 2017. http://www.prnewswire.com/news-releases/3d-display-market-is-expected-to-reach-1129-billion-global-by-2020---allied-market-research-294218651.html.
2. Fuensanta A. Vera-Díaz and Nathan Doble (2012). The Human Eye and Adaptive Optics, Topics in Adaptive Optics, Dr. Bob Tyson (Ed.), InTech,
3. Green, Daniel G., Powers, Maureen K., Banks, Martin S. (1980)."Depth of focus, eye size and visual acuity." Vision Research 20(10): 827-835.
4. Howard, Ian P., and Brian J. Rogers. *Binocular vision and stereopsis*. Oxford University Press, USA, 1995.
5. Reichelt, Stephan, et al. "Depth cues in human visual perception and their realization in 3D displays." *Proc. SPIE*. Vol. 7690. No. 1. 2010.
6. Mather, G. (2006): Foundation of Perception and Sensation, Psychological Press
7. M. Millidot, Dictionary of Optometry and Visual Science, 5th edition, Butterworth-Heinemann, 2000.
8. Mehrabi, Mostafa, et al. "Making 3D work: a classification of visual depth cues, 3D display technologies and their applications." *Proceedings of the Fourteenth Australasian User Interface Conference-Volume 139*. Australian Computer Society, Inc., 2013.
9. "GCSE Bitesize Science - The Eye : Revision, Page 2." *BBC*. BBC, n.d. Web. 07 Dec. 2017. http://www.bbc.co.uk/schools/gcsebitesize/science/triple_aqa/medical_applications_physics/the_eye/revision/2/.
10. "The Eye's Optical Mechanism." *The Visual Experience: Reading 2014*. N.p., n.d. Web. 07 Dec. 2017. http://arthistoryresources.net/visual-experience-2014/eye-optical-mechanism.html.
11. "Monocular Depth Cues Wiki." *FANDOM Powered by Wikia*. N.p., n.d. Web. 07 Dec. 2017. http://monoculardepthcues.wikia.com/wiki/Monocular_Depth_Cues_Wiki.
12. "Stereoscopy – A Brief History." *3D Alps*. N.p., 24 Jan. 2017. Web. 07 Dec. 2017. http://3dalps.wordpress.com/stereoscopy-a-brief-history/.
13. Epstein, Adam. "Americans Are over 3D Movies, but Hollywood Hasn't Got the Memo." Quartz Media, 24 Mar. 2017. Web. 07 Dec. 2017. http://qz.com/940399/americans-are-over-3d-movies-but-hollywood-hasnt-got-the-memo/.
14. Shibata, Takashi, et al. "The zone of comfort: Predicting visual discomfort with stereo displays." *Journal of vision* 11.8 (2011): 11-11.
15. Hong, Jisoo, et al. "Three-dimensional display technologies of recent interest: principles, status, and issues [Invited]." *Appl. Opt* 50.34 (2011): H87-H115.
16. Lambooij, Marc TM, Wijnand A. IJsselsteijn, and Ingrid Heynderickx. "Visual discomfort in stereoscopic displays: a review." *Stereoscopic Displays and Virtual Reality Systems XIV*. Vol. 6490. International Society for Optics and Photonics, 2007.
17. Gershun, Arun. "The light field. Moscow." *Journal of Mathematics and Physics* 18 (1936).
18. Adelson, Edward H., and James R. Bergen. "The plenoptic function and the elements of early vision." (1991).
19. Levoy, Marc, and Pat Hanrahan. "Light field rendering." *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996.
20. Gortler, Steven J., et al. "The lumigraph." *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996.
21. Geng, Jason. "Three-dimensional display technologies." *Advances in optics and photonics*5.4 (2013): 456-535.
22. Lee, Byoungho, et al. *Design and Implementation of Autostereoscopic Displays*. SPIE Press, 2016.
23. Ives, Frederic E. "A novel stereogram." *Journal of the Franklin Institute* 153.1 (1902): 51-52.
24. Jacobs, Adrian, et al. "2D/3D Switchable Displays." *Sharp Technical Journal* (2003): 15-18.
25. Emery, Daniel. "Nintendo Unveils 3DS Handheld Games Console." *BBC News*. BBC, 15 June 2010. Web. 07 Dec. 2017. http://www.bbc.com/news/10323971.
26. Perlin, Ken, Salvatore Paxia, and Joel S. Kollin. "An autostereoscopic display." *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000.

27. Roberts, David E., and Trebor Smith. "The history of integral print methods." *Lens Array Print Techniques* (2003): 1-21.
28. Hess, Walter. "Stereoscopic picture." U.S. Patent No. 1,128,979. 16 Feb. 1915.
29. Van Berkel, Cees, and John A. Clarke. "Characterisation and optimisation of 3D-LCD module design." *Proc. SPIE*. Vol. 3012. 1997.
30. "NLT Technologies Introduces High-Resolution Multi-View 2D/3D Display with HxDP Technology for Enhanced Stereoscopic Viewing Experience." *Renesas Electronics America*. N.p., 14 May 2012. Web. 07 Dec. 2017. http://www.renesas.com/en-us/about/press-center/news/2012/news20120514.html.
