

Holographic Glasses for Virtual Reality

JONGHYUN KIM, NVIDIA, USA and Stanford University, USA

MANU GOPAKUMAR, SUYEON CHOI, AND YIFAN PENG, Stanford University, USA

WARD LOPES, NVIDIA, USA

GORDON WETZSTEIN, Stanford University, USA

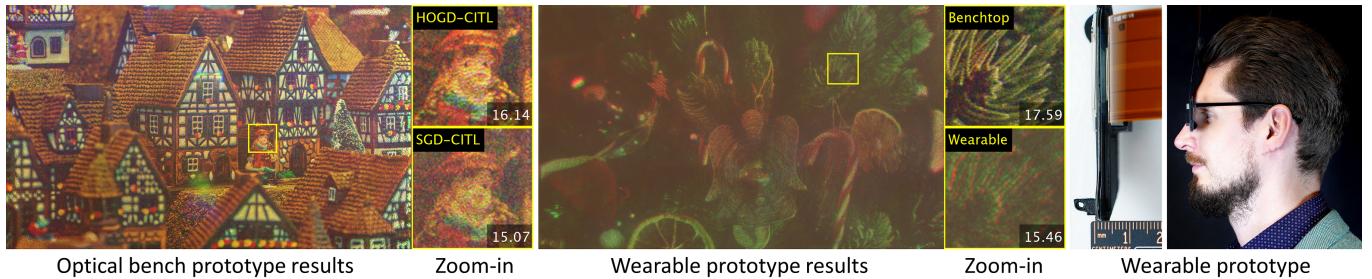


Fig. 1. *Left:* Captured results from the Holographic Glasses optical bench prototype. The bottom-right numbers indicate peak signal-to-noise ratios. *Center:* Captured results from the wearable prototype. *Right:* Images of the binocular wearable prototype, including a top-down view of one side of the prototype and side view on a human head. Holographic Glasses can provide full-color 3D holographic images with 2.5 mm thick optics.

We present Holographic Glasses, a holographic near-eye display system with an eyeglasses-like form factor for virtual reality. Holographic Glasses are composed of a pupil-replicating waveguide, a spatial light modulator, and a geometric phase lens to create holographic images in a lightweight and thin form factor. The proposed design can deliver full-color 3D holographic images using an optical stack of 2.5 mm thickness. A novel pupil-high-order gradient descent algorithm is presented for the correct phase calculation with the user's varying pupil size. We implement benchtop and wearable prototypes for testing. Our binocular wearable prototype supports 3D focus cues and provides a diagonal field of view of 22.8° with a 2.3 mm static eye box and additional capabilities of dynamic eye box with beam steering, while weighing only 60 g excluding the driving board.

CCS Concepts: • Hardware → Displays and imagers.

Additional Key Words and Phrases: near-eye display, holography

ACM Reference Format:

Jonghyun Kim, Manu Gopakumar, Suyeon Choi, and Yifan Peng, Ward Lopes, and Gordon Wetzstein. 2022. Holographic Glasses for Virtual Reality. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings (SIGGRAPH '22 Conference Proceedings)*, August 7–11, 2022, Vancouver, BC, Canada. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3528233.3530739>

1 INTRODUCTION

Virtual and augmented reality (VR/AR) systems promise unprecedented user experiences for computer graphics applications and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGGRAPH '22 Conference Proceedings, August 7–11, 2022, Vancouver, BC, Canada

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9337-9/22/08...\$15.00

<https://doi.org/10.1145/3528233.3530739>

beyond. A major barrier to widespread adoption of VR technology, however, is the bulky form factor of existing VR displays and the discomfort associated with that [PerkinsCoie 2019, 2021]. This problem stems from the magnifier principle of VR display optics, in which a lens enlarges the image of a small microdisplay. This design necessitates a relatively large distance between microdisplay and lens, making today's VR displays bulky and uncomfortable to wear.

The thickness of magnifier-based VR displays can be reduced by folding its optical path using (potentially holographic) "pancake" lenses [Bang et al. 2021; Maimone and Wang 2020; Narasimhan 2018] or waveguides [Kress 2020]. Other optical designs have also been explored to achieve thin displays form factors for VR/AR [Akşit et al. 2015; Lanman and Luebke 2013; Maimone et al. 2014; Ratcliff et al. 2020]. All of these approaches, however, are limited: either they restrict the display to presenting 2D images to each eye, potentially resulting in visual discomfort [Lambooij et al. 2009], or they offer only limited image resolution. Thus, no thin VR display exists today that also supports perceptually important focus cues.

We propose a holographic near-eye display system that offers 2D or 3D image presentation to each eye with a device thickness of just a few millimeters. Our system builds on recent ideas that use artificial intelligence techniques to improve the image quality and accelerate the computation of computer-generated holograms (CGH) (e.g., [Peng et al. 2020; Shi et al. 2021]). With our design and experimental prototypes, we hope to stimulate new research and engineering directions toward ultra-thin all-day-wearable VR displays with form factors comparable to conventional eyeglasses.

The specific contributions of our work include:

- A full-color 3D holographic near-eye display design with a thin and lightweight eyeglasses-like form factor that combines a geometric phase (GP) lens, a pupil-replicating waveguide, and a phase-only spatial light modulator (SLM) in a novel optical layout.

- Optical bench and wearable prototypes along with experimental demonstrations of 2D and 3D holographic images. The thickness of our wearable prototype, from the frontal surface to the SLM plane, is 2.5 mm and the weight of the binocular wearable prototype is 60 g (excluding the SLM driving board).
- The design and implementation of an algorithmic framework dubbed Pupil-HOGD for 3D holographic image synthesis with Holographic Glasses.
- An analysis of the design trade space for Holographic Glasses.

2 RELATED WORK

Thin optics for VR/AR displays. Commercial VR displays are based on the magnifier principle, where a lens enlarges the image of a microdisplay to create a virtual image that the user perceives at some distance [Cakmakci and Rolland 2006; Kress 2020]. Magnifiers are successful in offering wide fields of view with a large eyebox, but the large distance between lens and display results in bulky device form factors. Pancake lenses [Narasimhan 2018] fold the optical path to reduce this distance and holographic pancake lenses further reduce the thickness of the optical bulk [Cakmakci et al. 2021; Maimone and Wang 2020]. Most optical see-through AR displays also use optical folding, albeit using waveguides. Surveys of the optical principles underlying VR/AR displays outline further details of these systems [Cakmakci and Rolland 2006; Kress 2020; Xiong et al. 2021; Zhan et al. 2020].

