Pupil-aware Holography

PRANEETH CHAKRAVARTHULA and SEUNG-HWAN BAEK, Princeton University, USA
FLORIAN SCHIFFERS, Reality Labs Research, Meta, USA
ETHAN TSENG, Princeton University, USA
GRACE KUO, ANDREW MAIMONE, NATHAN MATSUDA, OLIVER COSSAIRT, and DOUGLAS LANMAN, Reality Labs Research, Meta, USA
FELIX HEIDE, Princeton University, USA

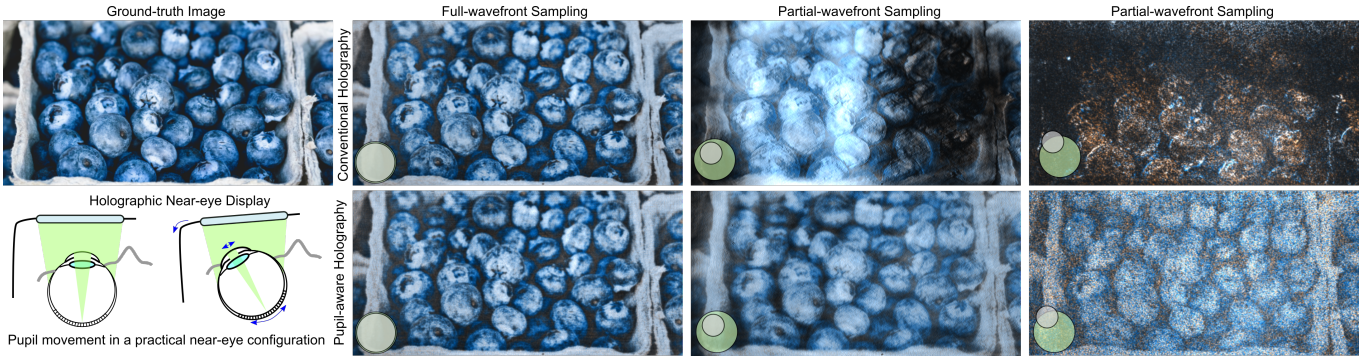


Fig. 1. *Pupil-aware holography on experimental hardware.* Conventional holographic displays often use a large lens to acquire the entire wavefront for high-fidelity reconstructions, but resulting in a tiny effective eyebox. However, near-eye holographic displays pose a unique problem in that the wavefront can only be partially sampled by the moving eye pupil with an unknown location and diameter at any given instant. This results in catastrophic failures such as complete loss of image on existing holographic displays (top-right). In this paper, we present the first pupil-aware near-eye holography framework that identifies and addresses this pupil-dependency problem, achieving robust reconstructions for arbitrarily partially sampled wavefronts, even at the edge of the eyebox (bottom-right). All of the results shown in this figure are acquired on an experimental hardware prototype.

Holographic displays promise to deliver unprecedented display capabilities in augmented reality applications, featuring a wide field of view, wide color gamut, spatial resolution, and depth cues all in a compact form factor. While emerging holographic display approaches have been successful in achieving large étendue and high image quality as seen by a camera, the large étendue also reveals a problem that makes existing displays impractical: the sampling of the holographic field by the eye pupil. Existing methods have not investigated this issue due to the lack of displays with large enough étendue, and, as such, they suffer from severe artifacts with varying eye pupil size and location.

We show that the holographic field as sampled by the eye pupil is highly varying for existing display setups, and we propose pupil-aware holography that maximizes the perceptual image quality irrespective of the size, location, and orientation of the eye pupil in a near-eye holographic display. We validate the proposed approach both in simulations and on a prototype holographic display and show that our method eliminates severe artifacts and significantly outperforms existing approaches.

