

Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays

A review of problem assessments, potential solutions, and evaluation methods

Gregory Kramida

Abstract—The vergence-accommodation conflict (VAC) remains a major problem in head-mounted displays for virtual and augmented reality (VR and AR). In this review, I discuss why this problem is pivotal for nearby tasks in VR and AR, present a comprehensive taxonomy of potential solutions, address advantages and shortfalls of each design, and cover various ways to better evaluate the solutions. The review describes how VAC is addressed in monocular, stereoscopic, and multiscope HMDs, including retinal scanning and accommodation-free displays. Eye-tracking-based approaches that do not provide natural focal cues – gaze-guided blur and dynamic stereoscopy – are also covered. Promising future research directions in this area are identified.

Index Terms—Vergence-Accommodation Conflict, Head-Mounted Displays



1 INTRODUCTION

The vergence-accommodation conflict (henceforth referred to as VAC), sometimes called “accommodation-convergence mismatch”, is a well-known problem in the realm of head- or helmet-mounted displays (HMDs), also referred to as head-worn displays (HWDs) [1]: it forces the viewer’s brain to unnaturally adapt to conflicting cues and increases fusion time of binocular imagery, while decreasing fusion accuracy [2]. This contributes to (sometimes, severe) visual fatigue (asthenopia), especially during prolonged use [3], [4], [5], which, for some people, can even cause serious side-effects long after cessation of using the device [6].

The current work is a checkpoint of the current state of the VAC problem as it relates to HMDs for augmented reality (AR) and virtual reality (VR). This review intends to provide solid and comprehensive informational foundation on supporting focal cues in HMDs for researchers interested in HMD displays, whether they are working on new solutions to the problem specifically or designing a prototype for a related application.

The remainder of this section presents a review of publications on the nature of the VAC problem and assess its severity and importance within different contexts. Section 2 describes a taxonomy of methods to address VAC in HMDs, comparing and contrasting the different categories. Section 3 covers specific designs for every method, addressing their unique features, advantages, and shortfalls. Subsequent Section 4 describes certain compromise approaches using eye tracking, which do not modify the focal properties of the display, but rather use software-rendered blur or alter the vergence cue instead. Section 5 addresses various ways and metrics that can be used to evaluate the effectiveness of

solutions. Finally, Section 6 identifies under-explored areas within the solution space.

1.1 The Vergence-Accommodation Conflict

The human visual system employs multiple depth stimuli, a more complete classification of which can be found in a survey by Reichelt et al. [5]. This survey finds that oculomotor cues of consistent vergence and accommodation, which are, in turn, related to retinal cues of blur and disparity, are critical to comfortable 3D viewing experience. Retinal blur is the actual visual cue driving the oculomotor response of accommodation, or adjustment of the eye’s lens to focus on the desired depth, thus minimizing the blur. Likewise, retinal disparity is the visual cue that drives vergence. However, there is also a dual and parallel feedback loop between vergence and accommodation, and thus one becomes a secondary cue influencing the other [4], [5], [7]. In fact, Suryakumar et al. in [8] measure both vergence and accommodation at the same time during the viewing of stereoscopic imagery, establishing that accommodative response driven from disparity and resultant vergence is the same as the monocular response driven by retinal blur. In a recent review of the topic, Bando et al. [6] summarize some of the literature about this feedback mechanism within the human visual cortex.

In traditional stereoscopic HMD designs, the virtual image is focused at a fixed depth away from the eyes, while the depth of the virtual objects, and hence the binocular disparity, varies with the content [9], [10], which results in conflicting information within the vergence-accommodation feedback loops. Fig. 1 demonstrates the basic geometry of this conflict.

The problem is not as acute in certain domains (such as 3D TV or cinema viewing) as it is in HMDs, provided that the content and displays both fit certain constraints. Lambooi et al. in [4] develop a framework of constraints for

- G. Kramida is with the Department of Computer Science, University of Maryland, College Park, MD, 20740.
E-mail: gkramida@umiacs.umd.edu

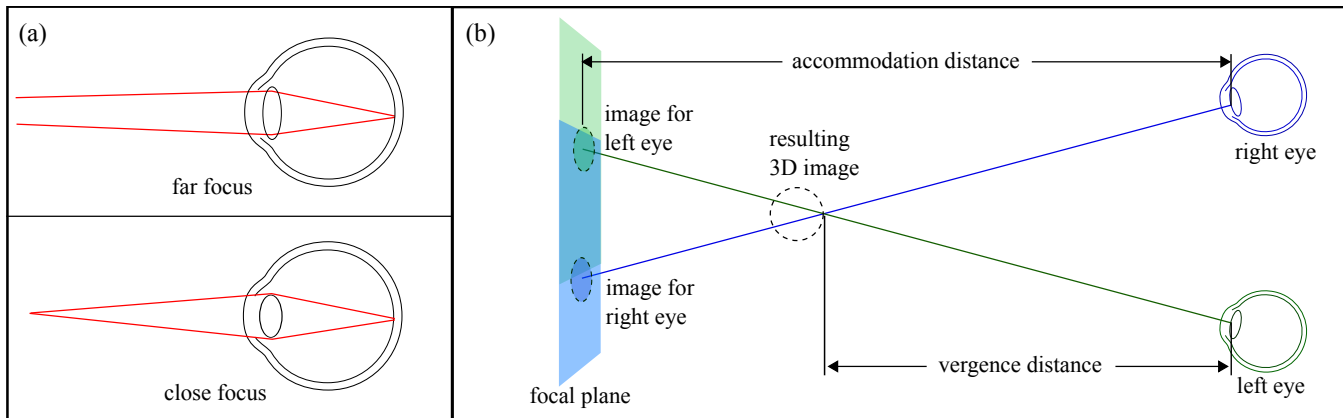


Figure 1. (A) Conceptual representation of accommodation within the same eye. Light rays from far-away objects are spread at a smaller angle, i.e. are closer to parallel, and therefore do not need to be converged much by the lens to be focused on the retina. Light rays from close-up objects fan out at a much greater angle, and therefore need to be redirected at a steeper angle to converge on the retina. The lens of the human eye can change its degree of curvature, and, therefore, its optical power, focusing the light from a different distance. (B) Conceptual representation of the VAC. Virtual display plane, or focal plane, is located at a fixed distance. The virtual objects can be located either in front or, if it is not at infinity, behind it. Thus the disparity cue drives the eyes to verge at one distance, while the light rays coming from the virtual plane produce retinal blur, which drives the eyes to accommodate to another distance, giving rise to the conflict between these depth cues.

such applications, the most notable of which in this context being that retinal disparity has to fall within the 1° “safety zone”, where the viewer’s eyes’ focus remains at or close to infinity. This indeed can be achieved in 3D cinematography, where virtual objects are usually located at a great depth and stereo parameters can be adjusted for each frame prior to viewing. Precise methodologies have been developed on how to tailor the stereo content to achieve this [11], [12], [13], [14].

However, these constraints have to be violated within the context of VR gaming [9], [10], [15] and the context of AR applications [16], where content is dynamic and interactive, and nearby objects have to be shown for a multitude of tasks, for instance – assembly, maintenance, driving, or even simply walking and looking around in a room.

I proceed by describing a taxonomy of methods used to address VAC in HMDs for AR and VR.

2 METHODS

Although the VAC problem remains generally unsolved in modern-day commercial HMDs, researchers have theorized about and built potential prototype solutions since early 1990s. Since the convergence cue in properly-configured stereo displays mostly corresponds¹ to natural world viewing, but the accommodation does not, vast majority of the effort on resolving VAC gears towards adjusting the focal cues to the virtual depth of the content.

The solution space can be divided along three categorical axes: extra-retinal vs. retinal, static vs. dynamic, and image-based vs ray-based. Each design uses either pupil-forming or non-pupil-forming optics, which come with their own advantages and disadvantages. Meanwhile, different see-through methods impose different constraints on various solutions. Fig. 2 for depicts a schematic view of the solution space and shows which categories describe the designs discussed in Section 3.

1. but not entirely, due to offset between virtual camera and pupil, as discussed later

2.1 Extra-Retinal Displays Versus Retinal Displays

The extra-retinal displays are a more traditional type of display in that they directly address a physical imaging surface or surfaces external to the eye. These displays typically use CRT², LCD³, DMD⁴, OLED⁵, or LCoS⁶ technology to form the image on a screen that emits rays in multiple directions.

Hence, extra-retinal displays can be observed from a range of angles. This range, however, may be limited by the use of pupil-forming optics, as discussed in the next section. Alternately, the eye box⁷ can be limited by the necessity to provide a sufficient number of rays to emulate a curved wavefront, as discussed in Section 2.3.

In contrast, retinal displays (RDs), which subsume retinal scanning displays (RSDs)⁸ and screen-based retinal projectors (RPs), are radically different from most image-forming displays in that they guide the rays to project the image directly to the retina.

RSDs scan a modulated low-power laser light beam via two or more pivoting mirrors (typically referred to as MEMS⁹ mirrors), through guiding optics, onto the pupil, forming the image on the retina rather than on an external physical screen. The reader is referred to [18] for a detailed review of laser-based display technology and to Section V of [17] for particulars on usage of MEMS mirrors in RSDs. RPs also project the image onto the retina. However, instead of scanning a beam onto the pupil, RPs shine collimated light

2. cathode-ray tube

3. liquid-crystal display

4. Digital Micromirror Devices are chips which host arrays of microscopic mirrors that can be individually rotated. As each mirror is rotated, the ratio of the on time to off time determines the shade of grey at the corresponding image point. See reference [17] for details.

5. organic light-emitting diodes

6. liquid crystal on silicon

7. the maximum volume within which the pupil center has to be for the intended viewing experience

8. also known as “virtual retinal displays”, or VRDs, a term widely used by researchers of the Human Interface Technology Laboratory of University of Washington (HITLab), among others

9. micro-electro-mechanical system

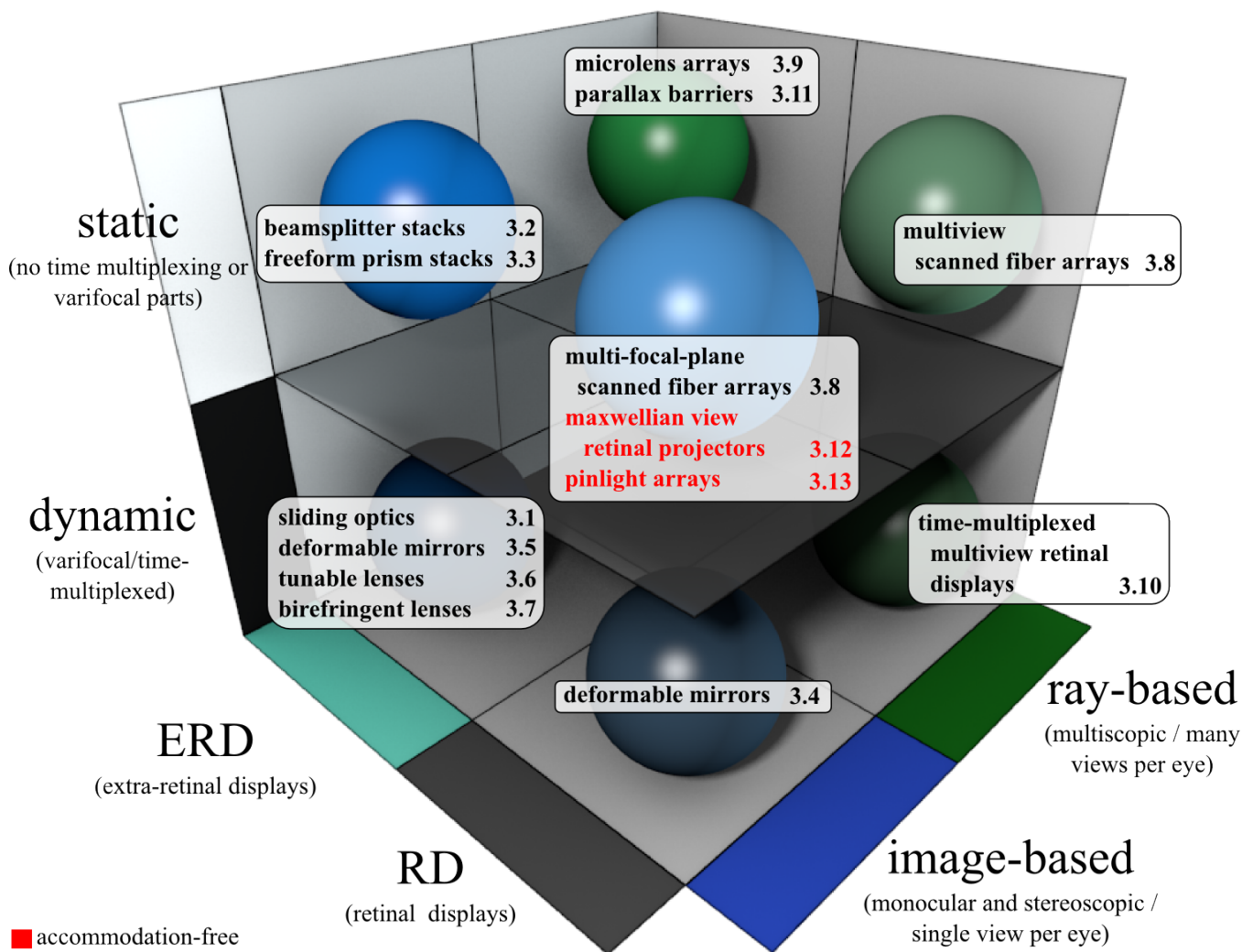


Figure 2. Classification tree of methods to provide focus cues in HMDs. Each method is followed by the number of the section where it is covered in detail.

through or off a modulation layer (typically LCD- or DMD-based), which forms an image. This light is then focused on a tiny spot (or multiple tiny spots) on the pupil, which results in the conjugate of the image being formed on the retina.

The primary advantage of RDs is that they potentially provide better sharpness, higher retinal contrast, and larger depth-of-focus (DOF)¹⁰ [19], [20]. There are several techniques to further extend the DOF, so that a greater depth range is in focus, as discussed in Section 3.12.

The primary disadvantage is that, while head-mounted RDs do not need a surface to form the image on, they always require complex pupil-forming assembly (discussed below). The resulting geometric constraints and considerations for eye rotation also impose limitations on the eye relief¹¹, angular field-of-view (FOV), and, sometimes, resolution [19], [21], [22]. Another challenge, posed specifically by RSDs, is the difficulty of achieving a high-enough scanning rate [1], [23]. Refer to Cakmakci et al. [1] for an overview of literature on scanning methods for RSDs and Section 3.8

for discussion of the newer scanned optical fiber method invented by Schowengerdt et al [24].

