Computational Augmented Reality Eyeglasses

Andrew Maimone* Henry Fuchs†

Department of Computer Science
University of North Carolina at Chapel Hill

Figure 1: A) Prototype optical-see through glasses with a wide field of view and occlusion support. (Driving electronics are external.) The device is capable of providing in-focus augmented images at multiple simultaneous focal depths using only a set of backlit stacked LCD panels placed closer than the eyes can accommodate: the distance of normal eyeglasses. B,C) Photos taken of a prototype display with a 3.7 mm lens aperture and a distance of 18 mm between the display and the camera's center of projection, approximating a human viewer wearing eyeglasses. The camera focus distance was set to 10 cm. Photo B) was taken with an image naively displayed on one of the layers, while photo C) was taken using the multilayer optimization technique described in this paper.

Abstract

In this paper we discuss the design of an optical see-through head-worn display supporting a wide field of view, selective occlusion, and multiple simultaneous focal depths that can be constructed in a compact eyeglasses-like form factor. Building on recent developments in multilayer desktop 3D displays, our approach requires no reflective, refractive, or diffractive components, but instead relies on a set of optimized patterns to produce a focused image when displayed on a stack of spatial light modulators positioned closer than the eye accommodation distance. We extend existing multilayer display ray constraint and optimization formulations while also purposing the spatial light modulators both as a display and as a selective occlusion mask. We verify the design on an experimental prototype and discuss challenges to building a practical display.

Keywords: augmented reality, three-dimensional displays

Index Terms: I.3.1 [Computer Graphics]: Hardware Architecture—Three-dimensional displays

1 Introduction

The limits of augmented reality (AR) are defined by displays. In the past few years, several promising commercial displays have been demonstrated which show the promise of delivering AR to widespread public use. In particular, Google Glass[1], a compact and lightweight optical see-through monocular display, provides two key benefits of AR: encumbrance-free and instant access to information by affixing it in visual field.

Encouraged by this first step, we are excited at the future prospect of practical optical see-through displays that support the full range of benefits of augmented reality. Namely *wide field of view* displays would allow meaningful connections to be established between the real world and overlays, and devices supporting *mutual occlusion* with a robust set of depth cues, notably the oft-missing *focal cues*, would support a more convincing sense that augmentations are present in the real world. At the same time, we wish to retain the compact and lightweight form factors that allow encumbrance-free, mobile, and extended use.

Impressive advances in optical see-through displays over the past decade have collectively spanned this set of important requirements – *encumbrance-free form factors* [4, 22], *wide fields of view* [6], *occlusion support* [15, 3, 23, 8], and *focal depth cues* [12] – yet no known individual device possesses all, or even more than one of these characteristics. As each of these features typically place many physical constraints on the device, it is very challenging to incorporate all of them in hardware.

In this paper, we introduce a novel strategy for optical see-through design that offers the potential to satisfy the aforementioned four key requirements: *encumbrance-free form factor*, *wide field of view*, *occlusion support* and *focal depth cues*. Our approach builds on work in the emerging field of *computational displays*, simple optical devices whose functionality and complexity generally lies in software. We rely on no reflective, refractive, or diffractive components, but rather on a set of stacked spatial light modulator layers *positioned closer than the typical eye accommodation distance* which are used to create a focused image through a series of optimized time-multiplexed patterns. Since the layers are also transparent, they can be purposed both for augmented image formation and to provide a selectively occluded view of the real scene in a time-multiplexed fashion.

We verify these key aspects of our design on an early prototype device and discuss optical considerations and limitations in image quality, computation speed, and spatial light modulator performance that must be addressed to produce a practical display.

Contributions Our primary contribution is the application of multilayer computational display techniques to head-worn optical see-through displays. Specific contributions follow:

*e-mail: maimone@cs.unc.edu
†e-mail: fuchs@cs.unc.edu

[1] http://www.google.com/glass

1. The analysis of multilayer display architectures for a new use case: a *transparent, near-eye* display that creates images that appear *far from the physical display layers*
2. The proposed use of the display layers both to form images and to provide *per-pixel, variable transparency occlusion* of the real scene, in a time-multiplexed fashion
3. Revised multilayer light field constraint and optimization criteria critical to achieve acceptable performance in near-eye display configurations

We also provide a simple scheme for perceptual image optimization that is compatible with existing optimization frameworks.

2 Related Work

2.1 Conventional Optical See-Through HMDs

Conventional optical see-through displays use a combination of reflective, refractive, and/or diffractive elements to place a focused augmented image from an electronic display in the observer's unobstructed viewing path. Previous work is categorized by its main design criteria below:

Non-Encumbering Several optical see-through designs have been proposed or demonstrated that approach an eyeglasses form factor. To achieve this compact size, freeform optics [4, 22] or holographic optical elements and waveguides [14] are often used. Such devices currently have fields of view limited to 40° or less. Commercial devices are available from Epson, Lumus, and Vuzix.

We also note compact devices in even smaller form factors. Contact lens displays [19, 7] have been demonstrated, but are in the very early stages of development: current displays consist of only a single or small number of pixels and do not address the issue of the eye focusing on the display. Small monocular displays have long been used for military applications and have recently been proposed for widespread public use by Google Glass, but their narrow field of view limits their use primarily to a small information display.

Wide Field of View Wide field of view optical see-through displays (those that cover much of the viewer's field of view) have also been proposed with freeform optics. One such design [6] uses a free form surface to achieve a $56^\circ \times 45^\circ$ field of view that can be tiled to achieve an ultra-wide $119^\circ \times 56^\circ$ field of view; however the lens elements are bulkier than ordinary eyeglasses at 1.7 cm thick.

Wide field of view hybrid contact lenses and glasses sets have also been proposed; Innovega[2] has demonstrated a contact lens containing an small inset lens that allows the user to focus on a glasses-based display while allowing light from the environment to enter the remainder of the pupil. Filters are used to separate light between the two regions. The manufacturer claims the design can achieve a 100° field of view [13]; however, the device is somewhat encumbering as it requires the user to wear both contact lenses and glasses, and the contact lens blocks light from the user's environment from reaching a region of the eye.

Mutual Occlusion Several optical see-through displays [15, 3, 23, 8] have been designed with a selective occlusion mask that allows augmented objects to appear opaque by blocking out regions of real objects. Such devices are currently bulky; the only known commercial device [23] has a form factor that resembles binoculars.

Focal Depth Cues Optical see-through displays have also been demonstrated that allow augmented objects to be placed at different focal positions, allowing more consistency between augmented object depth cues, as well as improved consistency between augmented object and real object cues. One such approach [12] achieves focus control through a liquid lens. Current optical approaches provide only global focus control, allowing only a single simultaneous focal depth; however, the appearance of multiple focal depths can be achieved through time-multiplexing [12].

[2] http://innovega-inc.com

Although impressive advances have been made in conventional optical see-through displays over the past decade, we are not aware of any existing device that combines a non-encumbering design, wide field of view, occlusion support, and focal depth cues (or more than one of these qualities) that we believe is important to take advantage of the full range of benefits of AR. Our proposed design incorporates all of these qualities using an alternative computational approach utilizing stacked light modulation layers.

2.2 Multilayer Computational Displays

An emerging class of multilayer computational displays produce multi-view imagery by displaying image patterns on a stack of display layers [9, 16, 25, 17, 26]. When the display stack is viewed from different positions, parallax causes different pixels on each layer to align and form the perceived image, providing motion parallax and stereo depth cues. The modulation state of each layer (e.g. attenuation, polarization, etc.) is optimized to produce the most numerically accurate images for the desired set of viewpoints. Pattern time-multiplexing over the flicker fusion threshold can be used to obtain additional degrees of freedom for layer optimization, improving image quality. Directional backlights may be used with the display layers to increase angular resolution [26].

To date, multilayer designs have been limited to desktop 3D displays. The present work builds on these existing displays, while exploring several new design aspects: near eye displays with layers placed closer than the typical eye accommodation distance, sight through the display with selective occlusion, and revised light field constraint and optimization criteria.

