

An Optical See-Through Head Mounted Display with Addressable Focal Planes

Sheng Liu*, Dewen Cheng†, Hong Hua‡

3DVIS Lab, College of Optical Sciences, University of Arizona

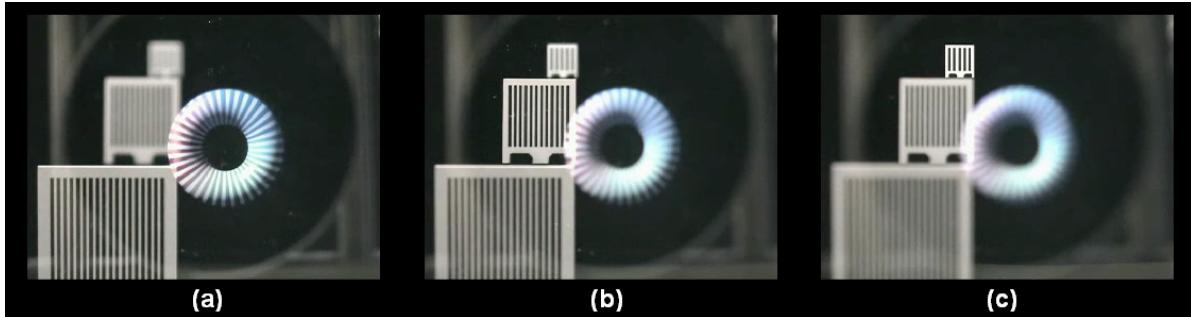


Figure 1. Move clip (VariFocalPlane.wmv) of a virtual torus rendered by the see-through HMD moving along the z-axis in the augmented space. The three subset clips were captured by a camcorder which was manually focused at (a) 16cm, (b) 33cm, and (c) 100cm, respectively.

ABSTRACT

Most existing stereoscopic head mounted displays (HMDs), presenting a pair of stereoscopic images at a fixed focal distance, lack the ability to correctly render the naturally coupled accommodation and convergence cues. Psychophysical studies have shown that such displays may cause many adverse consequences such as visual fatigue, diplopia, degraded oculomotor response, and depth perception errors. In this paper, we present a see-through HMD with addressable focal planes utilizing a novel active optical element—a liquid lens. The element, with a varying optical power from -5 to 20 diopters, is able to address the focal distance of the HMD from infinity to the near point of the eye. A monocular prototype was built from off-the-shelf elements and experimental results are presented to validate the proposed designs. We also describe both subjective and objective measurements of the accommodation responses of the viewer to the focal distances presented by the prototype.

KEYWORDS: Display hardware, head mounted display, accommodation, usability studies and experiments

INDEX TERMS: I.3.1 [Computer Graphics]: Hardware Architecture—Three-Dimensional Displays; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities; C.4 [Performance of Systems]: Performance Attributes.

Address: 1630 E. University Blvd., Tucson, AZ 85721, USA

* Email: sliu@optics.arizona.edu

† Email: dcheng@optics.arizona.edu

‡ Email: hhua@optics.arizona.edu

IEEE International Symposium on Mixed and Augmented Reality 2008
15 -18 September, Cambridge, UK
978-1-4244-2859-5/08/\$25.00 ©2008 IEEE

1 INTRODUCTION

Research interests in 3D displays have endured for decades, spanning the fields of flight simulation, scientific visualization, education and training, tele-manipulation and tele-presence, and entertainment systems. Many approaches to 3D displays have been proposed in the past, including head-mounted displays (HMDs) [11, 19, 21], projection-based immersive displays [3], volumetric displays [20, 7, 4], and holographic displays [12]. Among the various 3D display technologies, HMDs provide a good balance on affordability and unique capabilities. For instance, HMDs offer solutions to mobile displays for wearable computing, while in the domain of augmented reality, they are one of the enabling technologies for merging virtual views with physical scenes [2, 10].

Despite much significant advancement on stereoscopic HMDs over the past decades, there exist many technical and usability issues preventing the technology from being widely accepted for many demanding applications and daily usage. Many psychophysical and usability studies have suggested the association of various visual artifacts with the long-term usage of stereoscopic HMDs, such as apparent distortion in perceived depth [6, 27], visual fatigue [17, 22], diplopia [26] and degradation in oculomotor responses [14]. Although many factors may contribute to those artifacts from the engineering perspective, such as poor image quality, limited eye relief and inappropriate inter-pupillary distance (IPD) setup, one of the underlying causes attributes to the discrepancy between accommodation and convergence [17, 26]. Accommodation refers to the focus action of the eye where the shape of the crystalline lens is adjusted to see objects at different depths, while convergence refers to the convergent rotation of the eyes where the visual axis are brought to intersect at a 3D object in space. In a real-world viewing condition, these two oculomotor cues are tightly coupled with each other so that the convergence depth coincides with the accommodation depth. Most existing stereoscopic displays, however, create depth perception based on binocular disparities specifying objects at a range of convergence depths, forcing the

accommodation cue to be tied with a fixed focal distance. Contrary to the natural vision, all objects, regardless of their location in depth, are seen in focus if the viewer focuses on the image plane, or all objects are seen blurred if the user's accommodation varies with eye convergence. Many studies have investigated the adverse consequences of the forced decoupling of the two oculomotor depth cues. For instance, Watt *et al.* suggested that inappropriate focus cues in stereoscopic displays will adversely affect the depth perception, through both direct and indirect means by image blur and disparity scaling [27]. Such problems may become severe in two common types of viewing conditions: one in which virtual objects need to be presented across a wide range of distances to the user, from very close to far away (e.g. driving simulators); and one in which the display is used to augment a relatively close real-world scene with virtual objects and information (e.g. surgical training). The latter condition is very common in optical see-through stereoscopic displays for augmented reality applications.

1.1 Related work

Recently, varifocal and multi-focal plane stereoscopic displays have been proposed [23, 19, 18, 1, 24, 21, 15], aiming to address the accommodation-convergence discrepancy problem. Unlike traditional stereoscopic displays, these approaches enable the addressability of the accommodation cue in a stereoscopic display either by means of dynamic control of the focal distance through an active optical method, or by means of presenting multiple focal planes at equal dioptric spacing with tolerable errors for depth perception. Different from volumetric displays [20, 7, 4] in which all voxels within a volumetric space have to be rendered at a flickering-free rate regardless of the viewer's point of interest, these varifocal or multi-focal approaches are relatively more computational efficient, while being able to reduce the discrepancy artifacts to various extents. In general, these approaches can be categorized into time-multiplexed or spatial-multiplexed method.

