Holographic Near-Eye Displays for Virtual and Augmented Reality

ANDREW MAIMONE, Microsoft Research
ANDREAS GEORGIOU, Microsoft Research
JOEL S. KOLLIN, Microsoft Research

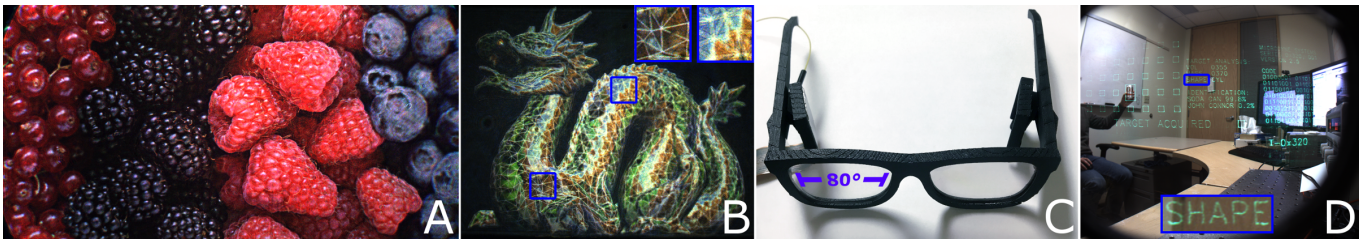

Fig. 1. *A)* Photograph of prototype holographic near-eye display presenting a full color image. *B)* Photograph of prototype displaying a true 3D hologram; each point of the 3D model appears at an individually addressable focal depth. Blue bordered inset regions show in-focus chest and out-of-focus tail. *C)* Photograph of prototype display in an eyeglasses-like form factor with a horizontal field of view of 80° (driving electronics external). *D)* Photograph of prototype in *C)* displaying an augmented image. Blue bordered inset region shows magnification of 1 pixel line width text. Berries image by Ana Blazic Pavlovic/Shutterstock. Dragon model by Stanford Computer Graphics Laboratory.

We present novel designs for virtual and augmented reality near-eye displays based on phase-only holographic projection. Our approach is built on the principles of Fresnel holography and double phase amplitude encoding with additional hardware, phase correction factors, and spatial light modulator encodings to achieve full color, high contrast and low noise holograms with high resolution and true per-pixel focal control. We provide a GPU-accelerated implementation of all holographic computation that integrates with the standard graphics pipeline and enables real-time (≥90 Hz) calculation directly or through eye tracked approximations. A unified focus, aberration correction, and vision correction model, along with a user calibration process, accounts for any optical defects between the light source and retina. We use this optical correction ability not only to fix minor aberrations but to enable truly compact, eyeglasses-like displays with wide fields of view (80°) that would be inaccessible through conventional means. All functionality is evaluated across a series of hardware prototypes; we discuss remaining challenges to incorporate all features into a single device.

CCS Concepts: • **Hardware** → *Displays and imagers*;

Additional Key Words and Phrases: holography, near-eye display, augmented reality, virtual reality, computational displays

ACM Reference format:
Andrew Maimone, Andreas Georgiou, and Joel S. Kollin. 2017. Holographic Near-Eye Displays for Virtual and Augmented Reality. *ACM Trans. Graph.* 36, 4, Article 85 (July 2017), 16 pages.
DOI: 10.1145/3072959.3073624

1 INTRODUCTION

It's an exciting time for near-eye displays. In the previous year, a new crop head mounted displays (Oculus Rift and HTC VIVE) have demonstrated sufficient resolution, tracking performance, and latency to begin to offer compelling virtual reality (VR) experiences to consumers. In the same year, the Microsoft HoloLens has demonstrated these same characteristics for the first time in a self-contained augmented reality (AR) device.

There's still a lot of work to be done. The lightweight optics found in current virtual reality near-eye displays do not scale to provide resolution on par with human visual acuity. Achieving such performance with conventional optics would require complex, multi-element optics with impractical size, cost and weight. Custom prescription lenses are required to correct visual impairments. Likewise, current optical solutions do not provide a path to obtain wide field of view (FOV) and high resolution see-through augmented reality displays in eyeglasses-like form factors that would enable true mobility and all-day use. We have also not encountered practical solutions to incorporate true 3D, multi-focal depth cues in these devices to eliminate the accommodation-convergence conflict and mimic the fidelity of our natural vision.

Solutions to these challenges require precise control of the wavefront of light. It would be very difficult to overcome all of these challenges in a passive optical device. We turn to *computational* solutions, in which the hardware is simplified and the complexity of wavefront control is pushed into software. One potential solution is *light field* displays, in which wavefront control is expressed through individually addressable ray bundles. Some impressive results have been achieved to date [Huang et al. 2015; Lanman and Luebke 2013]. However, light field displays have a few limitations. Ray bundles inefficiently encode wavefronts, allowing only coarse approximation of wavefront shape through sparse samples and causing high resolution loss. Due to diffraction, the selected size of the ray bundles also dictates a fundamental trade-off between the maximum possible spatial and angular resolution, limiting focal control in a high resolution device. We thus turn to *digital holography*, offering very efficient wavefront encoding by holographic superposition, smooth wavefront control, and no such resolution trade-off.

Authors' email: {amaimone,angeor,jkollin}@microsoft.com.

The benefits of digital holography have long been explored through decades of research, demonstrating powerful features such as variable focal control and optical aberration correction. Despite these capabilities, holography is often associated with noisy, low contrast, and mono color imagery, large benchtop form factors, high bandwidth requirements, and expensive computation – relegating it to the status of a perpetually "future" technology. However, this reputation is largely undeserved as many of these problems have already been addressed.

In this paper, we build on much of this existing work to describe how digital holography can be used to solve many of the key issues facing virtual and augmented reality displays. We describe a combination of algorithms that can be used to display high quality, full color phase-only holograms with true per-pixel focal control that exceed the capabilities of state-of-the-art displays. We show that these true 3D holograms can be generated directly from the output of the standard graphics pipeline as an independent post-processing step. We extend the known optical correction capabilities of holograms to not only correct aberrations but to allow vision correction capabilities and enable previously inaccessible eyeglasses-like form factors with wide (80°) fields of view. All holographic algorithms run at real-time rates (≥90 Hz) directly or through an approximate eye-tracked solution. We emphasize real-world performance by presenting all capabilities across a series of preliminary virtual and augmented reality display prototypes. We discuss remaining challenges to integrate all features into a single device.

1.1 Contributions

We provide optical and algorithmic designs for holographic virtual and augmented reality displays that are evaluated in simulations and a series of hardware prototypes. Specific contributions include:

(1) Variation of the double phase amplitude control technique for multi-focal Fresnel holograms that includes hardware design, SLM encoding, and phase adjustment factors
(2) Extension of the Zernike polynomial aberration correction technique that unifies aberration correction, vision correction, and per-pixel depth and provides for user calibration
(3) Real-time computation method based on linearly separable convolution and real-time approximations of spatially variant focus and aberration correction for eye tracked displays, with a framework for integration into the graphics pipeline
(4) Optical design for a lensless, compact, wide field of view optical see-through holographic display using an off-axis projector and off-axis holographic optical element eyepiece

1.2 Benefits and Limitations

The proposed holographic displays provide full color, high resolution, and high quality imagery and are capable of true per-pixel focal control and user vision correction. We provide optical design variations capable of large fields of view with very compact form factors. We provide a software implementation that is capable of real-time computation (either directly or through eye tracked approximations) and integration with the rendering pipeline.

However, there are several limitations of the presented results. We show various capabilities of near-eye holographic displays (wide FOV, compact form factors, multi-focus, etc.) but we have not yet achieved all these capabilities in a single device. We provide evaluation with simple tethered monoscopic prototypes with small exit pupils. A practical stereo display would require integration of a pupil expansion or steering device, which we discuss in Section 5.3. Some of the proposed techniques require eye tracking, which we simulate but we do not integrate an actual tracking device. We also do not consider system level AR/VR concerns such as head tracking and latency. Other minor limitations are listed in Section 5.3.

2 RELATED WORK

Our display designs are related to a wide body of display and holography research which we summarize below.

Holographic Near-Eye Displays. Several recent near-eye displays combine a holographic projector with various see-through eyepieces: holographic optical elements [Li et al. 2016], waveguides [Yeom et al. 2015], and lenses with beamsplitters [Chen and Chu 2015; Gao et al. 2016; Moon et al. 2014]. Our augmented reality displays share some similar characteristics to these displays; we too employ a holographic projector and demonstrate variable focus. Like the Li et al. [2016] display, we also employ a holographic optical element eyepiece, and like the Yeom et al. [2015] display we use holographic correction to fix optical aberrations. We extend these capabilities by incorporating true per-pixel focal depth, complex full-color shaded imagery, faster real-time calculation, more powerful aberration correction, vision correction, significantly wider fields of view, and more compact form factors.

Other holographic displays and projectors. There is a large body of closely related work in holographic projection that spans several decades; an overview can be found in Yaraş et al. [2010] or Tsang et al. [2016]. We draw on many related works; like Reichelt et al. [2013] and Qi et al. [2016] we use variations of the double phase encoding method [Hsueh and Sawchuk 1978] to obtain amplitude control with phase-only SLMs. However, our method varies slightly as we do not employ off-axis projection with encoded gratings [Qi et al. 2016] or custom phase combining hardware [Reichelt and Leister 2013]. Like Häussler et al. [2009] we also propose a reduced-bandwidth eye tracked display but in a near-eye form factor. Our holographic optical correction model is much like that of a set of related works [Freeman et al. 2010; Kaczorowski et al. 2015, 2016], but we measure aberrations from the perspective of the eye, rather than from a camera or display, include vision correction, provide integration with per-pixel depth control, and propose alternate acceleration routines based on eye tracking. As with Makowski et al. [2012] and Qu et al. [2015], we magnify our projectors using diverging light sources; we also combine this technique with traditional eyepiece magnification for near-eye display.

Near-Eye Light Field Displays. Near-eye light field displays provide some of the same capabilities as our near-eye holographic displays, such as variable focus [Hua and Javidi 2014; Huang et al. 2015], compact form factors [Maimone et al. 2014], or both [Jang et al. 2016; Lanman and Luebke 2013]. Compared to holographic displays, it is generally easier to get good image quality with light field displays due to the use of conventional, incoherent illumination

and the availability of high quality amplitude modulators. The computation required for these displays is also often more lightweight, with the exception of those employing a compressive optimization. The advantage of holographic displays is a very precise ability to control the wavefront of light, allowing fine focus control, aberration correction, and vision correction. Such displays also encode this wavefront control very efficiently through the holographic superposition principle, requiring only an amplitude and phase for each true 3D object point in the scene. Operating using diffraction, holographic displays are capable of providing both high spatial and angular resolution to the viewer. In contrast, light field displays provide coarser wavefront control that is encoded using a number of ray samples and have a diffraction limited trade-off between spatial and angular resolution. However, the use of holographic elements in light field displays [Lee et al. 2016] may help to overcome some of these limitations.

Multi-Focus Displays. Another body of related work creates multiple focus planes through special optical configurations. Akeley et al. [2004] describe a display in which different regions of an LCD panel are mapped to different focal planes through a series of beamsplitters. Liu et al. [2010] use a liquid lens that can be addressed to a specific focal plane or time-multiplexed between multiple planes. Love et al. [2009] provide a similar capability with a lens that can quickly switch between multiple states. Schowengerdt et al. [2010] scan an array of light sources at different depths from a lens to address multiple focal planes. Narain et al. [2015] describe how to generate optimal imagery for such multi-plane displays. Konrad et al. [2016] and Johnson et al. [2016] use focus tunable lenses that are matched to the depth of viewer fixation and also experiment with *monovision*, the use of an additional fixed focus lens in front of a single eye. Padmanaban et al. [2017] extend these capabilities by using tunable lenses that are matched to the gaze and refractive errors of the viewer. See Kramida [2016] for a survey of additional techniques. Like holographic displays, these displays provide adjustable focal cues. The techniques require little or no computation and tend to produce high quality images. However, they are limited in capability to produce a single or fixed number of focal planes and use more complex optical assemblies that are difficult to miniaturize and extend to wide fields of view.