Upcoming work in computational displays [20] also shows the possibility of stimulating correct eye accommodation by synthesizing narrow, high density light fields over the areas of the viewer's pupils. In the present work, we also synthesize such light fields, but with the primary goal of creating a focused image nearer than the eyes can accommodate, rather than presenting a scene with multiple focal depths on a distant display. We also use a different hardware approach to gain a thin form factor and see-through ability: light is filtered through a set of transparent spatial light modulators rather than focused with a lens. The present work also uses additional constraints on the input light field and considers the image formed on the retina during the optimization process to improve light field compression performance.

3 System Design and Analysis

Overview In a conventional optical see-through design, light rays from a 2D image source (e.g. a microdisplay) are *redirected* by reflection, refraction, and/or diffraction to form a 2D image that appears some distance in front of the viewer. In the proposed design, as in some existing multilayer desktop 3D displays [16, 26], we take a computational approach: the set of all possible light rays (i.e. a white omni-directional area light source) are *filtered* using optimized patterns displayed on spatial light modulators (SLM) so that only the desired rays remain (see Figure 2).

In this section, we describe the physical design and operation of our proposed device, provide an overview of multilayer optimization, describe two key insights that enable the use of multilayer techniques for near-eye applications, and discuss optical challenges for building a practical device.

3.1 Physical Design and Operation

Our proposed physical design in essentially a set of stacked transparent displays placed directly in front of the eyes, closer than the typical eye accommodation distance, and worn as eyeglasses. To create a focused augmented image, we use multilayer optimization techniques (see Sections 3.2, 3.3) rather than conventional optical components such as lenses and beam splitters. To provide a see-through capability, the displays are simply set to a transparent state, or are programmed to selectively occlude parts of the environment.

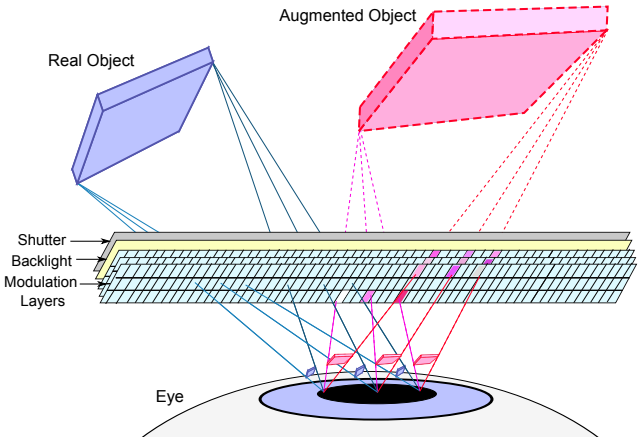

Figure 2: Components and ray formation. The display consists of a shutter and a backlit set of modulation layers. When the shutter is closed, each augmented image ray is reproduced as a white light ray from the backlight that is attenuated across the layers to produce the desired color. The final ray color is the product of the intensities of the intersected pixels across all layers. Multiple regions over the pupil are considered when setting the pixel intensities so that a distant focused image can be formed using layers placed very near the eye. When the shutter is open, light rays from real objects may reach the eye or may be selectively occluded by the layers.

Components The proposed design relies on the following components:

1. Two or more thin, high-speed *transmissive* spatial light modulators that control the intensity of passing light through attenuation, e.g. transparent LCDs.
2. A thin and transparent backlight that uniformly distributes light over its face and can be rapidly modulated, e.g. an edge-lit waveguide or transparent OLED lighting panel.
3. A thin high-speed shutter that can be switched from a globally transparent to a globally opaque state, e.g. a single cell liquid crystal shutter.

All components are sandwiched together with a small spacing between the spatial light modulator layers, as illustrated in Figure 2.

Physical Attributes The proposed design supports a compact form factor and wide field of view that approaches ordinary eyeglasses. The thickness of the device is determined by the thickness of the modulation layers and spacing (typically 3 mm to 10 mm) and the thickness of the integrated backlight and shutter (typically $\leq$1 mm each). Performance generally improves as the layer separation is increased, as described in Section 3.3.1. The field of view of the device is limited only by the size of the spatial light modulator layers, the distance between the eye and layers (typically eyeglasses distance of $\sim$15 mm), and the supported viewing angle of the modulation layers. The specific configuration used for our prototype display is discussed in Section 4.

Method of Operation The proposed device operates in two alternating phases, as illustrated in Figure 3. In the *augmented image formation* phase, light from the backlight passes through optimized patterns displayed on the spatial light modulation layers (see Section 3.2) to form an augmented image over the viewer's pupil while light from the real scene is blocked with the shutter. In the *occluded real image formation* phase, the shutter is opened and the backlight is disabled. Light from real objects enters the display and rays are selectively permitted to reach the eye or are blocked by the spatial light modulator layers. Real image light rays are blocked if they coincide with augmented image light rays for augmented objects that are intended to appear opaque. These two phases are rapidly alternated beyond the flicker fusion threshold so that the user perceives their union – a real scene with opaque augmented objects.

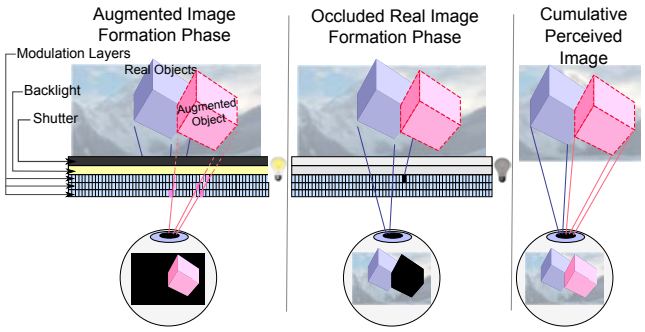

Figure 3: Method of operation. *Left:* In the first phase, an augmented image is formed through the backlight and spatial light modulation layers while the real scene is blocked by the shutter. *Center:* In the second phase, an occluded real image is formed by opening the shutter and blocking light rays from real objects that coincide with augmented objects. *Right:* The two phases alternated in rapid sequence form a perceived AR scene with opaque augmented objects.

Alternate Configuration Alternatively, the augmented image formation and occluded real image phases can occur simultaneously if an additional set of modulation layers are installed behind the backlight (i.e. farthest from the eye) for dedicated occluded real image formation. Augmented images are formed on the existing set of modulation layers, while leaving display regions that do not correspond to augmented images transparent. In this case, the backlight must provide selective light control over the display area (e.g. as with a transparent OLED display) so that only display regions corresponding to augmented objects are illuminated. The alternative approach offers a higher frame rate at the expense of a thicker and likely dimmer display, due to multiplicative light loss through each of the additional layers.

3.2 Multilayer Optimization

Overview In this section, we describe the general approach to multilayer optimization that has been applied to past displays [16]. In Section 3.3, we proceed to discuss innovations which open the approach to near-eye displays.

Each ray emitted by the display to form an augmented image begins as a white light ray in the backlight that is attenuated by each of the spatial light modulator layers as it travels toward the eye (see Figure 2). The final color of the ray entering the eye is the product of the attenuation values assigned to each of the intersected pixels across the layers. Thus, to reproduce a augmented image light ray, the corresponding pixel intensities on each layer should be set so that their product equals the desired ray color.

The multiplicative constraints placed on each pixel from all intersecting augmented image rays must be considered to determine the attenuation values for each layer. For a non-trivial light field, all constraints cannot be satisfied exactly using a practical display; however, additional degrees of freedom can be obtained by time-multiplexing a set of layer patterns such that the sum of their emitted light fields approximates the target light field. This set of patterns is displayed in rapid sequence beyond the flicker fusion threshold to give the impression of a persistent image.