2.2 Pupil-Forming Versus Non-Pupil-Forming Optics

There exists a common classification which splits HMDs' optical designs into *pupil-forming* and *non-pupil-forming*. *Non-pupil-forming* HMDs do not require any intermediary optics to relay the microdisplay, hence the user's own pupils act as pupils of the HMD [25]. Such displays are variations of a *simple magnifier* [1], sometimes referred to as "simple eyepiece" [25], [26], which magnify a physical screen to form a virtual image at a greater distance from the eye [26]. Fig. 3 shows the optical parameters of a simple magnifier. The primary benefit of a simple magnifier is that it requires fewer optical elements and, typically, a shorter optical path than the alternative pupil-forming designs. [1] For instance, although it features multiple lenses, the multiscopic design discussed in 3.9 achieves its eyeglasses form-factor using the same principle.

In contrast, *pupil-forming* designs use optics similar to a compound microscope or telescope: they feature an internal aperture and some form of projection optics, which magnify

10. also known as depth-of-field

11. distance between the eye and the display

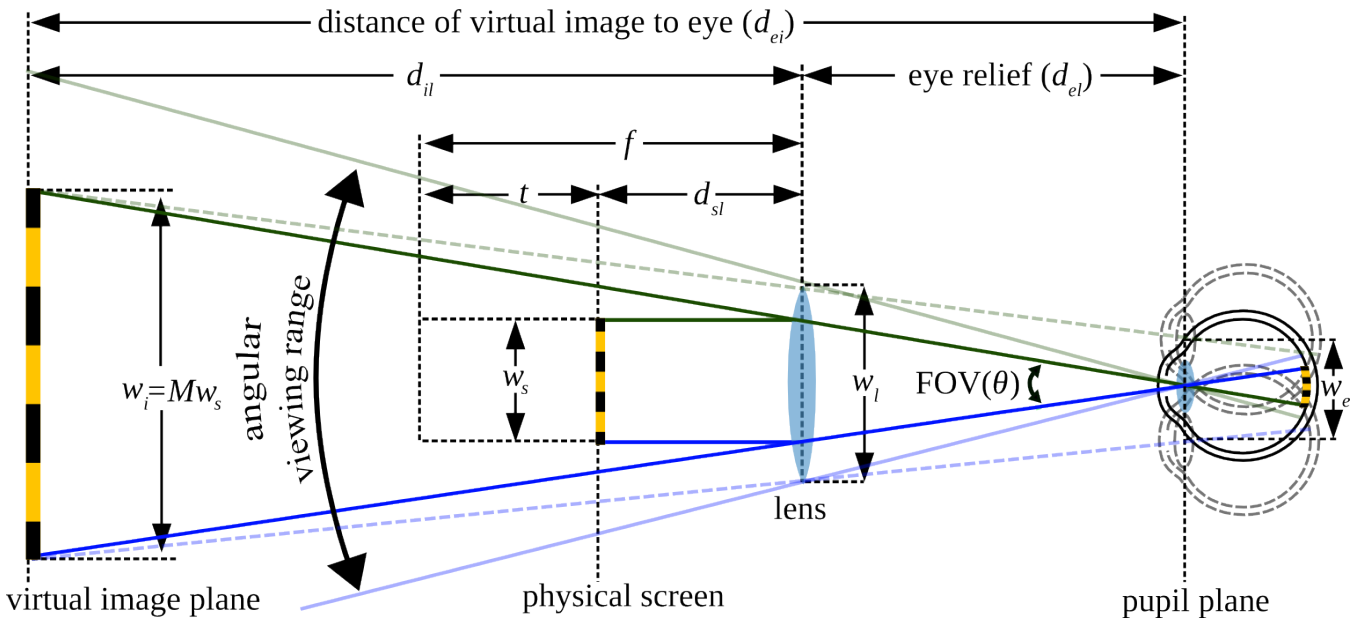


Figure 3. Optics of a simple magnifier, based on [27]. Subscripts e, i, l , and s represent “eye”, “(virtual) image”, “lens”, and “screen” respectively, so terms such as d_{il} explicitly denote “distance from image to lens”, w_l denotes “width of lens”; f is focal length; t is relevant for spatially-multiplexed MFP designs and represents the thickness of the display stack; M is the magnification factor from the physical to the virtual image. To allow viewing of the entire image, FOV must fit within the angular viewing range constrained by w_l and lateral offset of the pupil, which dictates the width of the eye box (w_e).

an intermediary image and relay it to the exit pupil [1], [25], [26]. These subsume the entire RSD category, since RSDs are essentially scanning projection systems [1]. The primary benefit of the more-complex projection systems is that, by allowing for a greater number of optical elements in the path to the exit-pupil, they can correct for optical aberrations [26] and even generate focal cues. For instance, some RSDs feature deformable mirrors, which focus the image at various depths, as discussed in greater detail in Section 3.4.

These benefits come at the cost of volume, weight, and complexity. Another drawback of pupil-forming optics is that increasing the optical path tends to reduce the FOV [28] and there is a trade-off between FOV and eye relief [21].

2.3 Image-Based Versus Ray-Based Methods

Independently from the see-through method (see Section 2.5) and the pupil-forming or non-pupil-forming optics, HMDs can be distinguished based on where they fall on the “extent of presence” axis of the taxonomy for mixed reality displays developed by Milgram and Kishino [29]. HMDs span the range including monocular, stereoscopic, and multiscopic displays. Although monocular heads-up displays cannot be used for VR or AR in the classical sense (they cannot facilitate immersive 3D [30]), if these are meant to display any information at a certain depth, i.e. a label at a specific point in space, just for one eye, the vergence-accommodation conflict still comes into play. Stereoscopic displays render a pair of views, one for each eye, with a disparity between the two views to facilitate stereo parallax. Monocular and stereoscopic designs both display one image per eye¹², hence this class of VAC solutions is referred to as

image-based. Multiscopic HMDs¹³, on the other hand, feature multiple views per eye. As discussed later, integration of rays from these views generates a *seemingly*-continuous light field, hence this class of approaches is referred to as *ray-based*.

Image-based methods can be further subdivided into three categories: discretely-spaced multiple-focal-plane methods, continuously-varying focus methods, and accommodation-free methods, further discussed in Section 2.3.1. One advantage of image-based methods is that computational requirements for rendering are typically less taxing when dealing with only one image per eye, even when separated into layers. Unlike their ray-based alternatives, image-based methods usually do not rely on optics with very small apertures that impose diffraction limits. Some challenges these designs pose include the difficulty to minify the design to an ergonomic form factor, the daunting refresh-rate requirements and blur problems in various multifocal designs, and the integration of fast, precise eye trackers for continuously-variable focus approaches to become practical.

The *ray-based* methods in multiscopic HMDs are fundamentally different. Multiscopic HMDs take their roots from autostereoscopic and multiscopic multi-view displays, which allow viewpoint-independent 3D viewing with a stationary screen via the integral imaging process. When this concept is applied to HMDs, multiple views are projected onto each eye to approximate a continuous light field.¹⁴

The underlying principle these displays use is called integral imaging and was first proposed by Gabriel Lipp-

13. also referred to as *light field displays*

14. Stereoscopic HMDs are, in some sense, a degenerate case of multiscopic HMDs with only one view per eye. Multi-view displays in general, however, may feature one-view per eye solely to achieve stereoscopy.

12. potentially separated into depth layers

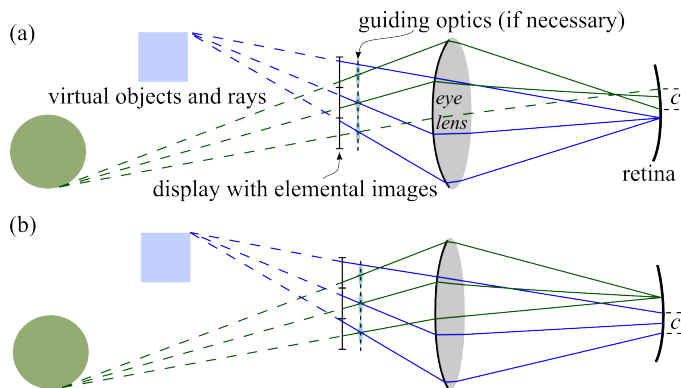


Figure 4. Principle of integral imaging. In (a) the eye accommodates closer, such that the ray set emanating from the blue object comes in focus on the retina, while the rays from the green object intersect before they reach the retina, thereby producing the circle of confusion marked as c . This emulates retinal blur. In (b), the eye accommodates farther away, such that the rays from the green object intersect at the retina causing it to appear sharp. Rays from the blue box, however, now intersect at a point behind the retina, resulting in the blue box being blurred.

mann in 1908 [31]. It involves generating multiple light rays corresponding to the same point in the virtual scene. This, in most cases, is equivalent to displaying multiple viewpoints of the same scene with a slight translational offset, which are called *elemental images*, as demonstrated in Fig. 4. The light rays corresponding to one point are guided in such a way that when they hit the retina, they emulate a cone of light fanning out from that point. The fan-out angle for each such set of rays causes the rays to intersect within the eye at different depths, which drives accommodation to bring this one scene point or another into focus.

To my knowledge, there is only one time-multiplexed multiscope HMD design published to date. It relies on a high-refresh-rate DMD screen and a galvanometer to generate the needed views. In contrast, the spatially-multiplexed multiscope designs achieve this using fine arrays (or layers of arrays) of microscopic optical elements, such as spatial light modulators, microlenses, and/or point light sources (“*pinlights*”) to properly angle the light rays from a screen subdivided into elemental images.

Ray-based designs may circumvent VAC and also be made more compact, but introduce other challenges. Since the generated light-field is angularly discrete, it is imperative for it to be dense enough to visually appear continuous: more elemental images are needed at smaller angular intervals. This comes at the expense of spatial resolution and may be complicated by diffraction limits, or, in the case of time-multiplexed viewpoints, places even more taxing requirements on the refresh rate than for time-multiplexed MFP image-based methods.

2.3.1 Image-Based Methods: Multi-Focal-Plane, Gaze-Driven, and Accommodation-Free

Image-based methods approach VAC in three different ways: they either (1) generate discretized addressable fo-

cal planes, as in multi-focal-plane (MFP)¹⁵ displays, (2) continuously vary the focal plane to trigger the desired accommodation response, or (3) present all content as if it were in focus, as in accommodation-free displays.

The earliest display prototypes supporting focal cues were built as proof-of-concept systems capable of displaying only simplistic images, often just simple line patterns or wireframe primitives. These either manipulate the focus to correspond to the vergence at a pre-determined set of points, or provide some manual input capability to the user to manipulate the X and Y coordinate of the desired focal target and adjust the focal plane in a continuous fashion to the depth at this target.

MFP HMD designs were proposed just prior to the turn of the century. These subsume depth-fused-3D displays and virtual retinal 3D displays, since they share the same principle: they show each virtual object at the focal plane closest to the depth of the object, thus emulating a volume using different regions of a single image. This approach can greatly shrink or eliminate the gap between the vergence and accommodation cues. By modulating the transparency of pixels on screens projected to different focal planes, virtual objects may be made to appear between the planes, and the depth effect is further enhanced. Several *depth-blending* models have been developed to control transparency of pixels; these models are discussed in greater detail in section 3.2.

If the user’s gaze point is known, depth of the virtual content may be determined, and tunable optics can continuously adjust the focal plane to this depth. Many advancements were made to integrate eye trackers in HMDs for this and other purposes, as discussed in Section 4. With this concept in mind, some tunable lens prototypes were designed to operate in either of two modes, a variable-focal-plane mode or a multi-focal-plane mode [34], [35].

In stark contrast to all other techniques this review describes, accommodation-free displays do not strive to provide correct focus or retinal blur to drive the accommodation cue. Instead, they display all content as if it were in focus at the same time, regardless of the eye’s accommodative state. To achieve this, most of these displays capitalize on Maxwellian view optics to expand the DOF. This method is covered in greater detail in Section 3.1.2.

2.4 Static (Space-Multiplexed) Versus Dynamic (Time-Multiplexed) Methods

Solutions falling in the *dynamic* category change the image (and, in certain cases, tune the optics) to provide focal cues,¹⁶ while the *static* ones do not.¹⁷ For image-based methods, only varifocal optics may be used for continuously-varying focus approaches, hence they fall into the dynamic category. In contrast, accommodation-free displays

15. I abstain from using the terms “voxel” [32] or “volumetric” [33] when describing MFP displays to avoid confusion with stationary volumetric voxel displays, which are contrasted to MFP HMDs in some of the literature [34].

16. or, in case of accommodation-free displays, to keep all content in focus

17. The terms “static” and “dynamic” used in this review refer strictly to the displays’ focusing method.

are strictly static, as they do not need to vary between viewpoints or depth layers. In both MFP and ray-based displays, the solution can be static or dynamic, depending on whether the focal planes or views are *space-* or *time-multiplexed*. In a nutshell, static approaches boast fewer (if any) moving parts, but often incur the need for a greater number of screens, compensating optics, or scanned projectors, while dynamic approaches involve fewer screens or projectors, but need to provide appropriate tunable optics or scanning mechanisms paired up with much faster refresh rates.

In extra-retinal MFPs, one challenge of using the space-multiplexed approach is the difficulty of stacking the surfaces for each focal plane in a light and compact way, potentially offsetting the compactness advantage gained by omitting focus-driving electronics. This can be, to an extent, addressed by freeform optics, but any slight increase in the FOV still comes at a multifold cost in weight and volume compared to designs with only one imaging surface before the eye. Yet another common problem of stacking screens is the loss of contrast and sharpness caused by blur in out-of-focus planes between the eye and the in-focus plane [36]. An additional challenge strictly in the AR domain is adding the optical-see-through capability, which is compounded by having to work with more surfaces between the eyes and the real world. These problems may potentially be circumvented in MFP RSDs using scanned fiber arrays (see Section 3.8), which form the image on the retina only rather than multiple surfaces, but only if optical fiber projectors are compact enough and can be cost-effectively produced. Also, the resolution of each projector is constrained by the diffraction limit imposed by the exit lenses it uses, resulting in a trade-off between the projector size and resolution [33]. Somewhat similarly, for microlens array HMDs, which also fall into the “static” category, diffraction of the microlenses imposes a constraint on pixel pitch and therefore the overall spatial resolution of the display, as discussed in Section 3.9.

The main advantage of dynamic designs is that they do not necessarily require multiple screens or projectors. This arguably yields a more compact design. There are several designs that add optical-see-through capabilities to MFP displays, either by a compensating freeform prism or via beamsplitters. However, a challenge for dynamic MFP displays is providing sufficiently high refresh rate *and* the response rate of tunable optics to which the image-swapping is synchronized, in order to achieve flicker-free quality at each of the focal planes. In dynamic ray-based RSD approaches, tunable optics are not necessary, but even greater refresh rates are required to display a sufficient number of views per frame at a high frame rate. These problems are alleviated to varied extent by advances in varifocal optics and screen technology.

2.5 AR See-Through Method

HMDs for VR are typically opaque, since they aim to fully immerse the user in a virtual environment (VE)¹⁸.

For AR, the displays fall into two general categories, optical-see-through (OST) and video-see-through (VST) [1].

18. Although it has been suggested to optionally display a minified video-feed of the outside world to prevent the user from running into real obstacles while exploring VEs

OST systems let through or optically propagate light rays from the real world and optically combine them with virtual imagery. Video see-through displays capture video of the real world and digitally combine it with virtual imagery before re-displaying it to the user.