Vision Correcting Displays. Past displays have demonstated the ability to correct the viewer's visual impairments through the use of light fields [Pamplona et al. 2012], or light fields and image processing techniques [Huang et al. 2014]. Inspired by these displays, we provide a similar vision correction capability using holograms. Compared to these methods, holographic displays are able to provide finer control of the wavefront needed for vision correction and offer less loss of resolution, but have a higher computation cost.

3 SYSTEM OVERVIEW

Our core approach to building near-eye displays is to incorporate most of the system complexity into a phase-only projector that is based on Fresnel holographic principles. We design a projector that is capable of precise and spatially variant control of image focus and optical aberrations. We use this control to generate true 3D holograms while providing a vision correction capability. We integrate this projector with an eyepiece that has highly favorable physical properties, but is highly aberrated, and correct all the optical defects in the hologram. We provide details in the following section.

3.1 Holographic Projection

3.1.1 Fresnel Holography. Our holographic projector is based on a variation of the well known principle of Fresnel holography; we provide a brief overview here. In a Fresnel holographic system (see Figure 2), light propagates from a source until encountering a *hologram plane* in which a *spatial light modulator* (SLM) device alters the amplitude and phase of the light. The light is diffracted in such a way that it is focused to one or more *object points* with varying intensities that form an image (or volume). We refer to the specific phase alterations that cause light to be focused to an object point as a *lens phase function*. Like a physical lens, these phase alterations cause light near the center of the hologram to be deflected by a small amount and light near the edge to be deflected by a large amount. The spatial frequency of the phase alterations dictates the degree of deflection; higher spatial frequencies cause greater deflections. If this holographic lens becomes too large, the spatial frequencies of the lens phase function will exceed those supported by the spatial light modulator. Thus we confine the lens phase function representing each object point to a local region and refer to it as a *sub-hologram*. Sub-holograms are sized so that they contain only spatial frequencies supported by the spatial light modulator. Sub-holograms representing different object points may overlap on the hologram plane.

Mathematically, we represent the lens phase function f_o for each point $\vec{p}$ in the sub-hologram representing object point $\vec{o}$ as

$$f(\vec{p})_o = e^{i(\phi_o + \frac{2\pi\|\vec{p}-\vec{o}\|}{\lambda})} \tag{1}$$

where λ is the light wavelength. Essentially f_o represents the phase difference of a light wave between object point $\vec{o}$ and each point $\vec{p}$ in the sub-hologram region. For convenience, f_o is represented as a complex-valued function where the phase angle of the function equals the phase difference. Note that ϕ_o is a phase offset factor that does not affect the appearance of the holographic image here but is useful for optimizing phase-only holograms (see Section 3.1.3). For now, ϕ_o may be set to zero.

To form a complete hologram H, we simply compute the sum of the lens phase functions f_o for each object point $\vec{o}$ while weighting them by the image intensities:

$$H(\vec{p}) = \sum_{j \in s_p} a_j f(\vec{p})_j \tag{2}$$

where s_p is the set of object points whose sub-holograms are defined at point $\vec{p}$ and the desired intensity of object point $\vec{j}$ is a_j^2. Since this method of calculation computes the hologram as the explicit integration of all the object points in the scene, we refer to it as the *point-wise integration method*. In computer graphic terms, computing hologram H can also be thought of as convolution of the target image with a spatially-varying, complex-valued kernel.

3.1.2 Magnification and Etendue. As mentioned in the previous section, the light deflection angles of the lens phase functions f_o

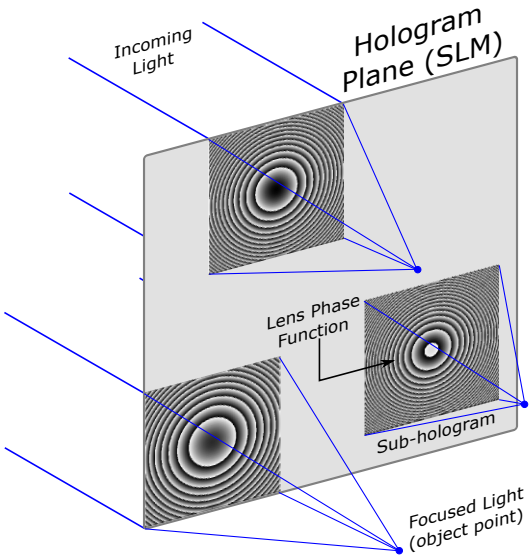

Fig. 2. Fresnel holography. A beam is focused to several object points at finite distances by a hologram. Each point is formed by overlapping regions known as sub-holograms, where each sub-hologram contains a phase-altering function that causes light to diffract inward much like a refractive lens.

in a Fresnel hologram are related to the spatial frequency of the phase alterations. In this way, a hologram acts like a series of superimposed diffraction gratings of various frequencies. The light deflection angle is related to the spatial frequency of a grating by the grating equation:

$$d(sin\theta_i + sin\theta_m) = \lambda \tag{3}$$

where d is the grating pitch, θ_i is the light incidence angle, θ_m is the diffraction angle, and λ is the wavelength. Using a typical microdisplay with pixel pitch of 8 μm (which by the Nyquist sampling theorem is able to create a grating with a minimum pitch of $d = 16\mu$m) that is illuminated with normally-incident $\lambda = 532nm$ green light, a holographic display can only bend light $\theta_m = \pm 1.9°$. Thus a holographic display with today's technology will need to be magnified greatly in order to support the wide fields of view (60-100° or more) expected by VR and AR displays. A holographic display can be magnified by conventional means (e.g. refractive lenses), but the required magnification is immense and thus bulky lenses are required. However, note that by Equation 3 that the supported diffraction angle is *relative* to the incident light angle. Therefore, one can magnify holograms in a *lensless* manner using a *diverging* point light source [Makowski et al. 2012; Qu et al. 2015]. This configuration allows the beam divergence to perform the coarse beam steering while the hologram steers light at a finer angle to focus the light into an image. We update the lens phase function of Equation 1 to handle diverging light as:

$$f_d(\vec{p})_o = e^{i(\phi_o + \frac{2\pi(\|\vec{p}-\vec{l}\|+\|\vec{p}-\vec{o}\|)}{\lambda})} \tag{4}$$

where $\vec{l}$ is the position of the diverging light source. This configuration is illustrated in Figure 3. Note that the field of view of the projector θ_p is significantly larger than the maximum diffraction angle θ_m. In general, the supported field of view is limited only by the divergence of the point light source and can be arbitrary large.

Etendue. Two key attributes of any display are the area over which the display emits light at the plane of the eye (the exit pupil or viewing eye box) and the range of angles that light is emitted over this area (the field of view). The product of these two quantities equals the *etendue* of the display system. When a display system is magnified the etendue is preserved, and thus one must trade off the size of the exit pupil and the field of view. For a conventional near-eye display consisting of a lens placed in front of a display panel, the etendue is equal to the product of the area of the display panel and the light emission angle of each pixel that is collected by the lens. In designs using large display panels and optics (e.g. the Oculus Rift or HTC Vive) the etendue is quite large and is generally not a concern.

The etendue of a holographic projector is defined by the product of the area of the spatial light modulator and the maximum diffraction angle (Equation 3). An unmagnified holographic projector has a field of view equal to twice the diffraction angle and an exit pupil equal to the area of the SLM. As the projector is magnified with diverging light, the field of view increases, but the effective size of the exit pupil decreases (see Figure 3). Note that the sub-hologram size used in computation matches the size of the exit pupil. If one attempted to further increase the size of the sub-hologram, it would require diffraction angles and therefore spatial frequencies higher than the SLM can reproduce. Thus, unlike a conventional display, the etendue of a holographic projector is linked to the device resolution. When the holographic projector is coupled with an eyepiece to form a complete near-eye display (see Section 3.2), the image may be further magnified but the etendue is preserved.

The size of the exit pupil and sub-hologram have several implications in a near-eye holographic display. If the exit pupil is too small, the image may disappear if the eyes move too far from the display, and an additional mechanism will be required to keep the exit pupil centered on the eye. Further, if the instantaneous exit pupil size is smaller than the pupil of the eye, the ability of the display to represent shallow depth of field is limited. We discuss these limitations further in Section 5.3. Finally, if the size of a sub-hologram is less than the size of the SLM, light can only be steered *locally* by phase modulation. This is a problem if the SLM affords only phase modulation and there are large differences in illuminance over the image. We address this issue in the next section. In summary, one must balance field of view, the exit pupil size, and the ability to display holograms with shallow depth of field in a holographic near-eye display.

3.1.3 Obtaining Amplitude. Equations 2 and 4 show that calculating Fresnel holograms is quite simple. However, to display such holograms we require SLMs with both amplitude *and* phase control, which are not generally available. Availability aside, *phase-only* holograms are often preferred because of their high light efficiency: light is only steered, not attenuated. Thus one of the main challenges of building holographic displays is calculating holograms that work with phase-only SLMs. We provide a brief overview of common alternatives below.

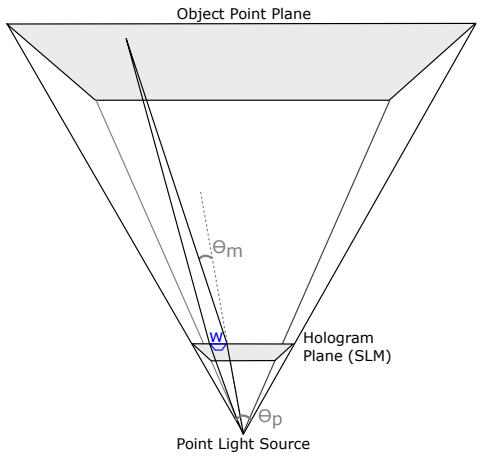

Fig. 3. Hologram magnification through a diverging light source. A holographic spatial light modulator capable of steering light a small angle θ_m can be used to make a holographic projector with a wide field of view θ_p by providing angularly varying incident light from a diverging light source. The sub-hologram size (and exit pupil size) correspondingly reduces from the full width of the spatial light modulator to width w.

Amplitude discard method. A trivial method for computing a phase-only hologram is to compute the full complex hologram and then simply discard the amplitude. Since the phase tends to encode high frequencies (e.g. object edges), this method may produce an acceptable result for drawing simple text and symbology that consist of thin lines. This is not a practical general solution; however, we use this method for performing simple tests of our displays.

Iterative phase optimization. The most common method to compute phase-only holograms is optimization. Observe in Equations 1 and 4 that we have control of the phase ϕ_o at each object point $\vec{o}$. As phase values ϕ_o are changed, the object points will interfere in a different way and produce a different set of required amplitude and phase values for hologram H, resulting in the same image intensity with different phase. Since we generally do not care about the phase of the image (it is not detected by the eyes), we consider values ϕ_o as free variables for hologram optimization. The goal of the optimization is to find the values ϕ_o that cause the object points $\vec{o}$ to interfere such that the required amplitude values of H are constant over the hologram. This method is generally performed by the Gerchberg-Saxton algorithm [Gerchberg and Saxton 1972], in which light is propagated back and forth between the object and hologram planes while enforcing amplitude constraints. Modern variations add error compensation and define "do not care" regions of the hologram to increase performance [Georgiou et al. 2008]. The downside of the iterative approaches is the computational expense of computing the hologram over many iterations (≈10-100). While numerically accurate, iteratively optimized holograms also tend to include undesirable high frequency noise or speckle.