In a time-multiplexed approach, the focal distance of a single-plane 2D display device is usually controlled either through a mechanical mechanism or an active optical element for focusing adjustment. For instance, Shiwa *et al.* achieved accommodation compensation by moving a relay lens in the HMD optical system [23] and Shibata *et al.* achieved similar function by axially translating the microdisplay mounted on a micro-controlled stage [24]. The dynamic control of the focal distance may be achieved by a hand-held device or by the user's eye gaze. These approaches should be more precisely categorized as a varifocal plane method rather than a multi-focal plane display in the sense that they provide dynamic addressability of the focal distance rather than render virtual objects on multi-focal planes at a flickering-free rate. In order to achieve the multi-focal plane capability, these approaches require a fast active optical element to replace the mechanical mechanism for focusing adjustment as well as a 2D display technology with multi-magnitude of faster refresh rate than existing displays. More recently, researchers proposed a see-through retinal scanning display (RSD) implemented with a deformable membrane mirror (DMM) device [18, 21]. In this design, a nearly-collimated laser beam is modulated and scanned across the field of view (FOV) to generate pixels on the retina. In the meantime, correct focusing cues can be rendered on a pixel-by-pixel basis by defocusing the laser beam through the DMM. However, in order to achieve a full color and flickering-free multi-focal plane stereo display, the practical development of such technology requires fast addressing speed on both the laser beam and the active optical element for up to MHz. In addition, rendering each pixel by a beam-scanning mechanism limits the

compatibility of such system with existing 2D display and rendering techniques.

In a spatial-multiplexed approach, multi-focal plane capability is achieved by the usage of multiple layers of 2D displays. For instance, Rolland *et al.* proposed to use a thick microdisplay stack, consisting of 14 layers of equally-spaced (in dioptric spacing) planar displays, to form focal planes in a HMD that divide the whole volumetric space from infinity to 2 diopters [19]. An optimal separation between two adjacent focal planes turns out to be 1/7 diopters by taking accounts of visual acuity, stereoacluity and pupil size of the human visual system (HVS). Although this approach allows for rendering multiple focal planes in parallel and reduces the speed requirements for the display technology, the practical implementation of such method is challenged by the lack of stack display technologies with high transmittance and by the demand for computational power to simultaneously render a stack of 2D images of a 3D scene based on the geometric depth. Recently, Akeley *et al.* presented a proof-of-concept display prototype by using a sparse stack of three LCD panels with an equal spacing of 0.67 diopter [1]. The implementation, however, was impractical to be miniaturized for a portable system.

1.2 Contribution

In this paper, we propose to develop a compact optical see-through HMD with addressable focal planes for improved depth perceptions. Inspired by the accommodative capability delivered by the crystalline lens in the HVS, a liquid lens device is implemented into a traditional stereoscopic HMD, enabling addressable accommodation cue from optical infinity to as close as the near point of the eye. Unlike the mechanical focusing methods [23, 24] and the RSD design based on a reflective DMM device [18, 21], the transmissive nature of the liquid lens allows for a compact and practical HMD layout, not only without moving components but also without compromising the accommodation range. Based on a proof-of-concept bench prototype, we further explore two types of addressability of the focal distances that the proposed system can deliver: one in a variable single-focal plane mode (see Figure 1); the other in a multi-focal plane mode by addressing the liquid lens, together with the microdisplay device, time-sequentially. In comparison to the time-multiplexed RSD approach based on a pixel-by-pixel rendering method [21], using a microdisplay device to present multiple full-color 2D images time-sequentially can remarkably decrease the requirement for the addressing speeds of each pixel and of the active optical components, from MHz to 100Hz level. Although the liquid lens used in our current bench prototype has a limited response speed of about 75ms, the new-generation liquid lens has a promising response time less than 10ms, which will make the time-multiplexed multi-focal plane approach more practical. Finally, we also describe two experiments, one subjective and one objective, to measure the accommodation response of the viewer using our bench prototype in the variable focal plane mode. Preliminary results suggest a potential improvement on depth perception in a HMD with addressable focal planes.

The rest of the paper is organized as follows. In section 2, we briefly describe the schematic design and implementation of the HMD prototype with addressable accommodation cue. In section 3, we investigate two types of addressability on the focal plane of the see-through HMD — one in a variable focal plane mode, the other in a multi-focal plane mode. In section 4, we carry out subjective and objective evaluations to examine the perceived depth and the accommodation response of the HVS versus the addressable accommodation cue. Finally, in section 5, future developments of the proposed system are discussed.

2 SYSTEM SETUP

To enable the addressability of focal planes in an optical see-through HMD, we propose to use an active optical element — liquid lens. Based on the electrowetting phenomenon [8, 16], the liquid lens demonstrates a varying optical power from -5 to 20 diopters by applying an AC voltage from 32 Vrms to 60 Vrms [13]. This anamorphic lens is capable of dynamically controlling the focal distance of a HMD, from infinity to as close as the near point of the eye.

Based on this simple concept, Figure 2 illustrates a schematic design of an optical see-through HMD with addressable focal planes. The system consists of four major components: a microdisplay, a focusing lens, a beamsplitter (BS), and a spherical mirror. The lens, drawn as a simplified singlet in Figure 2, is composed of an accommodation lens (e.g. the liquid lens) with varying optical power, ϕ_A , and an objective lens with a constant optical power, ϕ_O . The two lenses together form an intermediate image of the microdisplay on the left side of the spherical mirror. The spherical mirror then relays the intermediate image and redirects the light toward the viewer's eye through the BS. Since the liquid lens is the limiting aperture of the HMD optics, it is placed at the center of curvature (O_{SM}) of the spherical mirror so that a conjugate exit pupil is formed through the BS. Placing the eye at the conjugate pupil position, the viewer sees both the virtual image of the microdisplay and the real world through the BS.