When choosing how to address VAC in an AR HMD, it is important to consider the implications and trade-offs imposed by each see-through method. For this purpose, a more detailed comparison is provided in Table 1 in the appendices. Many of the designs covered in this review have been combined with both VST and OST methods in the past. However, in some designs providing OST capabilities may be impractical. In virtually all the designs, the OST method calls for additional beamsplitters, compensation lenses, and/or dynamic opacity masks, which may add to the design complexity, weight, and volume of the HMD, and may limit the FOV.

3 DESIGNS

This section covers various HMD designs that address VAC, the underlying technology and theoretical models, and the earlier experimental proof-of-concept bench prototypes. Each distinctive technology used in these HMD designs has a unique set of benefits and challenges. In many cases, more than one technology or principle can be combined in the same design to yield the best of both, such as the design by Hu et al. [37] using freeform optics and deformable mirrors described in Section 3.5.

3.1 Sliding Optics

Sliding optics designs feature a display and some relay optics on the same axis with the observer’s pupil. Either the display or the relay optics are moved mechanically along the axis. These designs are image-space telecentric systems, similar to focus control in conventional photo and video cameras. In such systems, when a relay lens is moved and the focal distance to the virtual image (d_{ei}) changes, the angular FOV(θ), remains constant [38].

The first experimentally-implemented sliding optics design is that of Shiwa et al. [39]. In their proof-of-concept prototype, images for both eyes were rendered in two vertically-separated viewports of a single CRT monitor. Relay lenses were placed in the optical paths between the exit lenses and the corresponding viewports. The relay lenses had the ability to slide back and forth along the optical path, driven by a stepper motor. The prototype initiated the focus point at the center of the screen and provided manual (mouse/key) controls to move it. As the user moved the focus point, the relay lens changed the focus to the depth of the content at this point. Shiwa et al. suggest that eye tracking should be integrated to set the focal depth to depth of the content at the gaze point.

Yanagisawa et al. [40] constructed a similar 3D display with an adjustable relay lens. Shibata et al. [41] built a bench system that, instead of changing the position of a relay lens, changed the axial position of the screen in relation to the static exit lens according to the same principle, varying the focus from 30 cm to 2 m (3.3 D to 0.5 D). Meanwhile, Sugihara et al. [42] produced a lightweight HMD version of the earlier-described bench system by Shiwa et al. [39].

Shiwa et al. [39] relied on the specifications of the optometer developed in [43]¹⁹ to determine the necessary speed of relay lens movement. Their mechanism took less than 0.3 seconds to change focus from 20 cm to 10 m (5 D and 0.1 D, respectively)²⁰, which they asserted was fast enough. A recent study on accommodation responses for various age groups and lighting conditions reaffirms this [44]²¹: the youngest, fastest-accommodating age group in the brightest setting showed an average peak accommodation velocity of only 1.878 ± 0.625 D/sec.

Although this may be fast-enough for on-demand accommodation, sliding optics are, unlike tunable lenses, too slow for a flicker-free MFP display [23]. On the other hand, adjusting the optics continuously to match the focus to the depth of the gaze point would require determining either the gaze point [35] or the accommodation state [45] of the eye in real time.

3.2 Multi-Focal-Plane Models and Depth Blending

The concept of using multiple focal planes in HMDs originates from a 1999 study by Rolland et al. [16], which examines feasibility of stacking multiple display planes, each focused at its own depth, and rendering different parts of the image to each plane simultaneously. The original idea is, at each plane, to render those pixels that most closely correspond to the depth of that plane, while leaving other pixels transparent. The viewers would then be able to naturally converge on and accommodate to an approximately correct depth, wherever they look. Rolland et al. [16], [46] derive length of the intervals between focal planes (dioptric spacing), the total number of planes required, and requirements for pixel density at each plane. They find that, based on stereoacuity of one arcmin, natural viewing requires a minimum of 14 planes between 50 cm and infinity, with interplanar spacing at $1/7$ D (Fig. 5).

They also suggest that if a fixed positive lens is positioned in front of the focal planes, physical thickness of the display can be greatly reduced. Their framework is analogous to fig. 3, so the thickness of the resulting display stack can be expressed as:

$$t = f - d_{sl} = \frac{f^2}{f + d_{il}} = \frac{f^2}{f + d_{ei} - d_{el}} \quad (1)$$

In the above equation, d_{ei} is the nearest distance to which the human eye can accommodate, while d_{sl} is the offset between the lens and first screen in the stack, which displays virtual objects at that distance. d_{sl} can be expressed as:

$$d_{sl} = \frac{1}{\frac{1}{f} + \frac{1}{d_{il}}} = \frac{fd_{il}}{f + d_{il}} \quad (2)$$

Based on these equations²², for a 30 mm focal length, 25 cm closest viewing distance, and 25 mm eye relief, d_{sl} would be 26.5 mm, and the stack thickness t would be 3.5 mm, resulting in an overall minimum display thickness of about

19. This optometer detected accommodation to within ± 0.25 D at the rate of 4.7 Hz.

20. 1 diopter (D) = 1/m

21. Subjects re-focused from a target at 4 m to one at 70 cm away

22. See appendix A for derivation

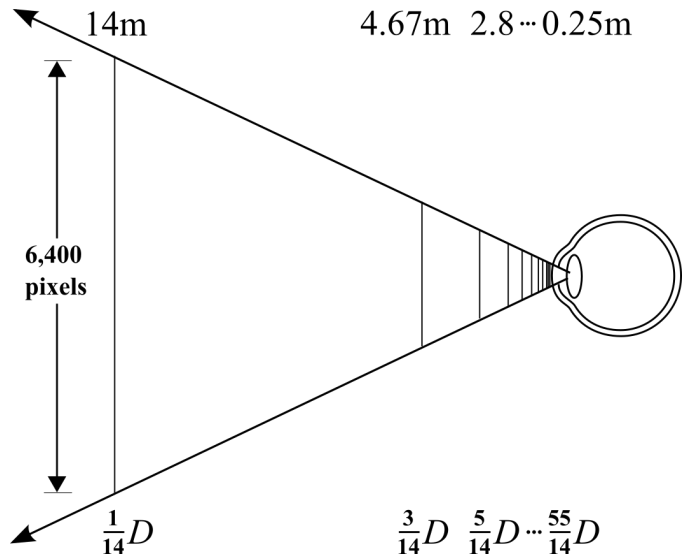


Figure 5. Stereoaucuity-based MFP display model by Rolland et al. [46]. The 6,400 x 6,400 resolution at each plane yields a minimum spatial resolution of 0.5 arcmin.

3 cm. Rolland et al. [16] conclude that HMDs using this model can be built using contemporary technology.

Liu et al. [35] pointed out that a practical application of this model is challenging, since no known display material had enough transmittance to allow light to pass through such a thick stack of screens. Suyama et al. [47], [48], [49], [50] describe a phenomenon they name “depth-fused 3-D”, or “DFD”, where two overlapped images at different depths can be perceived as a single-depth image. They built a bench prototype with two image planes, which they used to experimentally determine that as luminance ratio is changed, the perceived depth of the perceived content between the planes changes approximately linearly. Thus, by varying the intensity across each image plane, they were able to emulate the light field and generate what appears as 3D content between the two image planes.

Akeley et al. [51] designed and implemented a depth-fused MFP bench display with three focal planes. Viewports rendered on a single high-resolution LCD monitor are projected via mirrors onto beamsplitters at three different depths. Akeley et al. have implemented a depth-blending (also referred to as “depth-filtering”) algorithm, based on the luminance-to-perceived-depth relationship discussed above, to vary intensity linearly with the difference between virtual depth and the depth of the actual plane on which they are shown. Their user study showed that fusion time²³ is significantly shorter when consistent depth cues are approximated with this prototype than when only the nearest focal plane is used, especially for distant content.

Later, Liu and Hua [52] presented an elaborate theoretical model for designing depth-fused sparse MFP displays. It primarily focuses on two aspects: (1) the dioptric spacing between adjacent focal planes, based on the depth-of-field criterion rather than stereoacuity, and (2) a depth-weighted blending function to better approximate a continuous volume. They developed their own depth-blending model dif-

23. time it takes to fuse the left and right images

ferent from the linear model described by Akeley et al. [51] in that it takes into account the modulation transfer function at various intensity ratios, aimed at maximizing the contrast of the perceived depth-fused image.

In a later work, Ravikumar et al. [53] compare different depth blending functions. They have repeated Liu and Hua’s analysis and confirm that, in some cases, there is a deviation from linearity that yields greater contrast of the retinal image, but this deviation is opposite to what Liu and Hua suggested. However, Ravikumar et al. show that, after incorporating typical optical aberrations and neural filtering into their model, the linear blending rule is actually superior to the non-linear one in driving the eye’s accommodative response and in maximizing retinal-image contrast when the eye accommodates to the intended distance.

MacKenzie et al. [36], [54] experimentally establish requirements for plane separation in multifocal displays with depth blending. Both experiments used a spatially-multiplexed MFP bench-type display with linear depth blending. The first experiment tested the monocular case, establishing $[6/9D, 10/9D]$ as the acceptable range for plane separation. The second experiment tested the binocular case, where the authors found that accurate accommodation cues are triggered with spacing within the more-strict $[0.6D, 0.9D]$ interval, requiring a minimum of 5 planes between 28 cm and infinity.

The results of these experiments indicate that the plane separation requirements dictated by Liu and Hua’s model [52] are sufficient in practice for both the monocular and stereoscopic MFP displays. MacKenzie et al. [36] also note that for spatially-multiplexed displays, contrast (and, therefore, sharpness) is attenuated due to one or more planes between the eye and the target being defocused, an effect present at $8/9 D$ spacing and even more drastic at $6/9 D$ and beyond.

3.3 Static Freeform Prism Displays

By supporting off-axis display components, freeform optical elements allow a greater degree of freedom in designing compact eyewear, especially HMDs with complex optical properties. A brief survey on the use of freeform optics in HMDs is included in Appendix B. Stacking freeform prisms with multiple off-axis microdisplays results in a single MFP display with no moving parts.

Cheng et al. [55] design a spatially-multiplexed MFP display stacking a pair of custom-engineered freeform prisms. The freeform prisms reflect the light from two off-axis microdisplays into the eye, as shown in Fig. 6.

One problem with the MFP stacked freeform prism design is that it is much bulkier than ordinary eyeglasses. Fig. 6 shows that the proposed design features only two focal planes and is already thicker than 2 cm. The two focal planes are separated by $0.6D$, yielding a range from 1.25 to 5m. While such separation adheres to the prescribed formula, not being able to accommodate within 1.25 m possibly inhibits any tasks involving hand manipulation. To provide contiguous accommodation cues over the entire range, as dictated by Liu and Hua’s model [52] and experimentally confirmed by MacKenzie et al. [36], the design would need five focal planes, increasing the thickness yet further.

OST requirements would amplify the thickness problem: if freeform prisms guide the digital imagery, additional

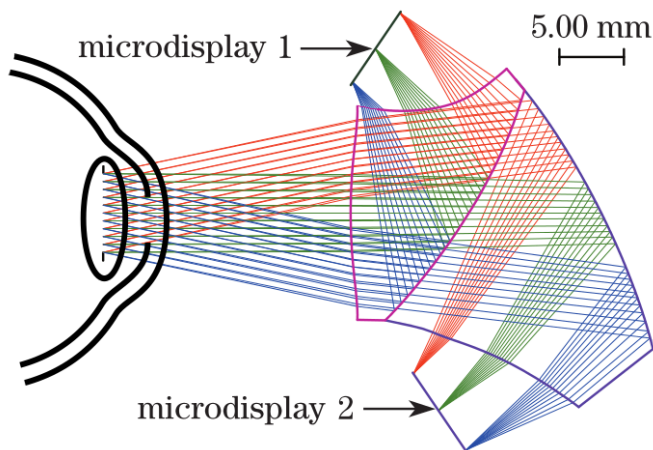


Figure 6. Design of a spatially-multiplexed MFP display using two freeform prisms, adapted from [55]. The design features a 40° monocular FOV.

prisms are required in order to compensate for the distortion of the environment image. Even the *single*-focal-plane tiled-prism OST design by Cheng et al. [56] featured rather bulky 17-mm-thick prisms. Moreover, freeform prism designs involve a significant FOV-to-compactness trade-off [57], which is only compounded by adding more prisms.

Yet another problem is that such designs are still prone to the same contrast and sharpness loss problem described by MacKenzie et al. [36], even in the case with only two surfaces, which would also be more acute if the number of focal planes were increased by adding more prisms.

Finally, there is distortion caused by the prisms reflecting the display at an angle. While the proposed design is optimized to reduce the keystone effect down to 10%, to fully cancel it would require computationally pre-wrapping the images before they are shown at the cost to resolution and latency. Hu et al. [37] address the same problem in their see-through time-multiplexed design that also uses freeform prisms²⁴, and achieve a distortion of only 5% at the edges of the 40° FOV monocular region that is critical for fusing the left and right images. However, their optical design is different, featuring only one display and a single prism to reflect it (aside from the compensation prism for real-world light). Provided the above static multifocal design is fully optimized, it remains to be shown if a spatially-multiplexed freeform-prism design with negligibly-low distortion is possible.

3.4 Deformable Membrane Mirrors in RSDs

A MOEMS²⁵ deformable membrane mirror (DMM) typically consists of a thin circular membrane of silicon nitride coated with aluminum (or similar materials) and suspended over an electrode. The surface of the mirror changes its curvature depending on the voltage applied to the electrode, thus directly re-focusing the laser beam being scanned onto the retina. In displays with tunable optics, DMMs can be used to alter the required accommodation to view the displayed objects without blur.

24. discussed in greater detail in Section 3.4

25. micro-opto-electro-mechanical system

DMMs stand out among other varifocal optics for several reasons. Whereas tunable lenses, birefringent lenses, and sliding optics can all be used in a telecentric way, DMMs require a more complex off-axis pupil-forming optical assembly, since the light has to be reflected from them, corrected for aberrations, and only then guided to the eye. However, their optical power can be adjusted really fast, allowing for time-multiplexing focal depths at 1-100 KHz rates, rivaled only by ferroelectric liquid crystals, the newer blue-phase liquid crystals, and electro-acoustic lenses.

In [58], McQuaide et al. at the Human Interface Technology Laboratory (HITLab)²⁶ use a monocular RSD with a DMM in the optical path to generate correct accommodation cues. They achieve a continuous range of focus from 33 cm to infinity (3 D to 0 D).

Schowengerdt et al. [59] took the DMM RSD design to the next level. They used a beamsplitter to separate the laser beam into left and right images, making the display stereoscopic, expanded the focal range to $[0D, 16D]$, exceeding the accommodation range of a human eye, and placed additional beamsplitters at the exit pupils, demonstrating that such displays can be used for AR.

These HITLab prototypes were bench proof-of-concept systems that displayed very basic images (line pairs). Their creators used autorefractors to experimentally demonstrate that observers' accommodative responses match the desired focal depth.