Temporal averaging. Another method to obtain amplitude in a phase-only device is temporal averaging. If object phase values ϕ_o are simply randomized, rather than optimized, the required amplitude on hologram H is a pseudo-random uniform noise pattern. If the amplitude is discarded on playback, the target image is perceived, but is heavily contaminated with noise since the required amplitude is missing. However, if phase values ϕ_o are set to a *different* set of random values, the perceived noise pattern is different. When several such holograms are played back in random succession on a high speed display, the noise tends to average and produce an image closer to the target. To facilitate the high bandwidth and switching speed requirements, *binary* phase holograms may instead be produced for playback on a high speed binary SLM. This approach is known as One Step Phase Retrieval (OSPR) [Cable et al. 2004]. The primary downside of the approach is that the low contrast of the noisy hologram sub-frames cannot be recovered by temporal averaging, and several hologram sub-frames must be calculated for each image frame.

Direct Amplitude Encoding. Another method to generate phase-only holograms is to explicitly encode amplitude and phase into two phase values. This is known as the *double phase* method [Hsueh and Sawchuk 1978]. Observe that the complex value c with amplitude a and phase p can be encoded into the sum of two values $c = c_a + c_b$, each value having variable phase but constant amplitude:

$$\begin{aligned} c &= ae^{ip} \\ p_a &= p - cos^{-1}a \\ p_b &= p + cos^{-1}a \\ c_a &= 0.5e^{ip_a} \\ c_b &= 0.5e^{ip_b} \end{aligned} \tag{5}$$

where a is in the range $[0, 1]$. Complex-valued hologram H can be converted to a double phase hologram by converting each complex value $H(\vec{p})$ to two phase values $P_a(\vec{p})$ and $P_b(\vec{p})$:

$$\begin{aligned} P_a(\vec{p}) &= \angle H(\vec{p}) - cos^{-1}|H(\vec{p})| \\ P_b(\vec{p}) &= \angle H(\vec{p}) + cos^{-1}|H(\vec{p})| \end{aligned} \tag{6}$$

where $\angle$ indicates the phase angle function. Thus $P_a(\vec{p})$ and $P_b(\vec{p})$ provide the values to display on a phase-only spatial light modulator in terms of the desired image intensity (from Equation 2) and the display geometry (from Equation 1 or 4). Since we require two phase values for each complex value in H, we expect a corresponding loss in resolution.

Another closely related method is *complex error diffusion* [Eschbach 1991; Tsang and Poon 2013] in which a complex value is forced to unity amplitude while the resulting error is pushed to neighboring values. This algorithm works like the classic error diffusion halftoning algorithm in image processing and tends to produce a similar result to the double phase method. The benefit of either of these direct methods is that the encoding is trivial to compute after the initial complex-valued hologram H is obtained and no iteration or optimization is required. They do not contain much of the high frequency noise associated with the iterative and temporal averaging methods since the phase is not randomized or optimized but is generally smooth or constant over the hologram. They also

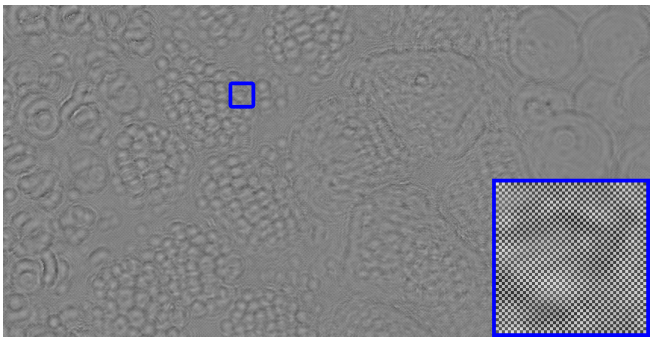

Fig. 4. Example of phase hologram. Hologram corresponds to the "berries" image in Figure 1A. Blue-bordered inset image shows magnified region, showing tiled pairs of phase values that encode amplitude and phase using the double phase method.

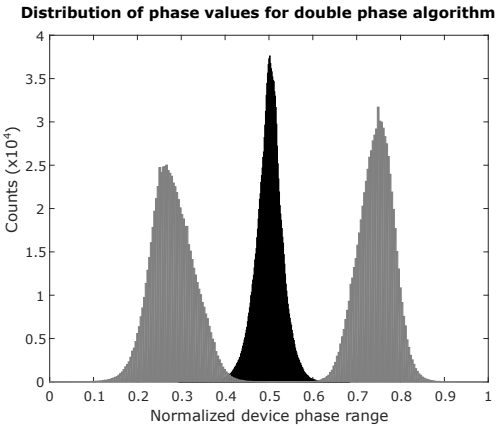

Fig. 5. Phase distribution for double phase method. We convert the original distribution of phase values (black) to a new set using the double phase method (gray) that utilizes the whole device phase range ($> 2\pi$) to avoid errors in phase wrapping.

allow arbitrary intensity distributions even if the sub-hologram size w is less than the size of the full hologram.

However, note that the display must facilitate physical integration of the two phase values into a single complex value. Two recent methods are to use a custom spatial light modulator that optically combines the adjcent phase-only pixels [Reichelt and Leister 2013] or to encode a grating in the hologram to tilt the projector through an aperture stop [Qi et al. 2016]. We use a simpler method that consists of an on-axis projector with a standard SLM, an aperture stop, and no explicitly encoded gratings. By placing adjacent pairs of phase values calculated according to Equation 5 in a checkerboard-like pattern on the SLM (see Figure 4) we implicitly create a high frequency grating that pushes unwanted energy (i.e. the result of applying amplitude modulation) to the edges of the exit pupil. We then filter this unwanted light using an aperture stop. The specific optical layouts of our prototype displays are described in Section 5.2.

We also encode our double phase holograms in a special way on the spatial light modulator to improve performance. Rather than use the standard 0 to 2π phase range of the SLM, we center the distribution of phase values to fill the full phase range of the device ($> 2\pi$) but scale the distribution if necessary to ensure the distribution of phase values tapers off to zero near the maximum and minimum phase rotation of the device (see Figure 5). This approach allows us to achieve superior contrast by ensuring smooth phase over the whole hologram and avoiding errors in wrapping phase from 2π back to 0 on imperfect, real-world SLMs. An example phase hologram using the techniques described in this section is displayed in Figure 4. Due to the smooth phase profile and a sub-hologram size w that is small compared to the full hologram size, some of the image structure can be observed in this phase image.

3.1.4 True 3D Holograms. One of the key advantages of holography is the ability to create true 3D images where each object point appears at a unique focal depth. To form such holograms, we use the color and depth buffers obtained from the standard rendering pipeline to form a point cloud from the viewer's eye position with the occluded points culled. This point cloud comprises the set of object points $\vec{o}$ used in the hologram calculation (see Equation 1). We assume that any view-dependent effects (e.g. occlusion and specular highlights) are handled through standard head-tracked rendering, as with a conventional near-eye display. Note that this arrangement does not consider *intra-pupil* view dependent effects, such as the difference in occlusion or the appearance of a specular highlight between the left and right edges of the eye pupil. We expect such effects would be quite subtle; nonetheless, they can be handled by appropriately sampling the scene and varying the amplitude over the lens phase functions [Zhang et al. 2011].

Calculating such a true 3D hologram is easy in principle with Fresnel holography; we simply set each object point $\vec{o}$ in the lens phase function of Equation 1 or 4 to the actual 3D position. In practice, it can be difficult to achieve good results when only phase modulation is available. When implementing planar holograms with the double phase method, best results are generally achieved when the object phase ϕ_o is constant over the hologram. This produces smooth phase at the hologram plane which keeps noise low and allows us to implement the previously described device encoding method. However, for true 3D holograms, the use of a constant object phase ϕ_o will result in widely varying object phase at the hologram plane, which we have found to produce poor results. However, we can correct object phase ϕ_o for improved consistency on the hologram plane according to:

$$\phi_o = -\frac{2\pi(\|\vec{p_{c_o}} - \vec{l}\| + \|\vec{p_{c_o}} - \vec{o}\|)}{\lambda} \tag{7}$$

where $\vec{p_{c_o}}$ corresponds to the center point of the sub-hologram of object point $\vec{o}$. With this correction factor, we are able to generate high quality holograms with widely varying depth as demonstrated on a prototype device in Section 5.2.2.

3.1.5 Aberration and Vision Correction. Another key strength of holographic displays is the ability to perform optical corrections to deficiencies in the display hardware (*aberration correction*) and in the eyesight of the viewer (*vision correction*). The optical correction

capabilities of Fresnel holograms are extremely powerful. As previously described, the lens phase functions of Equations 1 and 4 act like simple focusing lenses. However, in general, these functions can arbitrarily shape the wavefront that is used to form each object point, pre-compensating for any aberrations in the optics, and the correction functions for each object point are independent. In other words, the optical correction power of such holograms is *spatially variant* and offer practically unlimited degrees of freedom. This is akin to using an independently customized, complex lens *to form each point in the image*, and where each of these complex lenses can *physically overlap*. In practice, aberration correction is limited to the angular steering capability of the SLM (Equation 3); however, in general we find that even severe aberrations are the result of quite small angular deviations.

An "exact" solution for aberration and vision correction would involve inclusion of all the phase delays through the optics and eye to the exponential in Equation 4. However, we find this approach impractical as it is very high dimensional; each unique optical path through the system must be calculated. Freeman et al .[2010] instead use a lower dimensional model, the Zernike polynomials, as the basis for their lens phase functions. This set of polynomial functions, commonly used in optics to express aberrations, are convenient as they are orthogonal and each is used to characterize an intuitive optical defect (defocus, astigmatism, coma, etc.). As with the projectors of Kaczorowski et al. [2015; 2016], we measure the coefficients experimentally, as the optical characteristics (surface shapes, materials, etc.) of each component may not be known, and there may be some manufacturing and alignment variability.

We extend this method to virtual and augmented displays by using a single set of Zernike coefficients to encapsulate the light transport from the light source to the retina. In this way, we can treat the optics of the display and the eye as a "black box" and measure their aggregate effect through a user calibration process. We thus arrive at a single unified model for a new lens phase function $f_z(\rho,\theta)_o$ for object point $\vec{o}$ that encapsulates the display optics, aberration correction, vision correction, and variable focal depth imagery:

$$Z_3(\rho,\theta)_o = a_{3_o}(2\rho^2-1) \quad \text{focus} \tag{8}$$

$$Z_4(\rho,\theta)_o = a_{4_o}(\rho^2 cos2\theta) \quad \text{vertical astigmatism} \tag{9}$$

$$Z_5(\rho,\theta)_o = a_{5_o}(\rho^2 sin2\theta) \quad \text{oblique astigmatism} \tag{10}$$

$$Z_6(\rho,\theta)_o = a_{6_o}((3\rho^2-2)\rho cos\theta) \quad \text{horizontal coma} \tag{11}$$

$$Z_7(\rho,\theta)_o = a_{7_o}((3\rho^2-2)\rho sin\theta) \quad \text{vertical coma} \tag{12}$$

$$\phi_o = a_{3_o} \tag{13}$$

$$f_z(\rho,\theta)_o = e^{i(\phi_o+\sum_j Z_j(\rho,\theta)_o)} \tag{14}$$

where a_{j_o} are the measured Zernike coefficients during user calibration and $f_z(\rho,\theta)_o$ is parameterized by polar radius ρ and angle θ relative to the center of the sub-hologram. Additional Zernike polynomials can be added to correct higher order aberrations if needed. Note that a_{3_o} is adjusted from its measured value according to the focal depth of imagery. Also note that object phase ϕ_o is adjusted for consistent phase at the hologram plane as was similarly done for the simpler lens phase function model in Equation 7.