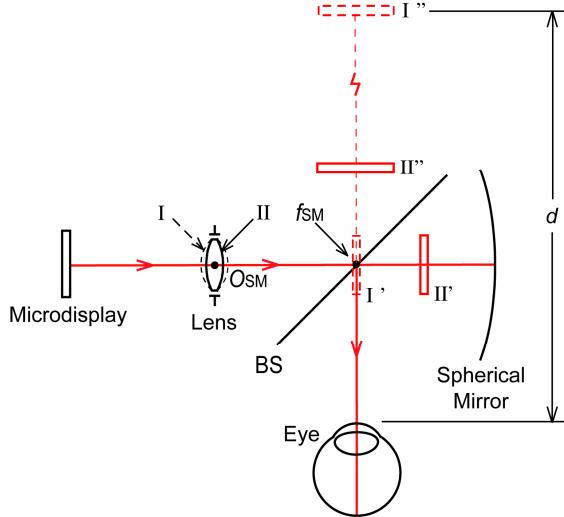


Figure 2. Schematics of an optical see-through HMD with addressable focal planes.

Indicated by the dashed and solid lines, respectively, in Figure 2, as the accommodation lens changes its optical power from high (I) to low (II), the intermediate image will be displaced towards (T') or away from (II') the focal plane (f_{SM}) of the spherical mirror. Correspondingly, the virtual image will be formed either far (I') or close (II') to the eye. Based on the first-order optics, the accommodation cue, d , of the HMD, which implies the distance from the eye to the virtual image plane, is determined by:

$$d = \frac{-uR}{2u + R + uR\phi} \quad (1)$$

where $\phi = \phi_O + \phi_A - \phi_O\phi_A t$ denotes the combined optical power of the focusing lens, t is the separation between the objective and accommodation lenses, u is the object distance from the microdisplay to the focusing lens and R is the radius of curvature

of the spherical mirror. All distances above are defined by the sign convention in optical designs [9]. A detailed derivation of Eq. (1) can be found in the Appendix.

Based on the schematic design, we implemented a proof-of-concept monocular bench prototype by using off-the-shelf components, as shown in Figure 3. The accommodation lens is a liquid lens (Arctic 320, Varioptic Inc.), which has a varying optical power from -5 to 20 diopters by applying an AC voltage from 32 Vrms to 60 Vrms. The liquid lens, with a clear aperture of 3mm, is attached to an off-the-shelf singlet (objective lens) with an 18mm focal length. The image source is a 0.59" full-color organic light emitting diode (OLED) microdisplay with 800x600 pixels and a refresh rate up to 85Hz (eMagin Inc.). The spherical mirror has a radius of curvature of 70mm and a clear aperture of 35mm. Based on those parametric combinations, the monocular bench prototype yields an exit pupil diameter of 3mm, an eye relief of 20mm, a diagonal field of view (FOV) of about 28°, and an angular resolution of 1.7 arcmins if assuming the absence of optical aberrations. The 28° FOV is derived by accounting for the chief ray angle in the image space. Although the FOV, exit pupil diameter and image quality are limited in the current prototype, the major purpose of this work is for a proof-of-concept demonstration and preliminary user studies. Those critical display performances are to be improved in future customized designs.

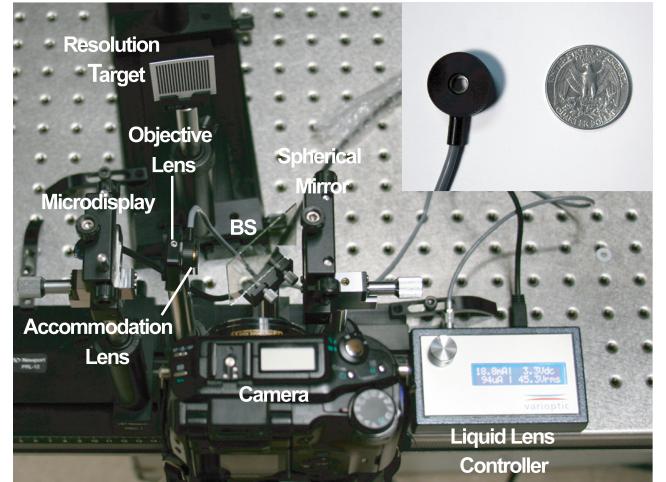


Figure 3. Monocular bench prototype of an optical see-through HMD with addressable focal planes. Inset: the liquid lens.

Figure 4 (a) plots the optical power of the liquid lens device as a function of applied voltages. The curve is simulated by importing specifications of the liquid lens under different driving voltages in an optical design software—CODE V [28]. Two examples are shown in Figure 4 (a): at 38Vrms of applied voltage, the liquid lens produces 0 diopter of optical power indicated by the flat shape of the liquid interface; while at 49Vrms the liquid lens delivers 10.5 diopters of optical power indicated by the strongly curved liquid interface.

Based on the parametric selections in the bench prototype and Eq. (1), Figure 4 (b) plots the accommodation cue of the HMD as a function of the applied voltage on the liquid lens. Labeled by two solid triangle markers in Figure 4 (b), driving the liquid lens at 38 Vrms and 49 Vrms corresponds to accommodation cues at 6 diopters and 1 diopter, respectively. Driving the liquid lens from 32 Vrms to 51 Vrms will enable the accommodation cue of the HMD prototype from 12.5cm (8 diopters) to infinity (0 diopter) correspondingly, covering almost the whole accommodative range of the HVS [25].

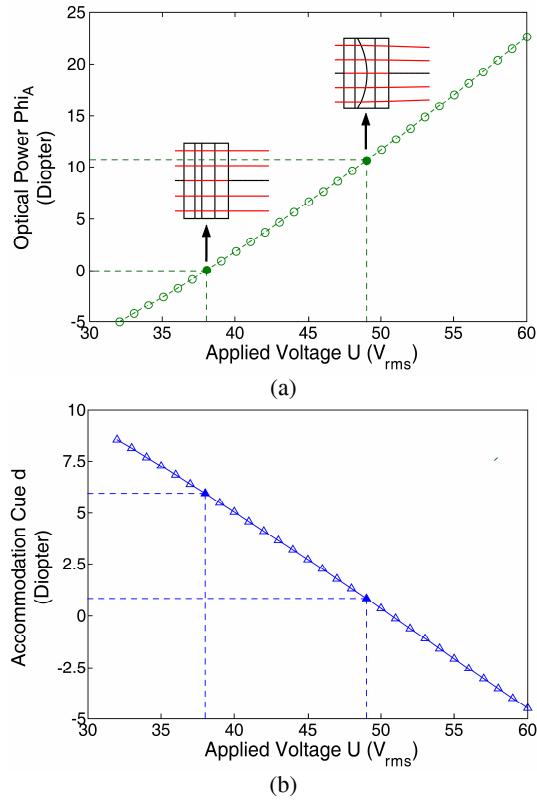


Figure 4. (a) Optical power of the liquid lens and (b) accommodation cue of the HMD prototype versus the applied voltage on the liquid lens.