0730-0301/2022/12-ART212 $15.00
https://doi.org/10.1145/3550454.3555508

Additional Key Words and Phrases: computational optics, holography

ACM Reference Format:
Praneeth Chakravarthula, Seung-Hwan Baek, Florian Schiffers, Ethan Tseng, Grace Kuo, Andrew Maimone, Nathan Matsuda, Oliver Cossairt,, Douglas Lanman, and Felix Heide. 2022. Pupil-aware Holography. *ACM Trans. Graph.* 41, 6, Article 212 (December 2022), 15 pages. https://doi.org/10.1145/3550454.3555508

1 INTRODUCTION

Augmented and virtual reality is emerging as a future technology with the potential to solve long standing challenges in human-computer interaction across domains, enabling applications as diverse as telepresence, surgical training, and automotive display. The key enabler of immersive augmented and virtual reality (AR/VR) are near-eye displays that are ultra compact and offer a wide field of view, high image quality and natural depth cues. Today, holographic displays are the only display technology that promises such unprecedented capabilities.

In a holographic display, an input wave field is modulated typically by a phase-only spatial light modulator (SLM). The modulated wavefront then propagates to a distance to create the desired image as an interference pattern. This image formation stands in contrast with today's ray-based displays such as LCD or LED displays. The

control over the entire wavefront of light enables holographic displays to create images with drastically reduced optical stacks by encoding much of the optics into the SLM phase pattern [Maimone and Wang 2020; Wakunami et al. 2016], thereby enabling the compact form factors and multi-focal capabilities required for near-eye display applications.

Recently, researchers have successfully demonstrated holographic displays that achieve image quality almost matching that of conventional displays [Chakravarthula et al. 2020; Peng et al. 2020; Shi et al. 2021]. These methods lift restrictions on holographic phase computation by using neural network predictors and cameras in-the-loop to calibrate the display setups. Moreover, today's holographic displays are subject to limited étendue, which means, that for 1K SLMs, one must heavily trade-off the eyebox (exit pupil) size for FoV. Emerging large étendue displays are also starting to address this limitation, either naively with smaller SLM pixel pitch < 4μm, such as the Holoeye GAEA-2 4K SLM, or using étendue-expanding diffractive elements [Kuo et al. 2020]. Although this new breed of displays holds the promise of practical high-quality holographic displays with a wide FOV and large eyebox simultaneously, and without the need for eye tracking or pupil steering, unfortunately, the large étendue comes at the cost of a high degree of noise that is pupil-dependent with varying pupil position and size – this pupil variance, unobserved in low étendue setups makes existing large étendue displays impractical. Specifically, even with an ideal modulator at 1 billion pixels and perfect phase and amplitude modulation, existing holographic displays are fundamentally subject to the problem of pupil variance.

The human eye pupil swivels over a large area due to factors ranging from involuntary saccades to voluntary gaze changes of the viewer [Bahill et al. 1975], mandating a wide eyebox. However, a large eyebox comes at the cost of image quality that is not uniform across the entire eyebox and drastically suffers with pupil sampling over the entire eyebox due to partial wavefront observations as shown in Fig 1. Depending on the size and orientation of the pupil, these perceptual artifacts can be widely different and quickly change. A Maxwellian-style display with a tiny eyebox in combination with eye tracking is one option to partly mitigate this problem [Maimone et al. 2017]. However, should the latency of the eye tracker do not match the eyebox movement, the user would perceive no image [Jang et al. 2019]. While recent large étendue approaches, in theory, lift the need for eye tracking, the pupil variance in these emerging systems has not been investigated in the past as existing displays did not provide enough étendue to observe this issue which is essential to make such displays practical.

In this work, we investigate the *unexplored problem of image quality dependence on eye pupil* for existing displays, and *propose a holographic phase retrieval method* that is aware of the different pupil states, lifting the dependency of the image quality on the exit pupil in existing holography systems. The proposed system does not rely on eye-tracking but instead optimizes the phase pattern on the SLM to produce consistent projections independently of the pupil location and size. To investigate pupil variance in the presence of large étendue, instead of adopting an étendue expanding element [2020], which mandates a complex calibration and alignment procedure, we use a system with limited étendue to emulate a small part of the FOV