While this research shows that DMMs can refocus the scanning display continuously, they can also be switched fast enough between a series of focal planes to create an illusion of a contiguous 3D volume, i.e. generate addressable focal planes in a varifocal fashion. Schowengerdt and Seibel [32] synchronized the membrane curvature changes with swapping between content at two different depths at every frame, generating a frame-sequential multiplanar image. Theoretical frameworks for such MFP displays are discussed in greater detail in Section 3.2. Schowengerdt and Seibel continue their work on RSDs providing focal cues, but move away from deformable mirrors in favor of arrays of miniature scanning fiber-optics projectors, which are discussed in Section 3.8.

3.5 Deformable Membrane Mirrors in Screen-Based Displays

Hu and Hua [37] designed a screen-based OST HMD with a DMM²⁷ and implemented a monocular prototype (Fig. 7). Their image generation subsystem (IGS) consisted of a DMD display, the DMM device, additional lenses to scale up the optical range of the DMM, and a polarization beam splitter guiding the beam to the eyepiece.

The DMM allowed switching between optical powers at up to 1kHz. Synchronized to the display switching between content at six different depths, it axially and time-sequentially moved the projected intermediate image, generating six evenly-spaced focal planes in the range of 0 D to 3 D after additional magnification. To increase compactness of the design, Hu and Hua [37] incorporated a freeform prism into the eyepiece, which allowed placing the IGS off-center,

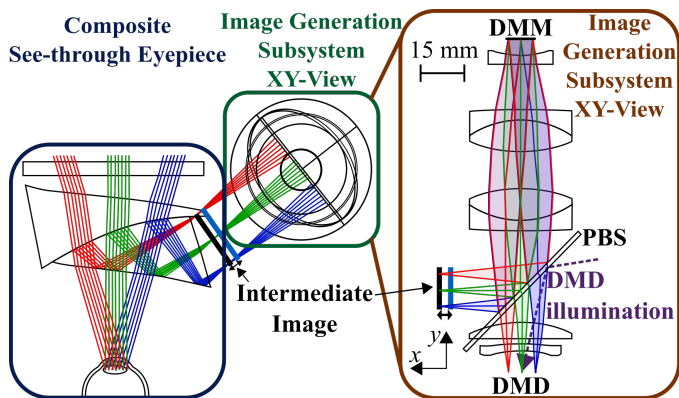


Figure 7. Design of a time-multiplexed MFP display with using two freeform prisms from [37]. In the image generation subsystem shown on the right, image from the DMD display passes through the polarization beamsplitter, and is re-focused by the DMM. The two lenses in-between are used to pre-magnify the image and to correct the lateral chromatic aberration. After reflection from the DMM and magnification, the image is reflected by the beamsplitter into the freeform prism which then guides it into the eye, as shown in the center and on the left.

and a compensation prism to cancel the distortion of the real imagery by the first prism. By optimizing the freeform eyepiece, they achieved a total 50° by 45° monocular FOV for the see-through imagery, with the central 40° low-distortion area for overlaying with virtual imagery, which they find suitable for proper stereoscopic fusion.

3.6 Tunable Lens Displays

Several technologies exist for making lenses with dynamic optical power: electro-optical, electromechanical, thermo-optical, and acoustomechanical. A survey of these technologies can be found in [60]. Many of such lenses can alter optical power fast enough that, if synchronized to images corresponding to different focal planes, they generate proper accommodation cues in the focal range between the planes. Therefore they can be used in time-multiplexed MFP displays.

The first to use a tunable lens for this purpose were Suyama et al. [47]. The dual-frequency liquid crystal (DFLC) lens in their design could be adjusted to any optical power between -1.2 D to +1.5 D at a rate of 60 Hz. Another static lens was placed between the exit pupil and the varifocal lens in order to keep FOV of the output image constant. A CRT display, switching between content at different depths, was synchronized to the lens. Suyama et al. captured images of the resulting monocular MFP prototype showing basic geometric shapes with a camera focused at different depths. They confirmed that correct object parts appear in focus on the reconstructed images.

Li et al. [61] employed liquid crystal lenses to develop and implement a glasses-thin prototype of adjustable eyewear for use by far-sighted people (presbyopes), whose optical power varies dynamically between 1.0 D and 2.0 D. Although this prototype is not a display, this work suggests that liquid-crystal lens displays could potentially be minified to eyeglasses-form-factor.

Later, Liu and Hua [23] built a proof-of-concept monocular prototype using an electrowetting varifocal lens. The lens

26. at University of Washington, www.hitl.washington.edu

27. see previous section for a brief description of DMM technology

they initially used could change between any two states in the range [-5 D, +20 D] within 74 ms (yielding a rate of 7 Hz), but they also tested the speed of several alternative lenses with response speeds up to 9 ms (56 Hz), which approach the 60 Hz frequency. With an additional magnification lens, optics of the entire display could vary focus between 8D and 0D (12.5 cm and infinity). They continued their research in [34], where they describe how they integrated the 9-ms-response liquid lens and made it oscillate between two different focal planes, synchronized to rendering the two corresponding views. This time, the update frequency was limited by the graphics card, achieving a rate of 37.5 Hz.

One problem with the electrowetting lens that Liu and Hua [34] identified is that, during settling time of the lens, when its driving signal is switched, there are longitudinal shifts of the focal planes, which yield minor image blur and less accurate depth representations. They hypothesized that this problem can be mitigated by a liquid lens with a response time at or above 60 Hz. Subsequently, Liu et al. [35] incorporated their liquid lens mechanism into an HMD. They tested it on ten subjects and determined the error rate in a basic depth estimation task, at the same time measuring the actual accommodation response with a near-infrared autorefractor. They showed that their approach yields better accommodation cues than conventional stereoscopy.

A critique of the tunable lens technique by Love et al. [62] is that a switchable-focal-plane display requires a minimum of four focal planes, not two, and, even provided a liquid lens frequency of 60 Hz, the display could yield a maximum refresh rate of only 12.5 Hz. Such low update frequencies would produce flicker and motion artifacts. However, newer blue-phase liquid crystal lenses are known to achieve sub-millisecond response times [63], and should be able to produce sufficient refresh rates in MFP prototypes with five or more focal planes.

3.7 Birefringent Lenses

Love et al. [62] built a MFT bench prototype that is time-multiplexed using light polarization. They used two birefringent lenses out of calcite interspersed with polarization switches. They took advantage of the fact that, while calcite is highly transparent, birefringent lenses have two different indices of refraction: one for light polarized along one crystalline axis and another for the light polarized along the orthogonal axis. Thus, for light with different polarization, the lenses would have different optical power.

The prototype featured only two lenses, each with two optical powers, which combined to produce one of four different focal planes. Polarization of light for each lens was controlled by a photoelectric liquid-crystal polarizer. Love et al. used the shutter technique for switching between volumetric slices of images with different polarization, and achieved a frame-rate of 45 Hz using two CRT monitors, one for each eye. The design demonstrated superior transmittance between focal planes. However, to my knowledge, there have been no published attempts to minify this bench design to fit into an HMD.

Aside from high transmittance of the calcite, an advantage of using such polarization switches for alternating between focal states is that response time for ferroelectric optics

is of the order of μs [64], providing a flicker-free experience given a fast-enough display. The clear limitation of this design is that it provides a fixed number of discrete focal states, although each additional lens doubles this number. OST designs using birefringent lenses would call for off-axis optical assemblies as complicated as Hu and Hua's assembly described in Section 3.4, but set-ups for VR and VST AR could remain telecentric.

3.8 Scanned Fiber Arrays

As an alternative to the more-common raster-scanned laser displays, Schowengerdt et al. [24] designed and built a full-color optical fiber projector scanning in a spiral pattern. The design allows minifying the projection head down to 1.1 mm \times 9 mm. A red, green, and blue laser beam are fed into an optical fiber via an RGB combiner. The fiber reaches the miniature projection head, where it runs through a hollow piezoelectric actuator, and terminates with a flexible cantilever. The actuator vibrates the fiber tip at rates of about 10 kHz, producing a circular scanning motion. By increasing amplitude of the drive signal over the course of each frame, the circular scan grows into a dense spiral, spanning angles up to 100°. At this actuator resonance rate, a single projector can scan 2000 pixels per each of the 250 rings per refresh cycle, at an overall refresh rate of 30 Hz.

The miniature size of these projectors allowed Schowengerdt et al. [65] to integrate them into a relatively-compact RSD HMD prototype. Schowengerdt et al. [33] then produced a bevelled array of these scanned fiber projectors, where each head is offset from the previous to project to its own focal depth. Fed through a single X-Y scanner and guiding optics, the beams would produce multiple focal planes, each in focus on the retina at a different accommodation state of the eye.

In [57], Schowengerdt et al. describe how the scanned fiber arrays can be used to make *multiscopic*, rather than multifocal HMDs. Rather than time-multiplexing the generated views as described in Section 3.10, they propose to multiplex in space, with each of the projectors producing its own view. The compact size of the projector heads allows to bundle a great number of them at slightly different angles within 3mm of each other. This technique can potentially produce a massively multiscopic HMD with a 120° FOV and a sufficiently high refresh rate to display smooth parallax motion.

The one problem with the scanned fiber arrays is that despite their compactness, achieving an eyeglasses-like form factor is challenging. Scores of 1.1 \times 9 mm projector heads are only part of the problem: ergonomically placing the lasers illuminating the optical fibers is also challenging.

3.9 Microlens Arrays

Lanman and Luebke [27] at Nvidia Research designed and built a multiscopic HMD prototype using a microlens array to magnify the images produced by an OLED screen. They subdivided the screen into multiple tiles, each tile showing a single elemental image. Due to the need for overlap between views, this kind of set-up greatly reduces the spatial

resolution of the display²⁸. The image tiles were magnified using a sheet of microlenses placed between the image and the eye, which allowed Lanman and Luebke to minify their prototype to an eye-glasses form factor.

The operation principle of this display is illustrated in Fig. 4. Rays from the same point in the virtual scene are relayed by multiple lenses to different locations on the pupil. The spread of these rays on the pupil varies with the offset of the point from one display section to the other. Rays from closer objects have a wider spread, while rays from more distant objects are closer to parallel, mimicking the natural viewing situation. The circles of confusion generated by ray bundles from multiple lenses emulate retinal blur. Hence, the eye tends to accommodate to objects within the virtual scene rather than the virtual image plane, but at the expense of greatly reduced spatial resolution, which, as Lanman and Luebke anticipated, may soon become acceptable given current technology trends.

However, increasing angular resolution of the display would call for increasing microlens density, which, in turn, causes increased diffraction and unwanted blur. An alternative design, using concave, instead of convex lenses has been proposed by Hong et al. [66]: it allows to extend the angular resolution of the display and increases the range of representable depths.

Another drawback of this design is that it may only support the VST operational model, since microlenses would distort the image of the physical environment, while the artificially back-lit screen would block it. To address this, Song et al. [67] proposed an optical see-through design using either microlenses or pinholes together with a pair of freeform prisms. The first prism guides light rays from the optical micro-structures, which are located off-axis, while the second prism compensates for distortion of light rays from the environment. Hua and Javidi [68] fabricated and tested a monocular prototype of a similar design with 33.4° FOV. Unfortunately, these designs are prone to the same excessive thickness problems and FOV limitations as any other freeform prism designs.

3.10 Time-multiplexed Multiview Retinal Displays

In [69], Kim et al. describe a multiscopic HMD prototype they built. It used a rotating galvanometer scanner synchronized to a DMD screen alternating between 32 slightly-different viewpoints of the scene for each frame at 30 frames per second, at an overall refresh rate of 960 Hz. The galvanometer changed the angle at which rays from the display fall on a relay lens for each viewpoint, which then directed the rays through a tight spot the observer’s pupil onto the retina, thereby realizing the Maxwellian view separately for each elemental image.

Kim et al. analyzed the light field produced by their experimental bench system by placing a camera at eye’s location and recording a sequence of lines shown by their display at different depths. They conclude that focal cues produced are good enough to control the eye’s accommodative response.

28. The physical 1280x720 pixel OLED display of the prototype in [27] yielded an effective spatial resolution of 146x78

The benefit of this design is that, unlike the microlens array displays discussed above, its screen does not need to be split into separate sections for elemental images, hence the multiscopic quality poses no limit to spatial resolution.

The main challenge with such a design is that the display update frequency limits how many elemental images can be displayed per frame at a fast-enough overall refresh rate. As recent developments in VR consumer-end products demonstrate, refresh rate of 75 Hz and upwards may be necessary to avoid the nausea-inducing motion blur effect [70], which would take the number of viewpoints down to 12, below the 32 that the authors consider to be the required minimum, and certainly below the number of views that can be produced by, for instance, the scanned fiber array or microlens array methods.

3.11 Parallax Barriers

Parallax-barrier multiscopic displays have recently been adapted for usage in HMDs by Maimone et al. [71]. Multiple SLM²⁹ screens were placed between the display and the eye. This stack acted as a *parallax barrier*, where light rays are modulated spatially and angularly as they pass through. The integral imaging concept of multiple rays per scene point was applied. However, instead of individual rays, *sums* of the perceived light rays were synthesized at precise locations on the pupil, so that the eye accommodates naturally to the depth of the displayed virtual object and its representation comes into focus on the retina, as discussed in greater detail below.

In this display design, the final color of each light ray emanating from the back light and entering the eye is the product of attenuation values at the pixels of each SLM screen that it intersects. Hence, Maimone et al. performed compressive optimization based on *content-adaptive* parallax barriers [72] to compute the proper attenuation values necessary for the rays from the multiple views to produce the correct light field. In such optimization, random noise is inherent in the output, which, unfortunately, overwhelms the angular variation between closely spaced elemental images, resulting in blurry output with no DOF. To resolve this problem, Maimone et al. discretized the target light field into a set of diffuse billboards, somewhat similar to a multifocal display, eliminating *local* angular variation within each billboard. This way, the noise produced at each of the billboards cancels out in the final image.

To further improve image fidelity, Maimone et al. [71] came up with a *retinal optimization* algorithm. It constrains groups of rays falling at the same spot on the pupil by the *perceived* sum of their intensities. Maimone et al. note that exact retinal optimization would require knowledge of the eye lens’ focal state in order to determine where exactly the rays will fall on the retina. Instead of determining the eye’s accommodation, they performed the optimization as if the eye is simultaneously focused on each object in the scene, at some expense to out-of-focus blur quality.

This design assumes there is no relative motion between the pupil and the display. In a natural setting where the gaze direction is unconstrained, in order to synthesize the ray sums correctly at each instance, eye tracking would have to

29. spatial light modulator

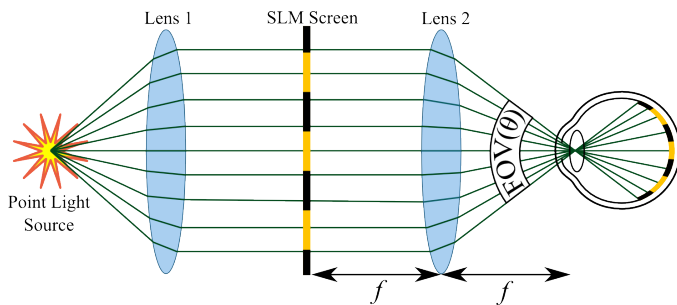


Figure 8. The Maxwellian view display principle. Diverging rays from the point light source are collimated by lens 1 and pass through the SLM screen. Transparency of every pixel is controlled, forming an image, which is then directed by converging lens 2 into a tiny spot on the pupil. The image conjugate to the screen is formed on the retina with an extremely large focal depth [22].

be integrated into the HMD³⁰. Another problem is that of pure computational efficiency: the optimization used took a few minutes for a single rendering of the scene. However, Maimone et al. note, faster methods can be adapted, such as the adaptive sampling framework developed by Heide et al. in [73], which uses only 3.82% of the rays in the full target light field.