3 EXPERIMENT RESULTS

Indicated by Figure 4, the addressability of the accommodation cue in the see-through HMD prototype is determined by the ways to address the liquid lens. In this section, we explore two operation modes—variable focal plane and multi-focal plane, in which the proposed system can be operated. In a variable focal plane mode, for example, the voltage applied on the liquid lens can be dynamically adjusted through a user interface to focus the display at different focal distances, from infinity to as close as the near point of the eye. This operation mode meets specific application needs, for instance, to match the accommodation cue of virtual and real objects in mixed and augmented realities. Additionally, the focus cues can be pre-programmed to animate a virtual object smoothly moving in a 3D space. In a multi-focal plane mode, the liquid lens may be fast-switched between multiple driving voltages, to provide multiple focal distances, such as I' and II' in Figure 2, time-sequentially. Synchronized with the focal plane switching, the microdisplay is updated accordingly to render virtual objects at distances approximately matching with the accommodation cue of the HMD. The faster the response speed of the liquid lens and the higher the refresh rate of the microdisplay device are, the more focal planes can be presented at a flickering-free rate. As indicated in Figure 4 (b), in the multi-focal plane mode, the dioptric spacing between adjacent focal planes and the overall range of accommodation cue can be controlled by changing the voltages applied to the liquid lens. Such unique capability enables the flexible management of accommodation cue suited for a variety of applications, which desire either focal planes spanning a wide depth range or dense focal planes within a relatively smaller depth range for better accuracy.



Figure 5. Two real cups, one at 40cm and the other at 100cm away from the eye, are realistically augmented with a virtual COKE can rendered at two different depths: (a)-(b) 40cm; and (c)-(d) 100cm. The digital camera was focused at (a), (d) 40cm and (b), (c) 100cm, respectively.

3.1 Variable focal plane approach

In order to demonstrate the addressability of focus cues in a variable focal plane mode, we set up three bar-type resolution targets along the visual axis of the HMD as references to the virtual image with variable focus cues. As shown on the left side of each sub-image in Figure 1, the three bar targets were placed 16cm (large-size), 33cm (mid-size) and 100cm (small-size), respectively, away from the exit pupil of the HMD (i.e. eye position). The periods of the bar targets are inversely proportional to their distance to the eye, so that the subtended angular resolution of the grating remains constant among all targets.

In the first example shown in Figure 1, a virtual torus was animated in a way that it moves along the visual axis of the HMD at a constant speed from either 16cm to 100cm or vice versa. In synchronization, the voltage applied to the liquid lens was adjusted linearly at a pace matching with the distance of the virtual torus to the eye. By varying the voltage from 38 Vrms to 49 Vrms, the accommodation cue of the HMD is varied correspondingly from 6 diopters to 1 diopter. A digital camcorder (DCR-TRV20 NTSC, SONY) was mounted at the exit pupil location of the HMD to capture demonstration videos. Figures 1(a) through 1(c) show three frames with the camcorder focusing on the resolution targets located at 16cm, 33cm and 100cm, respectively. A movie clip (VariFocalPlane.wmv) consisted of three sub-clips (a), (b) and (c) at the same time, the virtual torus in clip (a) only appears in focus when the applied voltage on the liquid lens is 38rms, since the camcorder in clip (a) was constantly focused at 16cm (6 diopters) distance. Similarly, the torus in clip (b) and (c) only appears in focus when the driving voltage is 45 Vrms and 49Vrms, respectively. The images and the video clearly demonstrate the change of accommodation cue of the virtual object.

The example above demonstrates the animation of a virtual object with variable accommodation cues. Such passive control of accommodation cues prevents natural user interactions with virtual and real worlds. Figure 5 shows a simple mixed-reality application in a variable focal plane mode. The real scene consists of two real cups, one located at 40cm from the viewer and the other at 100cm away. Focusing the digital camera at 40cm and 100cm, respectively, Figure 5(a) shows a virtual COKE can clearly rendered at a 40cm depth, while in Figure 5(c) the can was rendered at a 100cm depth. In comparison, the virtual COKE can appears blurred while the digital camera was focused at 100cm (Figure 5(b)) and 40cm (Figure 5 (d)), respectively. By applying a voltage at 46 Vrms and 49 Vrms, respectively, the virtual COKE can appears realistically mixed with the real cups at two different depths with correct focusing cues. In this example, while a user interacts with a virtual object, its focusing cue may be dynamically modified to match with its physical distance to the user, yielding a realistic augmentation with a real scene. Such capability may offer accurate depth perceptions in an augmented reality application. The experiment in Section 4 will discuss in detail about the depth perception and accommodative responses in the display prototype.

3.2 Multi-focal plane approach

Although the variable focal plane approach demonstrated potential applications in Section 3.1, there are endured research interests to the development of true-3D displays, in which depth perceptions are not limited by a single or a variable focal plane. As discussed in Section 1.1, compared to the volumetric displays, multi-focal plane displays leverage the accuracy on depth perception, the practicability for device implementation and the accessibility to computational resources as well as mainstream graphics rendering techniques. In this section, we explore the

feasibility of implementing time-multiplexed multi-focal planes in an optical see-through HMD.

3.2.1 Method

Similar to the RSD system with a variable focus DMM device [18, 21], adopting the liquid lens as an active optical element to control the accommodation cue offers the potential to develop a time-multiplexed multi-focal plane display. However, it is worth noting that there are a few major differences between our approach and the RSD technique. Firstly, we have used a transmissive liquid lens, rather than a reflective DMM device, which enables a compact and practical HMD layout without compromising the range of accommodation cues. Secondly, instead of addressing each pixel by a laser scanning mechanism in the RSD, our approach uses a 2D microdisplay device to present a high-resolution, full-color image time-sequentially to generate multiple focal planes, which remarkably relaxes the requirement for the addressing speed of the display device and the active optical components down to 100Hz level, whereas the pixel-sequential rendering approach in a RSD system requires MHz operation speed for both the DMM device and the scanning mechanism of multiple laser beams.