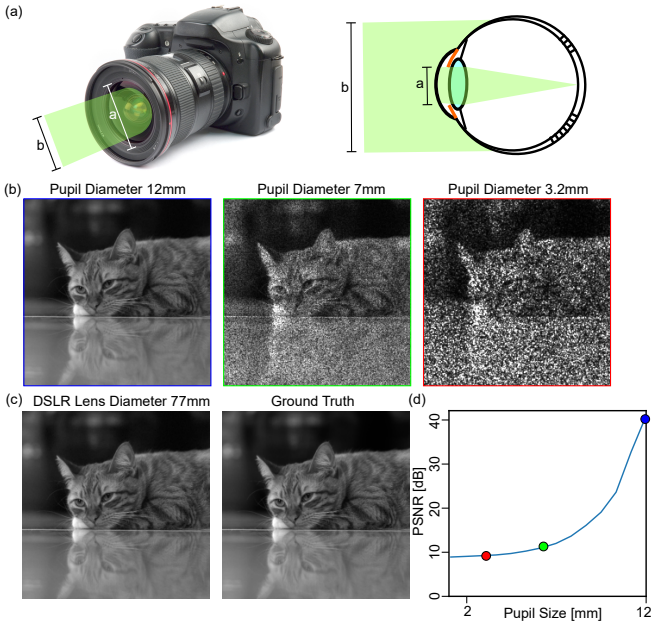


Fig. 2. (a) Compound DSLR lenses with large aperture a are capable of capturing the entire light bundle with diameter b emitted from a holographic display, as opposed to a human eye. (c) This results in high-fidelity holographic reconstructions on the imaging sensor. However, human eyes have limited- and variable-size pupils that only partially sample the incoming wavefront to the eyes. (b) & (d) As a result of the partial wavefront sampling due to an unknown pupil state, the projected image quality degrades, as a conventional computer-generated holography assumes that all of the wavefront is properly captured. Note that the low-contrast region of the cat's reflection on the floor is close to being invisible for smaller pupil diameters.

of a large étendue system (with large FOV and eyebox). This allows us to prototype the proposed method with affordable SLMs using a larger eyebox and smaller FOV system, and investigate pupil effects across the eyebox. Note that simulating a large étendue display on a low étendue display is a valid experiment as it just has a smaller FOV. However, the proposed algorithmic approach only considers the image fidelity across the eyebox and hence, irrespective of the FOV, holds for future large étendue displays as well.

Being able to study various pupil dependent artifacts at several pupil states, we propose a display algorithm that generates high quality images across the eyebox regardless of the eye pupil size and location. We introduce a differentiable pupil-aware image formation model and a corresponding optimization method that incorporate sampling over diverse pupil states, that, together, ensure pupil awareness when solving for SLM patterns that result in high quality reconstructions across the eyebox.

Pupil-aware holography brings an under-investigated problem of holographic displays to attention and proposes a method to overcome the problem of dynamic pupil sampling of the eyebox. The proposed approach takes a step forward to enable practical holographic displays for immersive augmented and virtual reality.

In particular, we make the following contributions

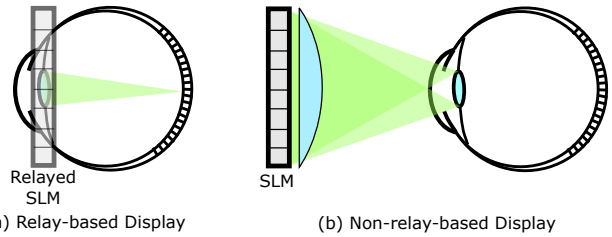
- We *analyze* the dependence of image quality on the eye pupil states and energy distribution across the eyebox for various holographic displays.
- We *introduce* a new differentiable pupil-aware holographic image formation method that accounts for different pupil positions and orientations within the eyebox.
- We devise a *content-aware phase optimization* method that allows us to learn optimal hologram phase patterns which maintain intensity across the eyebox without introducing image artifacts and severe speckle.
- We *analyze* our approach in simulation and demonstrate significant higher image quality across the eyebox compared to existing holographic display methods. Using a prototype system, we *validate experimentally* that the proposed pupil-aware holography approach achieves significant reduction in pupil-dependent artifacts.