Maimone et al. [71] tested their prototype display with a camera placed at the eye location and focused at different distances. Results showed that the design has promising occlusion qualities, while the focal cues in the generated images correctly correspond to the camera's focus.

Just as the microlens array described in Section 3.9, the parallax barrier method suffers from reduced spatial resolution. Likewise, if high-resolution screens with small pixel pitch were used to produce greater resolution, diffraction artifacts would become a problem, but in this case, caused by small pixel apertures in the SLM screens. Maimone et al. reflect that screens optimized to minimize diffraction [74] and devices with nonuniform pixel distributions [75] may alleviate this problem.

3.12 Maxwellian View Retinal Projectors

Most accommodation-free displays use the Maxwellian view principle. It is based on an experiment James Clerk Maxwell conducted in 1868 [76], where he increased the quantity of light reaching the retina of his eye. The principle of the Maxwellian view display is shown in Fig. 8.

Ando et al. [22] proposed the usage of accommodation-free displays in HMDs to address VAC. They constructed two bench OST Maxwellian view prototypes, one using DMD and another using LCD as the SLM screen and a converging HOE³¹. Later, von Waldkirch et al. [45] built a Maxwellian view retinal projector prototype where light from an LED³² source is first focused through a series of narrow apertures before reaching the retina, therefore greatly increasing the spatial coherence, in order to further

30. refer to Appendix C for a review on integration of eye-tracking into HMDs

31. or HOE, holographic optical element: an angle- and wavelength-selective optical element that can be used to reflect and converge or diverge a beam from a certain angle, while potentially also acting as a beamsplitter. For details, refer to [22] and Section V of [1]

32. light emitting diode

increase the DOF. They proposed a compact design with eye-glasses form-factor with mirrors to direct the image.

Later, von Waldkirch et al. [77] introduce an fluid lens oscillating at a high frequency into this design. Here, unlike in time-multiplexed MFP displays, the content is not changed depending on focal depth of the lens. However, the oscillation is so fast that the user perceives a temporal fusion of defocused and in-focus content, which extends the DOF yet further.

As noted by Ando et al. [22] and von Waldkirch [77], one challenge with Maxwellian view displays is that the convergence point of the rays needs to fall on the pupil, even with eye rotation and small pupil diameter, which poses geometric restrictions on the monocular FOV and causes vignetting effects.³³ Von Waldkirch et al. [19] assessed the constraints of using an RSD as an alternative to retinal projectors, but arrived at a similar trade-off between resolution, DOF, and FOV.

Yuuki et al. [78] realized a dense Maxwellian view: they placed a light absorption layer with pinhole patterns between a fly-eye lens sheet and an LCD panel, so that rays emanating through the holes are converged by the lenses in a dense grid of intersection points. When the pupil is in the same plane with the intersection points, the image is projected onto the retina with a large depth-of-field. Yuuki et al. simulated the behavior of this set-up at different viewing distances, and optimized the lens pitch for multiple usage scenarios, including the application of this design to HMDs.

3.13 Pinlight Arrays

Maimone and Lanman et al. [79] combined their efforts to come up with a similar, yet different design. They fabricated a *pinlight* head-mounted display, which uses a dense array of point light sources, projected through a barrier of liquid crystal modulators³⁴, onto the eye. The pinlights are simply cavities etched into a sheet of transparent plastic, which light up when much coarser diodes shine light into the plastic from the perimeter. Each pinlight illuminates a fixed section of the LCD with minimal overlap, forming a dense grid of miniature projectors. Resulting projections are out-of-focus and overlapped at the pupil plane, but form sharp image tiles at the back of the retina, as shown in Fig. 9.

This setup is, in a sense, the Maxwellian view in-reverse: rather than converging the rays at a point on the pupil to form a conjugate image on the retina, rays from pinlight projectors fan out ray bundles onto the whole pupil. After refraction by the pupil, the ray bundles form tiles on the retina without being conjugated, since their convergence point is much farther than the retina.

There are quite a few problems the pinlight design poses. For one, a single pinlight projector alone is unable to cover a wide-enough section of the retina. To provide a large FOV, multiple projections have to be tiled continuously and disjointly. However, due to the pupil being round, images produced by each pinlight projector are also round, and therefore cannot be tiled without gaps or overlap. Secondly, the model presumes the eye is fixed. If the eye rotates or moves, the projected sub-images will shift, corrupting the

33. See Appendix D for details.

34. essentially, an LCD panel

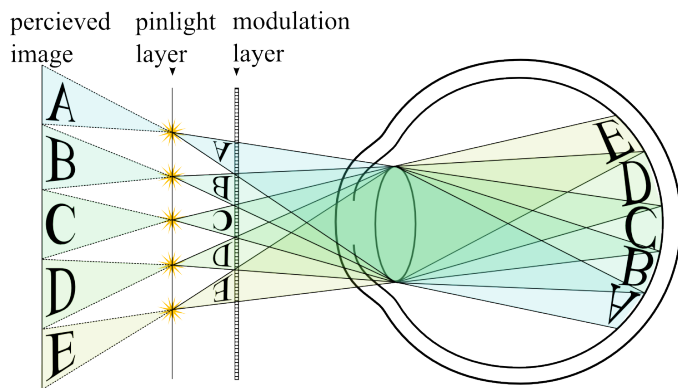


Figure 9. Conceptual diagram of pinlight displays. With proper placement of the layers, changes in optical power of the eye lens only modify the size of the projected tiles by about 3% and do not affect their sharpness [79].

overall image. Finally, it should be noted that, whereas a single pinlight projector is accommodation-free, the FOV of each projector changes with accommodation, again causing slight image misalignments.

Maimone et al. [79] propose several ways to address these challenges. The first is to incorporate eye-tracking into the display, as discussed in Section 4, and re-compute the image as the eye moves. The second forgoes integration of eye tracking by projecting multiple light rays corresponding to the same point in the virtual scene from different pinlights, dense enough to allow eye movement within a limited space. This pixel-sharing results in the overall spatial resolution.

Along the same line, Maimone et al. suggest the possibility of altering the design to allow angular variation around the eye, such that several rays corresponding to the same point in the scene reach the retina at the same time, generating a curved wavefront as the multiscopic HMDs discussed in Sections 3.11 and 3.9. This alteration would move the display from the accommodation-free to the static multiscopic category, albeit at even greater resolution expenses.

Increasing the resolution is not an easy task either, since it faces the same diffraction problems imposed by the small aperture of the SLMs as in the parallax barriers described in Section 3.11.

Aside from the accommodation-free quality of the pinlight projectors, the benefit of this design is that all of the components are transparent, which allows for the OST capability without the use of any extra cumbersome or FOV-limiting components. Hence, the described prototype boasts an eye-glasses form factor whilst maintaining a 110° FOV, never achieved before in any OST HMD of this size.

4 EYE TRACKING IN HMDs

Previously-described MFP methods display content at multiple depths (in a time- or space-multiplexed fashion), emulating the light field in a discrete fashion. As an alternative, it has been theorized that the adjustable optics in the varifocal methods can also be gaze-driven [35], [80], [81], adjusting focus specifically to the depth of the virtual point where the viewer looks at any given moment. Authors of several

works discussed in this review hypothesized about integrating an eye tracker into an HMD to accomplish this. Among them, Hua et al. [81] also designed compact and ergonomic eye-tracked HMDs (ET-HMDs) for this and other purposes.

As mentioned earlier, microlens, parallax barrier, and pinlight displays could also benefit from eye-tracking to circumvent the necessity of excessive micro-aperture density, which causes aberrations due to diffraction [27], [71], [79]. Pinlight displays may also benefit from eye-tracking to generate optimal projection tiling, as discussed in Section 3.13.

As with freeform prisms in HMDs, I found no survey literature covering integration of eye trackers in HMDs. For the purpose of providing a comprehensive guide for researchers and HMD designers, a brief survey on the topic is included in Appendix C.

Eye tracking has been applied in other ways to alleviate effects of VAC in HMDs. Several studies have used eye-trackers in conjunction with emulated (software-rendered) retinal blur, investigating the effects on accommodation. Alternative stereo vergence models driven by eye-tracking have also been explored.

4.1 Gaze-driven Retinal Blur

Hillarie et al. [82] were the first to implement gaze-dependent *rendered* depth-of-field (DOF) using eye tracking. They tested their approach in an immersive room with a large 90° curved screen. Their user study shows that this approach helps improve the sense of immersion.

Mantiuk et al. [83] extended the above work by testing whether gaze-guided DOF improves not the sense of immersion, but rather the sense of realism of the scene. They used a commercially-available glint-based eye tracker on a standard 22" LCD without stereo. Their algorithm determined focal blur by relative distance between the object gazed upon and other objects around it. Their experiment with 20 live subjects viewing animated and static virtual environments confirmed that the DOF effect guided by eye movements is preferential to a predefined DOF effect.

Vinnikov and Allison [84] followed suit and tested a similar system with stereoscopic bench display on a group of users viewing 3D scenes. Based on the results of a questionnaire, they concluded that simulated focal blur guided by eye tracking subjectively enhances the depth effect when combined with stereo. Finally, Duchowski et al. [85] conducted another gaze-contingent focal blur study with a stereo display and binocular eye-tracker. The depth blur amount in their system is deduced directly from vergence, i.e. from triangulating the intersection of the gaze vectors for each eye. Their user study showed that gaze-driven simulated DOF significantly reduces visual discomfort for people with high stereoacuity.

Although these studies suggest that gaze-driven software-rendered blur reduces visual discomfort, it alone cannot provide entirely correct focal cues: the light rays coming from a screen projected to a fixed distance still diverge at the same angle before reaching the eye lens, therefore, when accommodation matches vergence, the objects at the vergence distance still appear out-of-focus (although less so than others) [84], [86], [87]. Also, to my knowledge, these solutions have never been tested in HMDs.

4.2 Gaze-driven Dynamic Stereoscopy

An approach to addressing VAC radically different from ones aforementioned is that of adjusting vergence to the focal plane instead of the other way around, called *dynamic stereoscopy* (DS) or dynamic convergence. Pioneering work by State et al. [88] applied DS in an AR HMD prototype targeting medical applications. The prototype used static video cameras, but dynamically adjusted frame cropping to verge on the depth of the central object. Results from the user study indicate that DS does in fact mitigate VAC to some degree, but introduces other problems, discussed below.

Various DS models have been proposed that rely on saliency algorithms to determine the gaze point [9], [10], [15], [89]. Fisker et al. [90] were the first to use eye tracking integrated into an off-the-shelf VR HMD to do this. They discovered eye strain was more severe with their initial DS model turned on, which prompted them to improve their DS system by filtering and smoothing out the adjustments. Later, Bernhard et al. [91] experimented with eye tracking and an autostereoscopic display with a similar DS model and measured fusion time of the imagery as compared to static stereoscopy. They report improvement in fusion times with DS only for virtual objects placed in front of the focal plane, but no significant improvements at or beyond it.

The major problem with DS is what State et al. [88] referred to as the disparity-vergence conflict: adjusting the vergence to the focal plane means that, even though vergence no longer conflicts with accommodation, both cues now indicate the depth of the focal plane rather than the depth of the virtual object. In OST HMDs for AR this conflict induces a mismatch between vergence for real-world and virtual objects.

A preliminary experiment by Sherstyuk et al. [89] without eye-tracking suggests that DS may improve performance in VR tasks on nearby objects. However, further studies with improved DS models are required to determine whether the lack of a natural vergence cue will result in depth misjudgements and fusion delays in AR VST HMDs, where the disparity of the incoming video stream may also be adjusted, as well as in opaque HMDs for VR.

5 EVALUATION METHODS

There are four general evaluation strategies to evaluate VAC solutions: (1) subjective user studies, (2) direct measurement of oculomotor responses, (3) measurements of physiological fatigue indicators, and (4) assessment of brain activity via such tools as EEG³⁵ or fMRI³⁶. Each has its own merits and drawbacks; hence, a combination of several strategies is more robust than a single strategy alone.

5.1 Subjective User Studies

User studies are widely accepted and popular as a means to perceptually evaluate stereoscopic viewing experience [4]. These can be subdivided into two main types: performance-oriented, where a user's performance in a task using the evaluated system serves as a measure of effectiveness of the

display, and appreciation-oriented, where each user is asked of their subjective opinion of their viewing experience. Methodology for appreciation-based surveys of stereoscopic content has been developed in [92]. For general survey methodology, see [93].

Although questionnaires are technologically less involved than any of the other evaluation methods, they are prone to all common pitfalls of subjective measures, such as user bias, problems with quantification, and limited population samples, which exclude marginal cases.

5.2 Oculomotor Response Measurements

Infrared autorefractors provide an objective and precise measurement of the accommodation response. Despite accurate autorefractors now being widely available in handheld form factor [94], they are still both bulky and expensive, which sets a hurdle for their use with HMDs. An infrared autorefractor determines the optical power of the eye lens by measuring time-of-flight of infrared light it sends through the pupil, which is reflected from the inside surfaces and returns back to its sensors. The common implementation is a complex mechanism, which involves two optical paths (one for sending the IR beam and one for receiving it) separated by a beamsplitter [95].

Takaki [96], Shibata et al. [41], and McQuaide [58] used autorefractors to measure accommodation responses to their bench prototypes, while Liu et al. [35] are the only ones yet (to my knowledge) to test an HMD prototype with an autorefractor. Day et al. [86] used an autorefractor to experimentally evaluate effects of depth of field on accommodation and vergence, while MacKenzie et al. [36] used one to accurately measure eye accommodation responses to an MFP bench display similar to that of Akeley et al. [51] in order to establish the focal plane number and separation requirements for MFP displays.

To measure vergence, one can use a binocular eye tracker as Bernard et al. did in [91]. Suryakumar et al. [97] built a system with a custom photorefractor and a binocular eye tracker to measure vergence and accommodation to stereo imagery at the same time, which they later applied in their study on feedback between vergence and accommodation [8]. Various ways of integrating eye trackers with or into HMDs have already been discussed, but the integration of a custom photorefractor into an HMD is a complex task, and, to my knowledge, has not yet been attempted.

5.3 Fatigue Measurements

The drawback of directly measuring oculomotor response alone is that it does not assess the level of visual fatigue (asthenopia). While it may provide indications of how close the responses are to natural viewing, there are other neurological and psychological factors that may cause different individuals to experience different levels of discomfort while eliciting the same oculomotor responses. Studies suggest measuring blinking rate [98], heart rate and heart variability [99], [100], and blood pressure [99], may serve as an objective assessment of fatigue during stereo viewing. In addition, standard visual reflex timing measurements can be taken prior to and after the experiment [101].