Minimizing the thickness of VR/AR displays has been an active area of research. Lanman and Luebke [2013], for example, designed thin microlens-based near-eye light field displays. Radcliff et al. [2020] extended the FOV of these systems using curved microlenses. Variants of such a design that use pinhole arrays rather than microlenses have also been proposed [Akşit et al. 2015]. All of these near-eye light field displays, however, suffer from a spatio-angular resolution tradeoff that limits the resolution and number of views compared to other optical designs. Pinlight displays [Maimone et al. 2014] offer a wide FOV and a thin form factor for see-through AR displays but their resolution is limited by diffraction introduced by the required liquid crystal display. Lenslet VR combines microlenses with a folded optical path inspired by pancake lenses [Bang et al. 2021], but this system only offers 2D image generation capabilities whereas our system additionally aims at generating 3D images at multiple depths simultaneously. Our approach is unique in leveraging a holographic near-eye display design together with a thin waveguide to offer high-resolution 3D images within a few millimeters of thickness.

Holographic near-eye displays. Near-eye displays based on computer-generated holography have been extensively investigated in recent years (e.g., [Chakravarthula et al. 2019, 2020; Choi et al. 2021a,b; Glasner et al. 2014; Jang et al. 2018; Maimone et al. 2017; Padmanaban et al. 2019; Peng et al. 2020; Shi et al. 2021]), as recently reviewed by Chang et al. [2020]. These works primarily focus on improving the quality of holographic 2D or 3D images, or on speeding up the algorithm generating the holograms. Instead, our work addressed the problem of large VR display form factors with a novel systems design, building on some of the algorithmic advances from these previous works made. In particular, Gopakumar et al.’s recent

work [2021] allows the optical path from SLM to the user’s eye to be shortened by removing the need for an optical filter. Our system incorporates this idea, but further folds the illumination path and other optical components with the goal of minimizing the thickness of the display. Similar to other holographic displays, our system benefits from the ability to support both 2D and 3D imagery with focus cues at high image resolutions, which is necessary to mitigate the vergence-accommodation conflict and optimize user comfort in VR and AR [Lambooij et al. 2009; Shibata et al. 2011].

Near-eye displays with focus cues. Other display architectures also support focal cues, such as varifocal displays [Dunn et al. 2017; Konrad et al. 2016; Love et al. 2009; Padmanaban et al. 2017], multi-focal displays [Akeley et al. 2004; Chang et al. 2018; Liu et al. 2008; Rathinavel et al. 2018; Rolland et al. 2000], or light field near-eye displays [Hua and Javidi 2014; Huang et al. 2015; Jang et al. 2017; Lanman and Luebke 2013]. However, these technologies do not enable the thin device form factors our system does or they only offer a limited resolution. Detailed surveys of near-eye displays with focus cues include [Hua 2017; Koulieris et al. 2019; Kramida 2015].

3 HOLOGRAPHIC GLASSES DESIGN

3.1 System Components

Holographic Glasses comprise the following core components: a coherent light source that is coupled into a pupil-replicating waveguide, which provides the illumination for a phase-only SLM that is mounted on the waveguide in front of the user’s eye. This SLM creates a small image behind the device, which is magnified by a thin geometric phase (GP) lens. We describe these components next.

Pupil-replicating waveguide. A holographic near-eye display requires a (partially) coherent light source and an SLM, typically operating in phase-only mode. Most systems described in the literature (see Sec. 2) either use a beam splitter cube or off-axis illumination to create a planar or spherical source wave incident on the SLM. These designs require a lot of space, which we minimize by using a pupil-replicating waveguide designed for conventional optical see-through AR displays. This waveguide is very thin but not designed to provide planar or spherical waves, which we address by adopting modern approaches to CGH computation as described in the next section.

Holographic near-eye display. The phase-only SLM creates a 2D or 3D image at a small distance, which is slightly smaller than the focal length of the GP lens, behind the physical device (i.e., virtual-mode, see Supplementary). The SLM is mounted directly on the waveguide without an air gap to minimize the thickness of the display. This design is illustrated in Fig. 2.

Geometric phase (GP) lens and polarization control. A GP lens, also known as a Pancharatnam-Berry phase lens, is a polarization-dependent liquid crystal (LC) lens that works as a positive lens for a certain input beam polarization (see Supplementary). Because such GP lenses tend to be thin and lightweight, several previous near-eye display designs have been based on GP lenses [Lee et al. 2017; Moon et al. 2020; Yoo et al. 2020]. Nam et al. [2020] presented a GP lens-based holographic near-eye display and its aberration

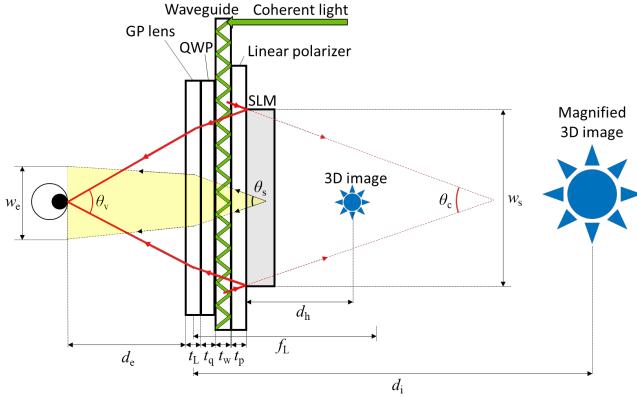


Fig. 2. Holographic Glasses schematic diagram.

compensation method. However, previous works were limited to AR applications with bulky form factors.

We use GP lenses as the eyepieces of Holographic Glasses. Some VR headsets adopt Fresnel lenses as their eyepieces [PerkinsCoie 2019, 2021], but we found that the saw-tooth structure of these optical elements create undesired interference artifacts when used with coherent light sources. Since most SLMs are also based on LCs and operate with linearly polarized input light, the polarization should be carefully considered throughout the optical path. Therefore, we mount a quarter wave plate (QWP) between the SLM and the GP lens to transform the linearly polarized input light to right-handed circularly polarized (RCP) light required by the GP lens. The lens in turn converts the RCP light to left-handed circularly polarized (LCP) light.