Content-Adaptive Parallax Barriers [16] provides a method for generating an optimized set of time-multiplexed layer patterns for a two layer display which is briefly summarized here. Using a two plane parameterization, the target 4D light field $\mathbf{L}$ to emit from such a display can be expressed as the sum of the T time-multiplexed tensor products of the 2D patterns displayed on a backlit pair of attenuation layers $\mathbf{f}$ and $\mathbf{g}$:

$$\mathbf{L}[i,j,k,l] = \sum_{t=1}^{T} \mathbf{f}_t[i,j] \otimes \mathbf{g}_t[k,l] \tag{1}$$

The N pixels of the 2D pattern for each of the T time-multiplexed patterns can be reordered as a $N \times T$ and $T \times N$ matrix for $\mathbf{f}$ and $\mathbf{g}$ respectively. Similarly, the 4D light field $\mathbf{L}$ can be reordered as $N \times N$ matrix to obtain the equivalent matrix product:

$$\mathbf{L} = \mathbf{FG} \tag{2}$$

The optimal time-multiplexed layer patterns $\mathbf{F}$ and $\mathbf{G}$ can then be obtained by solving the following optimization problem:

$$\begin{aligned} &\arg\min_{\mathbf{F},\mathbf{G}} \frac{1}{2} \|\beta\mathbf{L} - \mathbf{FG}\|_{\mathbf{W}}^2, \text{ for } 0 \leq \mathbf{F}, \mathbf{G} \leq 1, \\ &\frac{1}{2} \|\beta\mathbf{L} - \mathbf{FG}\|_{\mathbf{W}}^2 = \sum_{i \in W,L,FG} [\mathbf{W} \circ (\mathbf{L} - \mathbf{FG}) \circ (\mathbf{L} - \mathbf{FG})]_i, \end{aligned} \tag{3}$$

where $\mathbf{W}$ is a binary valued weight matrix that is used to select which emitted rays should be constrained (i.e. those that intersect the pupil), $\circ$ is the Hadamard (element-wise) product, and β is a scaling factor used to trade brightness for image fidelity. Note that $\mathbf{L}$ must be decomposed into $\mathbf{F}$ and $\mathbf{G}$ using non-negative values since elements represent light attenuation values. One such decomposition method is the weighted variant [2] of the iterative update rules by Lee and Seung [18]:

$$\mathbf{F} \leftarrow \mathbf{F} \circ \frac{[(\mathbf{W} \circ \mathbf{L})\mathbf{G}^\mathsf{T}]}{[(\mathbf{W} \circ (\mathbf{FG}))\mathbf{G}^\mathsf{T}]}, \mathbf{G} \leftarrow \mathbf{G} \circ \frac{[\mathbf{F}^\mathsf{T}(\mathbf{W} \circ \mathbf{L})]}{[\mathbf{F}^\mathsf{T}(\mathbf{W} \circ (\mathbf{FG}))]} \tag{4}$$

$\mathbf{F}$ and $\mathbf{G}$ can be initialized with random noise and will converge to a local stationary point (not necessarily the globally optimum) [18]. The result of this procedure is that an approximation of the target light field $\mathbf{L}$ is "compressed" into to a into a set of T time-multiplexed pairs of layer patterns $\mathbf{F}$ and $\mathbf{G}$. Current optimization performance is sub-real-time; optimization of a light field using the implementation of Tensor Displays [26] takes a few minutes under the configurations described in this paper, but faster methods have been proposed [11]. Figure 4 shows example layer patterns and the corresponding reconstructed view.

For additional information, see Tensor Displays [26], which extends this formulation to three or more modulation layers and nonuniform backlighting and provides analysis on the spatial and angular resolution limits for such displays.

3.3 Multilayer Optimization for Near-Eye Displays

In this section we discuss enhanced multilayer optimization techniques that open the approach to near eye displays.

3.3.1 Ray and Light Field Constraints

Unlike existing multilayer designs, the primary objective of the proposed display is to produce a distant and focused augmented image using display layers placed closer than the typical eye accommodation distance. To meet this objective, the display should ideally reproduce the same set of rays that would be emitted from an augmented object as if it were physically present at its apparent location in the scene (see Figure 2). To produce imagery that appears at focal depths other than the display layers, the eye cannot be treated as a single point; ray variation over the pupil must be considered.

Reproducing a light field with sufficient angular resolution to allow ray variation over the pupil is generally difficult for multilayer (and other) displays. Focus 3D [20] demonstrates that a set of modulation layers alone provides insufficient angular resolution for ray variation over the pupil for a typical desktop display, requiring the addition of a high angular resolution backlight. However, achieving a high angular resolution in the present work is less challenging

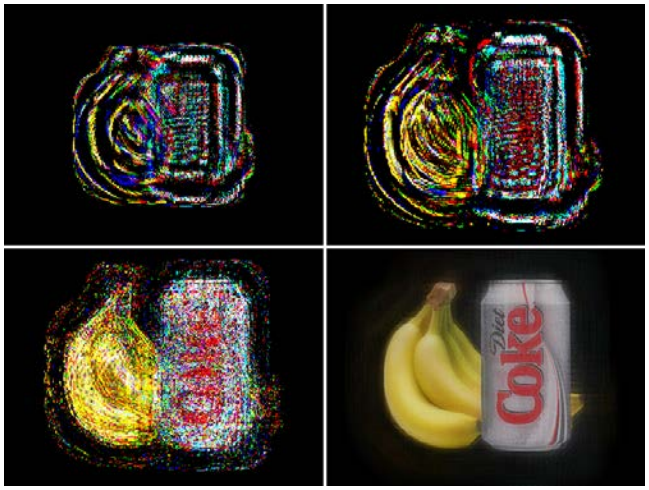

Figure 4: Example multilayer optimization. *Top row:* one of a set of 8 time-multiplexed patterns to display on front and rear modulation layers. *Bottom left:* Simulated reconstructed view from one constrained point on the pupil, summed over the time-multiplexed patterns. *Bottom right:* Simulated accumulation of the 121 reconstructed view zones over the pupil area: the image perceived by the user if the eye is focused at the intended focal depth of the image. Section 3.3 describes how noise is canceled across views. Optimization was performed with a brightness scaling factor $\beta = 0.1$, a focal distance of 15 cm, and the geometry of our prototype display (see Section 4).

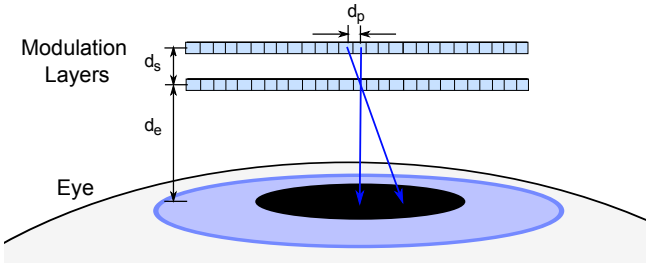

Figure 5: Display geometry to determine angular resolution for two layers. Theoretical maximum angular resolution r at the eye is computed as $\frac{d_s}{d_p d_e}$.

due to the short eye to display distance. In particular, the theoretical maximum angular resolution (in terms of the number of possible distinct rays per unit distance at the eye) for a two layer display is

$$r = \frac{d_s}{d_p d_e} \tag{5}$$

where d_p is the pixel pitch, d_e is the display to eye distance, and d_s is the layer separation distance, as shown in Figure 5. Since angular resolution is proportional to layer separation d_s, display performance will generally improve as the device thickness increases. For a near eye display, the minimum theoretical view spacing becomes a small multiple d_e/d_s of the pixel pitch, which is typically two orders of magnitude smaller than the size of a human pupil when a high density microdisplay is used. (Diffraction, however, will reduce the effective angular resolution as described in Section 3.4.) Due to this high angular resolution, optimization should be performed with a densely placed set of constraints over the pupil area, up to the maximum angular resolution. Unconstrained areas are undefined, degrading the image when illuminated by a uniform backlight.