31. Matusik, Wojciech, and Hanspeter Pfister. "3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes." *ACM Transactions on Graphics (TOG)* 23.3 (2004): 814-824.
32. Thomas, Clarence E. "System and methods for angular slice true 3-D display." U.S. Patent No. 7,513,623. 7 Apr. 2009.
33. Balogh, Tibor, and Péter Tamás Kovács. "Real-time 3d light field transmission." *Real-time image and video processing*. 2010.
34. Takaki, Yasuhiro. "Development of super multi-view displays." *ITE Transactions on Media Technology and Applications* 2.1 (2014): 8-14.
35. Lee, Beom-Ryeol, et al. "Increasing the depth of field in multiview 3D images." *Three-Dimensional Imaging, Visualization, and Display 2016*. Vol. 9867. International Society for Optics and Photonics, 2016.
36. Takaki, Yasuhiro. "High-density directional display for generating natural three-dimensional images." *Proceedings of the IEEE* 94.3 (2006): 654-663.
37. Takaki, Yasuhiro, Kosuke Tanaka, and Junya Nakamura. "Super multi-view display with a lower resolution flat-panel display." *Optics express* 19.5 (2011): 4129-4139.
38. Takaki, Yasuhiro, and Nichiyo Nago. "Multi-projection of lenticular displays to construct a 256-view super multi-view display." *Optics express* 18.9 (2010): 8824-8835.
39. Kanebako, Tsubasa, and Yasuhiro Takaki. "Time-multiplexing display module for high-density directional display." *Proc. SPIE*. Vol. 6803. 2008.
40. Lee, Jin-Ho, et al. "Optimal projector configuration design for 300-Mpixel multi-projection 3D display." *Optics express* 21.22 (2013): 26820-26835.
41. Klug, Michael, et al. "32.4: A Scalable, Collaborative, Interactive Light-field Display System." *SID Symposium Digest of Technical Papers*. Vol. 44. No. 1. Blackwell Publishing Ltd, 2013.
42. Burnett, Thomas. "Light Field Display Architecture and Properties." Display Summit 2017. Sterling, VA. Oct. 2017. *Insight Media*. Web. 3 Dec. 2017. http://www.displaysummit.com/proceedings-display-summit
43. Lanman, Douglas, et al. "Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization." *ACM Transactions on Graphics (TOG)* 29.6 (2010): 163.
44. Wetzstein, Gordon, et al. "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting." (2012).
45. Hirsch, Matthew, Gordon Wetzstein, and Ramesh Raskar. "A compressive light field projection system." *ACM Transactions on Graphics (TOG)*33.4 (2014): 58.
46. Stanley, Maurice, et al. "3D electronic holography display system using a 100 Mega-pixel spatial light modulator." *Optical Design and Engineering*. Vol. 5249. International Society for Optics and Photonics, 2004.
47. "How Does It Work ... SeeReal's Holographic Display Principle in a Nutshell." *SeeReal Technologies*. N.p., 08 May 2017. Web. 07 Dec. 2017. http://www.seereal.com/how-does-it-work-seereals-holographic-display-principle-in-a-nutshell/.
48. Blanche, P-A., et al. "Holographic three-dimensional telepresence using large-area photorefractive polymer." *Nature* 468.7320 (2010): 80-83.
49. Traub, Alan C. "Stereoscopic display using rapid varifocal mirror oscillations." *Applied Optics* 6.6 (1967): 1085-1087.
50. Lewis, Jordan D., Carl M. Verber, and Robert B. McGhee. "A true three-dimensional display." *IEEE Transactions on Electron Devices* 18.9 (1971): 724-732.
51. Downing, Elizabeth Anne. "Method and system for three-dimensional display of information based on two-photon upconversion." U.S. Patent No. 5,684,621. 4 Nov. 1997.
52. Koudsi, Badia, Hakki Refai, and James J. Sluss. "Scalable Upconversion Medium for Static Volumetric Display." *Journal of Display Technology*11.3 (2015): 266-272.

53. Schipper, Robert J., and John L. Coddington. "Three-dimensional display." U.S. Patent No. 3,097,261. 9 Jul. 1963.
54. Berlin Jr, Edwin P. "Three-dimensional display." U.S. Patent No. 4,160,973. 10 Jul. 1979.
55. Gately, Matthew, et al. "A three-dimensional swept volume display based on LED arrays." *Journal of Display Technology* 7.9 (2011): 503-514.
56. Jones, Andrew, et al. "An interactive 360 light field display." *ACM SIGGRAPH 2007 emerging technologies*. ACM, 2007.
57. Favalora, Gregg E., et al. "A multi-megavoxel volumetric 3-D display system for distributed collaboration." *proceedings of IEEE GlobeCom 2000 Workshop*. 2000.
58. Chinook, Chris. "LightSpace Back with Volumetric Display." *LightSpace Back with Volumetric Display*. Insight Media, 5 Apr. 2017. Web. 07 Dec. 2017. http://www.insightmedia.info/lightspace-back-with-volumetric-display/.