35. Electroencephalography

36. functional magnetic resonance imaging

5.4 Brain Activity Measurements

There are yet few studies that measure brain activity to measure fatigue caused by stereo viewing, and virtually none that evaluate VAC-alleviating HMD designs. Hagura and Nakajima performed a preliminary study using fMRI in combination with MEG³⁷ in order to detect fatigue caused by viewing random-dot stereograms [102]. More recently, Frey et al. performed a pilot study that sheds some light on how visual fatigue due to VAC can be measured using EEG [103].

6 CONCLUDING REMARKS

Head-mounted displays still have a far way to go before they are comfortable enough to be worn by any individual over extended periods of time. VAC remains a major factor contributing to the discomfort, especially for near tasks with VR or AR. I have presented a systematic review of different potential solutions, and now proceed to identify gaps and promising areas in this body of research.

For eyeglasses-form-factor OST HMDs, two solutions that appear to be most promising are pinlight and parallax-barrier displays. For those, integration of eye tracking and low-diffraction screens have been identified as most important future research directions. Where thickness is not as big an issue, freeform prism designs with off-axis DMM- or microlens-array-based image generation subsystems present a viable alternative.

Time-multiplexing imposes additional taxing requirements on the refresh rate, which is already so critical for HMDs [70]. However, the required refresh rates are lower for MFP displays than multiscopic HMDs. In depth-blended MFP HMDs, only five depth layers, as opposed to over thirty views, need to be time-multiplexed at every frame, which can easily be achieved using contemporary LCoS and DMD screens. Blue-phase liquid crystal lenses could provide sufficient frequency for switching between focal states in an MFP display, and it has been shown that liquid crystal lens displays can be made in an eyeglasses form factor.

RSDs present a solution to the ghosting problem, and both the multifocal and multiscopic scanned fiber array methods have great potential to eliminate VAC in OST HMDs. It remains to be shown if scanned fiber array designs can be minified further and yield an eyeglasses-form-factor.

In this vein, it also remains to be explored whether birefringent-lens MFP displays can be easily minimized to work in an HMD. Although quite challenging, waveguide stacks with more than two focal planes are yet another under-explored area. Requirements for focal plane stacks have been evaluated based on the criteria of how closely the accommodation response resembles actual live viewing [36], but fatigue levels have not been measured for designs that do not fully adhere to these criteria.

With recent advances in ET-HMDs³⁸ (see Appendix C) and integration of commercial eye trackers into Oculus Rift VR HMDs available from SMI [104], HMDs with gaze-guided tunable optics should be implemented and tested.

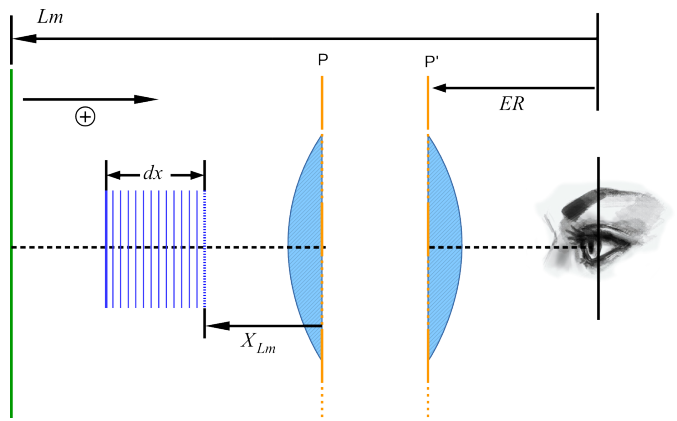


Figure 10. Basic layout of a magnifier, adapted from Rolland et al. [46]. L_m corresponds to d_{ei} in our nomenclature, X_{L_m} is d_{sl} , ER is d_{el} , and dx , the stack thickness, is t . Note that the x axis is inverted relative to direction of these vectors.

In those, gaze-driven software blur, described in Section 4.1, may be required to provide the necessary retinal cues; otherwise, the entire image will appear in focus at all times. Similarly, gaze-driven software blur may serve as an alternative approach to integral imaging for providing correct focal cues in pinlight displays, or it could be tested in Maxwellian view displays.

Integration of camera sensors directly into the screen, as in [105], may preserve compactness of ET-HMDs. As an alternative to eye tracking, a photorefractor, similar to the one described in [97], may be used inside an HMD to measure the eyes' accommodative state.

I anticipate that combinations of various optical designs presented in this review with eye tracking will yield much lighter, more ergonomic designs with greater spatial resolution in the near future, while also greatly alleviating, if not eliminating, side-effects of the VAC.

ACKNOWLEDGMENTS

I express sincere thanks to Dr. Simon J. Watt for sharing the [36] article, to Dr. Sujal Bista and Dr. Alexander Kramida, for helping with understanding the optics behind microlens-array displays. I also thank Dr. Kramida for helping me edit this article, and to Prof. Amitabh Varshney for providing useful feedback during the entire process.

APPENDIX A DERIVATION OF DISPLAY STACK PARAMETER EQUATIONS

This appendix shows step-by-step derivation of the equations for MFP displays from [16] and [46]. Refer to Fig. 10 for expalnation of variable designations.

The first is the imaging equation,

$$\frac{1}{x'} = \frac{1}{x} - \frac{1}{f}, \quad (3)$$

where x and x' are distances of a single screen from the principal plane P and of the corresponding virtual image from the principal plane P', respectively; x' falls within the range $[d_{ei}, \infty]$, while x varies within $[d_{sl}, f]$, with t

37. Magnetoencephalography

38. eye-tracked HMDs

representing the total span of the latter interval (see Fig. 3). We apply the close limits and solve for d_{sl} , or stack thickness:

$$\frac{1}{d_{ei}} = \frac{1}{d_{sl}} - \frac{1}{f}, \quad (4)$$

$$d_{sl} = \frac{fd_{il}}{f + d_{il}} \quad (5)$$

APPENDIX B A BRIEF SURVEY ON FREEFORM OPTICS IN HMDs

Our eyes do not come retrofitted within threaded circular nests. If that were the case, designing a light-weight, super-compact, wide-FOV HMD with conic optics would be trivial. Hence, although stacked display designs with conic optics can be said to have “evolved” into freeform optics displays, as Rolland and Thompson semi-humorously note in [106], the advent of automatic fabrication of freeform optics under computer numerical control (CNC) in the past two decades constitutes a revolution in HMD designs and other optics applications. Indeed, freeform optics provide HMD researchers with a much greater freedom and flexibility than they had with conventional rotationally-symmetric surfaces. We first provide some historical context about freeform optics in HMDs as the precursor of resulting VAC solutions.

The first VR HMDs using freeform prisms were presented by Canon [107], [108]. Yamazaki et al. [109] improved this design, building stereoscopic VR HMD prototype with 18-mm-thick freeform prisms and a 51° binocular horizontal FOV. This was followed by an outburst of related research at the Optical Diagnostics and Applications Laboratory (the O.D.A. Lab), University of Rochester. Cakmakci et al. [110] target eyeglasses form factor and OST operational principle as the most important qualities in future wearable displays. They put forth a set of optical and ergonomic requirements for such a display. They proposed to use freeform optics to realize the original idea in the patent by Bettinger [111]: placing the imaging unit off the optical axis, so that the physical environment is not obscured. They also described a way to minimize the number of optical elements.

The same group proposed a design that features a radially-symmetric lens in conjunction with a freeform mirror to guide the light to the exit pupil. They published the particulars of the fabrication process in [112]. Their prototype was extremely light and compact, but featured a monocular FOV of only $27^\circ \times 10^\circ$ FOV. Cakmakci et al. [113], [114] evaluate their design and discuss how Gaussian Radial Basis Functions (RBFs) yield an advantage over Zernike polynomials when used to optimize surfaces for freeform optics. Kaya et al. [115] describe a method for determining the RBFs basis size required to achieve desired accuracy for optics applications.

Cheng et al. [116] focused on the problem of providing a large FOV in OST HMDs using freeform prisms. Their prototype featured an unprecedented 53.5° diagonal FOV per eye, while maintaining the microdisplay compact, the contrast high, and the vignetting low. Later, Cheng et al. [56] proposed to use tiled freeform prisms for a binocular OST HMD prototype to achieve a much wider FOV, $56^\circ \times 45^\circ$ per eye, or, potentially, a $119^\circ \times 56^\circ$ total binocular resolution.

Two years later, Wang et al. [117] published the particulars of the fabrication process and evaluation of the resulting prototype. In parallel, Gao et al. [118] modified the design so that, theoretically, it could display opaque virtual imagery even in outdoor environments.

Meanwhile, Cheng et al. [55] made an HMD prototype using a stack of freeform prisms, thereby creating an MFT display supporting the accommodation cue, which is covered in greater detail in Section 3.3. In a more recent work, Hu and Hua [37] used freeform prisms to direct and magnify an image projected from a DLP microdisplay via a deformable-membrane mirror, controlling focus by changing the voltage applied on the mirror (see Section 3.4 for details).

APPENDIX C INTEGRATION OF EYE TRACKERS INTO HMDs

There have been several early instances of integrating eye tracking hardware into off-the-shelf VR headsets. Beach et al. [120] proposed to track gaze to provide a hands-free interface to the user. In parallel, Duchowski [121] integrated an existing eye tracker with a bench display, stipulating it may allow for *foveated rendering*, or outputting greater detail exactly at the users gaze point in a “just in time” fashion. Later, Duchowski et al. [122] integrated the ISCAN tracker into an off-the-shelf VR HMD to train and evaluate visual inspection of aircraft cargo bays [122]. Hayhoe et al. [123] used the same headset integrated with an off-the-shelf magnetic tracker for the head and another near-infrared (NIR) eye tracker in order to study saccadic eye movements on subjects performing simple tasks in virtual environments. Quite recently, SensoMotoric Instruments (SMI) integrated their eye tracker into the Oculus Rift VR headset, and now offers this upgrade to customers [104].

Vaissie and Rolland [124], [125] made the first efforts in designing a fully-integrated eye-tracked HMD (ET-HMD). They propose that ET-HMDs can be used to place the virtual cameras at virtual locations that correspond to pupil locations, rather than eyeball centers of the user [126].

Hua [127] developed a prototype of a fully-integrated OST ET-HMD using an infrared tracker and an ISCAN circuit board, and Curatu et al. [128] adapted the design to head-mounted projective displays (HMPD). In [129], Hua et al. developed corneal reflection eye-tracking methods and algorithms for ET-HMDs which are more tolerant to slippage than the alternatives. In [130], they devised an eye illumination model for such tracking. David et al. [105] designed an ET-HMD integrating the NIR sensors (used for eye tracking) with the LCoS microdisplay components on a single chip.

In [80] and [81], Hua et al. designed and built a new OST ET-HMD prototype with the glint tracking and NIR illumination, using a freeform prism to combine four optical paths – the virtual image, light from the real environment, eye illumination, and eye imaging – while keeping the overall design relatively compact. It featured a 50° low-distortion area for virtual image overlay³⁹. Hu and Hua [131] further refined the design, increasing the eye clearance, providing a

39. compare to $21.4^\circ \times 16.1^\circ$ in Hua’s earlier work [127]

Issue	Optical-See-Through (OST) HMDs	Video-See-Through (VST) HMDs
Latency	Suffer from latency between the real and the virtual imagery, but interaction with real objects is not hindered	Enforce synchronization of real and virtual imagery, but user suffers from overall latency during interaction with the real world
Resolution Limit	Only virtual imagery suffers from resolution restrictions	Both real and virtual imagery suffer from resolution restrictions
Distortion	Need to overcome distortion of real imagery by optimization of both display and compensating optics*	Need to correct image for optical aberrations in cameras
	In both, display optics aberrations may need compensation by by warped / chroma-corrected rendering	
FOV Limit	FOV of the virtual display is constrained by size limits on the compensating optics, but the real imagery does not need to be obscured*	FOV for real and virtual imagery may approach the human FOV, but the display is usually sealed off to increase contrast
Viewpoint Displacement	Viewpoint remains at eye position	Offset between camera and eye position introduces depth misjudgement and disorientation**
Occlusion	Most designs cannot fully occlude real imagery by virtual content, which introduces depth misjudgement and ghosting effects, while SRDs may provide sufficient occlusion by overwhelming the see-through view with a brighter augmented image	Do not suffer from ghosting or occlusion problems, since the real and virtual content are both rendered by the display, with full control over opacity**
Complexity and Size	Involve complex optical paths*: design and fabrication are more complex, and resulting display may be bulkier	Typically have fewer optical elements and are cheap and easy to manufacture

Table 1
(Appendix E)

Comparison of OST and VST HMDs, based on [1], [28], [119], with a few updates based on recent designs.

*Pinlight (see Section 3.13) and parallax-barrier (see Section 3.11) HMDs go around the FOV limits of the virtual display by presenting real imagery in a way that forgoes compensating optics. They also require few optical elements, which allows for eye-glasses-like assembly.

**The alternative of folding video-capture optical path using mirrors placed in front of the eye, as described in [28], eliminates the offset at the expense of increased optical path (reduced FOV) and ghosting problems.

better angular coverage for eye imaging, and allowing the insertion of a hot mirror to separate the eye-tracking path.

APPENDIX D LIMITS ON FOV FOR MAXWELLIAN VIEW DISPLAYS

The monocular FOV, denoted as θ , can be expressed in terms of the focal distance to the converging lens in front of the eye (f), distance from eye rotation center to the pupil center (R), and maximum allowed eye rotation angle (the angle at which the line of sight goes through the border of the displayed image), δ :

$$\theta = 2 \tan^{-1} \left[\frac{\tan \delta (R + f)}{f} \right] \quad (6)$$

Meanwhile, with D as the pupil diameter, the critical eye rotation angle at which rays from the Maxwellian view still intersect the boundary of the pupil can be expressed as:

$$\phi = \tan^{-1} \left(\frac{D}{2R} \right), \quad (7)$$

To maximize the field of view during eye rotation, the condition $\phi > \delta$ must be satisfied. Thus, substitution of equation 6 into 7 yields a limit on the FOV posed by focal length, pupil diameter, and eye radius. In well-lit conditions, an average pupil constricts to about $D = 4$ mm [132], while $R \approx 12$ mm, or half of the average eye diameter. Given these values and a focal length of 30 mm, the limit on the monocular FOV is about 26° , and, in very bright conditions, even smaller, due to smaller pupil diameter.