However, as Focus 3D [20] also observes, the actual performance of a multilayer display using compressive optimization may be sig-

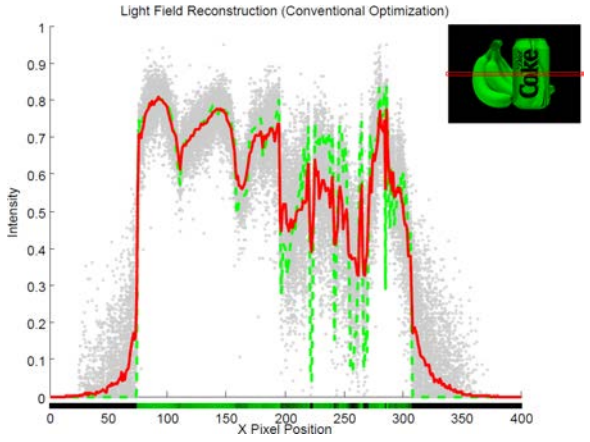

Figure 6: Noise cancellation of the conventional optimization of Section 3.2. Plot provides a 2D slice of a 4D reconstructed "planar" light field: the data from a single scanline (shown below x-axis and in top right inset image) over all of the constrained view zones on the pupil. The x-axis corresponds to the pixel position across the scanline and the y-axis represents the pixel intensity. At each x-value, the set of gray dots provides the intensity values over the set of all views. The data from all view zones should ideally match as the light field represents a diffuse plane. Note gray values are very noisy compared to the intended signal (green dashed line). However, when the views are averaged over the pupil (solid red line), the reconstructed signal more closely matches the intended signal.

nificantly less that the theoretical limits defined by the display's geometry, limiting sharpness and depth of field. Our empirical testing echoed this result; when compressing general light fields we found performance to be exceptionally poor; extreme noise in reconstructed views (see Figure 4, bottom left) overwhelmed angular variation between closely spaced views, causing images formed on the retina to be very blurry with no depth of field.

Our first key insight to improve image quality for near-eye display configurations was to convert input light fields to those without local angular variation, i.e. to "planarize" the light field into a set of diffuse billboards at varying depths. With this restriction, the reconstructed views are very noisy as before; however, each represents a noisy version of the *same* image. Therefore, when all views over the pupil are summed (i.e. light is collected by the lens of the eye), much of the noise *cancels*, significantly improving image quality.

This process is demonstrated quantitatively in Figure 6. The plot shows that individually constrained viewpoints over the pupil (gray dots), are very noisy, typically showing a variation of approximately $\pm 10\%$ or more of the entire intensity range when compared to the intended signal (green dashed line). However, when the views are summed over the pupil area (solid red line), much of the noise cancels and the result is much closer to the intended signal.

Of course, eliminating local angular variation in the input light field comes at a cost: objects grouped in the same plane will have the same focal depth, and thus the focal depth cues (accommodation and retinal blur) will be lost *within* each grouping. However, as shown in Figure 7, it is possible to preserve these focal cues *between* groups; that is, to create multiple planar regions over the display that each appear at their own focal depth. We believe the ability to create multiple focal depths will reduce the eye fatigue caused by the accommodation-convergence conflict, and will increase the sense of presence of augmented objects, as they will have consistency between depth cues that more closely resembles natural objects. There is no performance penalty for creating multiple planar regions at different depths if the regions share no ray constraints (i.e. they do not overlap over the circle of confusion of the pupil area), a common case for sparse augmented overlays. We leave the analysis of multiple overlapping planes as a topic for future work.

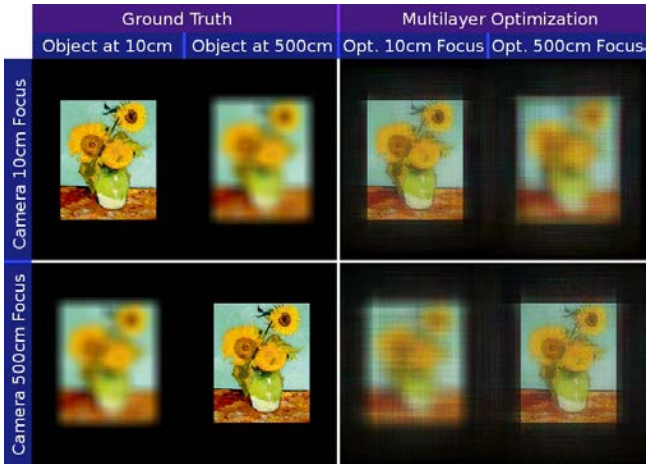

Figure 7: Multiple focal positions. Synthetic objects at different focal depths can be displayed simultaneously on the display. *Top row:* Light field refocused with synthetic camera at close (10 cm) focus. *Bottom row:* Light field refocused with synthetic camera at far (500 cm) focus. *Left column:* Original light field, provided as input to the optimization process. *Right column:* Simulated reconstructed views using multilayer optimization. Optimization was performed with 2 layers, 8 time-multiplexed frames, brightness scaling factor $\beta = 0.1$ and the geometry of our prototype display (see Section 4).

3.3.2 Retinal Optimization

Although restricting light fields to diffuse planes for noise cancellation significantly improves image quality, the result is still substandard (see Figure 9, bottom left). Our second key insight to improve image fidelity for near-eye display configurations was to optimize the *perceived* image rather than attempt to reconstruct the original *light field*. To form an image, rays enter the eye, are refracted by the lens, and strike the retina. For a given eye focal state, the individual intensities of the rays falling on a spot on the retina are not important, but rather their perceived sum; however, preserving individual ray intensities places additional constraints on the optimization. By constraining groups of rays only by their sum of intensities, we expect that there will be more freedom to meet other constraints.

Performing this *retinal optimization* will, however, require knowledge of the focal state of the eye; it is necessary to know the state of the eye's lens to determine where rays will fall on the retina. Rather than attempt to measure this focal state, we instead assume that image quality is most important when an object is in focus. Therefore, we propose to perform optimization as if the eye is simultaneously focused on each object in the scene, improving in focus performance at the expense of out-of-focus blur quality.

To perform retinal optimization, the reconstructed light field matrix $\mathbf{FG}$, computed during each iteration of the rules specified in Equation 4, should be replaced with matrix $\mathbf{R}$, in which the reconstructed intensity of each ray is replaced with the average intensity of the set of rays falling on the same retinal spot. Matrix $\mathbf{R}$ is computed according to:

Algorithm 1 Compute Retinal Reconstruction $\mathbf{R}$

for each $(x,y) \in \boldsymbol{E}$ **do**
  $\boldsymbol{S} = RetinaRays(x,y)$
  $t = \frac{1}{|S|} \sum_{(i,j)\in \boldsymbol{S}} (\mathbf{FG})[i,j]$
  for each $(i,j) \in \boldsymbol{S}$ **do**
    $\mathbf{R}[i,j] = t$
  end for
end for

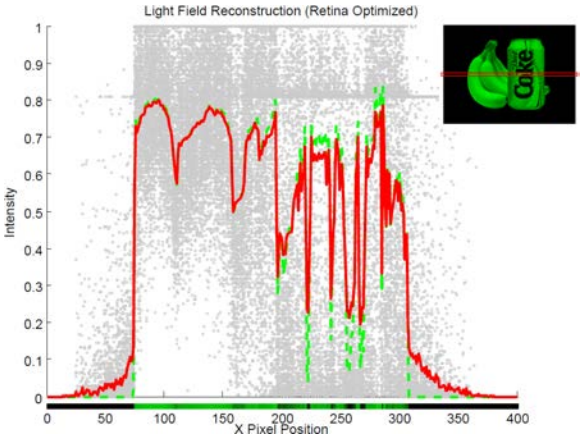

Figure 8: Improved noise cancellation for retinal optimization. Plot provides a 2D slice of a 4D reconstructed "planar" light field (shown below x-axis and in top right inset image) described in Figure 6, except that retinal optimization was performed as described in Section 3.3.2. Note individual reconstructed views (gray dots) have extreme noise compared to intended the signal (green dashed lines); however, the average of views over the pupil (red solid line) more closely matches the intended signal than with conventional optimization (Figure 6).

where $\boldsymbol{E}$ is the set of unique spots on the retina, and $RetinaRays(x,y)$ returns the set of rays that fall on retinal spot (x,y). Note that conventional optimization should be performed before retinal optimization to provide a reasonable initial solution.