3.2 Holographic Glasses

System overview. Figure 2 shows the schematic diagram of Holographic Glasses. Coherent and collimated light with wavelength λ is coupled into and travels through a waveguide with thickness t_w and refractive index n_w . The light is then coupled out of the waveguide with a converging angle θ_c . After the light is polarized at the linear polarizer (thickness t_p), the light is then modulated at the SLM. The SLM has pixel pitch p_s , pixel number $N_x \times N_y$, width w_s , and height h_s . The modulated light travels back towards the user with a diffraction angle θ_s and meets the waveguide again. The grating will couple some light back into the waveguide (due to reciprocity) based on the diffraction efficiency, which is designed to be low. The remaining uncoupled light passes through a QWP with thickness t_q , and is then refracted at the GP lens with thickness t_L and focal length f_L . The chief rays create a viewpoint at an eye relief of d_e from the GP lens and offer a horizontal FOV θ_v . The eye box w_e depends on the diffraction angle θ_s , d_e , θ_v , and w_e can be calculated as follows:

$$d_e = f_L \frac{w_s - 2n_w t_w \tan(\theta_c/2)}{w_s + 2f_L \tan(\theta_c/2) - 2n_w t_w \tan(\theta_c/2)}, \quad (1)$$

$$\theta_v = 2 \tan^{-1} \left(\frac{w_s - n_w t_w \tan(\theta_c/2)}{2d_e} \right), \text{ and} \quad (2)$$

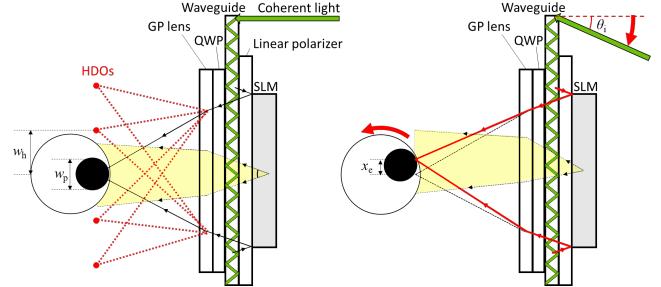


Fig. 3. *Left:* High-diffraction orders (HDOs) in Holographic Glasses. If the interval of HDO w_h is smaller than the pupil diameter w_p , then the HDOs should be considered during the phase generation (HOGD). *Right:* Dynamic eye box method in Holographic Glasses. The viewpoint can be modified on the plane by changing the incident light angle (θ_i).

$$w_e = 2 \tan(\theta_s/2) \frac{n_w t_w f_L + d_e f_L - n_w d_e t_w}{f_L}, \quad (3)$$

where the thickness of the linear polarizer t_p , the QWP t_q , and the GP lens t_L are neglected. The SLM creates a virtual 3D image located behind the SLM at a distance d_h . Then the GP lens magnifies the 3D image and moves it back to the perceivable distance d_i .

Waveguide illumination and phase uniformity. The out-coupled light has larger-than-normal deviations in phase uniformity based on differing path lengths through the waveguide, which is difficult to solve analytically. However, this scrambled phase can be compensated with a camera-in-the-loop (CITL) approach, by calculating the loss based on the captured image during phase optimization [Choi et al. 2021a,b; Peng et al. 2020]. The waveguide also scrambles the polarization, which can be corrected by the linear polarizer in front of the SLM.

High diffraction orders (HDOs) and dynamic eye box. Holographic Glasses have two distinct features which are not observed in conventional VR displays. The first one is HDOs. As shown on the left of Fig. 3, the periodic structure of the SLM pixels creates HDOs which, in the location of the pupil plane, repeat with an interval w_h . If w_h is smaller than the pupil diameter w_p , then we have to consider HDOs during the phase calculation process (high-order gradient descent, or HOGD) as introduced by Gopakumar [2021]. When the HDOs are well separated the user's eye pupil works as a "natural" optical filter. When the HDOs are overlapping, then one can use the pupil-HOGD algorithm (see Sec. 4.2) to create an optimized image.

The second feature stems from the pupil-replicating waveguide. Since the waveguide is designed to reproduce a light field with a range of incident light angles (θ_i), the direction of the entire SLM illumination can be controlled by the input beam direction. A small deflection of the input beam results in the dynamic eye box as shown on the right of Fig. 3. With an additional gaze tracker, the system could follow the gaze and move around the center lobe by simply changing the direction of the input beam.

Design trade-offs. Figure 4 shows the design trade space of Holographic Glasses. As shown in the two left graphs, the FOV θ_v gets

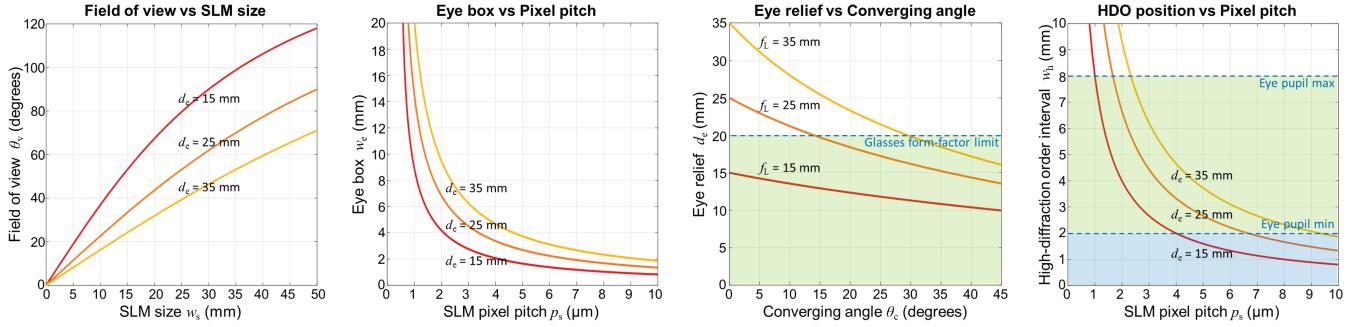


Fig. 4. Design trade space of Holographic Glasses. *Left:* Field of view (FOV, θ_v) vs. SLM size (w_s). Larger SLMs at a closer distance (i.e., shorter eye relief d_e) can cover larger FOV. *Center-left:* Eye box (w_e) vs. SLM pixel pitch (p_s). A smaller pixel pitch creates a larger diffraction angle and larger eye box. *Center-right:* Eye relief (d_e) vs. converging angle (θ_c). The converging angle of the out-coupled light can make eye relief smaller than the focal length of the eyepiece (f_L). The glasses-form factor should satisfy $d_e < 20$ mm condition. *Right:* High-diffraction order interval (w_h) vs. SLM pixel pitch (p_s). The smaller pixel pitch moves HDOs away from the center.