Starting with the simplest case of a dual focal plane design, the driving signal of the liquid lens and the rendering method of virtual objects, are shown in Figure 6 (a) and (b), respectively. Different from the variable focal plane approach in Section 3.1, the liquid lens is fast-switched between two driving voltages, as shown in Figure 6(a). The accommodation cue of the HMD is consequently fast-switched between far and near distances. In synchronization with the driving signal of the liquid lens, far and near virtual objects are rendered on two separate image frames and displayed sequentially, as shown in Figure 6(b). Apparently, to create a flickering-free appearance of the virtual objects rendered sequentially at two depths, from the hardware perspective, this dual-focal plane method requires not only the microdisplay and graphical card to have a frame rate two-times higher than their regular counterparts, but also the liquid lens with compatible response speed. In general, the maximally achievable frame rate, f_N , of a multi-focal plane display is given by:

$$f_N = \frac{f_{\min}}{N} \quad (2)$$

where N is the total number of focal planes and f_{\min} is the lowest response speed (in Hz) among the microdisplay, the active optical elements, and the graphics card. The waveforms in Figure 6 were drawn under the assumption that all of these elements can operate at ideal speed.

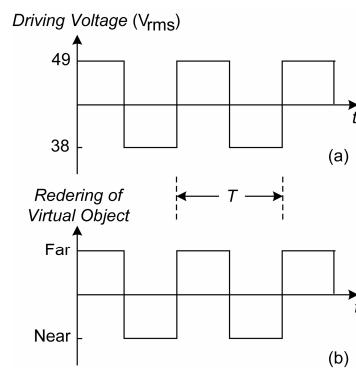


Figure 6. Driving mechanism of the liquid lens and rendering of virtual objects to produce a dual-focal plane display.

3.2.2 Results

In our prototype implementation, the liquid lens was driven by a square wave switching between 49 Vrms and 38 Vrms, respectively, and the accommodation cue of the HMD was consequently fast-switched between 100cm and 16cm. The period, T , of the driving signal was adjustable in the rendering program. Ideally, T should be set to match the response speed of the slowest device in the system, which determines the frame rate of the dual-focal plane display. For example, if T is set to be 200ms matching the speed (f_{\min}) of the slowest device in the system, the speed of the dual focal plane display will be 5Hz and the virtual objects at the two depth ranges will appear alternately. If T is set to 20ms (50Hz) faster than the slowest device (e.g. the highest refresh rate of the graphics card is 75 Hz in the current setup), the virtual objects will be rendered at a speed of about $f_{\min}/2 = 37.5\text{Hz}$. The control unit of the liquid lens device allows for a high-speed operation mode, in which the driving voltage is updated at every $600\mu\text{s}$ to drive the liquid lens. However, the response speed of the current liquid lens Arctic 320, shown as the red curve with diamond markers in Figure 8, is in the level of 75ms. The maximum refresh rate of the microdisplay is 85Hz and of the graphics card is 75Hz. Obviously, the liquid lens is the limiting factor on speed in the current prototype.

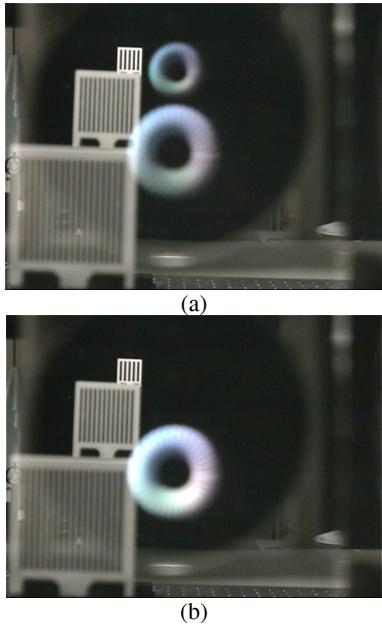


Figure 7. Two toruses at different depths are rendered sequentially by the dual-focal plane HMD prototype at a frame rate of (a) 37.5Hz (Dual-focalPlane20ms.wmv) and (b) 0.25Hz (Dual-focalPlane4s.wmv).

Along with two video clips, Figures 7(a) and (b) demonstrate the experimental results of the dual-focal plane display at two different periods, T , of 20ms and 4s, respectively. In both examples, the focus of the camcorder was manually adjusted slowly from 100cm to 16cm. As shown in the video clip 7(a) (Dual-focalPlane20ms.wmv), the two virtual toruses, one rendered at a depth of 100cm and the other at 16cm, appeared simultaneously with noticeable flickering effect, partially due to the relatively low frame rate (75Hz) of the graphics card and the sampling rate of the camcorder (30Hz). However, along with the focus change of the camera, both virtual objects appear to be

either in focus or blurred at the same time. The focus cues of the two virtual objects were hardly visible due to the slow response of the liquid lens. The video clip (Dual-focalPlane4s.wmv) in Figure 7 (b) demonstrates the result with $T=4\text{s}$. The two virtual objects appear alternately but with a faithful focus cue in synchronization with the focus setting of the camcorder. This clearly demonstrates the need for much faster display and active optical element to develop practically usable multi-focal plane displays with less flickering effects.

3.2.3 Discussions

The observation in Figure 7 was confirmed by the measurement of the response speed of the liquid lens currently used in the prototype (Arctic 320). The time response is plotted in Figure 8, shown as the red curve with diamond markers [13]. The optical power of the liquid lens was measured on the time scale by a Shark-Hartman wavefront sensor with 5ms temporal resolution. Corresponding to a step function of the driving voltage from 0 to 60 Vrms, the response curve of the Arctic 320 suggests a 74ms rise time to reach 90% of the maximum optical power, which apparently is the slowest device in the current prototype and is insufficient to render the dual-focal plane display at 37.5Hz with faithful accommodation cues.

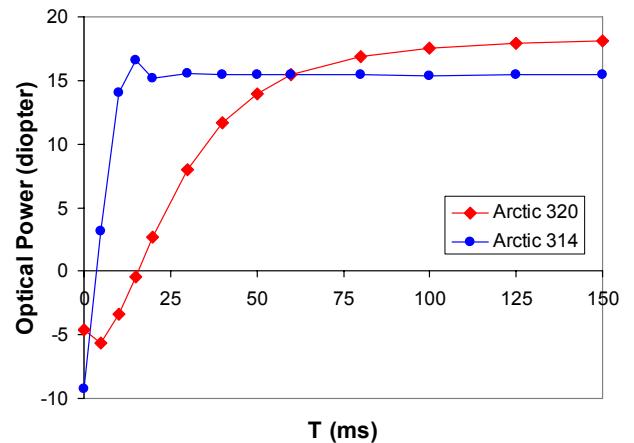


Figure 8. Time response of the liquid lens: module Arctic 320 and module Arctic 314.