Figure 8 provides a quantitative insight into retinal optimization. Note that the individual constrained zones over the retina (gray dots) are extremely noisy – often spanning the entire intensity range; however, when averaged over the retina (solid red line), they more closely match the target signal (green dashed line) then when conventional optimization is performed (see Figure 6). Figure 9 provides a qualitative comparison of conventional and retinal optimization. Note that that when the eye is focused at the depth of the imagery, the retinal optimized image (bottom center) appears much sharper than the conventionally optimized image (bottom left). When the eye is focused elsewhere, the conventional optimization provides a smoother, more accurate retinal blur (top left), but the retinal optimization still provides a reasonable out of focus cue (top center). We believe that a much sharper focused image and the expense of a moderate degradation of focal blur is a very acceptable trade-off.

3.3.3 Perceptual Optimization

Perceptual, rather that least squares, optimization for multilayer displays has been suggested [16, 26] and implemented [21] through the use of complex non-linear objective functions. We propose a simple perceptual optimization scheme that relies on the conventional optimization formulation (see Section 3.2) through the use of the existing ray weighting mechanism. In past work, weight matrix $\mathbf{W}$ (see Equation 4) was *binary*-valued and used to select which of the set of all possible rays emitted by the display should be constrained. We propose the use of a *real*-valued matrix $\mathbf{W}$, in which a zero value indicates an unconstrained ray, and a non-zero value in the range $(0..1]$ indicates the perceptual importance of the ray.

The weight of each ray could be mapped to a variety of factors: a perceptual metric (e.g. contrast with adjacent rays), object importance, or the distance from the center of the field of view. A simple illustrative example of weighting is provided in Figure 10, in which edges were deemed most perceptually valuable. In the perceptually weighted case, object edges (detected with a Sobel filter) were assigned a high weight ($\sim$0.75) while all other rays were assigned a low weight ($\sim$0.25). Note that the perceptually weighted case (center) appears to have higher contrast than the uniformly weighted case (left), at the expense of more noise in low contrast areas.

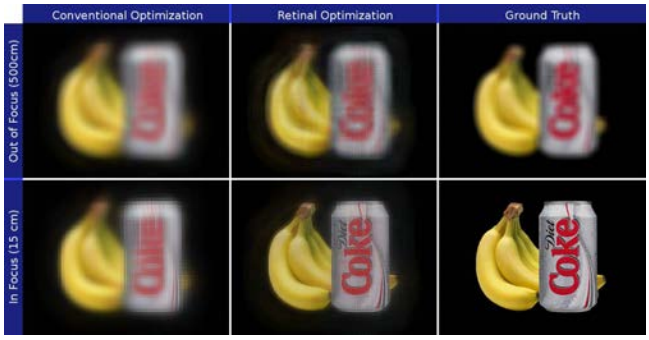

Figure 9: Qualitative comparison of conventional and retinal optimization. *Left column:* Simulated views reconstructed with conventional optimization. *Center column:* Simulated views reconstructed with retinal optimization. *Right column:* Original light field, provided as input to optimization. *Top row:* Light field synthetically refocused at far distance (500 cm). *Bottom row:* Light field refocused at close distance (10 cm). Note improved sharpness of retinal optimized image (bottom center) as compared to conventionally optimized image (bottom left) for in focus case. Optimization was performed with 2 layers, 8 time-multiplexed frames, brightness scaling factor $\beta = 0.1$ and the geometry of our prototype display (see Section 4).

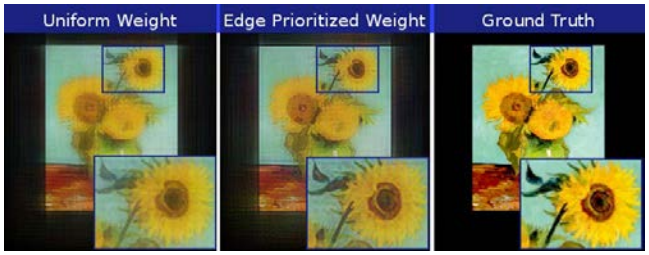

Figure 10: Perceptual Optimization. *Left:* Simulated light field reconstruction using uniform ray weighting. *Center:* Simulated reconstruction using ray weighting based on contrast with adjacent rays. *Right:* Original light field, used as input to the optimization process. Inset images show magnified region of topmost flower. Note improved contrast of green stem and leaves when using edge prioritized weights (center) as compared to use of uniform weights (left). Optimization was performed with 2 layers, 8 time-multiplexed frames, brightness scaling factor $\beta = 0.1$, a focal distance of 15 cm and the geometry of our prototype display (see Section 4).

3.3.4 Occlusion Masks

As described in Section 3.1, the modulation layers can be purposed both to form augmented images and provide an occluded view of the real environment. Occlusion masks are also specified as light fields and are optimized into sets of time-multiplexed layer patterns so that they appear in focus and at the correct focal depth.

Occlusion masks are typically used to block real light rays behind augmented objects so that the augmented objects appear opaque. In this scenario, the occlusion mask light field is a copy of the augmented image light field where object pixels are assigned an intensity of 0 (opaque) and all other pixels are assigned an intensity of 1 (transparent). Object pixels can also be assigned non-zero values for partial transparency. An example occlusion mask is presented in Figure 14E.

3.4 Optical Considerations

In this section we discuss a set of optical considerations that must be addressed to create a practical multilayer near-eye display.

Eyebox A key parameter of a near-eye display is the eyebox, the region over which the eyes can be positioned to see a complete image. In the proposed design, the expected location of the eyes with respect to the display is defined in software and can be adjusted for each user; however, in the present work we have assumed that there is no relative motion between the eyes and the display (i.e. the eyebox matches the eye pupil size). In a practical display, the eyebox should be larger to account for eye movement and shifting of the device on the head. One possibility for increasing the eyebox size is to enlarge the constrained region during optimization, although some image degradation is expected as there are more constraints to satisfy and less noise cancellation will occur for images with retinal optimization. Another possibility is to track the eyes; a promising approach is to place a small camera on the edge of the backlight waveguide that can see the eyes through total internal reflection, as proposed by Travis et al. [24].

Spatial Light Modulator Performance The proposed design places demanding constraints on spatial light modulator performance that are not met by current hardware. As shown in our simulations (see Section 5.1), generally 8 or more time-multiplexed frames are needed for adequate performance, requiring 480 Hz modulators for smooth motion at 60 Hz; however, current LCD panels are limited to 240 Hz rates. High light efficiency is also an important consideration for a see-through design. Unfortunately, much light is lost through the polarizers of stacked LCDs. Finally, to provide high spatial resolution over a wide field of view, ultra high density modulators are needed in moderate sizes (e.g. the size of eyeglasses lenses); however, current ultra high density displays are limited to small microdisplays.

We believe these limitations will be alleviated by upcoming display technologies. A recent spatial light modulator technology, digital microshutters [10], provides a claimed 10 fold light efficiency advantage over LCDs, binary switching rates in the kHz range, and has been demonstrated in a mobile form factor. Samsung also now manufactures LCDs with high transmittance for transparent applications[3]. Mobile-sized, very high density LCD displays are also beginning to become available; Japan Display Inc. has demonstrated a 2.3 inch 651 dpi panel[4]. Microdisplays are also now available with extremely high densities; Silicon Microdisplays sells a 3,000 dpi microdisplay available in commodity products[5].

Diffraction Light that passes through an aperture expands angularly (diffracts) to a degree inversely related to the aperture size. Diffraction is burdensome in the proposed design as the viewer looks through a spatial light modulator that consists of an array of very small pixel apertures, causing degradation of both the real and augmented images, as described below.