larger when the SLM size w_s gets bigger, and the eye box w_e gets larger when the SLM pixel pitch p_s gets smaller. So both the FOV and the eye box are limited by the characteristics of the SLM. To achieve a wearable form factor, the eye relief d_e should be less than 20 mm, as shown in the center-right graph of Fig. 4. A shorter focal length of the eyepiece f_L and a larger converging angle θ_c are beneficial. Although a short eye relief is crucial for a wearable form factor and a large FOV, a short eye relief also results in a small eye box and brings the high orders closer to the center. Additionally, the pupil works as a Fourier filter with a dynamic aperture based on scene brightness. If the HDO interval w_h is smaller than the smallest pupil diameter (2 mm, blue region), then the HDOs are always observed by the user. If w_h is within the pupil diameter range (green region), the HDOs could be perceived by the user based on the pupil diameter. When HDOs are separated enough (>8 mm), then we do not have to worry about the HDOs or the pupil diameter. Since state-of-the-art SLMs cannot satisfy this condition, we need a generalized phase generation algorithm considering the pupil diameter effect (see Sec. 4.2). This condition can be relaxed with an SLM with a smaller pixel pitch or by using the scene brightness to control the user's pupil diameter.

4 IMPLEMENTATION

4.1 Hardware Implementation

We implemented two Holographic Glasses prototypes – an optical bench prototype and a wearable prototype. The optical bench prototype was implemented without the pupil replicating waveguide and used a beam splitter instead to evaluate the effect of the waveguide and the holographic near-eye display. All captured results were taken by a FLIR Grasshopper 3 CCD (GS3-U3-123S6C-C, 4,096×3,000, 30 FPS) with a Canon EF 50 mm f/1.4 lens or a 35 mm f/2.0 lens. Color results are captured as separate exposures for each wavelength and combined in post-processing. We used 3 laser diodes (645 nm, 525 nm, 456 nm) for the coherent light source (see Supplementary A.1.1). Although the GP lens works most effectively with input RHS polarized light, it still introduces a certain amount of unwanted stray light. Thus, we attached another circular polarizer



Fig. 5. Holographic Glasses prototype photos. *Top-left:* Optical bench prototype. *Top-right:* Monocular wearable prototype for testing. *Bottom:* Binocular wearable prototype used in Fig. 1.

film to the GP lens for both prototypes to mitigate this negative effect. The total thickness and the weight are listed in Table 1.

Table 1. Weight and thickness of prototype components in g/mm. B and W denote benchtop and wearable prototype, and BS, WG, and LP mean beam splitter, waveguide, and linear polarizer, respectively. Note that this table is for the monocular prototype and the total thickness is from the frontal surface to the SLM plane including the gaps between components.

	GP lens	QWP	BS/WG	LP	SLM	Total
B	0.7/0.5	5.4/3.2	45/20	4/3.3	10.3/-	65.4/33
W	0.7/0.5	0.0/0.1	10.0/0.8	0.0/0.1	10.3/-	21.0/2.5

4.1.1 Optical bench prototype. The top-left photograph of Fig. 5 shows the optical bench prototype, which is composed of a HoloEye SLM (LETO-3, 1,920×1,080, $p_s=6.4\text{ }\mu\text{m}$), a Thorlabs 1-inch linear polarizer (LPVISE100-A), a Thorlabs beam splitter (BS016), a Thorlabs achromat QWP (AQWP10M-580), and an Edmund Optics GP lens (34-464, $f_L=50\text{ mm}$). All optical components are carefully aligned with a 3D-printed holder. The total weight without the 3D printed part and the SLM driving board was 65.4 g, and the thickness from the frontal surface to the SLM plane was 33 mm, including small gaps between optical components. Note that our optical bench prototype is substantially smaller than previous benchtop prototypes because of the GP lens and the virtual-mode holography.

4.1.2 Wearable prototype. The top-right photograph of Fig. 5 shows the monocular wearable prototype. The incident light angle θ_i is controlled by a mirror in front of the in-coupler. We used basically the same optical components of the bench prototype but facilitated a thinner form factor. The wearable prototype is composed of the same SLM, a Thorlabs dichroic film linear polarizer (LPVISE2X2), a Dispelix pupil replicating waveguide (DPX 50), an Edmund Optics film-type QWP (WP140HE), an Edmund Optics GP lens (14-778, $f_L=35\text{ mm}$). The total weight of the monocular wearable prototype without the 3D printed part and the SLM driving board was 21.0 g. The bottom photograph in Fig. 5 shows the binocular wearable prototype used in Fig. 1 and the total weight of the binocular wearable prototype *with* the 3D printed part and the temples was 60 g. The thickness from the frontal surface to the SLM plane was 2.5 mm.

4.2 Software Implementation

The phase holograms to display were synthesized with the HOGD and HOGD-CITL algorithms [Gopakumar et al. 2021], which are built on the SGD and CITL algorithms proposed by Peng et al. [2020]. The key is to model the propagation of high diffraction orders to enable high image quality without optically filtering out high orders. This HOGD propagation, as detailed in Eq. 4, is crucial for realizing a high image quality when the pupil collects light from multiple diffraction orders. In these equations, ϕ is the SLM phase pattern, p_s is the SLM pixel pitch, α is the set of orders to be optimized, λ is the wavelength, z is the distance between the SLM and target plane, and M_p is a pupil mask.

$$\begin{aligned} u(\phi; z) &= \iint U(f_x, f_y; \phi) A(f_x, f_y; z) e^{i2\pi(f_x x + f_y y)} df_x df_y, \\ U(f_x, f_y; \phi) &= \sum_{i,j \in \alpha} \mathcal{F}\{e^{i\phi}\} \left(f_x + \frac{i}{p_s}, f_y + \frac{j}{p_s} \right), \\ A(f_x, f_y; z) &= \mathcal{H}(f_x, f_y; z) \text{Sinc}(\pi f_x p_s) \text{Sinc}(\pi f_y p_s) M_p(f_x, f_y) \\ \mathcal{H}(f_x, f_y; z) &= \begin{cases} e^{i\frac{2\pi}{\lambda} \sqrt{1-(\lambda f_x)^2 - (\lambda f_y)^2} z}, & \text{if } \sqrt{f_x^2 + f_y^2} < \frac{1}{\lambda}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (4)$$