User Calibration Procedure. The goal of the user calibration procedure is to obtain the coefficients a_{j_o} for the lens phase function $f_z(\rho,\theta)_o$ at each object point $\vec{o}$. Each function represents the wavefront that must be emitted so that light is focused to a small spot on the retina after traveling through all the surfaces in the display optics and the eye. Knowledge of the optics of the display and eye are not required but rather the aggregate effect of these components is measured through user feedback. During calibration, the display presents a single object point $\vec{o}$ to the user which is represented in the hologram as $f_z(\rho,\theta)_o$. The user adjusts the coefficients a_{j_o} of $f_z(\rho,\theta)_o$ until the most tightly focused spot is seen. This process is repeated for additional object points $\vec{o}$. The procedure is fairly intuitive since the effect of adjusting various coefficients is orthogonal and each corresponds roughly to a familiar geometric operation (scale, squash, rotate, etc.). The measured coefficients are applicable for any image content.

Note that it is impractical to measure $f_z(\rho,\theta)_o$ for all object points $\vec{o}$ since there are expected to be millions. However, as aberrations tend to be smoothly varying, one can measure coefficients a_{j_o} at a sparse set of object points and interpolate the coefficients for the remaining points. For a highly aberrated system (e.g. the prototype described in Section 5.2.4), 15 measured points were sufficient to correct the display. For a display without significant aberrations (e.g. Figure 1A), we measured coefficients for a single object point at the center of the display. Also note that for true 3D holograms we perform the calibration for a small number of focal positions (e.g. four) and interpolate the coefficients over the complete focal range.

3.1.6 High Speed Hologram Calculation. Although the Fresnel holographic model we have built in this section is very powerful, it is also a very complex calculation. Note that using the point-wise integration method of Equation 2 to compute hologram H is essentially a quadruple summation over the two axes of the hologram plane and each sub-hologram. While not intractable, runtimes to generate a hologram with 2 million object points on a high end GPU run in the range of one to tens of seconds depending on the sub-hologram size (see Figure 6, bottom).

To generate holograms in real-time on today's hardware, faster algorithms are needed. As noted previously, the point-wise integration method of Equation 2 is equivalent to a convolution of the target image with a spatially varying kernel. One approach to speed up the calculation is to remove the spatial variance of the lens phase function $f_z(\rho,\theta)_o$ so that the function is the same for all object points $\vec{o}$. In this case, we can substitute the point-wise integration method of Equation 2 for standard spatially *invariant* convolution:

$$H = A * f_z \tag{15}$$

where $*$ denotes complex-valued convolution, A contains the square root of the image intensities, and f_z is the spatially invariant lens phase function. It is well known in holography that Equation 15 can be computed quickly using the Fast Fourier Transform:

$$H = FFT^{-1}(FFT(A) \odot FFT(f_z)) \tag{16}$$

where $\odot$ indicates point-wise multiplication. Additionally, if f_z is linearly separable into vectors $\vec{f_{z_x}}$ and $\vec{f_{z_y}}$, Equation 15 can also be computed quickly with direct separable convolution as:

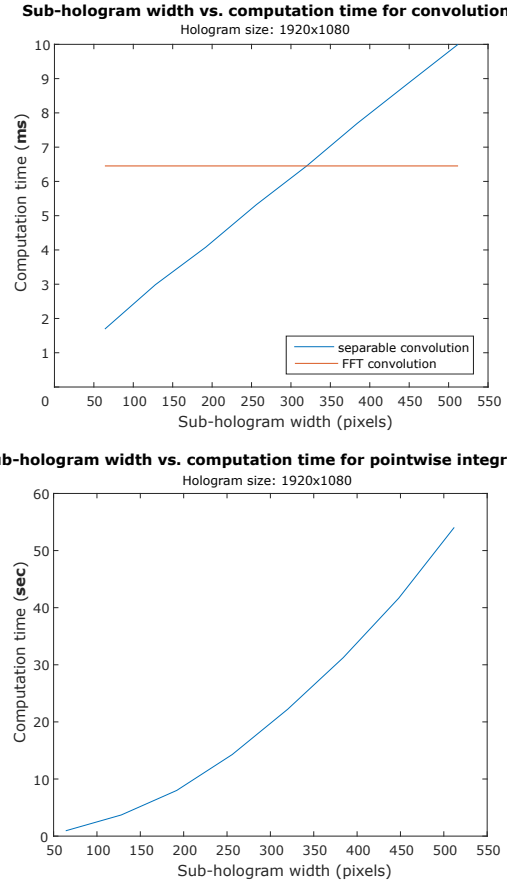

Fig. 6. Measured runtimes for various hologram computation methods on an NVIDIA GTX 980 Ti GPU. *Top:* Separable and FFT-based convolution. *Bottom:* The point-wise integration method. All runtimes measure the total computation time needed to generate a 1920×1080 hologram with 1920×1080 object points and encode the result using the double phase method. Note that the convolution methods are measured in milliseconds while the point-wise integration method is measured in seconds.

$$H = (A * \vec{f_{z_x}}) * \vec{f_{z_y}} \tag{17}$$

This simplification is common in image processing but to our knowledge has not been used to compute holograms. Either method will increase computation speed by several orders of magnitude; calculation of a hologram of 2 million object points takes just a few milliseconds on a GPU (see Figure 6, top). Note that the runtime of the separable method scales linearly with the sub-hologram size while the FFT method does not depend on this size. The separable method is faster for smaller sub-hologram sizes and the crossover point between the two algorithms is near the middle of the range of sub-hologram sizes used for our prototype displays.

In summary, the prerequisite for high speed computation is a *spatially invariant* lens phase function f_z, which implies that the focus and aberration correction is constant over the image. (However, f_z can change on a frame-by-frame basis, which we explore in the next section.) Additionally, if f_z is linearly separable, then separable direction convolution can be used. Lens phase functions of any focal power and certain aberrations (e.g. horizontal astigmatism or vertical astigmatism) are linearly separable. If f_z is not linearly separable, then FFT-based convolution can be used for arbitrary f_z; both methods run at real-time rates.

3.1.7 Increasing Performance in an Eye-Tracked Display. Although the convolution method of Equation 15 can be computed quickly, it does not support *per-pixel* focal control as lens phase function f_z is not spatially variant. A common approximation is to restrict depth to a set of fixed planes [Chen and Chu 2015]. In our case, we could compute each plane with a separate convolution and lens phase function f_z. However, this approach in general will result in a slowdown of a factor N for N planes and will not support smooth depth changes. Another option is the wavefront recording plane method of Tsang et al. [2015]; however, this method has only been demonstrated to work in a shallow depth range (0.02 m).

Likewise, the convolution based method does not allow *spatially variant* aberration correction. Kaczorowski et al. [2016] show spatially variant aberration correction is possible at interactive rates with an approximate piece-wise method. However, this approach is not ideal for our application as we explore displays with severe, quickly changing aberrations and our amplitude encoding method does not perform well with abrupt phase discontinuities.

Instead we look to the trend of providing eye tracking to displays as in such technologies as foveated rendering [Guenter et al. 2012]. Similarly, one can reduce the resolution of the hologram outside the tracked fovea position; Hong et al. [2016] performed this method by reducing the density of object points $\vec{o}$ outside the fovea. Alternatively, we suggest including all object points but decreasing the sub-hologram width w outside the fovea (cutting high frequencies), as our calculation speed scales quadratically with the sub-hologram size in the point-wise integration method (see Figure 6, bottom). This method may also reduce temporal aliasing as all object points are always included in the calculation. However, we have not found this method to run at high video rates on current GPUs.

A high speed option is to approximate spatially variant focus and aberration control by providing the correct lens function *where the user is looking* rather than computing or approximating the full spatially variant solution. This can be computed simply as the convolution of Equation 15 where we substitute fixed lens phase function f_z for the temporally varying function $f_{z_{o_t}}$, which corresponds to the lens phase function for object point $\vec{o}_t$, the closest object point $\vec{o}$ to the tracked fovea position. The advantage of this method is that computation time is nearly the same as for the fixed lens phase function case; we must only recompute $f_{z_{o_t}}$ each frame, which is very fast when performed on the GPU. Another advantage is a smooth and continuous falloff of aberration correction outside the foveal region without sharp phase discontinuities in the hologram from discretized approximations. The limitation is that we must ensure the that the falloff in resolution from correcting aberrations only near the fovea occurs no more quickly than the falloff of the visual acuity of the eye. Additionally, although we are able to adjust focus and address the accommodation-convergence conflict with this method, we are unable to present true depth of field if there are

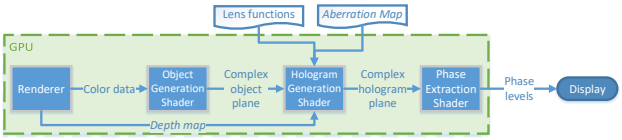

Fig. 7. Holographic software pipeline. The scene is drawn off-screen to color and depth buffers using a standard renderer. The scene is then converted to a hologram using a series of pixel shaders and sent to a phase-only SLM.

object points with varying depths closely surrounding the tracked fovea. However, since the eye is only able to focus on a single depth simultaneously, it is possible to present the correct optical focus while approximating depth of field blur in image space. We implemented the tracked spatially varying lens phase function method on our holographic display prototypes but simulated eye position with a mouse cursor in lieu of an actual eye tracker.

3.1.8 Rendering pipeline and GPU Implementation. When generating content for a 3D display, it is preferential to render imagery using the standard graphics pipeline and APIs (DirectX, OpenGL, etc.), as many tools are available, and independently convert to a display-specific representation in a post-processing step. Our implementation provides a true 3D, multi-focal hologram directly from the GPU color and depth buffers of a rendered 3D scene using a series of pixel shaders (see Figure 7). The first shader converts the color data to the set of complex object points $\vec{o}$. The next stage converts these object points into the complex hologram plane using the depth buffer and lens phase functions $f_z(\rho, \theta)_o$, which are computed on-the-fly. The complex hologram is converted to a phase-only representation in a final shader and is sent to the SLM. Note that we provide three implementations of the hologram generation shader that correspond to the point-wise integration, separable convolution, and FFT convolution methods described in this paper. The separable convolution shader operates in two passes, one for each of the linearly separable axes. Two versions of the phase extraction shader were also implemented: one for the amplitude discard method, and one for the double phase method. As a result of this implementation, we are able to generate holograms directly from rendered polygonal models. Figure 8 shows a photograph of a prototype display where the model was rendered, converted to a hologram, and displayed in real-time in excess of the display refresh rate.

3.2 Completing a Near-Eye Display

In Section 3.1, we described how a point light source and phase-only SLM are used as a basis for a holographic projector. However, as with other near-eye displays, a projector unit (or image panel) is just one component of a display system; we must also add an *eyepiece* to magnify the image, relay it to the eye, and optionally provide a see-through capability. In this section, we describe the addition of a suitable eyepiece to complete a VR or AR display.

3.2.1 Virtual Reality Displays. The first goal of building our near-eye displays is to show the best intrinsic capabilities of holograms: good image quality, true 3D multi-focal imagery, and vision correction capabilities. We focus on this set of capabilities in our VR displays, which we implement using simple benchtop designs and conventional eyepieces. We defer more ambitious optics and form factors for our subsequent AR prototypes.