During augmented image formation, diffraction limits angular resolution. With the simplifying assumption of round pixel apertures, light diffracts through a pixel aperture to form an Airy disk, the central element of which is bound by the angular extent:

$$\theta_d \approx arcsin\frac{1.22\lambda}{p}, \tag{6}$$

where λ is the wavelength of the light and p is the diameter of the pixel aperture. For 550 nm light and the configuration of our prototype display (a 770 dpi spatial light modulator placed 18 mm from the eyes), the diameter of the central element of the Airy disk becomes 0.73 mm at the user's eyes. This spot size is significantly less than the theoretical angular resolution described in Section 3.3.1, but still is considerably smaller than the size of a human pupil, allowing angular variation over the eye.

During real image formation, light from the real environment will diffract to varying degrees by wavelength when passing through the pixel apertures. When light from a point in the scene is diffracted at the layers, it will no longer remain a point when focused by the lens of the eye, causing blurriness and rainbow effects.

To mitigate diffraction while preserving resolution, spatial light modulators with higher fill factors and field sequential color backlighting (rather than smaller color subpixels) may be used. Spatial light modulations with structures optimized to minimize diffraction have been proposed [5], as have devices with nonuniform pixel distributions to avoid diffraction peaks [1].

4 Implementation

This section describes the characteristics of our prototype display, driving hardware, and optimization software.

Hardware As shown in Figures 1 and 11, our prototype display is constructed from two small stacked spatial light modulators (SLMs) and a backlight (per eye) and a 3D printed eyeglasses frame. Monochrome SLMs were used in conjunction with a color sequential RGB backlight, allowing for larger physical pixels at the same effective resolution to reduce diffraction. LCD spatial light modulators were obtained from a Epson Powerlite 5000 3LCD projector and have a 26.4 mm $\times$ 19.8 mm active area, an 800$\times$600 resolution and an 85 Hz refresh rate. To prevent a set of crossed polarizers from blocking light through the display, the integrated rear polarizer was removed from the rear LCD, and polarizers were installed at the front and rear of the unit so that there was a alternating set of linear polarizers surrounding each LCD. The LCDs were stacked 8 mm apart center-to-center with spacers in between. To construct the backlight, five RGB LEDs (for field sequential color) were installed on an edge lit waveguide that was extracted from the backlight of a larger LCD panel and cut down to size. The LCD and backlight assemblies were installed in a custom eyeglass frame, designed in Google SketchUp and printed on a Rostock MAX 3D printer. The LCDs were driven by the original projector hardware (housed externally), and the backlight was connected to an external power supply. An Intel Core i7 PC with an NVIDIA GeForce GTX 580 GPU running Ubuntu Linux was used both to generate the display patterns and drive the prototype display over a VGA link.

Note that for convenience we did not install a shutter on the rear of the unit to switch between the augmented and real image formation phases (see Section 3.1), but rather placed a opaque sheet behind the display when photographing the display in the augmented image phase (see Section 5.2). A liquid crystal shutter from a pair of LCD shutter glasses would be a likely candidate to be installed in a future device, and would add approximately 1 mm thickness and some additional light attenuation.

The total thickness of the prototype device is approximately 14 mm: 9 mm for the LCDs and spacing, 1 mm for the backlight, and 4 mm of LCD housing that could be removed on a production device. With an eye to display distance of 18 mm (used for testing in Section 5.2), the display provides a 65° diagonal field of view.

Software All light fields used as input for layer optimization for simulations or prototype display consisted of one or more diffuse planes located at a distance of 10 cm to 500 cm from the eye, as indicated in each figure. A light field resolution of 800$\times$600 was used for results shown on the prototype display and a resolution of 400$\times$300 was used for the simulations. The optimization was constrained at 11$\times$11 equally spaced zones over a 4 mm pupil. Unless otherwise indicated, all results use 8 time-multiplexed frames, a brightness scaling factor of $\beta = 0.1$ (see Section 3.2) and the geometry of our prototype display described above. All results also use an eye to display distance of 18 mm, which corresponds to the distance between the display and the camera's center of projection when the display was tested (see Section 5.2).

Optimization was performed using the GPU-accelerated nonnegative tensor factorization routines of Tensor Displays [26], modified

[3] http://www.samsung.com/us/business/commercial-display-solutions/LH22NLBVLVC/ZA

[4] http://www.j-display.com/english/news/2012/20120604.html

[5] http://www.siliconmicrodisplay.com/st1080.html

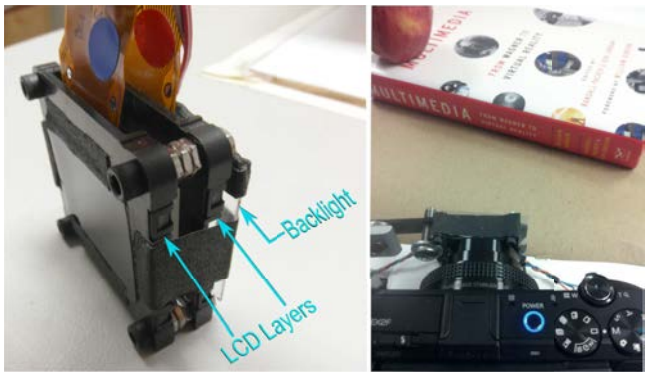

Figure 11: Prototype hardware and testing configuration. *Left:* Prototype display unit for one eye, consisting of two stacked LCD microdisplays and an RGB backlight. *Right:* Testing configuration for experiments described in Section 5.2. The camera setup was configured to approximate a human wearing eyeglasses: the camera lens was placed nearly in contact with the display so that the display to camera center of projection distance was 18 mm, and a 3.7 mm aperture was used.

to accommodate near-eye displays as described in Section 3.3. Using the parameters described above, simulation results were performed at a rate of approximately 100 iterations per minute; acceptable results were reached at approximately 150 iterations and near convergence at approximately 500 iterations. Recent work in adaptive display optimization has shown the possibility of increasing optimization performance significantly [11].

5 Experimental Assessment

In this section we provide an assessment of our proposed design in simulation and on the prototype display described in Section 4.

5.1 Simulation

Configuration The input light field of the simulations is a diffuse plane located at a distance of 15 cm from the eye, which corresponds to the augmented overlay tested on the prototype display (see Section 5.2). The simulations were performed using the geometry of the prototype display (see Section 4), with any additional layers beyond two being equally spaced within the same total layer separation distance (8 mm). Simulations were performed across two key display parameters: the number of display layers, and the number of time multiplexed frames.

Simulated images were generated by reconstructing the light field from the 11×11 constrained viewing positions while summing over the set of time-multiplexed frames. These reconstructed views were then averaged to produce the perceived image when the eye is focused at the apparent depth of the imagery (15 cm).

Qualitative Results Figure 12 shows a qualitative assessment of performance over various display configurations. We observe that a recognizable image can be formed even in a very modest configuration (2 display layers, 2 time-multiplexed frames), although perceived sharpness and contrast is low and there is a glow extending outside the image. As the number of layers and frames increase, the perceived sharpness and contrast improve, and the glowing recedes. In a more advanced configuration (3 layers, 12 frames), image quality is quite acceptable; however, there is still a noticeable loss of contrast as compared to the ground truth image.