In these equations, u is the propagated wavefront, U is the frequency representation of the SLM wavefront with high orders, H is the ASM kernel in the frequency domain, and A is this ASM kernel with attenuation due to the pixel pitch of the SLM and the aperture of the pupil. The aperture of the pupil is accounted for with the masking term, M_p , which extends the HOGD algorithm [Gopakumar et al. 2021] to an algorithm we call *Pupil-HOGD*. In this algorithm,

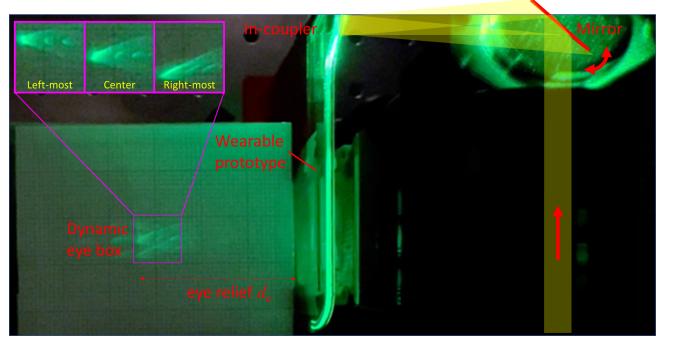


Fig. 6. Dynamic eye box capability experiment in the wearable prototype. The entire illumination direction can be controlled by a folding mirror located on the top right, which results in shifting the viewpoints at the eye relief plane. Note that this is the superposition of 3 captured images for the left-most, center and right-most dynamic viewpoints, showing a net 8 mm horizontal eye box at the 35 mm eye relief plane.

the pupil mask, M_p , enables phase patterns to be optimized while accounting for the wavefront filtering that is performed by the pupil. Here, pupil filtering is modeled for the case that $d_e = f_L$, where a pupil, with a diameter of w_p , acts as a circular filter in the Fourier domain with a diameter of $w_p/(\lambda f_L)$. With this, the pupil aperture can be modeled simply by constructing the mask M_p to be 1 for frequencies inside this circular filter and 0 otherwise.

Beyond modeling the pupil aperture, our work extends the HOGD and HOGD-CITL algorithms to generate phase patterns for RGBD content using the masked multiplane loss from Choi et al. [2021a] as outlined in Eq. 5. In these equations, a_{target} is the target amplitude and $D(x, y)$ is the target depth map. The desired scene is decomposed over J target planes using masks $m^{(j)}$, that are generated by quantizing the target depth map to the nearest depth $z^{(j)}$, in the multiplane decomposition

$$\begin{aligned} \arg \min_{\phi, s} \sum_{j=1}^J & \left\| \left(s \cdot |u(\phi; z^{(j)})| - a_{\text{target}} \right) \circ m^{(j)} \right\|_2^2 \\ m^{(j)}(x, y) &= \begin{cases} 1, & \text{if } |z^{(j)} - D(x, y)| < |z^{(k)} - D(x, y)|, \forall k \neq j \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (5)$$

Minimizing this objective produces phase patterns to display desired 3D content. For the multiplane HOGD algorithm, this objective is directly optimized with the Adam optimizer in Pytorch. As discussed by Gopakumar et al. [2021], the HOGD-CITL algorithm enhances this optimization by pairing the gradients of the simulated propagation with captured outputs of the holographic display. These algorithms are run on an Nvidia RTX3090 graphics card with alternating optimization steps for content at each target plane to limit memory usage. The calculation time of a single phase pattern generation with SGD-CITL and HOGD-CITL for 500 iterations was 10 minutes and 20 minutes with the PC, respectively.

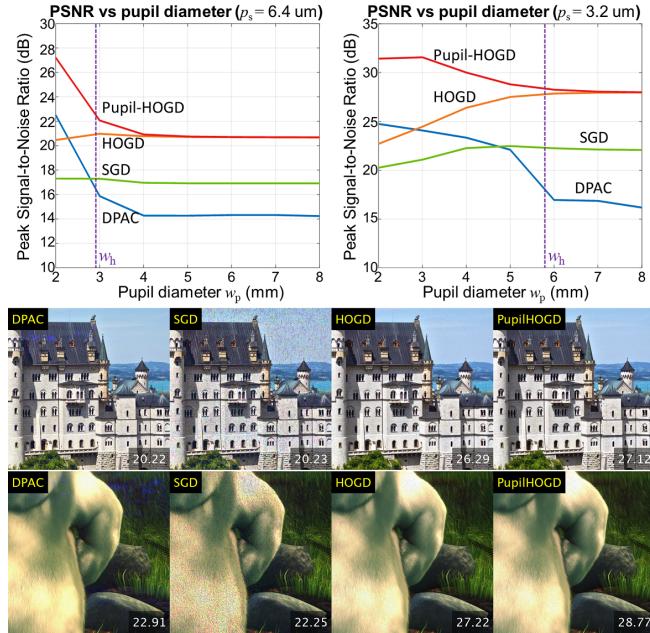


Fig. 7. Top: Simulated comparisons of image quality produced with different computer-generated holography algorithms for a range of different pupil sizes and two different SLM pixel pitches. Here w_h indicates the pupil size at which the high orders start to be captured by the pupil. For the smaller pixel pitch, w_h is increased. For both pixel pitches, our Pupil-HOGD algorithm, which models the pupil aperture, produces the best image quality at all pupil sizes. Bottom: The simulated images of the various algorithms with a 5 mm pupil diameter and $3.2 \mu\text{m}$ pixel pitch.

5 DISPLAY ASSESSMENT

5.1 Display Characteristics

The FOV of the benchtop prototype and wearable prototypes were 16.1° and 22.8° diagonal, respectively. The static eye box of each prototype was 3.4 mm and 2.3 mm. The dynamic eye box of the wearable prototype was measured from the top with a 1 mm grid diffuser. Figure 6 shows the superposition of captured images for left-most, center, and right-most dynamic viewpoints in the wearable prototype. By tilting the folding mirror based on the user’s gaze, the dynamic viewpoints can cover 8 mm horizontal eye box at the 35 mm eye relief plane, as shown in the close-up photos. The mirror part can be minimized with a lens shifter or multiple laser diodes in front of the in-coupler.