1.1 Scope and Limitations

This paper does not contribute new hardware but the analysis of an unexplored problem in holography and an algorithm to address it irrespective of the setup. For the first time, we achieve robust image fidelity and eyebox energy across the eyebox. To study pupil variance, we use a large eyebox display as a test bed, albeit with limited field of view. Irrespective of the FoV, the current approach only considers the image fidelity across the eyebox and hence holds for future large etendue displays with smaller SLM pixel pitch or larger area. We validate the applicability of the proposed method to future large etendue systems in simulation.

Human eyes distort the phase of incident light waves, deviating from an ideal thin lens, given the conditions of each human observer [Chakravarthula et al. 2021], which we do not consider in this work. We do also not incorporate the holistic perception of the human visual system into the proposed method but only make a first step by addressing the display with varying pupil states. Incorporating further aspects of the perception of human visual system into our method may prove as exciting future work.

We highlight that a major benefit of holography is its promise of form factor reduction for both AR [Maimone et al. 2017] and VR displays [Kim et al. 2022; Maimone and Wang 2020] and the ability to correct for several aberrations, both in display optics [Chakravarthula et al. 2020; Peng et al. 2020] as well as the user's eyes [Chakravarthula et al. 2021] without additional optics adding to the bulk of the display. While 3D volume projections are a further important benefit of using holographic displays and an extensive investigation of 3D pupil-aware capabilities is an important direction, we restricted ourselves to 2D and multifocal 3D cases. The form factor supported by holographic displays for multifocal and varifocal displays is significantly smaller than other alternatives [Akşit et al. 2017; Chakravarthula et al. 2018; Dunn et al. 2017]. While studying pupil dependent image quality, depth and parallax effects for 3D holograms is indeed part of our future work, our findings show that the existing holography methods are impractical without considering pupil awareness, even for 2D holographic displays.



(a) Relay-based Display (b) Non-relay-based Display

Fig. 3. Existing holographic near-eye display designs either (a) directly sample diffracted light from the SLM by relaying the virtual SLM to the pupil plane (left), or (b) by projecting the diffracted light to a narrow valid region on the pupil plane using a lens (right). These designs, combined with a computer-generated hologram (CGH) algorithm form high-quality visual images if the entire wavefront is fully sampled by the eye pupil. Unfortunately, the eye pupil position, orientation and size dynamically change, resulting in partial and incomplete sampling of the diffracted light wavefront, causing severe artifacts in all existing holographic near-eye display systems.

2 RELATED WORK

2.1 Holographic Displays

Holographic displays promise to be the ultimate virtual and augmented reality displays of the future, potentially achieving high projection quality with depth cues in a thin eye-glasses form factor [Chakravarthula et al. 2020; Maimone et al. 2017; Peng et al. 2020]. The most successful approaches have achieved the highest image quality only with large benchtop prototype displays [Shi et al. 2021]. At the same time, researchers have shown that a combination of holographic optical elements and dynamic holography can achieve sunglasses-like form factor for both AR and VR applications [Maimone et al. 2017]. All of these methods have in common that they are constrained by the low étendue of existing SLMs, limiting either the FoV or eyebox of the displays. To address this issue, researchers have proposed methods for more effective use of étendue of holographic displays by incorporating eyetracking [Lu et al. 2020] and dynamically moving the small eyebox or statically expanding the eyebox [Xia et al. 2020]. Such an eyebox expansion can be achieved by pupil steering where the incident illumination on the SLM is changed in demand [Jang et al. 2019] or by static eyebox expansion where several copies of the exit pupil are created [Xia et al. 2020]. Unfortunately, these pupil steering methods mandate precise and low-latency eyetracking, along with complex and bulky optics. Without eyetracking present, no existing methods have considered the non-uniform sampling of eye pupil and the complex artifacts that arise across the eyebox. In this work, we analyze pupil variance in existing work and we propose a pupil-aware hologram generation approach for wide eyebox holographic displays.