REFERENCES

- [1] O. Cakmakci and J. Rolland, "Head-worn displays: a review," *Display Technology, Journal of*, vol. 2, no. 3, pp. 199–216, 2006.
- [2] S. R. Bharadwaj and T. R. Candy, "Accommodative and vergence responses to conflicting blur and disparity stimuli during development." *Journal of vision*, vol. 9, no. 11, pp. 4.1–18, Jan. 2009.
- [3] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue." *Journal of vision*, vol. 8, no. 3, pp. 33.1–30, Jan. 2008.
- [4] M. Lambooi, W. IJsselstein, M. Fortuin, and I. Heynderickx, "Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review," *Journal of Imaging Science and Technology*, vol. 53, no. 3, p. 030201, 2009.
- [5] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister, "Depth cues in human visual perception and their realization in 3D displays," in *SPIE Defense, Security, and Sensing*, B. Javidi, J.-Y. Son, J. T. Thomas, and D. D. Desjardins, Eds. International Society for Optics and Photonics, Apr. 2010, pp. 76 900B–76 900B–12.
- [6] T. Bando, A. Iijima, and S. Yano, "Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review," *Displays*, vol. 33, no. 2, pp. 76–83, Apr. 2012.
- [7] G. K. Hung, K. J. Ciuffreda, and M. Rosenfield, "Proximal contribution to a linear static model of accommodation and vergence." *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, vol. 16, no. 1, pp. 31–41, Jan. 1996.
- [8] R. Suryakumar, J. P. Meyers, E. L. Irving, and W. R. Bobier, "Vergence accommodation and monocular closed loop blur accommodation have similar dynamic characteristics." *Vision research*, vol. 47, no. 3, pp. 327–37, Feb. 2007.
- [9] A. Sherstyuk and A. State, "Dynamic eye convergence for head-mounted displays," *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology - VRST '10*, p. 43, 2010.
- [10] T. Oskam, A. Hornung, H. Bowles, K. Mitchell, and M. Gross, "OSCAM - optimized stereoscopic camera control for interactive 3D," *ACM Transactions on Graphics*, vol. 30, no. 6, pp. 189:1–189:8, 2011.
- [11] A. Shamir and O. Sorkine, "Visual media retargeting," in *ACM SIGGRAPH ASIA 2009 Courses*. ACM, 2009, p. 11.
- [12] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear disparity mapping for stereoscopic 3d," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 4, p. 75, 2010.

- [13] C.-W. Liu, T.-H. Huang, M.-H. Chang, K.-Y. Lee, C.-K. Liang, and Y.-Y. Chuang, "3d cinematography principles and their applications to stereoscopic media processing," in *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 2011, pp. 253–262.
- [14] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H.-P. Seidel, and W. Matusik, "A luminance-contrast-aware disparity model and applications," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 6, p. 184, 2012.
- [15] U. Celikkan, G. Cimen, E. B. Kevinc, and T. Capin, "Attention-aware disparity control in interactive environments," in *User Modeling and User-Adapted Interaction*, vol. 29, no. 6-8, 2013, pp. 685–694.
- [16] J. P. Rolland, M. W. Krueger, and A. A. Goon, "Dynamic focusing in head-mounted displays," pp. 463–470, 1999.
- [17] C.-d. Liao and J.-c. Tsai, "The Evolution of MEMS Displays," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 4, pp. 1057–1065, Apr. 2009.
- [18] K. V. Chellappan, E. Erden, and H. Urey, "Laser-based displays: a review," *Applied optics*, vol. 49, no. 25, pp. F79–98, Sep. 2010.
- [19] M. von Waldkirch, P. Lukowicz, and G. Troster, "Defocusing simulations on a retinal scanning display for quasi accommodation-free viewing," *Optics Express*, vol. 11, no. 24, p. 3220, Dec. 2003.
- [20] E. Viirre, H. Pryor, S. Nagata, and T. A. Furness, "The virtual retinal display: a new technology for virtual reality and augmented vision in medicine," *Studies in health technology and informatics*, vol. 50, pp. 252–7, Jan. 1998.
- [21] J. S. Kollin and M. R. Tidwell, "Optical engineering challenges of the virtual retinal display," in *SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1995, pp. 48–60.
- [22] T. Ando, K. Yamasaki, M. Okamoto, and E. Shimizu, "Head-mounted display using a holographic optical element," in *Three-dimensional television, video, and display technologies*, B. Javidi and F. Okano, Eds. Springer Science & Business Media, Mar. 2002, pp. 67–100.
- [23] S. Liu, D. Cheng, and H. Hua, "An optical see-through head mounted display with addressable focal planes," in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE, Sep. 2008, pp. 33–42.
- [24] B. T. Schowengerdt, C. M. Lee, R. S. Johnston, C. D. Melville, and E. J. Seibel, "1-mm Diameter, Full-color Scanning Fiber Pico Projector," *SID Symposium Digest of Technical Papers*, vol. 40, no. 1, p. 522, 2009.
- [25] J. P. Rolland and H. Hua, "Head-Mounted Display Systems," *Encyclopedia of Optical Engineering*, pp. 1–14, 2005.
- [26] C. E. Rash, M. B. Russo, T. R. Letowski, and E. T. Schmeisser, *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues*, 2009.
- [27] D. Lanman and D. Luebke, "Near-eye light field displays," *ACM Transactions on Graphics*, vol. 32, no. 6, pp. 1–10, Nov. 2013.
- [28] J. P. Rolland, R. L. Holloway, and H. Fuchs, "A comparison of optical and video see-through head-mounted displays," in *Photonics for Industrial Applications*, H. Das, Ed. International Society for Optics and Photonics, Dec. 1995, pp. 293–307.
- [29] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE TRANSACTIONS on Information and Systems*, vol. E77-D, no. 12, pp. 1321–1329, Dec. 1994.
- [30] R. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.
- [31] G. Lippman, "épreuves réversibles photographiques intégrales," *Académie des sciences*, pp. 446–451, 1908.
- [32] B. T. Schowengerdt and E. J. Seibel, "True 3-D scanned voxel displays using single or multiple light sources," *Journal of the Society for Information Display*, vol. 14, no. 2, pp. 135–143, 2006.
- [33] B. T. Schowengerdt, M. Murari, and E. J. Seibel, "Volumetric Display using Scanned Fiber Array," *SID Symposium Digest of Technical Papers*, vol. 41, no. 1, pp. 653–656, 2010.
- [34] S. Liu and H. Hua, "Time-multiplexed dual-focal plane head-mounted display with a liquid lens," *Optics Letters*, vol. 34, no. 11, p. 1642, May 2009.
- [35] S. Liu, H. Hua, and D. Cheng, "A novel prototype for an optical see-through head-mounted display with addressable focus cues," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 3, pp. 381–93, Jan. 2010.
- [36] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt, "Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control." *Journal of vision*, vol. 10, no. 8, p. 22, Jan. 2010.
- [37] X. Hu and H. Hua, "High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics." *Optics express*, vol. 22, no. 11, pp. 13 896–903, Jun. 2014.
- [38] A. Wilson, "Telecentric lenses achieve precise measurements," *Vision Systems Design*, vol. 6, no. 7, Jul. 2001. [Online]. Available: <http://www.vision-systems.com/articles/print/volume-6/issue-7/features/product-focus/telecentric-lenses-achieve-precise-measurements.html>
- [39] S. Shiwa, K. Omura, and F. Kishino, "Proposal for a 3D display with accommodative compensation: 3DDAC," *Journal of the Society for Information Display*, vol. 4, no. 4, pp. 255–261, Dec. 1996.
- [40] N. Yanagisawa, K.-t. Kim, J.-Y. Son, T. Murata, and T. Orima, "Focus-distance-controlled 3D TV," in *Photonics China '96*, E. G. Lean, Z. Tian, and B. G. Wu, Eds. International Society for Optics and Photonics, Sep. 1996, pp. 120–125.
- [41] T. Shibata, T. Kawai, K. Ohta, M. Otsuki, N. Miyake, Y. Yoshihara, and T. Iwasaki, "Stereoscopic 3-D display with optical correction for the reduction of the discrepancy between accommodation and convergence," *Journal of the Society for Information Display*, vol. 13, no. 8, p. 665, 2005.
- [42] T. Sugihara and T. Miyasato, "System development of fatigueless HMD system 3DDAC (3D Display with Accommodative Compensation) : System implementation of Mk.4 in light-weight HMD," *IEICE technical report. Image engineering*, vol. 97, no. 467, pp. 33–36, Jan. 1998.
- [43] T. Takeda, Y. Fukui, K. Ikeda, and T. Iida, "Three-dimensional optometer III." *Applied optics*, vol. 32, no. 22, pp. 4155–68, Aug. 1993.
- [44] T. E. Lockhart and W. Shi, "Effects of age on dynamic accommodation." *Ergonomics*, vol. 53, no. 7, pp. 892–903, Jul. 2010.
- [45] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Spectacle-based design of wearable see-through display for accommodation-free viewing," in *Pervasive Computing*. Springer, 2004, pp. 106–123.
- [46] J. P. Rolland, M. W. Krueger, and A. Goon, "Multifocal Planes Head-Mounted Displays," *Applied Optics*, vol. 39, no. 19, pp. 3209–3215, Jul. 2000.
- [47] S. Suyama, M. Date, and H. Takada, "Three-Dimensional Display System with Dual-Frequency Liquid-Crystal Varifocal Lens," *Japanese Journal of Applied Physics*, vol. 39, no. Part 1, No. 2A, pp. 480–484, Feb. 2000.
- [48] S. Suyama, H. Takada, K. Uehira, S. Sakai, and S. Ohtsuka, "A New Method for Protruding Apparent 3-D Images in the DFD (Depth-Fused 3-D) Display," *SID Symposium Digest of Technical Papers*, vol. 32, no. 1, p. 1300, 2001.
- [49] S. Suyama, S. Ohtsuka, H. Takada, K. Uehira, and S. Sakai, "Apparent 3-D image perceived from luminance-modulated two 2-D images displayed at different depths." *Vision research*, vol. 44, no. 8, pp. 785–93, Apr. 2004.
- [50] H. Takada, S. Suyama, M. Date, and Y. Ohtani, "Protruding apparent 3D images in depth-fused 3D display," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 233–239, 2008.
- [51] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," in *ACM SIGGRAPH 2004 Papers on - SIGGRAPH '04*, vol. 23, no. 3. New York, New York, USA: ACM Press, Aug. 2004, p. 804.
- [52] S. Liu and H. Hua, "A systematic method for designing depth-fused multi-focal plane three-dimensional displays." *Optics express*, vol. 18, no. 11, pp. 11 562–73, May 2010.
- [53] S. Ravikummar, K. Akeley, and M. S. Banks, "Creating effective focus cues in multi-plane 3D displays." *Optics express*, vol. 19, no. 21, pp. 20 940–52, Oct. 2011.
- [54] K. J. MacKenzie and S. J. Watt, "Vergence and accommodation to multiple-image-plane stereoscopic displays: Real world responses with practical image-plane separations?" *Journal of Electronic Imaging*, vol. 21, no. 1, pp. 752 417–752 417–10, Feb. 2012.
- [55] D. Cheng, Q. Wang, Y. Wang, and G. Jin, "Lightweight spatially-multiplexed dual focal-plane head-mounted display using two freeform prisms," *Chin. Opt. Lett.*, vol. 11, no. 3, pp. 031 201–, 2013.
- [56] D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using free-form optics tiling," *Optics letters*, vol. 36, no. 11, pp. 2098–100, Jun. 2011.
- [57] B. T. Schowengerdt, R. S. Johnston, C. D. Melville, and E. J. Seibel, "3D Displays using Scanning Laser Projection," in *SID Symposium*