5.2 Pupil-HOGD algorithm

Figure 7 shows simulated comparisons of image quality produced with different computer-generated holography algorithms for different pupil sizes. For both pixel pitches, the novel Pupil-HOGD algorithm, which models the pupil aperture, produces the best image quality at all pupil sizes. When the high orders are captured by the pupil, the HOGD algorithm performs second best because it can optimize these high orders but does not properly model the partial filtering of these orders due to the pupil. The SGD algorithm

performs third best when the high orders are captured since it can utilize the full central order but cannot model the high orders. Notably, even when the pupil completely filters out the high orders, the SGD algorithm has limited performance because it does not model the partial filtering of the central order. The double phase amplitude coding (DPAC) algorithm performs well when the pupil size is small because it concentrates its signal in low frequencies and sends unwanted light to high frequencies, but performs the worst when the pupil is large.

5.3 Captured Results

We captured 2D and multi-plane 3D results with the benchtop and wearable prototypes. Since the size and the aperture of the camera lens do not exactly match a human pupil, we tested two algorithms: SGD and HOGD. The lens aperture was open wide to test the effect of HDOs. The focus was at 1.5 m for 2D results. The phase patterns were optimized with CITL. Figure 8 shows the captured benchtop 2D and 3D results. The 2D image quality and the color representation were very good considering the small form factor of the benchtop prototype. The red arrow shows the artifact because of the stray light from the SLM due to the GP lens loss. The HOGD-CITL algorithm presented better image quality and higher contrast, as expected.

Figure 8 shows captured multi-plane 3D results from the benchtop prototype. The house, duck, angel, and leopard scenes are located at 50 cm, 80 cm, 1.5 m, and 3 m from the camera, respectively. The results show correct in-focus images in the different planes which can induce the user’s accommodation reflex. However, the out-of-focus blur was different from the natural blur. This can be optionally improved with additional regularization of the out-of-focus regions, as discussed in [Choi et al. 2021a].

The 2D results from the wearable prototype showed worse image quality and contrast, as shown in Fig. 1. It is mainly because of the scrambled phase of the waveguide illumination. Even with the limited image quality, our wearable prototype showed the first captured holographic display results with a scrambled-phase illumination. Further, the mismatch between the waveguide and the implemented coherent light source provides non-uniform illumination in amplitude, which could be improved with a different grating design.

6 DISCUSSION

6.1 Limitations and Future Work

Field of view (FOV). The implemented FOV of the wearable prototype was 22.8° diagonal, which is far smaller than commercially available VR/AR displays. However, the FOV was mainly limited by the size of the available SLM and the focal length of the GP lens, both of which could be improved with different components. For example, the focal length can be halved without significantly increasing the total thickness by stacking two identical GP lenses and a circular polarizer [Moon et al. 2020]. With a 2-inch SLM and a 15 mm focal length GP lens, we could achieve a monocular FOV of up to 120° .

Eye box. The static eye box of the holographic glasses is mainly determined by the pixel pitch of the SLM. It is inversely proportional to the pixel pitch, and a sufficient eye box for a single user (4~8 mm)



Fig. 8. Top: Captured 2D results from the optical bench prototype with the HOGD-CITL algorithm. The red arrow shows the artifact due to the GP lens loss. The close-up photo shows the comparison of HOGD-CITL and SGD-CITL. Center/Bottom: Captured multi-plane results from the optical bench prototype with 50 cm (Center) and 3 m (bottom) focus settings. The house and leopard scenes are located at a distance of 50 cm and 3 m, respectively.

can be achieved with a $1\sim 2 \mu\text{m}$ pixel pitch. Compared to the $3.74 \mu\text{m}$ pixel pitch of Holoeye's 4k GAEA SLM, that condition requires additional engineering efforts. However, the dynamic eye box technique could successfully address this challenge. The dynamic eye box range is not limited by the pixel pitch but is determined by the material and thickness of the waveguide. Unlike previous dynamic eye box approaches based on Maxwellian view displays, [Jang et al. 2018, 2017; Kim et al. 2019], Holographic Glasses create a relatively wide static eye box, so we do not have to precisely redirect the viewpoint to the pupil center at all times. Also, the required moving part is smaller than that used in the previous work. The input beam direction could be controlled by a fast steering mirror or a lens shifter in front of the in-coupler, which relieves the system requirements. For example in our wearable prototype, the 20° saccade requires 5.3° steering of the input beam within 50 ms saccadic suppression [Ibbotson and Cloherty 2009; Matin 1974], so the required angular

velocity of the mirror would be $103^\circ/\text{s}$. The state-of-the-art fast steering mirror satisfies this requirement by a large margin. Finally, the waveguide has less angular selectivity and can cover a wider range of input k -vectors than the holographic optical element image combiners. The maximum incident light angle $\theta_{i,\max}$ is determined by the waveguide refractive index and the thickness t_w , and it is usually enough to cover the user's eye movement.

Wearable prototype image quality. In this paper, we used a commercially available diffractive optical element-based waveguide which is designed for incoherent AR displays. It provides non-uniform illumination as well as a somewhat scrambled phase, which degrades the quality of wearable prototype results in Fig. 1. This could be improved by extracting the illumination intensity and phase map using CITL model training [Choi et al. 2022, 2021a; Peng et al. 2020], but is out of the scope of this paper. With a waveguide designed for Holographic Glasses, the image quality of the wearable prototype could be improved. One can design a holographic optical element-based waveguide with a converging output beam ($\theta_c > 0$) as presented by Maimone and Wang [2020], which can also present higher efficiency with the wavelength matching. Further, we presented only 2D results from the wearable prototype. However, the same techniques used for the benchtop prototype could be used to create 3D results in the wearable prototype since the waveguide would not change the principle. Further, the 2D image plane distance is not a designed parameter but a certain (1.5 m) distance that we chose during the experiment, which shows that the generated phase pattern can form a holographic image in the desired plane.

Pupil diameter measurement and control. The pupil-HOGD algorithm is the correct way to deliver crisp holographic images to the user's eye. However, precise measurement of the pupil diameter may be necessary to achieve the best image quality. A commercially available infrared gaze tracker can capture and measure the pupil diameter. However, the pupil diameter is also modified by the scene intensity, due to the pupillary reflex [Watson and Yellott 2012]. For example, if the target scene is much brighter than the previous frame, then the pupil-HOGD would not work since the new frame will reduce the pupil size. On the other hand, we do not have to precisely measure the user's pupil diameter every time, if we can calibrate the user's pupillary reflex as a function of the scene intensity. Previous research papers including [Pamplona et al. 2009] have introduced a physiologically-based model for pupil light reflex. This model can be calibrated once per user and integrated into the pupil-HOGD as a function of the scene brightness for an optimized viewing experience. And finally, the simulated results show that the Pupil-HOGD with less than 0.5 mm pupil diameter error always provides better image quality than the naive HOGD algorithm, which significantly relaxes the required accuracy and latency (see Supplementary C.2).