2.2 Étendue Expansion for Holographic Displays

Existing approaches for the expansion of étendue in holographic displays employ a diffractive mask with higher pixel pitch than the SLM employed [Buckley et al. 2006; Kuo et al. 2020; Park and Askari 2019] to expand the largest diffraction angle that existing SLM devices can support. However, these existing methods were limited to generating sparse projections such as a few foci or letters.

Buckley et al. [2006] used a diffractive phase mask in front of a binary phase SLM to create a static display where the twin images were removed and the viewing angle was increased simultaneously. Park et al. [2019], on the other hand, used a known diffractive amplitude mask to achieve significant increase in étendue, but only could generate sparse focal spots simultaneously. Recently, Kuo et al. [2020] extended this idea to generate dense, photo-realistic étendue expanded holograms at the native resolution of the SLM. To make for an efficient testbed for the proposed method, instead of increasing étendue, which is an active direction that today requires manufacturing, alignment and calibration of diffractive elements, we use a focusing lens in front of the SLM to achieve a large eyebox. While similar hardware configurations have been proposed in the past [Maimone et al. 2017] to achieve a large FOV, we propose a variant with large eyebox, which allows us to investigate pupil variance and validate the proposed method when only observing a partial eyebox.

2.3 Computer Generated Holography Algorithms

Traditional Phase Retrieval Methods. Hologram generation algorithms find the appropriate SLM pattern for a given target image displayed by the system. This is a particularly challenging task with existing phase-only SLM devices available today. Holographic phase retrieval approaches can be broadly classified into single-step methods and iterative methods. While single-step methods such as amplitude-discard or double-phase amplitude coding [Maimone et al. 2017] may provide sufficient image quality for some applications, iterative algorithms such as the popular Gerchberg-Saxton method [Gerchberg 1972] significantly improve the image fidelity. Some of the early iterative methods for holographic phase retrieval include error reduction using iterative optimization [Gerchberg 1972; Lesem et al. 1969], together with an assumption on a non-zero support of the real-valued signal. One of many extensions of such iterative algorithm is the popular hybrid input-output (HIO) method [Fienup 1982], and others with various relaxations [Bauschke et al. 2003; Luke 2004]. Researchers have explored using alternating direction methods for phase retrieval [Marchesini et al. 2016; Wen et al. 2012], non-convex optimization [Zhang et al. 2017] and overcoming the non-convex nature of the phase retrieval problem by lifting, i.e., relaxation, to a semidefinite [Candes et al. 2013] or linear [Bahmani and Romberg 2017; Goldstein and Studer 2018] program. Recent iterative phase retrieval methods using first-order stochastic gradient descent with complex Wirtinger gradients [Chakravarthula et al. 2019; Peng et al. 2020] have been able to achieve high image quality on holographic displays. Such iterative optimization approaches were recently used to produced high-quality 3D holograms [Chakravarthula et al. 2022; Choi et al. 2021].

Neural Phase Retrieval Methods. Neural networks and deep learning approaches have recently been proposed as tools for optical design and holographic phase retrieval. Researchers have tackled holographic microscopy by solving phase retrieval problems using neural networks [Eybposh et al. 2020; Rivenson et al. 2018]. In a similar fashion, neural networks have been investigated for learning holographic wave propagation from a large training dataset. For example, Horisaki et al. [2018] trained a U-net on a pair of SLM phase and intensity patterns, and predicted SLM phase patterns during inference. Recently, Eybposh et al. [2020] proposed an unsupervised training strategy and predicted the SLM phase patterns in real-time that produced 2D and 3D holographic projections. Peng et al. [2020] and Chakravarthula et al. [2020] have recently demonstrated camera-in-the-loop (CITL) calibration of hardware using neural networks and high fidelity holographic images on prototype displays. Shi et al. [2021] have demonstrated high-resolution real-time holography with an light-weight neural phase retrieval network that may be suitable for inference on mobile hardware in the future.