- Digest of Technical Papers*, vol. 43, no. 1. Wiley Online Library, 2012, pp. 640–643.
- [58] S. C. McQuaide, E. J. Seibel, J. P. Kelly, B. T. Schowengerdt, and T. A. Furness III, "A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror," *Displays*, vol. 24, no. 2, pp. 65–72, Aug. 2003.
- [59] B. T. Schowengerdt, E. J. Seibel, J. P. Kelly, N. L. Silverman, and T. A. Furness III, "Binocular retinal scanning laser display with integrated focus cues for ocular accommodation," in *Electronic Imaging 2003*. International Society for Optics and Photonics, May 2003, pp. 1–9.
- [60] A. Wilson, "Tunable Optics," *Vision Systems Design*, vol. 15, no. 7, jul 2010. [Online]. Available: http://www.vision-systems.com/articles/print/volume-15/issue-7/Features/Tunable_Optics.html
- [61] G. Li, D. L. Mathine, P. Valley, P. Ayrás, J. N. Haddock, M. S. Giridhar, G. Williby, J. Schwiegerling, G. R. Meredith, B. Kippen, S. Honkanen, and N. Peyghambarian, "Switchable electro-optic diffractive lens with high efficiency for ophthalmic applications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 16, pp. 6100–4, Apr. 2006.
- [62] G. D. Love, D. M. Hoffman, P. J. W. Hands, J. Gao, A. K. Kirby, and M. S. Banks, "High-speed switchable lens enables the development of a volumetric stereoscopic display." *Optics express*, vol. 17, no. 18, pp. 15716–25, Aug. 2009.
- [63] Y. Li and S.-T. Wu, "Polarization independent adaptive microlens with a blue-phase liquid crystal." *Optics express*, vol. 19, no. 9, pp. 8045–8050, Apr. 2011.
- [64] M. Koden, S. Miyoshi, M. Shigeta, K. Nonomura, M. Sugino, T. Numao, H. Katsuse, A. Tagawa, Y. Kawabata, P. Gass *et al.*, "Ferroelectric liquid crystal display," *SHARP TECHNICAL JOURNAL*, pp. 47–50, 1997.
- [65] B. T. Schowengerdt, H. G. Hoffman, C. M. Lee, C. D. Melville, and E. J. Seibel, "57.1 : Near-to-Eye Display using Scanning Fiber Display Engine," *SID Symposium Digest of Technical Papers*, vol. 41, no. 1, pp. 848–851, 2010.
- [66] J. Hong, S.-W. Min, and B. Lee, "Integral floating display systems for augmented reality," *Applied optics*, vol. 51, no. 18, pp. 4201–4209, 2012.
- [67] W. Song, Y. Wang, D. Cheng, and Y. Liu, "A high-resolution optical see-through head-mounted display with eyetracking capability." *Chin. Opt. Lett.*, vol. 12, no. 6, pp. 060 010–, 2014.
- [68] H. Hua and B. Javidi, "A 3D integral imaging optical see-through head-mounted display." *Optics express*, vol. 22, no. 11, pp. 13 484–91, Jun. 2014.
- [69] D.-W. Kim, Y.-M. Kwon, Q.-H. Park, and S.-K. Kim, "Analysis of a head-mounted display-type multifocus display system using a laser scanning method," *Optical Engineering*, vol. 50, no. 3, p. 034006, Mar. 2011.
- [70] I. Goradia, J. Doshi, and L. Kurup, "A review paper on oculus rift & project morpheus," 2014.
- [71] A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Oct. 2013, pp. 29–38.
- [72] D. Lanman, M. Hirsch, Y. Kim, and R. Raskar, "Content-adaptive parallax barriers," in *ACM SIGGRAPH Asia 2010 papers on - SIGGRAPH ASIA '10*, vol. 29, no. 6. New York, New York, USA: ACM Press, Dec. 2010, p. 1.
- [73] F. Heide, G. Wetzstein, R. Raskar, and W. Heidrich, "Adaptive image synthesis for compressive displays," *ACM Transactions on Graphics*, vol. 32, no. 4, p. 1, Jul. 2013.
- [74] H.-C. Chiang, T.-Y. Ho, and C.-R. Sheu, "Structure for reducing the diffraction effect in periodic electrode arrangements and liquid crystal device including the same," Patent 6977 705, Dec. 20, 2005.
- [75] C. Benoît-Pasanau, F. Goudail, P. Chavel, J.-P. Cano, and J. Ballet, "Minimization of diffraction peaks of spatial light modulators using voronoi diagrams," *Opt. Express*, vol. 18, no. 14, pp. 15 223–15 235, Jul 2010.
- [76] G. Westheimer, "The Maxwellian View," *Vision Research*, vol. 6, no. 11-12, pp. 669–682, Dec. 1966.
- [77] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Oscillating fluid lens in coherent retinal projection displays for extending depth of focus," *Optics Communications*, vol. 253, no. 4-6, pp. 407–418, Sep. 2005.
- [78] A. Yuuki, K. Itoga, and T. Satake, "A new Maxwellian view display for trouble-free accommodation," *Journal of the Society for Information Display*, vol. 20, no. 10, pp. 581–588, Oct. 2012.
- [79] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays: Wide Field of View Augmented Reality Eyeglasses using Defocused Point Light Sources," in *ACM SIGGRAPH 2014 Emerging Technologies on - SIGGRAPH '14*. New York, New York, USA: ACM Press, Jul. 2014, pp. 1–1.
- [80] H. Hua and C. Gao, "A compact eyetracked optical see-through head-mounted display," in *IS&T/SPIE Electronic Imaging*, A. J. Woods, N. S. Holliman, and G. E. Favalora, Eds. International Society for Optics and Photonics, Feb. 2012, p. 82881F.
- [81] H. Hua, X. Hu, and C. Gao, "A high-resolution optical see-through head-mounted display with eyetracking capability." *Optics express*, vol. 21, no. 25, pp. 30993–8, Dec. 2013.
- [82] S. Hillaire, A. Lecuyer, R. Cozot, and G. Casiez, "Using an Eye-Tracking System to Improve Camera Motions and Depth-of-Field Blur Effects in Virtual Environments," in *2008 IEEE Virtual Reality Conference*. IEEE, 2008, pp. 47–50.
- [83] R. Mantiuk, B. Bazyluk, and A. Tomaszewska, "Gaze-Dependent Depth-of-Field Effect Rendering in Virtual Environments," in *Serious Games Development and Applications*, ser. Lecture Notes in Computer Science, M. Ma, M. Fradinho Oliveira, and J. a. Madeiras Pereira, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, vol. 6944, pp. 1–12.
- [84] M. Vinnikov and R. S. Allison, "Gaze-contingent depth of field in realistic scenes," in *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14*. New York, New York, USA: ACM Press, Mar. 2014, pp. 119–126.
- [85] A. T. Duchowski, D. H. House, J. Gestring, R. I. Wang, K. Krejtz, I. Krejtz, R. Mantiuk, and B. Bazyluk, "Reducing visual discomfort of 3D stereoscopic displays with gaze-contingent depth-of-field," in *Proceedings of the ACM Symposium on Applied Perception - SAP '14*. New York, New York, USA: ACM Press, Aug. 2014, pp. 39–46.
- [86] M. Day, D. Seidel, L. S. Gray, and N. C. Strang, "The effect of modulating ocular depth of focus upon accommodation microfluctuations in myopic and emmetropic subjects." *Vision research*, vol. 49, no. 2, pp. 211–8, Jan. 2009.
- [87] L. O'Hare, T. Zhang, H. T. Nefs, and P. B. Hibbard, "Visual discomfort and depth-of-field." *i-Perception*, vol. 4, no. 3, pp. 156–69, Jan. 2013.
- [88] A. State, J. Ackerman, G. Hirota, J. Lee, and H. Fuchs, "Dynamic virtual convergence for video see-through head-mounted displays: maintaining maximum stereo overlap throughout a close-range work space," in *Proceedings IEEE and ACM International Symposium on Augmented Reality*, 2001, pp. 137–146.
- [89] A. Sherstyuk, A. Dey, C. Sandor, and A. State, "Dynamic eye convergence for head-mounted displays improves user performance in virtual environments," in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games - I3D '12*. New York, New York, USA: ACM Press, Mar. 2012, p. 23.
- [90] M. Fisker, K. Gram, K. K. Thomsen, D. Vasilarou, and M. Kraus, "Automatic Convergence Adjustment in Stereoscopic Using Eye Tracking," in *EG 2013 - Posters*, 2013, pp. 23–24.
- [91] M. Bernhard, C. Dell'mour, M. Hecher, E. Stavrakis, and M. Wimmer, "The effects of fast disparity adjustment in gaze-controlled stereoscopic applications," *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14*, pp. 111–118, 2014.
- [92] J.-S. Lee, L. Goldmann, and T. Ebrahimi, "Paired comparison-based subjective quality assessment of stereoscopic images," *Multimedia tools and applications*, vol. 67, no. 1, pp. 31–48, 2013.
- [93] R. M. Groves, F. J. Fowler Jr, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, *Survey methodology*. John Wiley & Sons, 2013.
- [94] W. Wesemann and B. Dick, "Accuracy and accommodation capability of a handheld autorefractor," *Journal of Cataract & Refractive Surgery*, vol. 26, no. 1, pp. 62–70, 2000.
- [95] T. Dave, "Automated refraction: design and applications," *Optom Today*, vol. 48, pp. 28–32, 2004.
- [96] Y. Takaki, "High-Density Directional Display for Generating Natural Three-Dimensional Images," *Proceedings of the IEEE*, vol. 94, no. 3, pp. 654–663, Mar. 2006.
- [97] R. Suryakumar, J. P. Meyers, E. L. Irving, and W. R. Bobier, "Application of video-based technology for the simultaneous measurement of accommodation and vergence." *Vision research*, vol. 47, no. 2, pp. 260–8, Jan. 2007.

- [98] H. Heo, W. O. Lee, K. Y. Shin, and K. R. Park, "Quantitative measurement of eyestrain on 3D stereoscopic display considering the eye foveation model and edge information." *Sensors (Basel, Switzerland)*, vol. 14, no. 5, pp. 8577–604, Jan. 2014.
- [99] H. Oyamada, A. Iijima, A. Tanaka, K. Ukai, H. Toda, N. Sugita, M. Yoshizawa, and T. Bando, "A pilot study on pupillary and cardiovascular changes induced by stereoscopic video movies." *Journal of neuroengineering and rehabilitation*, vol. 4, no. 1, p. 37, Jan. 2007.
- [100] C. J. Kim, S. Park, M. J. Won, M. Whang, and E. C. Lee, "Autonomic nervous system responses can reveal visual fatigue induced by 3D displays." *Sensors (Basel, Switzerland)*, vol. 13, no. 10, pp. 13 054–62, Jan. 2013.
- [101] S. Mun, M.-C. Park, S. Park, and M. Whang, "SSVEP and ERP measurement of cognitive fatigue caused by stereoscopic 3D." *Neuroscience letters*, vol. 525, no. 2, pp. 89–94, Sep. 2012.
- [102] H. Hagura and M. Nakajima, "Study of asthenopia caused by the viewing of stereoscopic images: measurement by MEG and other devices," in *Electronic Imaging 2006*, B. E. Rogowitz, T. N. Pappas, and S. J. Daly, Eds. International Society for Optics and Photonics, Feb. 2006, pp. 60 570K–60 570K–11.
- [103] J. Frey, L. Pommereau, F. Lotte, and M. Hachet, "Assessing the zone of comfort in stereoscopic displays using eeg," in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, Apr. 2014, pp. 2041–2046.
- [104] (2014) Eye tracking hmd update package for oculus rift dk2. SensoMotoric Instruments. [Online]. Available: <http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/eye-tracking-hmd-upgrade.html>
- [105] Y. David, B. Apter, N. Thirer, I. Baal-Zedaka, and U. Efron, "Design of integrated eye tracker-display device for head mounted systems," in *SPIE Photonic Devices + Applications*, E. L. Dereniak, J. P. Hartke, P. D. LeVan, P. E. Longshore, and A. K. Sood, Eds. International Society for Optics and Photonics, Aug. 2009, pp. 741 910–741 910–10.
- [106] J. Rolland and K. Thompson, "Freeform optics: Evolution? No, revolution! — SPIE Newsroom: SPIE," *SPIE Newsroom*, 2012.
- [107] S. Yamazaki, A. Okuyama, T. Ishino, A. Fujiwara, and Y. Tamekuni, "Development of super compact hmd with sight line input," in *Proceedings of 3D Image Conference*, vol. 95, 1995, pp. 70–75.
- [108] H. Hoshi, N. Taniguchi, H. Morishima, T. Akiyama, S. Yamazaki, and A. Okuyama, "Off-axial hmd optical system consisting of aspherical surfaces without rotational symmetry," in *Electronic Imaging: Science & Technology*. International Society for Optics and Photonics, 1996, pp. 234–242.
- [109] S. Yamazaki, K. Inoguchi, Y. Saito, H. Morishima, and N. Taniguchi, "Thin wide-field-of-view hmd with free-form-surface prism and applications," in *Electronic Imaging'99*. International Society for Optics and Photonics, 1999, pp. 453–462.
- [110] O. Cakmakci, A. Oranchak, and J. Rolland, "Dual-element off-axis eyeglass-based display," in *Contract Proceedings 2006*, G. G. Gregory, J. M. Howard, and R. J. Koshel, Eds. International Society for Optics and Photonics, Jun. 2006, pp. 63 420W–63 420W–7.
- [111] D. S. Bettinger, "Spectacle-mounted ocular display apparatus," U.S. Patent 4 806 011, Feb. 21, 1989.
- [112] O. Cakmakci and J. Rolland, "Design and fabrication of a dual-element off-axis near-eye optical magnifier," *Optics Letters*, vol. 32, no. 11, p. 1363, 2007.
- [113] O. Cakmakci, B. Moore, H. Foroosh, and J. P. Rolland, "Optimal local shape description for rotationally non-symmetric optical surface design and analysis," *Optics Express*, vol. 16, no. 3, p. 1583, 2008.
- [114] O. Cakmakci, K. Thompson, P. Vallee, J. Cote, and J. P. Rolland, "Design of a free-form single-element head-worn display," in *OPTO*, L.-C. Chien, Ed. International Society for Optics and Photonics, Feb. 2010, pp. 761 803–761 803–6.
- [115] I. Kaya, O. Cakmakci, K. Thompson, and J. P. Rolland, "The assessment of a stable radial basis function method to describe optical free-form surfaces," in *Optical Fabrication and Testing*. Optical Society of America, 2010, p. OTuD2.
- [116] D. Cheng, Y. Wang, H. Hua, and M. M. Talha, "Design of an optical see-through head-mounted display with a low f-number and large field of view using a freeform prism." *Applied optics*, vol. 48, no. 14, pp. 2655–2668, 2009.
- [117] Q. Wang, D. Cheng, Y. Wang, H. Hua, and G. Jin, "Design, tolerance, and fabrication of an optical see-through head-mounted display with free-form surface elements." *Applied optics*, vol. 52, no. 7, pp. C88–99, Mar. 2013.
- [118] C. Gao, Y. Lin, and H. Hua, "Occlusion capable optical see-through head-mounted display using freeform optics," in *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Nov. 2012, pp. 281–282.
- [119] J. P. Rolland and H. Fuchs, "Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 3, pp. 287–309, Jun. 2000.
- [120] G. Beach, C. Cohen, J. Braun, and G. Moody, "Eye tracker system for use with head mounted displays," in *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218)*, vol. 5. IEEE, 1998, pp. 4348–4352.
- [121] A. T. Duchowski, "Incorporating the viewer's point of regard (POR) in gaze-contingent virtual environments," in *Photonics West '98 Electronic Imaging*, M. T. Bolas, S. S. Fisher, and J. O. Merritt, Eds. International Society for Optics and Photonics, Apr. 1998, pp. 332–343.
- [122] A. T. Duchowski, E. Medlin, A. Gramopadhye, B. Melloy, and S. Nair, "Binocular eye tracking in VR for visual inspection training," in *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '01*. New York, New York, USA: ACM Press, Nov. 2001, p. 1.
- [123] M. M. Hayhoe, D. H. Ballard, J. Triesch, H. Shinoda, P. Aivar, and B. Sullivan, "Vision in natural and virtual environments," in *Proceedings of the symposium on Eye tracking research & applications - ETRA '02*. New York, New York, USA: ACM Press, Mar. 2002, p. 7.
- [124] L. Vaissie and J. P. Rolland, "Eyetracking in head-mounted displays: analysis and design," Technical Report TR98-007, University of Central Florida, Tech. Rep., 1998.
- [125] Laurent Vaissie and Jannick P. Rolland, "Head mounted display with eyetracking capability," U.S. Patent 6 433 760, Aug. 13, 2002.
- [126] L. Vaissie, J. P. Rolland, and G. M. Bochenek, "Analysis of eyepoint locations and accuracy of rendered depth in binocular head-mounted displays," in *Electronic Imaging '99*, J. O. Merritt, M. T. Bolas, and S. S. Fisher, Eds. International Society for Optics and Photonics, May 1999, pp. 57–64.
- [127] H. Hua, "Integration of eye tracking capability into optical see-through head-mounted displays," in *Photonics West 2001 - Electronic Imaging*, A. J. Woods, M. T. Bolas, J. O. Merritt, and S. A. Benton, Eds. International Society for Optics and Photonics, Jun. 2001, pp. 496–503.
- [128] C. Curatu, H. Hua, and J. Rolland, "Projection-based head-mounted display with eye tracking capabilities," in *Optics & Photonics 2005*, J. M. Sasian, R. J. Koshel, and R. C. Juergens, Eds. International Society for Optics and Photonics, Aug. 2005, pp. 58 750J–58 750J–9.
- [129] H. Hua, P. Krishnaswamy, and J. P. Rolland, "Video-based eye-tracking methods and algorithms in head-mounted displays," *Optics Express*, vol. 14, no. 10, p. 4328, 2006.
- [130] H. Hua, C. W. Pansing, and J. P. Rolland, "Modeling of an eye-imaging system for optimizing illumination schemes in an eye-tracked head-mounted display," *Applied Optics*, vol. 46, no. 31, p. 7757, 2007.
- [131] X. Hu and H. Hua, "Optical Design of An Eyetracked Head-mounted Display using Freeform Waveguide," in *Classical Optics 2014*. Washington, D.C.: OSA, 2014, p. ITh3A.4.
- [132] Y. Yang, K. Thompson, and S. A. Burns, "Pupil location under mesopic, photopic, and pharmacologically dilated conditions," *Investigative ophthalmology & visual science*, vol. 43, no. 7, pp. 2508–2512, 2002.



Gregory Kramida received a BA in Graphic Design, a BS, and an MS in Computer Science from University of Maryland, College Park, in 2010, 2011, and 2014 respectively. He is currently a PhD student at the same university. His research interests include computer vision, with focus on 3D reconstruction of dynamic scenes, as well as related applications in augmented reality.