System integration. In this paper, we only covered monocular results. Real-time phase pattern calculation was not implemented, but maybe possible as discussed in the literature, e.g. [Peng et al. 2020; Shi et al. 2021]. Speckle could also be reduced, for example with partially coherent light sources [Peng et al. 2021]. The eye tracking system and pupil diameter measurement system were not

integrated into the prototype. The 3-color coherent light source can be minimized and included in a small form factor with further engineering efforts. Such efforts, however, are outside of this paper's scope.

6.2 Conclusion

In this paper, we presented a 3D holographic near-eye display in a glasses form factor. Holographic Glasses is based on a novel optical design combining a GP lens, a pupil replicating waveguide, and a virtual-mode holographic display. With an idea utilizing the user's pupil as a natural Fourier filter as well as the Pupil-HOGD algorithm considering the HDOs and the pupil size, a true glasses-form factor holographic VR was presented for the first time.

ACKNOWLEDGMENTS

The authors would like to thank David Luebke for helpful discussions and advice. Stanford researchers on this project were in part supported by a Kwanjeong Scholarship, a Stanford SGF, Intel, NSF (award 1839974), a PECASE by the ARO (W911NF-19-1-0120), and Sony.

REFERENCES

- Kurt Akeley, Simon J Watt, Ahna Reza Girshick, and Martin S Banks. 2004. A stereo display prototype with multiple focal distances. *ACM transactions on graphics (TOG)* 23, 3 (2004), 804–813.
- Kaan Akşit, Jan Kautz, and David Luebke. 2015. Slim near-eye display using pinhole aperture arrays. *Applied optics* 54, 11 (2015), 3422–3427.
- Kiseung Bang, Youngjin Jo, Minseok Chae, and Byoungcho Lee. 2021. Lenslet VR: Thin, Flat and Wide-FOV Virtual Reality Display Using Fresnel Lens and Lenslet Array. *IEEE TVCG* 27, 5 (2021), 2545–2554.
- Ozan Cakmakci, Yi Qin, Peter Bosel, and Gordon Wetzstein. 2021. Holographic pancake optics for thin and lightweight optical see-through augmented reality. *Optics Express* 29, 22 (2021), 35206–35215.
- Ozan Cakmakci and Jannick Rolland. 2006. Head-worn displays: a review. *Journal of Display Technology* 2, 3 (2006), 199–216.
- Praneeth Chakravarthula, Yifan Peng, Joel Kollin, Henry Fuchs, and Felix Heide. 2019. Wirtinger holography for near-eye displays. *ACM Trans. Graph.* 38, 6 (2019), 1–13.
- Praneeth Chakravarthula, Ethan Tseng, Tarun Srivastava, Henry Fuchs, and Felix Heide. 2020. Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. *ACM Transactions on Graphics (TOG)* 39, 6 (2020), 1–18.
- Chenliang Chang, Kiseung Bang, Gordon Wetzstein, Byoungcho Lee, and Liang Gao. 2020. Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective. *Optica* 7, 11 (2020), 1563–1578.
- Jen-Hao Rick Chang, BVK Vijaya Kumar, and Aswin C Sankaranarayanan. 2018. Towards multifocal displays with dense focal stacks. *ACM Transactions on Graphics (TOG)* 37, 6 (2018), 1–13.
- Suyeon Choi, Manu Gopakumar, Yifan Peng, Jonghyun Kim, Matthew O'Toole, and Gordon Wetzstein. 2022. Time-multiplexed Neural Holography: A Flexible Framework for Holographic Near-eye Displays with Fast Heavily-quantized Spatial Light Modulators. In *Proceedings of the ACM SIGGRAPH*.
- Suyeon Choi, Manu Gopakumar, Yifan Peng, Jonghyun Kim, and Gordon Wetzstein. 2021a. Neural 3D holography: Learning accurate wave propagation models for 3D holographic virtual and augmented reality displays. *ACM Transactions on Graphics (TOG)* 40, 6 (2021), 1–12.
- Suyeon Choi, Jonghyun Kim, Yifan Peng, and Gordon Wetzstein. 2021b. Optimizing image quality for holographic near-eye displays with michelson holography. *Optica* 8, 2 (2021), 143–146.
- David Dunn, Cary Tippets, Kent Torell, Petr Kellnhofer, Kaan Akşit, Piotr Didyk, Karol Myszkowski, David Luebke, and Henry Fuchs. 2017. Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors. *IEEE Trans. on Visualization and Computer Graphics* 23, 4 (2017), 1322–1331.
- Daniel Glasner, Todd Zickler, and Anat Levin. 2014. A reflectance display. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 1–12.
- Manu Gopakumar, Jonghyun Kim, Suyeon Choi, Yifan Peng, and Gordon Wetzstein. 2021. Unfiltered holography: optimizing high diffraction orders without optical filtering for compact holographic displays. *Optics Letters* 46, 23 (2021), 5822–5825.
- Hong Hua. 2017. Enabling focus cues in head-mounted displays. *Proc. IEEE* 105, 5 (2017), 805–824.
- Hong Hua and Bahram Javidi. 2014. A 3D integral imaging optical see-through head-mounted display. *Opt. Express* 22, 11 (2014), 13484–13491.
- Fu-Chung Huang, Kevin Chen, and Gordon Wetzstein. 2015. The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Transactions on Graphics (TOG)* 34, 4 (2015), 1–12.
- M. Ibbotson and S. Cloherty. 2009. Visual perception: saccadic omission, suppression or temporal masking? *Current Biology* 19, 12 (2009).
- Changwon Jang, Kiseung Bang, Gang Li, and Byoungcho Lee. 2018. Holographic near-eye display with expanded eye-box. *ACM Transactions on Graphics (TOG)* 37, 6 (2018), 1–14.
- Changwon Jang, Kiseung Bang, Seokil Moon, Jonghyun Kim, Seungjae Lee, and Byoungcho Lee. 2017. Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 1–13.
- Jonghyun Kim, Youngmo Jeong, Michael Stengel, Kaan Akşit, Rachel Albert, Ben Boudaoud, Trey Greer, Joohwan Kim, Ward Lopes, Zander Majercik, Peter Shirley, Josep Spjut, Morgan McGuire, and David Luebke. 2019. Foveated AR: Dynamically-Foveated Augmented Reality Display. *ACM Trans. Graph.* 38, 4, Article 99 (2019).
- 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*. 1211–1220.
- George Alex Koulieris, Kaan Akşit, Michael Stengel, Rafal K Mantiuk, Katerina Mania, and Christian Richardt. 2019. Near-eye display and tracking technologies for virtual and augmented reality. In *Computer Graphics Forum*, Vol. 38. Wiley Online Library, 493–519.
- Gregory Kramida. 2015. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics* 22, 7 (2015), 1912–1931.
- Bernard C Kress. 2020. Optical architectures for augmented-, virtual-, and mixed-reality headsets. (2020).
- Marc Lambooij, Marten Fortuin, Ingrid Heynderickx, and Wijnand IJsselsteijn. 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of imaging science and technology* 53, 3 (2009), 30201–1.
- Douglas Lanman and David Luebke. 2013. Near-eye light field displays. *ACM Transactions on Graphics (TOG)* 32, 6 (2013), 1–10.
- Yun-Han Lee, Guanjun Tan, Tao Zhan, Yishi Weng, Guigeng Liu, Fangwang Gou, Fenglin Peng, Nelson V Tabiryan, Sebastian Gauza, and Shin-Tson Wu. 2017. Recent progress in Pancharatnam-Berry phase optical elements and the applications for virtual/augmented realities. *Optical Data Processing and Storage* 3, 1 (2017), 79–88.
- Sheng Liu, Dewen Cheng, and Hong Hua. 2008. An optical see-through head mounted display with addressable focal planes. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE, 33–42.
- Gordon D Love, David M Hoffman, Philip JW Hands, James Gao, Andrew K Kirby, and Martin S Banks. 2009. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics express* 17, 18 (2009), 15716–15725.
- Andrew Maimone, Andreas Georgiou, and Joel S Kollin. 2017. Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.* 36, 4 (2017), 1–16.
- 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. (SIGGRAPH)* 33, 4, Article 89 (Jul 2014), 11 pages.
- Andrew Maimone and Junren Wang. 2020. Holographic optics for thin and lightweight virtual reality. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 67–1.
- E. Matin. 1974. Saccadic suppression: a review and an analysis. *Psychological bulletin* 81, 12 (1974).
- Seokil Moon, Seung-Woo Nam, Youngmo Jeong, Chang-Kun Lee, Hong-Seok Lee, and Byoungcho Lee. 2020. Compact augmented reality combiner using Pancharatnam-Berry phase lens. *IEEE Photonics Technology Letters* 32, 5 (2020), 235–238.
- Seung-Woo Nam, Seokil Moon, Byoungcho Lee, Dongyeon Kim, Seungjae Lee, Chang-Kun Lee, and Byoungcho Lee. 2020. Aberration-corrected full-color holographic augmented reality near-eye display using a Pancharatnam-Berry phase lens. *Optics Express* 28, 21 (2020), 30836–30850.
- Bharathwaj Appan Narasimhan. 2018. Ultra-Compact pancake optics based on Thin Eyes super-resolution technology for virtual reality headsets. In *Digital Optics for Immersive Displays*, Vol. 10676. International Society for Optics and Photonics, 106761G.
- Nitish Padmanabhan, 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. <https://doi.org/10.1073/pnas.1617251114> arXiv:<https://www.pnas.org/content/114/9/2183.full.pdf>
- Nitish Padmanabhan, Yifan Peng, and Gordon Wetzstein. 2019. Holographic near-eye displays based on overlap-add stereograms. *ACM Trans. Graph.* 38, 6 (2019), 1–13.
- Vitor F Pamplona, Manuel M Oliveira, and Gladimir VG Baranowski. 2009. Photorealistic models for pupil light reflex and iridal pattern deformation. *ACM Transactions on*