For future developments, specifically on the dual-focal plane see-through HMD with satisfactory frame rates, we examined the hardware requirements for all of the active components being used in the proposed system. Shown on the left column of Table 1, potential limiting factors to the maximum speed of the dual focal planes display are listed, including the liquid lens, the microdisplay and the graphical card. For example, if the liquid lens Arctic 320 is used in the HMD, the maximum achievable frame rate of the dual-focal plane display is only 7Hz. However, the new generation liquid lens, Arctic 314, from Varioptic has promised 5~10 times faster response speed than the Arctic 320 [13]. As shown by the blue curve with circles in Figure 8, the new liquid lens module demonstrates a 9ms rise time to reach 90% of its maximum optical power. With this new liquid lens, the highest achievable frequency of the dual focal planes display would be 56Hz, if the lens is the limiting factor of speed in the whole system. This frame rate almost approaches the flickering-free frequency of 60Hz. In future developments, we plan to implement the new liquid lens into the bench prototype for a dual-focal planes optical see-through HMD. We also plan to explore many other technology candidates for the fast switching active optical components, such as a bi-focal liquid crystal lens device.

The SVGA resolution full color OLED microdisplay currently used in the prototype has a refresh rate up to 85Hz, which can produce dual-focal planes at 42.5Hz in maximum. Other display technologies, such as the digital micromirror device (DMD) and ferroelectric liquid crystal on silicon (FLCOS) displays, should be explored to achieve high refresh rate.

Table 1. Hardware evaluations for the development of a dual focal plane stereoscopic display.

Limiting factor	Hardware speed (ms)	Maximum speed of the dual focal planes display (Hz)
Liquid Lens, Arctic 320	74	7
Graphics Card, 75Hz	13.3	37.5
OLED Microdisplay, 85Hz	11.8	42.5
Liquid Lens, Arctic 314	9	56
Flicker-free frequency	8.4	60

4 EVALUATION EXPERIMENTS

Numerous usability studies have been conducted on traditional stereoscopic displays with a fixed focal distance [6, 27, 17, 22, 26, 5]. One of the widely reported adverse effects by using stereoscopic displays is the depth perception error partially attributed to the conflicting accommodation-convergence cues [5, 6, 27]. To better understand how the depth perception is affected by and how the HVS responds to the addressable focal planes in the see-through HMD prototype, we performed two types of experiments: one is a subjective evaluation, in which we explored the perceived depth of the virtual display by a controlled depth judgment task; the other is an objective measurement, in which we quantitatively measured the accommodative response of a user to the virtual target presented at different depths. Both experiments were carried out in a variable focal plane mode.

4.1 Subjective evaluation

4.1.1 Experimental setup and task description

The major purpose of the subjective evaluation is to understand the relationship of the perceived depth of virtual objects versus the accommodation cue rendered by the active optical method. We devised a depth judgment task to evaluate the perceived depth of a virtual object. In this task, a subject is asked to position a reference real object based on the perceived depth of a virtual object.

Figure 9 illustrates the schematic setup for the subjective evaluation. The monocular bench prototype described in Section 2 was employed as the testbed. The total FOV of the HMD prototype is divided into top and bottom regions, each of which subtends about 12-degree diagonal FOV. The top region displays a virtual target while a real target covers the bottom part. Because the focus cue is the major visual stimuli in the subjective evaluation, square-wave gratings were chosen as the virtual and real targets: circular gratings as the virtual target and bar gratings for the real one. The real target was mounted on a translational stage and can be moved along the visual axis of the HMD. The accommodation cue of the virtual target, indicated as the distance from the eye to the virtual image display, was controlled by applying three different voltages to the liquid lens: 49 Vrms, 44

Vrms and 40 Vrms which correspond to accommodation cues of 1 diopter, 3 diopters and 5 diopters, respectively.

For a virtual target displayed at one of the three depths, the experimenter helps the subject to move the real target back and forth along the visual axis until the subject determines that they appear to be at the same depth. For better accuracy, once the initial depth judgment was made, the subject made fine depth judgment by repeatedly moving the real target from the previous judgment position backward and forward at the decreasing amount of increment until the incremental step was small enough. The final position of the real target was recorded as the perceived depth of the virtual target. A good match between the perceived depth and the actual accommodation cue rendered by the liquid lens will suggest a realistic depth perception of the virtual object. During the experiment, the subject positions his or her head on a chin rest and only views the targets through one eye with the other eye covered to eliminate other depth cues rather than the accommodation cue we are interested in.

Five subjects with either normal vision or corrected vision were recruited, with 1 female and 4 males. The average age of the subjects is 28. Each subject performed three depth judgment trials. The virtual targets were presented in a random order from the three accommodation cues. Before starting the experiment, the experimenter explained each subject the experiment procedure.

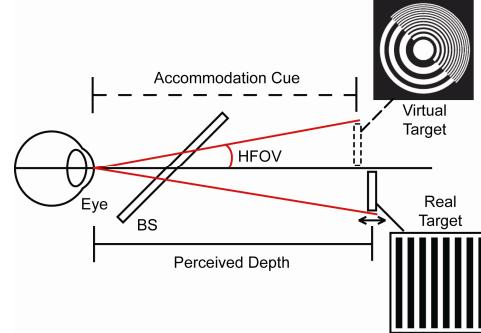


Figure 9. Schematic illustration of the subjective evaluation experiment.

4.1.2 Results

Figure 10 plots the relationship between perceived depth of the virtual target and the accommodation cue rendered by the liquid lens. As the accommodation cue varies, shown as 1 diopter, 3 diopters and 5 diopters along the x-axis, the average perceived depth among the 5 subjects is 1.1 diopter, 2.9 diopters and 5.2 diopters, respectively, shown as blue diamonds in Figure 10. The standard deviation of the perceived depths at the three accommodation cue conditions are 0.1 diopter, 0.4 diopter, and 1.3 diopters, respectively. The error bars in Figure 10 indicate the minimum and maximum readings of the perceived depths among all subjects. Overall, the average values of the perceived depths match well with the rendered accommodation cues. This preliminary result suggests a potentially improved depth perception by appropriately rendering accommodation cues in a HMD with addressable focal planes, by comparing to the usability studies on traditional stereoscopic displays which have suggested distorted and compressed perceived depths by solely rendering binocular disparity cue [5, 14, 26]. The errors in the user evaluation may be explained by many contributing factors, such as image aberrations, ambient noise, and subject's fatigue level. Similar results were also reported on a user experiment with an addressable focal plane HMD in which the addressable focal plane was achieved by the active DMM device [18, 21].