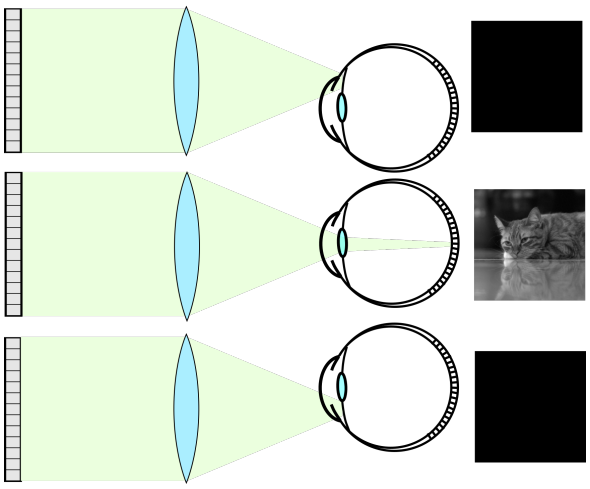


Fig. 4. Conventional holographic displays [Kim et al. 2022; Maimone et al. 2017] focus the SLM modulated light down to a point on the pupil plane, resulting in a very tiny eyebox (middle). This approach severely restricts the "valid" region where the eyes can see the holographic image, thereby resulting in no viewable imagery outside of the small eyebox region (top and bottom) that is mandated by the limited étendue of today's holographic systems.

The proposed pupil-aware holography approach shares with prior work that we solve an optimization problem to find optimal SLM patterns for a target image. However, we show that existing method, although achieving high image quality, are not practical due to pupil variance. We propose a pupil-aware holographic display approach where we formulate the loss function to include several pupil states within a large eyebox and compensate for perceptual artifacts across the eyebox.

3 PUPILS IN NEAR-EYE HOLOGRAPHIC DISPLAYS

As a result of the limited bandwidth and pixel count of SLMs in the past, most existing approaches in near-eye holographic displays had to maximize the FOV at the cost of a small eyebox. Large FOV together with large eyebox has recently been achieved by using étendue expanding methods [Kuo et al. 2020]. We describe here the complications that arise with pupil variance and thereby partial wavefront sampling for all of these existing display types. To this end, we consider pupil variance for holographic displays in two possible configurations where: (1) the SLM is relayed directly onto

the eye pupil (relay-based displays), and (2) the SLM is away from the eye (non-relay-based displays).

Pupils in Relay-based Displays. In holographic displays with a relay, as illustrated in Fig. 3(a), the wavefront at the SLM is relayed onto (or sometimes close to) the eye pupil plane and focused onto the retina. In this configuration, near-field or far-field holograms can be generated by controlling the z distance of the target holographic image from the relayed SLM. If the complex wavefront at the SLM is H, the observed image intensity after the eye pupil sampling is given by

$$I = |\mathcal{P}(H \odot M)|^2, \tag{1}$$

where $\mathcal{P}$ is a wave propagation operator, $\odot$ is element-wise Hadamard product and M is an amplitude mask, modeling the pupil sampling of the wavefront at the SLM plane, by the eye pupil. The position and shape of the mask depends on the pupil location, size and orientation. The propagation operator $\mathcal{P}$ is a Fresnel propagation function for a near-field hologram and a Fraunhofer or Fourier propagation function for a far-field hologram, respectively.

As the propagating wavefront is apodized, a significant portion of the light from the SLM does not enter the eye pupil. A far-field holographic image mimics light coming from optical infinity and hence the pupil sampling generally results in reduced intensity in the reconstructed images. However, masking a near-field holographic wavefront results in complete loss of information from masked wavefront and hence the reconstructed images appear cropped, see Figure 8 and Figure 12. Not only does this mean that pupil sampling comes at the expense of SLM bandwidth, it also results in under-sampling of spatial frequencies (or the angular rays) from the SLM. This produces poor perceived quality in both near- and far-field configurations. Specifically, in the far-field configuration, the small pupil sizes result in extreme speckle artifacts due to the missing frequencies outside of the sampled wavefront as shown in Figure 2 and Figure 8. Note that each pixel on the retina receives partial wavefront sampled by the pupil. For the near-field configuration, partial sampling of the wavefront by a pupil state causes both cropping and diffraction artifacts from the wave propagation as shown in Figure 8.