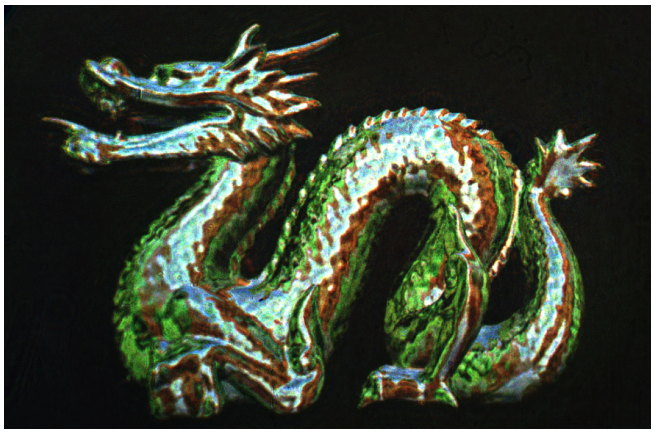

Fig. 8. Photograph of holographic display. The model was rendered using the standard graphics pipeline and converted to a hologram in post-processing pixel shaders. In this example, the hologram resolution is 1920×1080 pixels and the rendering and holographic computation loop ran at a color field rate of 157 Hz. Dragon model by Stanford Computer Graphics Laboratory.

A basic virtual reality display can be formed by placing a magnifying eyepiece in front of a holographic projector, forming a simple imaging system. The projector forms an aerial image (or volume) at the location of the object points, which the eyepiece magnifies and relays to the eye. Since phase-only spatial light modulators are generally small microdisplays, the eyepiece must provide a high degree of magnification. Alternatively, magnification may take place using a diverging beam incident on the SLM, as described in Section 3.1.2. For our virtual reality displays, we use a combination of these approaches. As also described in Section 3.1.2, as we increase the field of view of our display, we decrease the size of the exit pupil and extend the depth of field, making focal effects more subtle. Therefore, we built two virtual reality displays: one to demonstrate a wide field of view, and one to demonstrate multi-focal and vision correction capabilities. As we describe in Section 5.3, this is not a fundamental restriction as SLMs allowing both capabilities are available today. We assess and describe the specific configurations of our virtual reality displays in Sections 4.1 and 5.2.

3.2.2 Augmented Reality Displays. The other goal of building our near-eye displays is to show the optical corrective power of holograms, which enable form factors that are inaccessible through conventional means. We use our augmented reality display prototypes to show that wide fields of view and very compact form factors are together possible.

Eyepieces for optical see-through AR displays present a unique set of challenges to the display designer, as it is difficult to maintain both see-through and augmented optical paths in a compact, wide field of view device. Various eyepieces have been proposed, such as freeform prisms, waveguides, and off-axis reflectors. Off-axis reflectors, which employ an image projector that is tilted away

from the optical axis of a reflective focusing element, have demonstrated the largest fields of view. However, to reduce the form factor, the focusing power of these eyepieces must be increased and the projector must be tilted farther off-axis. Furthermore, curved reflective elements tend to have undesirable bulbous form factors, so it is instead desirable to use flat elements with Fresnel-grooved or fixed holographic structures. Unfortunately, all these factors increase optical aberrations, especially as the field of view increases. Thus, such optical systems tend to have large helmet-sized form factors or narrow fields of view. However, our holographic projector affords powerful aberration correction capabilities, enabling us to instead select the most desirable physical properties – *compact, flat, high magnification, wide FOV, and highly off-axis* – and use the holographic projector to correct all the associated aberrations.

The eyepiece we selected for our augmented reality displays is a thin, flat, and wide field of view holographic optical element (HOE). The HOE consists of a fixed hologram that focuses light from a highly off-axis beam into an on-axis beam in front of the eye, performing a similar function as an elliptically curved reflector. As a *volume* (or *Bragg*) hologram, the HOE acts only over a narrow set of angles and wavelengths and is thus able to redirect and focus the light of the projector while providing a clear, highly transparent see-through view to the user. We assess this eyepiece under two configurations. In the first configuration, we combine the benchtop holographic projector of our virtual reality displays with the HOE eyepiece, showing the ability to correct severe aberrations in the compact eyepiece while providing a full color display and complex imagery. In the second configuration, we replace the benchtop projector with a miniature projection engine that also employs highly off-axis optics, and attach it to an eyeglasses-like frame. We show that this compact projector can correct even more severe aberrations in simple holographic imagery. We assess and describe the specific configurations of these displays in Sections 4.1 and 5.2.

4 IMPLEMENTATION

4.1 Hardware

We constructed four prototype displays which we assess in Section 5.2. Each prototype included a HOLOEYE PLUTO (model HES-6010-VIS) liquid crystal on silicon (LCOS) reflective phase-only spatial light modulator with a resolution of 1920×1080 pixels. The pitch of the SLM is 8 μm, and the active area is 15.36×8.64 mm. The SLM was illuminated by a single optical fiber that was coupled to three laser diodes emitting at 448, 524, and 638 nm. The laser light was linearly polarized to match the requirement of the SLM. For full color operation, the displays operated in a color field sequential operation at a 20 Hz frame rate (60 Hz color field rate), limited by the refresh rate of the SLM. An Arduino microcontroller was used to control laser power and synchronize the three color lasers to the display. The spatial light modulator was connected to a PC with a NVIDIA GTX 980 Ti GPU, which performed all of the rendering and holographic calculations. The display was photographed and filmed using a Point Grey Blackfly S industrial color camera with a resolution of 2448×2048 pixels. The camera focus was set to near infinity for results not showing multi-focal capabilities. To avoid artificially boosting contrast, the non-linear image processing functions of the camera (black level, gamma adjust) were disabled and no post-exposure adjustments were made to the image files.

4.2 Software

We performed rendering operations in OpenGL and holographic calculations in the OpenGL Shading Language (GLSL) running on Windows 10. The GLFFT library[1] was used for FFT calculations. All input images and holographic output used a resolution of 1920×1080.

For results featuring aberration correction, Zernike coefficients were measured at a grid of 15 points over the display, and MATLAB was used to interpolate to the full display resolution. Defocus, vertical astigmatism, oblique astigmatism, vertical coma, and horizontal coma coefficients were measured as these were sufficient to correct the aberrations in our display prototypes. For results featuring variable depth, the Zernike focus coefficient was measured at several known focus depths, and these measurements were used to build a linear function mapping inverse depth to the focus coefficient. For results featuring distortion correction, a sparse set of image-camera correspondences were measured and interpolated to the full display resolution using MATLAB. For full color results, these operations were performed over all three color channels.

All the computation times listed in this paper indicate the time to run the full rendering and holographic computation loop; thus results that include the rendering of a complex 3D model (e.g. the dragon in Figure 8) have lower performance than those in which a simple 2D image is displayed (e.g "berries" photograph of Figure 9). Note that all rates listed for color results indicate the color field rate, where three color fields are needed to represent a full color image. Also note that computation times were measured with vertical sync disabled to measure rates above the display refresh rate of 60 Hz; however, photographs were captured with vertical sync enabled to synchronize the display to the three color lasers.

5 EXPERIMENTAL ASSESSMENT

5.1 Software Simulation

To evaluate the theoretical performance of the algorithms, we first performed a numerical reconstruction to simulate the holographic image. A phase hologram was generated using the separable convolution method with a sub-hologram size of 177 pixels and double phase encoding. The lens phase function was selected to match the prototype configuration described in Section 5.2.1. Reconstruction was performed by inverting the holographic generation process, i.e. deconvolution of the hologram with the lens phase function, or equivalently, convolution of the hologram with the complex conjugate of the lens function. Note that this simulation process does not capture all the intricacies of playback on a specific real-world SLM (e.g. accounting for the dead space between pixels or phase inaccuracies), but provides a good first approximation.

The results are shown in Figure 9. Qualitatively, the simulation (Figure 9, center) provides a faithful representation of the target image (Figure 9, left) and provides higher contrast and less high frequency noise than typical results for random phase, iteratively

[1] https://github.com/Themaister/GLFFT

generated holograms. However, we notice the resolution loss associated with encoding a single complex value into two phase values as well as some subtle ringing artifacts. Quantitatively, the PSNR of the reconstruction was 30.57 dB for red, 30.05 dB for green, and 27.26 dB for blue. We find this promising as a PSNR ≥30 dB is generally a target for lossy video compression. See Yoshikawa et al. [2016] for quantitative analysis of holographic image quality.

5.2 Hardware Prototypes

To further evaluate our approach for real-world use, we built a series of prototype displays. Each prototype is designed to best showcase one of more of the strengths of holography in near-eye VR and AR displays. The prototypes are described in the following sections.

5.2.1 Wide Field of View Virtual Reality Display. Our first prototype is designed to demonstrate a wide FOV virtual reality scenario with the best image quality. We use a simple, benchtop optical design (see Figure 10, left) and a conventional eyepiece (Meade Series 5000 21 mm MWA), deferring compactness for the subsequent prototypes. Magnification of the 15.36 mm wide SLM to a 70° wide FOV was achieved using a diverging beam and high magnification eyepiece. An intermediate aperture plane with a stop is used to block the higher diffraction orders and facilitate the double phase encoding method. A Goyo 6 mm focal length C-mount camera lens was used to photograph the display. Hologram computation was performed using separable convolution and double phase encoding.

Figure 8 shows a frame of a video sequence in which a 3D model was rendered, converted to a hologram at a fixed depth, and displayed in real-time. The result shows the ability of our software to integrate with the standard graphics pipeline. We also note this result shows the ability to create high contrast blacks and black regions larger than the sub-hologram size. For this result, the rendering and computation loop ran at a color field rate of 157 Hz, including the time to render the model. The supplemental video includes an animated sequence of the model rotating.

Figure 9 (right) shows a photograph of the prototype display. The hologram was computed and displayed in real-time. The prototype shows a reasonable representation of the expected result (Figure 9, center), although the colors are a bit harsher. In the magnified region, we observe the resolution is also close to the simulation. Note that the camera is close to the resolution of the spatial light modulator and employs Bayer color filters, resulting in the blocky pixel and color structure visible in the magnified region. For this result, the computation loop ran at a color field rate of 260 Hz.

5.2.2 Multifocal Virtual Reality Display. Our second prototype is designed to display multi-focal and vision correction capabilities. The second prototype (Figure 10, right) is physically similar to the first but uses a lower magnification eyepiece (Edmund Optics Erfle 32 mm) and has a narrower 40° diagonal FOV. A Fujinon 12.5 mm focal length C-mount camera lens was used to photograph the display. For results showing per-pixel depth, the point-wise integration method and double phase encoding were used. For the results showing eye tracked depth and vision correction, separable convolution and double phase encoding were used. All results use a sub-hologram size of 401 pixels.

Figure 11 shows the ability of our display to display true 3D holograms where each pixel appears at a unique and individually controllable depth. By adjusting only the focus ring of the camera, we were able bring various parts of the model in and out of focus (see supplemental video). Since each pixel acts as spherical wave with virtually no discretization, we observe very smooth and natural focus over the depth of the model. This hologram requires a very complex computation so processing was performed offline.

Although we are encouraged by the prospect of true 3D holograms with per-pixel depth, we were not yet able to demonstrate real-time calculation. As described in Section 3.1.7, a useful approximation is to set the entire scene focus to match the depth of the viewer's tracked gaze. In lieu of an eye tracker, we simulate this configuration by setting the global scene focus to the depth of the scene at the position of the mouse cursor. The results are shown in Figure 12, which are calculated and displayed in real-time. We demonstrate that the whole scene appears in focus when the hologram focus matches the camera focus, and out of focus otherwise. See the supplemental video for additional results. For this result, the rendering and computation loop ran at 91 Hz; there is virtually no speed penalty for changing the global scene focus each frame.