Quantitative Results Figure 13 shows a quantitative measure of performance, expressed in peak signal-to-noise ratio (PSNR), over various display configurations. The PSNR calculations exclude the black background area (see Figure 12, bottom left) to prevent overinflating the PSNR in unconstrained regions and overpenalizing the PNSR in the glowing areas that extend beyond the

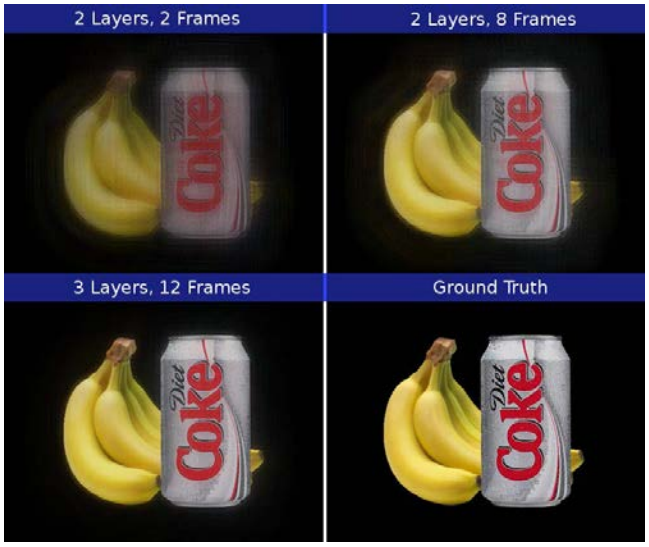

Figure 12: Qualitative performance measure across parameter space. Each image is a simulated view reconstruction with a varying number of display layers and time-multiplexed frames, as indicated above each image. All optimizations used brightness scaling factor $\beta = 0.1$, a focal distance of 15 cm, and the geometry of our prototype display (see Section 4), with any additional display layers over two equally spaced to fit within the original layer separation distance.

image (since the true background when displayed is not black, but rather the real scene behind the display). The plot shows that the proposed approach is clearly lossy; however, with sufficient layers and frames performance on the order of lossy video compression ($\geq$30 db) can be achieved.

5.2 Prototype

Configuration To test our prototype (described in Section 4), we placed a camera behind one of the backlit LCD units and took photographs through the display (see Figure 11, right). To determine the relative position and orientation of the camera and the display, calibration patterns were displayed on the layers and extrinsic parameters were computed using OpenCV. The camera center of projection to display distance was determined to be 18 mm when the lens was placed nearly in contact with the display, approximating the typical distance between eyes and eyeglasses. The

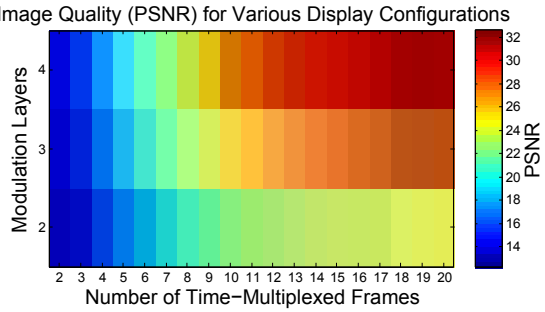

Figure 13: Quantitative performance measure across parameter space. Plot shows reconstruction accuracy, expressed as peak signal-to-noise ratio (PSNR), compared to a reconstruction of the original light field used as input to the optimization process (Figure 12, lower right). All optimizations used brightness scaling factor $\beta = 0.1$ and the geometry of our prototype display (see Section 4), with any additional display layers over two equally spaced to fit within the original layer separation distance.

camera was set to its maximum aperture of 3.7 mm, within the range of human pupil diameters. To produce color images with our monochrome LCDs, we took three color sequential exposures – each with the backlight set to one of the RGB primaries and the corresponding channel displayed on the LCDs.

We made several accommodations for the limitations of our LCD panels. To simulate a faster display that would allow more time-multiplexed frames within the flicker fusion threshold, we set the camera to the least sensitive setting (ISO 80) and used a long exposure of $1/4$ sec. Additionally, since our older panels have a slow response time, we displayed each pattern for two display frames to give the liquid crystals more time to stabilize. Another shortcoming of the LCDs is their poor viewing angle – a problem when viewed at extremely close range – which caused pixels to appear increasing dark toward the bottom left corners of the LCDs. To compensate, use used a higher brightness scaling factor of $\beta = 0.4$, which shifted displayed colors toward white at the expense of compression performance. We believe these problems will be ameliorated with the use of modern projector LCD panels, which have faster refresh rates, lower response times, and likely wider viewing angles through the use of vertically aligned (VA) liquid crystals.

Augmented Reality Scene Results Our test first measures the ability of our prototype to display an optical see-through augmented reality scene. An augmented image (see Figure 12, bottom right) was optimized to appear at a distance of 15 cm from the camera, in front of a real apple and book placed on a desk (see Figure 11, right). The camera was set for focus at 15 cm, which corresponds roughly to the depth of the spine of the book. The scene was shot in two combined sets of RGB exposures – one of the augmented image and one of the occluded real scene – to simulate the time-multiplexed operation described in Section 3.1. To keep the combined result within $0-255$ intensity range, the exposures E_1, E_2 were combined with the following formula:

$$I = 255 - \frac{(255-E_1)(255-E_2)}{255} \tag{7}$$

The results are shown in Figure 14, which demonstrates the ability of our prototype display to meet three key objectives. First, the display is able to create a reasonably focused image using layers placed closer than the camera focus setting. Note that the optimized image (B) is much clearer than an image naively displayed on the layer farthest from the camera (A) when the camera is focused into the scene. Second, the display exhibits occlusion of real objects (F); the augmented banana and soda can appear solid and in front of the real apple and book. Finally, the synthesis of the augmented scene (F) shows the ability to time-multiplex an occluded real scene and augmented image using a shared set of display layers.

Despite these promising results, the prototype also shows several shortcomings. The augmented image (B), although recognizable, is not nearly as sharp as the simulated result (see Figure 12, top right). The generated occlusion mask (E) has blurry edges, causing the augmented image to appear surrounded by a black halo (F). Finally, the view through the display (C) is somewhat blurry due to diffraction through the LCDs and the partial transparency of the backlight. We believe that we can improve prototype results by utilizing LCD panels with wider viewing angles and replacing the backlight with a completely transparent OLED lighting panel.

Multiple Focus Results Our second test measures the ability of our prototype to display objects at multiple focal depths simultaneously. The left half of the display shows a pattern optimized for a 10 cm focal depth, while simultaneously the right half shows the same pattern optimized for a 500 cm focal depth. Two sets of RGB exposures were taken which varied only by the camera's focus setting, which was adjusted between 10 cm and 500 cm.

The results, shown in Figure 15, demonstrate the ability of the prototype to display images at varying focal depths simultaneously: each pattern is in reasonable focus when the camera is focused at the intended focal depth, and out of focus when the camera is focused elsewhere. We find these results promising; however, in the future we would like to demonstrate this ability with overlapping objects and more challenging high spatial frequency patterns.

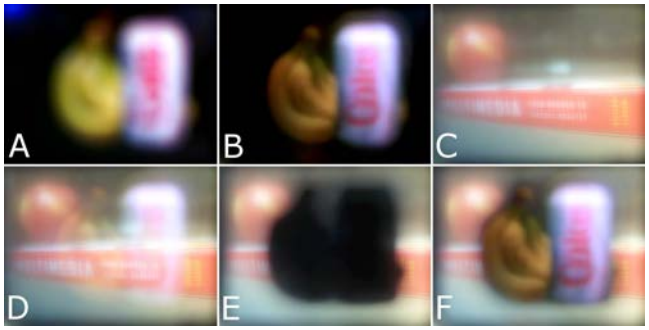

Figure 14: Augmented reality scene on prototype display. *A)* Photograph of prototype showing image naively displayed on farthest layer. *B)* Photo of prototype displaying a series of optimized time-multiplexed patterns. *C)* Photo of real environment taken through display with backlight off. *D)* photographs (B) and (C) combined to simulate an augmented display without occlusion support. *E)* Photo of real environment with an optimized occlusion mask displayed on device. *F)* photos (B) and (E) combined to simulate an augmented scene with occlusion support. All photos were taken at ISO 80 for $1/4$ sec with focus set at approximately 15 cm. Note that color augmented photos were taken as three exposures, one with the RGB backlight and display layers set for each primary color channel, and combined.

Comparison to Simulations We observe that the quality of the simulated results is significantly better than those achieved on our prototype display and believe that this discrepancy could be reduced through more accurate modeling of the display in the optimization and simulation software. Since the number of constrained viewing zones (11×11, limited by GPU memory) in less than the theoretical maximum angular resolution, optimization may benefit from an even denser set of view constraints. Recent work in display optimization [11] provides a time and memory efficient optimization method for very dense sets of view constraints. Simulation and optimization may also benefit from a better model of the spatial light modulators that considers off axis viewing performance, diffraction, and the varying display intensities across layers. Finally, our hardware prototype could benefit from improved layer calibration; we found camera-based layer calibration challenging due to the fine pixel pitch of the utilized spatial light modulators and the close proximity of the camera to the display (yielding slightly unfocused images at the nearest focal setting of the lens).