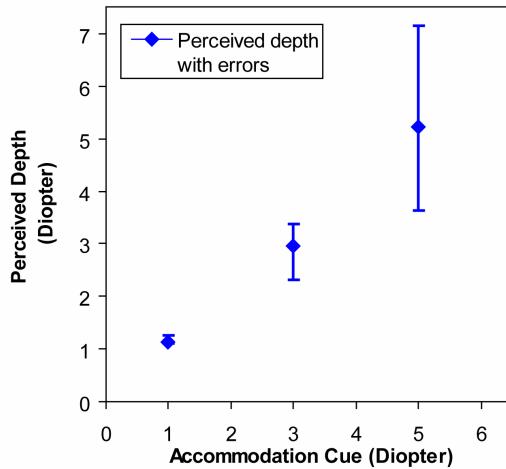


Figure 10. Subjective evaluation results of the perceived depth of the virtual display versus the accommodation cue rendered by the liquid lens.

4.2 Objective measurement

The major purpose of the objective evaluation is to quantify accommodative response of the HVS to the depth cues presented through our display prototype. In this experiment, the accommodative responses of the eye were measured by a near-infrared (NIR) auto-refractor (RM-8000B, Topcon). The auto-refractor has a measurement range of the refractive power from -20 to 20 diopters and a measurement speed of about 2s. The eye relief of the auto-refractor is about 50mm. In the objective measurement, the auto-refractor was placed right in front of the BS, so that the exit pupil of the auto-refractor coincides with that of the see-through HMD prototype. Throughout the data acquisition procedure, the ambient lights were turned off to prevent their influences on accommodation responses.

During the test, a subject with normal vision was asked to focus on the virtual display, which was presented at 1 diopter, 3 diopters and 5 diopters, respectively, in a three-trial test. At each trial, after the subject set his or her focus on the virtual display, the accommodative response of the subject's eye was recorded at every 2s for up to 9 measurement points. The results for one subject were plotted in Figure 11 for the three trials corresponding to three focal distance of the virtual display and the data points were shown as three sets of blue diamonds. The red solid lines in Figure 11 correspond to the accommodation cues rendered by the liquid lens. Although the measured accommodative response of the user fluctuates with time, the average value of the 9 measurements in each trial is 0.97 diopters, 2.95 diopters, and 5.38 diopters, with standard deviations of 0.33 diopters, 0.33 diopters and 0.42 diopters, respectively. The averages of the accommodative responses of the user match well with the accommodation cues stimuli presented by the see-through HMD prototype.

4.3 Discussion

Although accommodation response may be caused by a variety of visual stimuli, such as convergence, brightness and spatial frequency, the major observations in the experiments described above are elicited mostly by the accommodation cue of the virtual display, by simply controlling the position of the focal plane in the monocular prototype without rendering other types of visual stimuli. The preliminary results from both the subjective and objective evaluations in this paper suggest that in a variable focal

plane HMD, a user can yield appropriate accommodative responses to the variable accommodation cue and the perceived depths can match well with rendered accommodation cues. Referring to the variable focal plane method in Section 3.1, the proposed HMD prototype can potentially yield realistic rendering of virtual objects throughout an augmented space by presenting extra accommodation cue in traditional stereoscopic displays. Referring to the multi-focal plane method in Section 3.2, the development of the prototype based on a faster liquid lens would possibly allow for decreased discrepancy between accommodation and convergence.

There have been increasing research interests to study characteristics of accommodation response in stereoscopic displays [5, 14, 27]. The addressable focal plane HMDs can be applicable to such human factor studies, whenever dynamic manipulation of the focus cues is desired. As an example, convergence induced accommodation has been broadly explored in traditional stereoscopic displays. Due to the lack of providing accommodation cues in traditional stereoscopic displays, however, accommodation induced convergence has not been well studied and understood. An addressable focal plane HMD can offer opportunities for such psychophysical studies to better understand the underlying mechanism of stereoscopic vision.

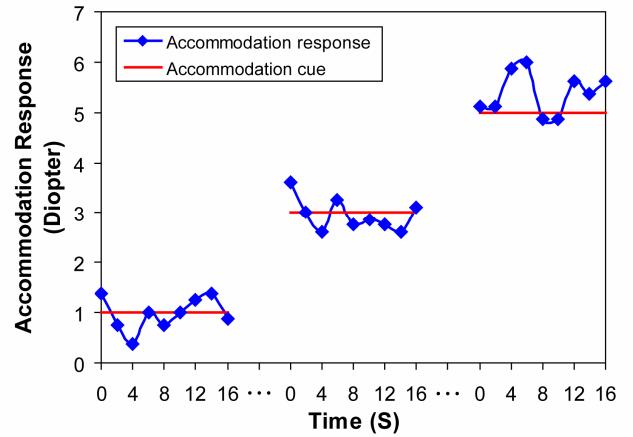


Figure 11. Objective measurements of the accommodative responses to the accommodation cues presented by the see-through HMD prototype.

5 CONCLUSION AND FUTURE WORK

In this paper, we presented the design of a bench prototype for an optical see-through HMD with addressable focal planes, enabled by a liquid lens device. The accommodation cues in a proof-of-concept bench prototype can be rendered, from optical infinity to as close as the near point of the eye, either in a variable focal plane mode or in a multi-focal plane manner. The applicability of the variable focal plane approach was validated by two mixed-reality examples. We also examined the hardware requirements for the development of a flickering-free dual-focal plane see-through HMD. Finally, we reported two experiments to evaluate the perceived depth and the accommodative responses of the eye as the accommodation cue is varied in the display prototype. The preliminary results from both evaluations suggest that in a variable focal plane HMD, a user can yield appropriate accommodative responses to the variable accommodation cue and the perceived depths can match well with rendered accommodation cues.

Addressable focal planes in an optical see-through HMD enable numerous possibilities for future research. From a technology

development perspective, we will explore many candidate technologies, such as high refresh rate DMD or FLCOS displays as well as high speed liquid lens or liquid crystal lens to develop a multi-focal plane display at flickering-free rate. From the usability perspective, we will use the addressable focal plane HMD for various human factor studies, such as to explore accommodation induced convergence in stereoscopic displays.