Pupils in Non-relay-based Displays. In order to avoid undersampling of the wavefront by the pupil, non-relay-based displays use additional optics such as an eyepiece to converge the SLM modulated light into the eye pupil, instead of relaying the virtual SLM onto the eye. Note that all SLM pixels contribute to the eyebox in this display configuration, and therefore the image formation model can be formalized by

$$I = |\mathcal{P}_2\{M \odot \mathcal{P}_1(H)\}|^2, \tag{2}$$

where $\mathcal{P}_1$ is the wave propagation function from SLM to the eye pupil, $\mathcal{P}_2$ is the propagation from pupil to retina and M is the masking due to the aperture of the pupil. While all SLM pixels are exploited in this configuration due to the étendue conservation, using a field of view expanding eyepiece comes at the cost of the eyebox being shrunk into a tiny viewable region around the focal spot of the eyepiece optics. The user would only see the hologram if their pupil lies directly on the focal spot region. However, once the pupil moves away, the user would cease to view the image entirely, as depicted in Figure 4. Consequently, the eyebox needs to be expanded or replicated using additional optical elements [Jang et al. 2017] or steered in accordance with the pupil movement [Kim et al. 2022] for such displays.

Overall, all the discussed *existing display configurations do not adequately account for eye pupil movement and varying pupil size*, even if the eyebox is sufficiently large. We propose a method in the next section that addresses these limitations.

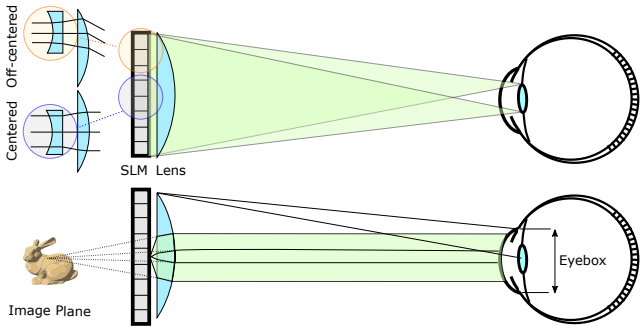


Fig. 5. *Wide eyebox variant holographic display.* To evaluate the proposed pupil-aware holographic phase retrieval, we use a combination of a large physical lens and local lenslets as displayed on the SLM (top) together to bias the rays towards the eye, thereby making a sufficiently large field of view and a wide eyebox prototype display testbed. The SLM phase of each lenslet is a conjugate of the large physical ray-biasing lens. Therefore, the lenslet array is nothing but a spatially shifted conjugate phase array of the large physical lens and their combination results in plane waves leaving the SLM-lens system (bottom), thereby creating a virtual image at optical infinity. While other configurations are possible, the proposed design ensures smooth transitions at the boundaries of local lenslets displayed on the SLM, and ray biasing without chromatic artifacts.

4 PUPIL-AWARE HOLOGRAPHY

In this section, we introduce the proposed pupil-aware holographic phase retrieval algorithm. To formalize the image formation method and validate the proposed algorithm, we rely on a wide eyebox holographic display as a test-bed. To effectively use étendue, we use a single lens co-planar to the SLM to off-load a significant amount of SLM bandwidth (i.e., the supported spatial frequencies), see Fig. 5. While similar hardware configurations have been used in the past, where a lens is used as an eyepiece, to achieve a large FOV [Maimone et al. 2017], we devise a variant for a larger eyebox which allows us to evaluate effects due to pupil variance. In the following, we first describe this setup. Then, we derive a pupil-aware image formation model for this setup, and optimize an SLM phase pattern for image quality and energy distribution throughout the eyebox.