6 Conclusion and Future Work

Inspired by recent work in desktop 3D displays, we have proposed a novel design for optical see-through head-worn displays that eschews traditional optical components for optimized stacks of spatial light modulators placed directly in front of the eyes. We believe this design will support a wide range of augmented reality applications through the fulfillment of four key requirements: an *encumbrance-free form factor*, a *wide field of view*, *occlusion support* and preservation of the *focal depth cues*.

Our early prototype display has demonstrated preliminary capability to display augmented overlays at multiple focal depths with occlusion support in a compact wide field of view design. However, the current device does not rival the image quality of modern head-worn displays and several limitations must be addressed to enable practical use. In subsequent work, we plan to study hardware characteristics to bridge the gap between real-world and simulated results, reduce optimization time to enable real-time use, address human factors such as eyebox size, and mitigate diffraction effects.

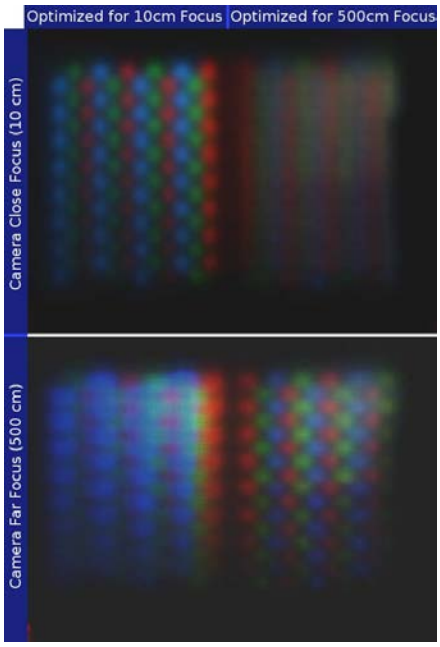

Figure 15: Multifocal scene on prototype display. An optimized set of patterns was generated to simultaneously display a tiled RGB pattern at two focal depths, 10 cm (*left* half of photographs) and 500 (cm) (*right* half of photographs). *Top:* Photograph with camera focus set to 10 cm. *Bottom:* Photograph with camera focus set to 500 cm.

We are excited at the prospect of future displays that are as convenient as eyeglasses, but enable the placement of high fidelity augmented overlays anywhere in the field of view at any place and time, narrowing the gap between computer graphics and human vision.

Acknowledgements

The authors thank Gordon Wetzstein for providing the tensor factorization code used as the basis for our optimizer and Xubo Yang for helpful suggestions and assistance calibrating an earlier prototype. This research was supported in part by the BeingThere Centre, a collaboration between UNC Chapel Hill, ETH Zurich, and NTU Singapore, supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the Interactive Digital Media Programme Office.

References

[1] 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*, 18(14):15223–15235, 2010.

[2] V. D. Blondel, N.-D. Ho, and P. van Dooren. Weighted nonnegative matrix factorization and face feature extraction. *Image and Vision Computing*, 2008.

[3] O. Cakmakci, Y. Ha, and J. P. Rolland. A compact optical see-through head-worn display with occlusion support. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 16–25, 2004.

[4] O. Cakmakci, K. Thompson, P. Vallee, J. Cote, and J. P. Rolland. Design of a free-form single-element head-worn display. *Proc. SPIE 7618, Emerging Liquid Crystal Technologies V*, pages 761803–761803–6, 2010.

[5] H.-C. Chaing, T.-Y. Ho, and C.-R. Sheu. Structure for reducing the diffraction effect in periodic electrode arrangements and liquid crystal device including the same. US Patent, 2005. US 6977705.

[6] D. Cheng, Y. Wang, H. Hua, and J. Sasian. Design of a wide-angle, lightweight head-mounted display using free-form optics tiling. *Opt. Lett.*, 36(11):2098–2100, Jun 2011.

[7] J. De Smet, A. Avci, R. Beernaert, D. Cuypers, and H. De Smet. Design and wrinkling behavior of a contact lens with an integrated liquid crystal light modulator. *Display Technology, Journal of*, 8(5):299–305, 2012.

[8] C. Gao, Y. Lin, and H. Hua. Occlusion capable optical see-through head-mounted display using freeform optics. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 281–282, 2012.

[9] H. Gotoda. A multilayer liquid crystal display for autostereoscopic 3d viewing. *In SPIE Stereoscopic Displays and Applications XXI*, 7524:1–8, 2010.

[10] N. Hagood, R. Barton, T. Brosnihan, J. Fijol, J. Gandhi, M. Halfman, R. Payne, and J. L. Steyn. 35.5l: Late-news paper: A direct-view mems display for mobile applications. *SID Symposium Digest of Technical Papers*, 38(1):1278–1281, 2007.

[11] F. Heide, G. Wetzstein, R. Raskar, and W. Heidrich. Adaptive Image Synthesis for Compressive Displays. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 32(4):1–11, 2013.

[12] X. Hu and H. Hua. Design of an optical see-through multi-focal-plane stereoscopic 3d display using freeform prisms. In *Frontiers in Optics Conference*, page FTh1F.2. Optical Society of America, 2012.

[13] Innovega. *Eyewear Industry Presentation*, 2012. http://innovega-inc.com/events.php.

[14] I. Kasai, Y. Tanijiri, T. Endo, and H. Ueda. A practical see-through head mounted display using a holographic optical element. *Optical Review*, 8:241–244, 2001.

[15] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 133–, 2003.

[16] D. Lanman, M. Hirsch, Y. Kim, and R. Raskar. Content-adaptive parallax barriers: optimizing dual-layer 3d displays using low-rank light field factorization. *ACM Trans. Graph.*, 29(6):163:1–163:10, 2010.

[17] D. Lanman, G. Wetzstein, M. Hirsch, W. Heidrich, and R. Raskar. Polarization fields: Dynamic light field display using multi-layer LCDs. *ACM Trans. Graph.*, 30(6), 2011.

[18] D. D. Lee and H. S. Seung. Learning the parts of objects by nonnegative matrix factorization. *Nature*, 401:788–791, 1999.

[19] A. R. Lingley, M. Ali, Y. Liao, R. Mirjalili, M. Klonner, M. Sopanen, S. Suihkonen, T. Shen, B. P. Otis, H. Lipsanen, and B. A. Parviz. A single-pixel wireless contact lens display. *Journal of Micromechanics and Microengineering*, 21(12):125014, 2011.

[20] A. Maimone, G. Wetzstein, M. Hirsch, D. Lanman, R. Raskar, and H. Fuchs. Focus 3d: Compressive accommodation display. *To appear in:* ACM Transactions on Graphics.

[21] B. Masia, G. Wetzstein, C. Aliaga, R. Raskar, and D. Gutierrez. Perceptually-optimized content remapping for automultiscopic displays. In *SIGGRAPH 2012 Posters*, SIGGRAPH '12, page 63:1, 2012.

[22] J. P. McGuire, Jr. Next-generation head-mounted display. *Proc. SPIE 7618, Emerging Liquid Crystal Technologies V*, pages 761804–761804–8, 2010.

[23] P. Santos, T. Gierlinger, O. Machui, and A. Stork. The daylight blocking optical stereo see-through hmd. In *Workshop on Immersive projection technologies/Emerging display technologiges*, IPT/EDT '08, pages 4:1–4:4, 2008.

[24] A. R. L. Travis, T. A. Large, N. Emerton, and S. N. Bathiche. Wedge optics in flat panel displays. *Proc. of the IEEE*, PP(99):1 –16, 2012.

[25] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4), 2011.

[26] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):1–11, 2012.