59. Milgram, Paul, and Fumio Kishino. "A taxonomy of mixed reality visual displays." *IEICE TRANSACTIONS on Information and Systems*77.12 (1994): 1321-1329.
60. "After Mixed Year, Mobile AR to Drive $108 Billion VR/AR Market by 2021." Blog post. DigiCapital, Jan. 2017. Web. 04 Dec. 2017. https://www.digi-capital.com/news/2017/01/after-mixed-year-mobile-ar-to-drive-108-billion-vrar-market-by-2021/#.WiV9plWnHuo
61. Pohl, Daniel, Gregory S. Johnson, and Timo Bolkart. "Improved pre-warping for wide angle, head mounted displays." *Proceedings of the 19th ACM symposium on Virtual reality software and technology*. ACM, 2013.
62. Carnegie, Kieran, and Taehyun Rhee. "Reducing visual discomfort with HMDs using dynamic depth of field." *IEEE computer graphics and applications*35.5 (2015): 34-41.
63. Massof, Robert W., et al. "37.1: Invited Paper: Full-Field High-Resolution Binocular HMD." *SID Symposium Digest of Technical Papers*. Vol. 34. No. 1. Blackwell Publishing Ltd, 2003.
64. Lanman, Douglas, and David Luebke. "Near-eye light field displays." *ACM Transactions on Graphics (TOG)* 32.6 (2013): 220.
65. Huang, Fu-Chung, Kevin Chen, and Gordon Wetzstein. "The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues." *ACM Transactions on Graphics (TOG)* 34.4 (2015): 60.
66. Ryana, Louise, Kevin J. MacKenziea, and Simon J. Watta. "Multiple-focal-planes 3D displays: A practical solution to the vergence-accommodation conflict?." *3D Imaging (IC3D), 2012 International Conference on*. IEEE, 2012.
67. Maimone, Andrew S. *Computational see-through near-eye displays*. Diss. The University of North Carolina at Chapel Hill, 2015.
68. Mirza, Kayvan, and Khaled Sarayeddine. "Affordable AR Displays: Focus on Optical See-Through Waveguide Technologies for AR Glasses." *Optinvent*. N.p., 13 Dec. 2016. Web. 07 Dec. 2017. http://www.optinvent.com/waveguide-ar-displays/
69. Langley, Hugh. "Trying Vuzix Blade 3000, the Smartglasses for Work and Play." *Wareable*. N.p., 06 Jan. 2017. Web. 07 Dec. 2017. https://www.wareable.com/smartglasses/vuzix-blade-3000-review .
70. Kress, Bernard C., and William J. Cummings. "Optical architecture of HoloLens mixed reality headset." *Digital Optical Technologies 2017*. Vol. 10335. International Society for Optics and Photonics, 2017.
71. Hua, Hong, and Bahram Javidi. "A 3D integral imaging optical see-through head-mounted display." *Optics express* 22.11 (2014): 13484-13491.
72. Hu, Xinda, and Hong Hua. "Design and assessment of a depth-fused multi-focal-plane display prototype." *Journal of Display Technology* 10.4 (2014): 308-316.
73. Liu, Sheng, and Hong Hua. "A systematic method for designing depth-fused multi-focal plane three-dimensional displays." *Optics express* 18.11 (2010): 11562-11573.
74. Liu, Sheng, and Hong Hua. "Time-multiplexed dual-focal plane head-mounted display with a liquid lens." *Optics letters* 34.11 (2009): 1642-1644.
75. Rolland, Jannick P., Myron W. Krueger, and Alexei Goon. "Multifocal Planes Head-Mounted Displays." Applied Optics, vol. 39, no. 19, 2000, pp. 3209-3215.
76. Ravikumar, Sowmya, Kurt Akeley, and Martin S. Banks. "Creating effective focus cues in multi-plane 3D displays." *Optics express* 19.21 (2011): 20940-20952.

77. Llull, Patrick, et al. "Design and optimization of a near-eye multifocal display system for augmented reality." *Propagation through and Characterization of Distributed Volume Turbulence and Atmospheric Phenomena*. Optical Society of America, 2015.
78. Maimone, Andrew, and Henry Fuchs. "Computational augmented reality eyeglasses." *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*. IEEE, 2013.
79. Maimone, Andrew, et al. "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources." *ACM SIGGRAPH 2014 Emerging Technologies*. ACM, 2014.
80. Durbin, Joe. "Testing the Future of AR Optics with Avegant Light Fields." *UploadVR*. UploadVR, 31 Mar. 2017. Web. 07 Dec. 2017. https://uploadvr.com/testing-avegant-light-fields/.
81. Robertson, Adi. "Why Do People Keep Giving Magic Leap Money?" *The Verge*. The Verge, 22 Oct. 2017. Web. 07 Dec. 2017. https://www.theverge.com/2017/10/22/16505430/magic-leap-augmented-reality-temasek-funding-investment-why