Finally, we use our second prototype display to demonstrate vision correction. We placed a cylindrical lens with a focal length of 200 mm in front of our camera to induce fairly severe vertical astigmatism (see Figure 13B). When viewing a hologram with this astigmatic vision, we observe the associated blurring (Figure 13D), especially among lines with a primarily horizontal component, as compared to a hologram viewed with normal vision (Figure 13C). However, when we apply vision correction to our hologram (Figure 13E) the result is virtually the same as the when no astigmatism was induced (Figure 13C). For these results, the computation loop ran at a color field of rate of 92 Hz. We note that the astigmatism is less severe in the hologram (Figure 13D) than a general camera view (Figure 13B) because the exit pupil of the display is smaller than the camera aperture. Also note that the linearly separable convolution method used for these results can only fix vision problems that can be corrected with a linearly separable phase function. For general-purpose vision correction, FFT based convolution can substituted, which runs at a similar speed as demonstrated in Section 5.2.3. Finally note that the that the cylindrical lens also introduced some geometric distortion (vertical image stretch) which we did not correct but it would be trivial to do so.

5.2.3 Benchtop Augmented Reality Display. Our remaining prototypes are designed to show the potential of holography for optical see-through AR applications. In particular, we demonstrate the ability to correct severe optical aberrations that help to enable previously inaccessible form factors. Our first AR prototype (see Figure 14, left) is designed to show that a holographic projector can be used with a high magnification, wide field of view, highly off-axis, and flat eyepiece. Our eyepiece consists of a holographic optical element that focuses light from a projector approximately 60° off-axis to the on-axis position of the eye, while providing a clear see-through view. The HOE operates at three wavelengths, enabling a full color display and a horizontal field of view of > 82° (limited by the field of view of our camera). We used the eyepiece

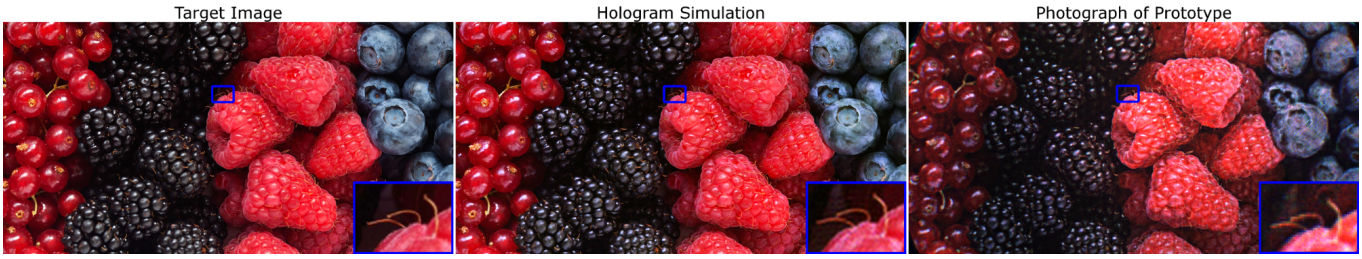

Fig. 9. Image quality comparison. *Left:* Target image. *Center:* Numerical reconstruction of hologram. *Right:* Photograph of prototype display described in Section 5.2.1. Blue bordered inset images show magnified regions. Berries image by Ana Blazic Pavlovic/Shutterstock.

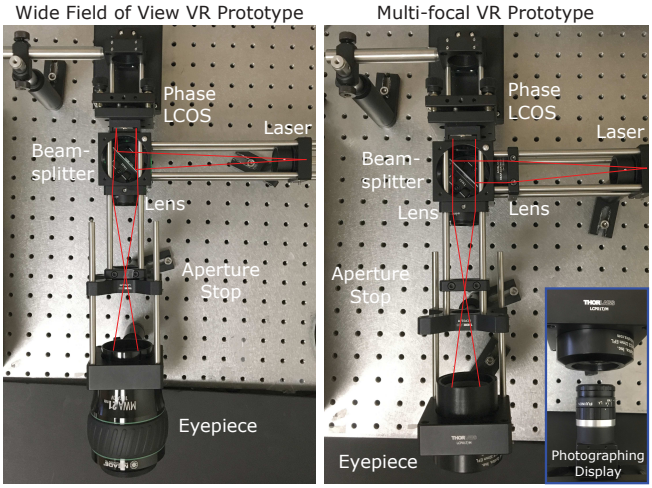

Fig. 10. Virtual reality holographic display prototypes. *Left:* Prototype display designed to demonstrate a wide field of view. *Right:* Display designed to demonstrate multi-focal capabilities. The blue bordered inset image shows how the display was photographed using a camera at the eye position.

with the same benchtop holographic projector as was used for the multifocal VR Display (see Section 5.2.2). A Goyo 6 mm focal length C-mount camera lens was used to photograph the display. Hologram computation was performed using point-wise integration and the amplitude discard method for the full screen aberration correction results, FFT convolution and the amplitude discard method for the tracked aberration correction results, and point-wise integration and the double phase encoding method for the display of a full color model. All results use a sub-hologram size of 401 pixels.

Figure 15 shows the ability of our display to correct eyepiece aberrations. Figure 15A shows the display with no aberration correction, i.e. the projector focus is set to the correct value for the center of the display, but no other corrections are applied. Note that there is severe and spatially varying astigmatism over the display; no region of the display is in both proper horizontal and vertical focus. Figure 15B shows the display after full screen, spatially varying aberration correction is applied. We observe sharp lines in both directions over the entire field of view. We note, however, that the spatially varying aberration correction is an expensive operation and was computed offline. However, in a eye tracked display, we are able to correct aberrations in the region that the viewer is looking. We simulate this effect by correcting aberrations in the position of a tracked cursor (see Figure 15C). The tracked method runs in real-time at a rate of 89 Hz; see the supplemental video for an animated result. Finally, in Figure 15D we show a full color, augmented image of a complex model. This result uses both optical aberration and geometric distortion correction and was computed offline, although the tracked aberration correction method could be applied.

We are encouraged by the ability to show clear, full color, augmented holograms in a challenging optical configuration. The see-through view is also clear and transparent. However, the contrast is poorer than the VR displays, and some light leaks through the "clear" regions of the display. The display edges and corners are also dim, and there is a dimmer region in the center of the display due to uneven illumination and HOE diffraction efficiency. These factors could be improved with a more carefully constructed HOE.

5.2.4 Compact Augmented Reality Display. Our previous AR prototype, while using a very compact eyepiece, still retained a holographic projector with a large benchtop form factor. Our final augmented reality prototype is designed to show that the holographic projector itself can also be miniaturized, enabling the optics to be incorporated into an eyeglasses-like form factor, while retaining a wide 80° horizontal FOV. We achieved miniaturization by eliminating the aperture stop and all lenses so that the light needed only to diverge between the laser and eyepiece (rather than diverge, converge through the stop, and re-diverge), greatly reducing the optical path length. We also switched to a highly off-axis (≈60°) SLM illumination scheme that reduced the required space and eliminated the need for a beamsplitter optic. Finally, the optical path was folded around the head using a set of mirrors until it reached the HOE, which was cut down to the shape of an eyeglasses lens and inserted into a 3D printed eyeglasses frame (see Figure 14 (right) and Figure 1C). We did not incorporate the SLM driving electronics or light source driving electronics into the glasses frame; we note, however, that Google Glass has demonstrated that LCOS panel and light source driving electronics can be housed in a (sub-)eyeglasses-sized frame. The display was photographed with a Goyo 6 mm focal length C-mount camera lens. Hologram computation was performed using point-wise integration and the amplitude discard method. We only tested the display in one color (green), but it is not a limitation of the design.

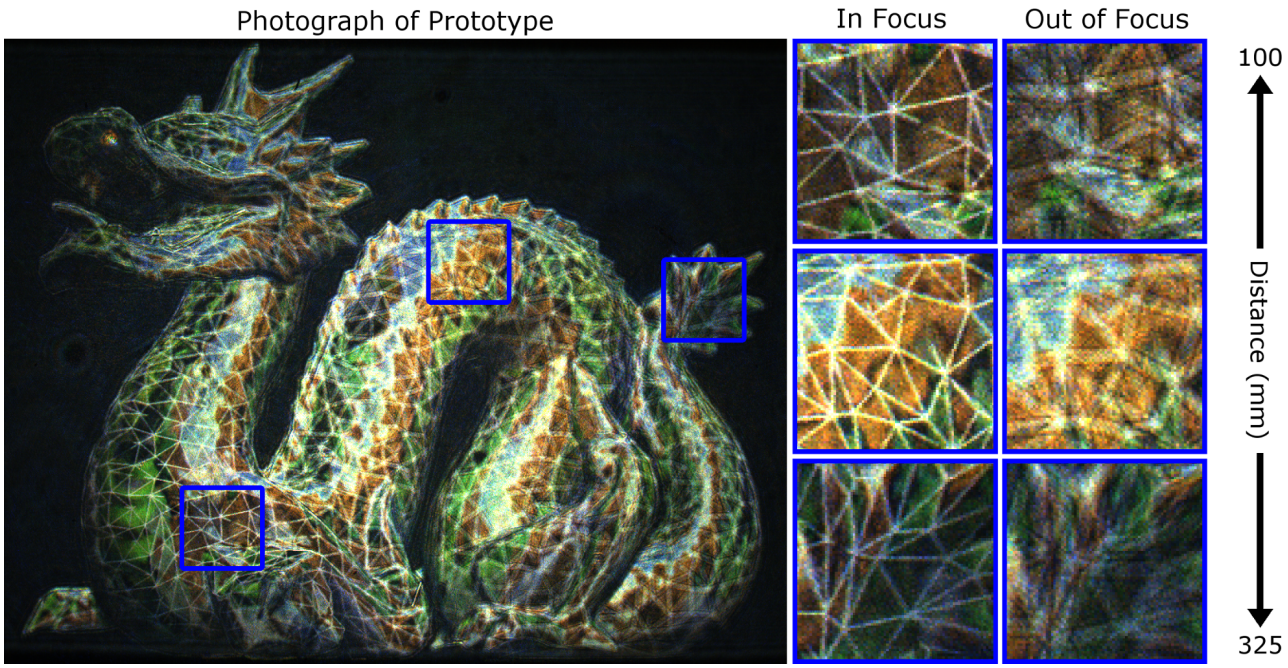

Fig. 11. Photograph of holographic display exhibiting true depth of field. Each pixel appears at a unique and individually addressable focal depth. *Left:* Photograph of display with camera focus set to depth of the dragon's chest region. Notice that the chest appears in focus, while the farther body and tail appear out of focus. *Center column:* Magnified regions of the three blue bordered regions of the dragon (chest, middle body, tail) when the camera was focused at that region. *Right Column:* Magnified regions of the three blue bordered regions when the camera was focused away from that region. Dragon model by Stanford Computer Graphics Laboratory.

Figure 16 shows the ability of our compact augmented reality prototype to show clear, high resolution images. Although the projected hologram is unrecognizable without aberration correction (Figure 16, left), the aberration corrected image (Figure 16, center) shows sharp horizontal and vertical lines over the whole field of view. Figure 16 (right) shows the display is capable of retaining much of the resolution of the SLM. Note that the text can be resolved despite a line width of only SLM pixel wide, with a one pixel space between the body and dot of the letter *i*. Also note that text is sharp in both the center and the corners of the display. Figure 1D shows a simple augmented scene. These results were calculated offline; however, real-time operation is possible using the previously described tracked aberration correction method.

We are encouraged by the initial results in this diminutive form factor. However, we note some drawbacks of the display. The edges and corners of the display are quite dim due to uneven illumination and HOE diffraction efficiency. The consequence of removing the aperture stop in the projector is that we can no longer block the higher diffraction orders, resulting in some faint ghosting and the inability to use the double phase encoding method for more complex imagery. However, we expect that a *virtual* aperture stop can be added to the projector without increasing size (see Section 5.3).