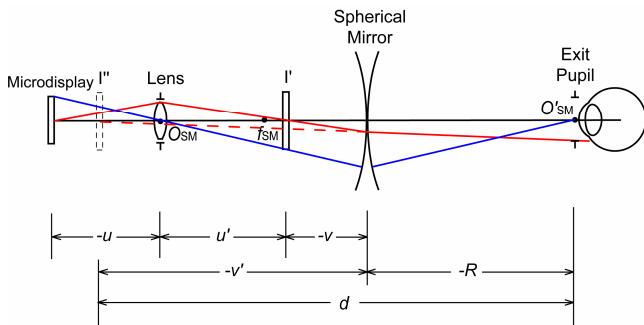


Figure 12. Unfolded optical path of the HMD schematics.

APPENDIX: DERIVATION OF EQUATION 1

Figure 12 shows the unfolded optical path of the schematic diagram in Figure 2. The final accommodation cue, d , is determined by a two-step imaging procedure: from the microdisplay to the intermediate image I' , and from the intermediate image I' to the final virtual image I'' . Based on the first order optics, the location of the first intermediate image I' is:

$$u' = \frac{u}{1+u\phi} \quad (3)$$

Furthermore, the intermediate image I' becomes the object for the spherical mirror with the object distance v given by:

$$v = R + u' = \frac{R + u + uR\phi}{1 + u\phi} \quad (4)$$

After the reflection by the spherical mirror, the virtual image I'' is formed at distance v' to the vertex of the spherical mirror given by:

$$v' = -\frac{R^2 + uR + uR^2\phi}{R + 2u + uR\phi} \quad (5)$$

Finally, the accommodation cue, d (assuming $d > 0$), is obtained as:

$$d = -v' - R = \frac{-uR}{2u + R + uR\phi} \quad (6)$$

ACKNOWLEDGEMENT

We would like to thank Professor Jim Schwiegerling for providing the auto-refractor used in our evaluation experiment. This work was partially funded by the National Science Foundation grant 0534777 and the TRIF Imaging Fellowship of the University of Arizona.

REFERENCES

- [1] K. Akeley, S. J. Watt, A. R. Girshick and M. S. Banks. A stereo display prototype with multiple focal distances. In Joe Marks, editor, *Proceedings of SIGGRAPH '04* (Los Angles, California, August 08–14, 2004), pages 804–813, ACM SIGGRAPH, ACM Press, September 2004.
- [2] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21:34–47, Nov/Dec 2001.
- [3] C. Cruz-Neira, D. J. Sandin and T. A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142. ACM SIGGRAPH, ACM Press, September 1993.
- [4] E. Downing, L. Hesselink, J. Ralston and R. A. Macfarlane. A three-color, solid-state, three-dimensional display. *Science*, 273:1185–1189, August 1996.
- [5] H. Davis, D. Buckley, R. E. G. Jacobs, D. A. A. Brennan and J. P. Frisby. Accommodation to large disparity stereograms. *Journal of AAPOS*, 6:377–384, Dec 2002.
- [6] G. K. Edgar, J. C. D. Pope and I. R. Craig. Visual accommodation problems with head-up and helmet-mounted displays. *Displays*, 15:68–75, November 1994.
- [7] G. E. Favalaro, J. Napoli, D. M. Hall, R. K. Dorval, M. G. Giovinco, M. J. Richmond and W. S. Chun. 100 million-voxel volumetric display. In *Proc. SPIE*, volume 4712, pages 300–312, August 2002.
- [8] D. Graham-Rowe. Liquid lenses make a splash. *Nature-Photonics*, September: 2–4, 2006.
- [9] J. E. Greivenkamp. Field guide to geometrical optics. Bellingham, WA, SPIE Press, 2004.
- [10] H. Hua. Merging the Worlds of Atoms and Bits: Augmented Virtual Environments. *Optics and Photonics News*, 17:26–33, October 2006.
- [11] H. Hua and C. Y. Gao. Design of a bright polarized head-mounted projection display. *Applied Optics*, 46:2600–2610, May 2007.
- [12] J. F. Heanue, C. B. Matthew, and L. Hesselink. Volume Holographic Storage and Retrieval of Digital Data. *Science*, 265:749–752, August 1994.
- [13] <http://www.varioptic.com>.
- [14] T. Inoue and H. Ohzu. Accommodative responses to stereoscopic three-dimensional display. *Applied Optics*, 36:4509–4515, July 1997.
- [15] H. Kurabayashi, M. Date, S. Suyama and T. Hatada. A method for reproducing apparent continuous depth in a stereoscopic display using "Depth-Fused 3D" technology. *Journal of the Society for Information Display*, 14:493–498, May 2006.
- [16] S. Kuiper, B. H. W. Hendriks. Variable-focus liquid lens for miniature cameras. *Applied Physics Letters*, 85:1128–1130, August 2004.
- [17] M. Mon-Williams, J. P. Wann and S. Rushton. Binocular vision in a virtual world : visual deficits following the wearing of a head-mounted display. *Ophthalmic Physiol. Opt.*, 13:387–391, October 1993.
- [18] S. C. McQuaide, E. J. Seibel, J. P. Kelly, B. T. Schowengerdt BT and T. A. A. Furness. A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror. *Displays*, 24:65–72, August 2003.
- [19] J. P. Rolland, M. W. Krueger, and A. Goon. Multifocal planes head-mounted displays. *Applied Optics*, 39:3209–3215, July 2000.
- [20] A. Sullivan. A solid-state multi-planar volumetric display. In *SID Symposium Digest of Technical Papers*, 34:1531–1533, May 2003.
- [21] B. T. Schowengerdt and E. J. Seibel. True 3-D scanned voxel displays using single or multiple light sources. *J. Soc. Info. Display*, 14:135–143, February 2006.
- [22] J. Sheedy and N. Bergstrom. Performance and comfort on near-eye computer displays. *Optometry and Vision Science*, 79:306–312, May 2002.

- [23] S. Shiwa, K. Omura and F. Kishino. Proposal for a 3-D display with accommodative compensation: 3DDAC. *Journal of the Society for Information Display*, 4:255-261, December 1996.
- [24] 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*, 13:665-671, August 2005.
- [25] B. A. Wandell. Foundations of Vision. Sinauer Associates Inc., 1995.
- [26] J. P. Wann, S. Rushton, and M. Mon-Williams. Natural problems for stereoscopic depth perception in virtual environments *Vision Res.*, 35:2731-2736, October 1995.
- [27] S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks. Focus cues affect perceived depth. *J. Vision*, 5:834-862, December 2005.
- [28] <http://www.opticalres.com>.