- Graphics (TOG)* 28, 4 (2009), 1–12.
- Yifan Peng, Suyeon Choi, Jonghyun Kim, and Gordon Wetzstein. 2021. Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration. *Science advances* 7, 46 (2021), eabg5040.
- Yifan Peng, Suyeon Choi, Nitish Padmanaban, and Gordon Wetzstein. 2020. Neural holography with camera-in-the-loop training. *ACM Transactions on Graphics (TOG)* 39, 6 (2020), 1–14.
- PerkinsCoie. 2019. *2019 Augmented and Virtual Reality Survey Report*. <https://www.perkinscoie.com/images/content/2/1/v4/218679/2019-VR-AR-Survey-Digital-v1.pdf>
- PerkinsCoie. 2021. *XR Industry Insider 2021 XR Survey: Industry Insights into the Future of Immersive Technology*. <https://www.perkinscoie.com/content/designinteractive/xr2021/>
- Joshua Ratcliff, Alexey Supikov, Santiago Alfaro, and Ronald Azuma. 2020. ThinVR: Heterogeneous microlens arrays for compact, 180 degree FOV VR near-eye displays. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 1981–1990.
- Kishore Rathinavel, Hanpeng Wang, Alex Blate, and Henry Fuchs. 2018. An extended depth-at-field volumetric near-eye augmented reality display. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2857–2866.
- Jannick P Rolland, Myron W Krueger, and Alexei Goon. 2000. Multifocal planes head-mounted displays. *Applied Optics* 39, 19 (2000), 3209–3215.
- Liang Shi, Beichen Li, Changil Kim, Petr Kellnhofer, and Wojciech Matusik. 2021. Towards real-time photorealistic 3D holography with deep neural networks. *Nature* 591, 7849 (2021), 234–239.
- Takashi Shibata, Joohwan Kim, David M Hoffman, and Martin S Banks. 2011. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision* 11, 8 (2011), 11–11.
- Andrew B Watson and John I Yellott. 2012. A unified formula for light-adapted pupil size. *Journal of vision* 12, 10 (2012), 12–12.
- Jianghao Xiong, En-Lin Hsiang, Ziqian He, Tao Zhan, and Shin-Tson Wu. 2021. Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light: Science & Applications* 10, 1 (2021), 1–30.
- Chanhyung Yoo, Jianghao Xiong, Seokil Moon, Dongheon Yoo, Chang-Kun Lee, Shin-Tson Wu, and Byounghee Lee. 2020. Foveated display system based on a doublet geometric phase lens. *Optics Express* 28, 16 (2020), 23690–23702.
- Tao Zhan, Kun Yin, Jianghao Xiong, Ziqian He, and Shin-Tson Wu. 2020. Augmented reality and virtual reality displays: Perspectives and challenges. *Iscience* (2020), 101397.