5.3 Assessment

We have demonstrated near-eye holographic displays that offer high resolution, full color, good image quality, variable focus, and vision correction. We have also shown a powerful aberration correction capability that enabled a wide field of view in a very compact device. However, we note several remaining issues to address in future work to create practical virtual and augmented reality displays:

Pupil Expansion. One of the major limitations of our prototypes is the small exit pupil (or eye box). A practical stereo display requires a pupil expansion device or steering device. One possibility is to shift the exit pupil by switching light sources or by using a beam-steering element, which has been demonstrated successfully in large format holographic displays [Häussler et al. 2009]. When using the device with human viewers, the small exit pupil and constant holographic phase can also contribute to the appearance of low frequency speckle as the eyes move. We noticed that this effect disappeared when we experimented with a pupil expansion device.

Display Etendue. As described in Section 3.1.2, the limited etendue of our holographic phase modulator places a field of view and depth of field trade-off in our displays. For this reason, we were unable to show both a wide FOV and shallow depth of field in a single device. However, we note the commercial availability of new 4K resolution phase modulation panels (e.g. HOLOEYE GAEA and Jasper JD8714) that have the same area as our phase panel but over twice the linear pixel density. Such panels will allow both a wide field of view and shallower depth of field in a single device, while also increasing resolution by more than a factor of four.

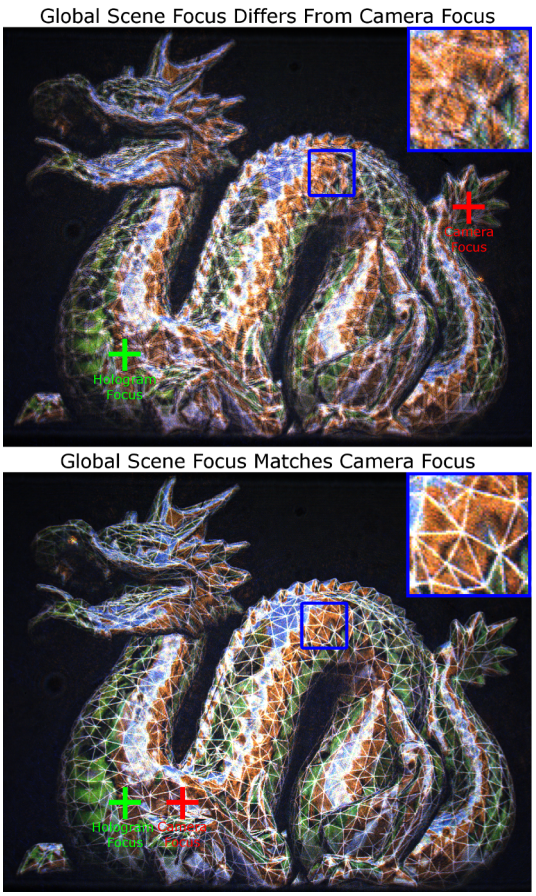

Fig. 12. Photographs of prototype showing real-time tracked global focus. The focus of the entire scene was changed to match the depth at the tracked green cursor position. *Top:* The scene depth at the tracked position (green crosshair) does not match the depth of focus of the camera (red crosshair). The scene appears out of focus. *Bottom:* The scene depth at the tracked position (green crosshair) matches the depth focus of the camera (red crosshair). The scene appears in focus. Blue bordered inset images show magnified regions. Dragon model by Stanford Computer Graphics Laboratory.

Higher Diffraction Orders. Our compact AR display did not include an aperture stop plane to block the higher diffraction orders, which did not allow us to implement the double phase encoding technique for more precise amplitude control. We expect that a "virtual" aperture stop could be implemented in the holographic optical element itself by increasing the angular selectivity of the volume hologram through the use of a thicker holographic medium. This would only affect the thickness of the eyepiece by tens of microns.

Eye Tracking. Although we proposed accelerating holographic calculations based on gaze-tracking data, we did not incorporate an eye tracker in our prototypes. We note the availability of subminiature eye tracking cameras (e.g. Omnivision OVM6211, 3 mm × 3 mm) that could be incorporated into compact devices. We also note the commercial availability of eye trackers embedded in head mounted displays by vendors such as SensoMotoric Instruments.

Display Refresh Rate. Our phase spatial light modulator was only capable of a 60 Hz refresh rate, which caused flicker when the display was viewed in a color sequential operation. Although phase LCOS modulators tend to be slower than amplitude modulators due to the need for a thicker liquid crystal layer, we note the availability of 180 Hz phase modulators (e.g. HOLOEYE LETO and Thorlabs EXULUS-HD1) which should allow flicker-free display.

5.4 Conclusion

Virtual and augmented reality near-eye displays offer great potential to take us to new places, provide instant and spatially-aware access to information, and begin to integrate computer graphics with human vision. Realizing these goals effectively requires great advances in display technology and the satisfaction of numerous conflicting requirements in a single hardware device. *Computational* displays are an attractive solution that push complexity to software where it is easier to fulfill a large number of optical constraints. We have described how digital holography provides an appealing computational solution to near-eye display, offering high resolution imagery, per-pixel focal control, vision correction, wide FOV, and compact form factors. These capabilities were demonstrated across a series of preliminary hardware prototypes. In future work, we plan to integrate all these capabilities into a single hardware device while expanding the exit pupil to create a practical stereo display. In this way, we hope to become one step closer to truly mobile near-eye displays that match the range of capabilities of human vision.

ACKNOWLEDGMENTS

We thank Adrian Travis, Bernard Kress, and Steve Robbins for useful discussions and advice. We also thank Triple Take Holographics for constructing the holographic optical elements used in this paper and the Microsoft Applied Sciences group for producing additional holographic elements used during development. The Microsoft Research Hardware Lab provided 3D printing support.

REFERENCES

Kurt Akeley, Simon J. Watt, Ahna Reza Girshick, and Martin S. Banks. 2004. A Stereo Display Prototype with Multiple Focal Distances. In *ACM SIGGRAPH 2004 Papers (SIGGRAPH '04)*. ACM, New York, NY, USA, 804–813.

A. J. Cable, E. Buckley, P. Mash, N. A. Lawrence, T. D. Wilkinson, and W. A. Crossland. 2004. 53.1: Real-time Binary Hologram Generation for High-quality Video Projection Applications. *SID Symposium Digest of Technical Papers* 35, 1 (2004), 1431–1433.

J.-S. Chen and D. P. Chu. 2015. Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications. *Opt. Express* 23, 14 (Jul 2015).

Reiner Eschbach. 1991. Comparison of error diffusion methods for computer-generated holograms. *Appl. Opt.* 30, 26 (Sep 1991).

Jonathan P. Freeman, Timothy D. Wilkinson, and Paul Wisely. 2010. Visor projected HMD for fast jets using a holographic video projector. *Proc. SPIE* 7690 (2010), 76901H–76901H–12.

Qiankun Gao, Juan Liu, Jian Han, and Xin Li. 2016. Monocular 3D see-through head-mounted display via complex amplitude modulation. *Opt. Express* 24, 15 (Jul 2016), 17372–17383.

A Georgiou, J Christmas, N Collings, J Moore, and W A Crossland. 2008. Aspects of hologram calculation for video frames. *Journal of Optics A: Pure and Applied Optics* 10, 3 (2008), 035302.

R. W. Gerchberg and W. Owen Saxton. 1972. A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik* 35 (1972), 237–246.

Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder. 2012. Foveated 3D Graphics. *ACM Trans. Graph.* 31, 6, Article 164 (Nov. 2012), 10 pages.

R. Häussler, S. Reichelt, N. Leister, E. Zschau, R. Missbach, and A. Schwerdtner. 2009. Large real-time holographic displays: from prototypes to a consumer product. *Proc. SPIE* 7237, 72370S–72370S–9.

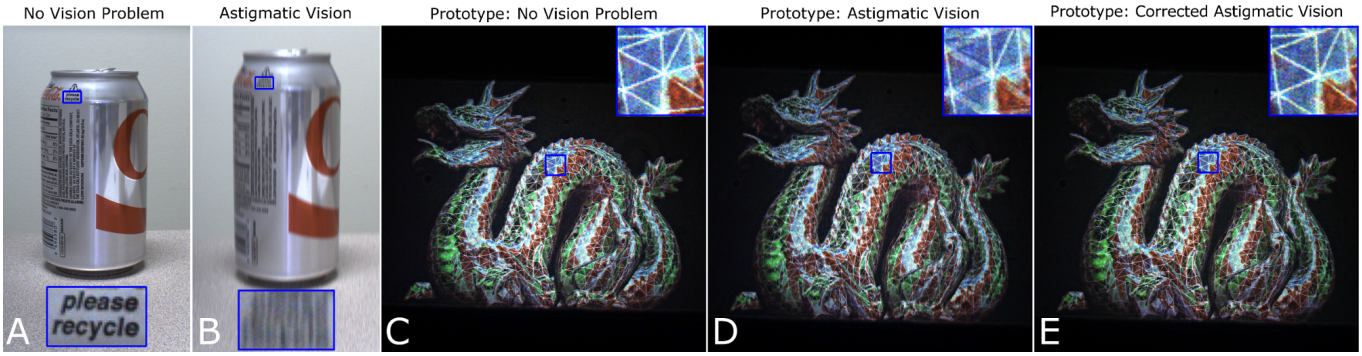

Fig. 13. Holographic Vision Correction. *A)* For reference, camera view of a physical object with no induced vision problem. *B)* For reference, camera view of a physical object with astigmatic vision introduced by an additional cylindrical lens. *C)* Photograph of prototype display taken with camera without vision problem. *D)* Photograph of prototype display taken with camera with astigmatic vision. *E)* Photograph of prototype taken with camera with astigmatic vision and vision correction applied. Dragon model by Stanford Computer Graphics Laboratory.

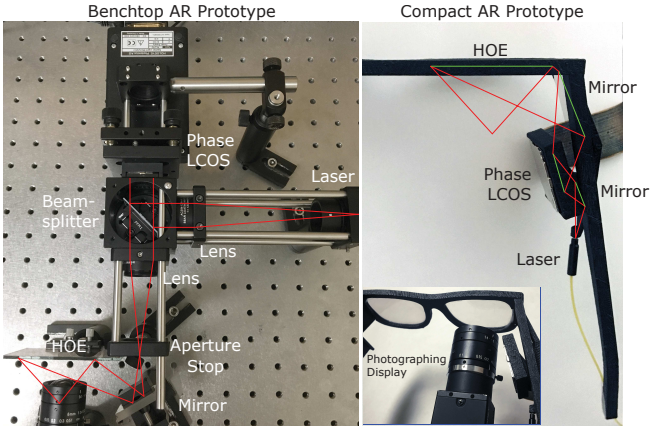

Fig. 14. AR prototypes. *Left:* Full color benchtop prototype. *Right:* Compact prototype with miniaturized projector. The blue bordered inset image shows how the display was photographed.

Jisoo Hong, Youngmin Kim, Sunghee Hong, Choonsung Shin, and Hoonjong Kang. 2016. Gaze contingent hologram synthesis for holographic head-mounted display. *Proc. SPIE* 9771 (2016), 97710K–97710K–6.

C. K. Hsueh and A. A. Sawchuk. 1978. Computer-generated double-phase holograms. *Appl. Opt.* 17, 24 (Dec 1978), 3874–3883.

Hong Hua and Bahram Javidi. 2014. A 3D integral imaging optical see-through head-mounted display. *Opt. Express* 22, 11 (Jun 2014).

F. Huang, K. Chen, and G. Wetzstein. 2015. The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-Eye Light Field Displays with Focus Cues. *ACM Trans. Graph. (SIGGRAPH)* 4 (2015). Issue 34.

F. Huang, G. Wetzstein, B. Barsky, and R. Raskar. 2014. Eyeglasses-free Display: Towards Correcting Visual Aberrations with Computational Light Field Displays. *ACM Trans. Graph. (Proc. SIGGRAPH)* 33, 4 (2014), 1–12.

Changwon Jang, Chang-Kun Lee, Jinsoo Jeong, Gang Li, Seungjae Lee, Jiwoon Yeom, Keehoon Hong, and Byoungho Lee. 2016. Recent progress in see-through three-dimensional displays using holographic optical elements. *Appl. Opt.* 55, 3 (Jan 2016), A71–A85.

Paul V. Johnson, Jared AQ. Parnell, Joohwan Kim, Christopher D. Saunter, Gordon D. Love, and Martin S. Banks. 2016. Dynamic lens and monovision 3D displays to improve viewer comfort. *Opt. Express* 24, 11 (May 2016), 11808–11827.

A. Kaczorowski, G. S. Gordon, A. Palani, S. Czerniawski, and T. D. Wilkinson. 2015. Optimization-Based Adaptive Optical Correction for Holographic Projectors. *Journal of Display Technology* 11, 7 (July 2015), 596–603.

Andrzej Kaczorowski, George S. D. Gordon, and Timothy D. Wilkinson. 2016. Adaptive, spatially-varying aberration correction for real-time holographic projectors. *Opt. Express* 24, 14 (Jul 2016).

Robert Konrad, Emily A. Cooper, and Gordon Wetzstein. 2016. Novel Optical Configurations for Virtual Reality: Evaluating User Preference and Performance with Focus-tunable and Monovision Near-eye Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1211–1220.

G. Kramida. 2016. Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 22, 7 (July 2016), 1912–1931.

Douglas Lanman and David Luebke. 2013. Near-eye Light Field Displays. *ACM Trans. Graph.* 32, 6, Article 220 (Nov. 2013), 10 pages.

Seungjae Lee, Changwon Jang, Seokil Moon, Jaebum Cho, and Byoungho Lee. 2016. Additive Light Field Displays: Realization of Augmented Reality with Holographic Optical Elements. *ACM Trans. Graph.* 35, 4, Article 60 (July 2016), 13 pages.

Gang Li, Dukho Lee, Youngmo Jeong, Jaebum Cho, and Byoungho Lee. 2016. Holographic display for see-through augmented reality using mirror-lens holographic optical element. *Opt. Lett.* 41, 11 (Jun 2016).

S. Liu, H. Hua, and D. Cheng. 2010. A Novel Prototype for an Optical See-Through Head-Mounted Display with Addressable Focus Cues. *IEEE Transactions on Visualization and Computer Graphics* 16, 3 (May 2010), 381–393.

Gordon D. Love, David M. Hoffman, Philip J.W. Hands, James Gao, Andrew K. Kirby, and Martin S. Banks. 2009. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Opt. Express* 17, 18 (Aug 2009), 15716–15725.

Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke, and Henry Fuchs. 2014. Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses Using Defocused Point Light Sources. *ACM Trans. Graph.* 33, 4, Article 89 (July 2014), 11 pages.

Michal Makowski, Izabela Ducin, Karol Kakarenko, Jaroslaw Suszek, Maciej Sypek, and Andrzej Kolodziejczyk. 2012. Simple holographic projection in color. *Opt. Express* 20, 22 (Oct 2012).

Eunkyong Moon, Myeongjae Kim, Jinyoung Roh, Hwi Kim, and Joonku Hahn. 2014. Holographic head-mounted display with RGB light emitting diode light source. *Opt. Express* 22, 6 (Mar 2014), 6526–6534.

Rahul Narain, Rachel A. Albert, Abdullah Bulbul, Gregory J. Ward, Martin S. Banks, and James F. O'Brien. 2015. Optimal Presentation of Imagery with Focus Cues on Multi-plane Displays. *ACM Trans. Graph.* 34, 4, Article 59 (July 2015), 12 pages.

Nitish Padmanaban, Robert Konrad, Tal Stramer, Emily A. Cooper, and Gordon Wetzstein. 2017. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences* 114, 9 (2017), 2183–2188.

Vitor F. Pamplona, Manuel M. Oliveira, Daniel G. Aliaga, and Ramesh Raskar. 2012. Tailored Displays to Compensate for Visual Aberrations. *ACM Trans. Graph.* 31, 4, Article 81 (July 2012), 12 pages.

Yijun Qi, Chenliang Chang, and Jun Xia. 2016. Speckleless holographic display by complex modulation based on double-phase method. *Opt. Express* 24, 26 (Dec 2016), 30368–30378.

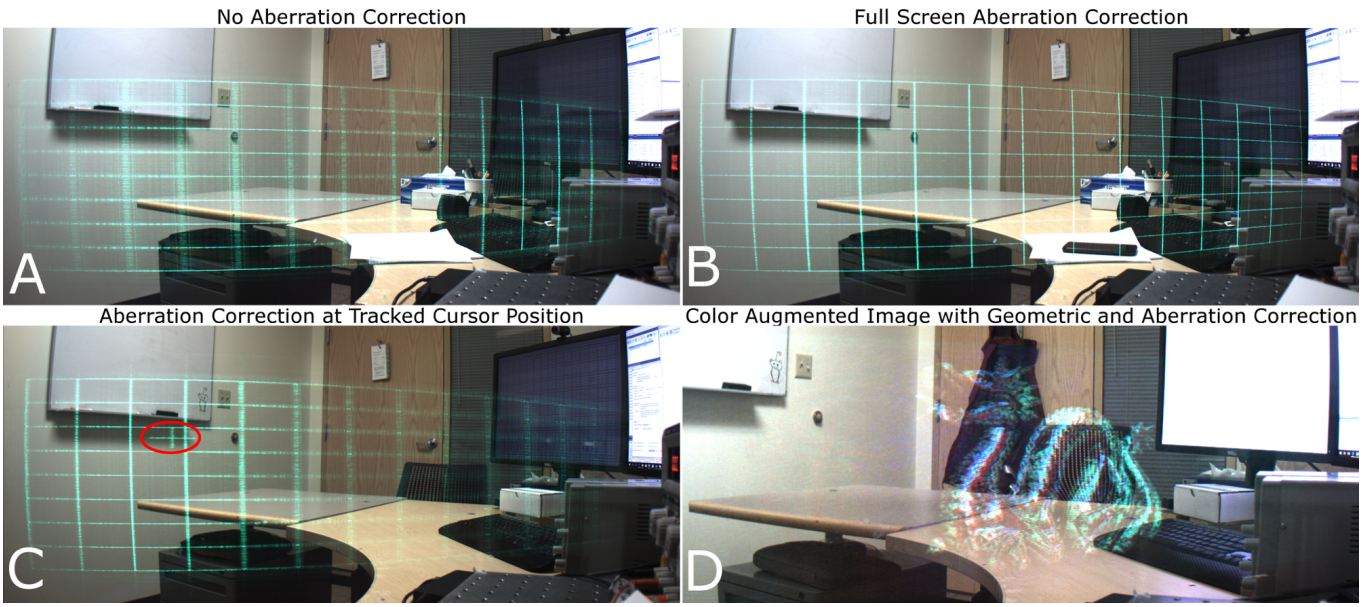

Fig. 15. Benchtop AR prototype results. *A)* Display with no aberration correction applied. *B)* Display with full screen, spatially varying aberration correction applied. *C)* Display with real-time aberration correction applied around a tracked cursor (circled in red). *D)* Full color augmented image with aberration and geometric corrections applied. Dragon model by Stanford Computer Graphics Laboratory.

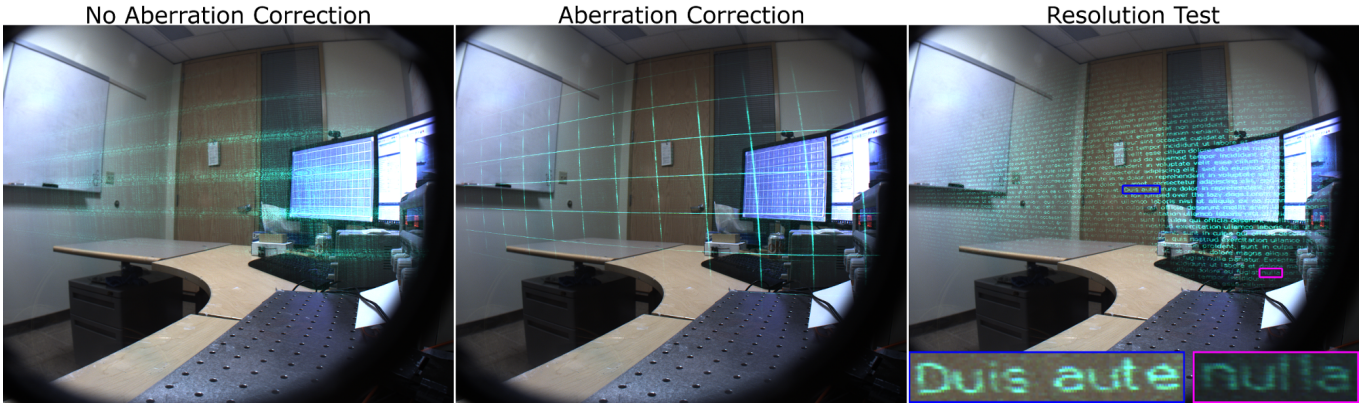

Fig. 16. Compact AR prototype results. *Left:* Display without aberration correction. *Center:* Display with aberration correction. *Right:* Display showing high resolution capability; the line weight of the text is one pixel wide. The bordered inset images show magnified regions.

Weidong Qu, Huarong Gu, Hao Zhang, and Qiaofeng Tan. 2015. Image magnification in lensless holographic projection using double-sampling Fresnel diffraction. *Appl. Opt.* 54, 34 (Dec 2015), 10018–10021.

Stephan Reichelt and Norbert Leister. 2013. Computational hologram synthesis and representation on spatial light modulators for real-time 3D holographic imaging. *Journal of Physics: Conference Series* 415, 1 (2013), 012038.

Brian T. Schowengerdt, Mrinal Murari, and Eric J. Seibel. 2010. 44.1: Volumetric Display using Scanned Fiber Array. *SID Symposium Digest of Technical Papers* 41, 1 (2010), 653–656.

P.W.M. Tsang and T. C. Poon. 2013. Novel method for converting digital Fresnel hologram to phase-only hologram based on bidirectional error diffusion. *Opt. Express* 21, 20 (Oct 2013).

P.W.M. Tsang and T.-C. Poon. 2015. Fast generation of digital holograms based on warping of the wavefront recording plane. *Opt. Express* 23, 6 (Mar 2015), 7667–7673.

P. W. M. Tsang and T. C. Poon. 2016. Review on the State-of-the-Art Technologies for Acquisition and Display of Digital Holograms. *IEEE Transactions on Industrial Informatics* 12, 3 (June 2016), 886–901.

Fahri Yaraş, Hoonjong Kang, and Levent Onural. 2010. State of the Art in Holographic Displays: A Survey. *J. Display Technol.* 6, 10 (Oct 2010).

Han-Ju Yeom, Hee-Jae Kim, Seong-Bok Kim, HuiJun Zhang, BoNi Li, Yeong-Min Ji, Sang-Hoo Kim, and Jae-Hyeung Park. 2015. 3D holographic head mounted display using holographic optical elements with astigmatism aberration compensation. *Opt. Express* 23, 25 (Dec 2015), 32025–32034.

Hiroshi Yoshikawa, Takeshi Yamaguchi, and Hiroki Uetake. 2016. Image quality evaluation and control of computer-generated holograms. *Proc. SPIE* 9771 (2016), 97710N–97710N–9.

Hao Zhang, Neil Collings, Jing Chen, Bill Crossland, Daping Chu, and Jinghui Xie. 2011. Full parallax three-dimensional display with occlusion effect using computer generated hologram. *Optical Engineering* 50, 7 (2011), 074003–074003–5.