Bandwidth-preserving Ray Biasing. For a holographic display, étendue (or space-bandwidth product) is the product of the area of the SLM and the maximum diffraction angle θ supported by the SLM. For a given wavelength of light λ, this angle is determined by the pixel pitch p as $\theta = sin^{-1}(\lambda/2p)$. In other words, the supported field of view, without magnification optics, is only twice the diffraction angle, which is no more than 4° for an SLM of $8\mu m$ pixel pitch, while the eyebox is the size of the SLM. To achieve a field of view of about

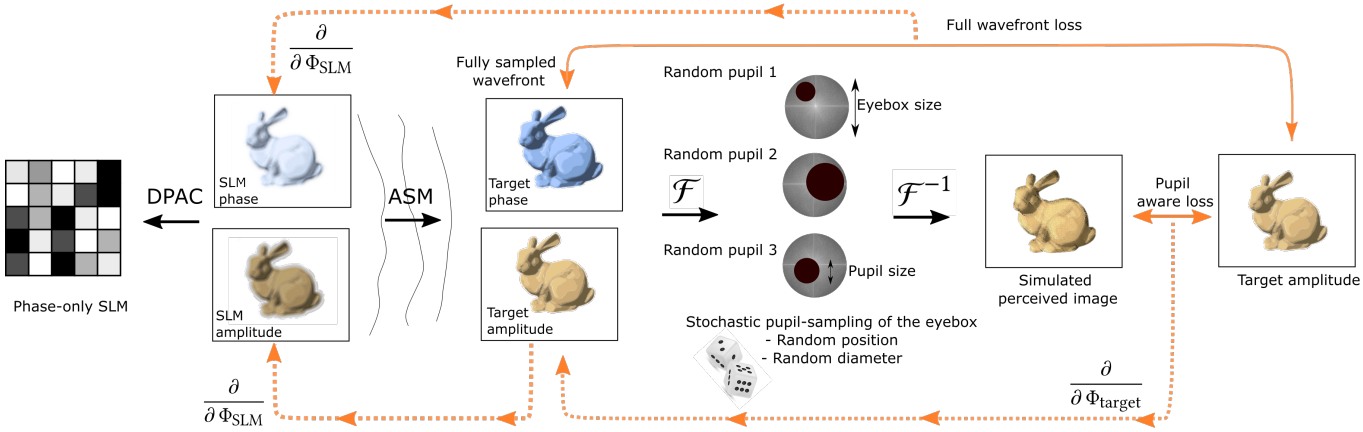


Fig. 6. *Pupil-aware Holography.* We optimize for phase-only holograms to produce high-fidelity reconstructions and energy distribution across the eyebox. To this end, we rely on a differentiable image formation model that explicitly considers the eye pupil sampling of the eyebox, allowing us to backpropagate an intensity reconstruction loss to the SLM phase pattern for each step of our iterative optimization. By a stochastically sampled pupil-aware optimization, the proposed method is able to achieve both image quality and energy distribution over the entire eyebox.

90°, the SLMs would need to have pixels of size less than $0.3\mu m$ – today's SLM pixels are more than an order of magnitude larger.

Under these limitations, unlike the existing display setups which maximize the FOV at the expense of the eyebox, we use a lens to bias rays towards the eyebox to get the desired ratio of a large eyebox and a limited FOV, given our limited available étendue. Specifically, we use a large physical lens coplanar with an SLM showing an array of overlapping lenslets that are conjugate to the physical lens used. As shown in Figure 5, a local lenslet on the SLM diverges the incident plane waves whereas the large physical lens collects these diverging waves and converges them back, effectively making them parallel and propagating towards the focal point of the physical lens. Note that the eyebox of the display is located at the focal point of the large physical lens and the choice of the physical lens and the size of local lenslets determine the field of view and the eyebox size of the display. As the ray-biasing lens supports the FoV, the majority of the SLM bandwidth thereby is used to effectively increase the eyebox of the system. For example, off-loading the bandwidth of an SLM of pixel pitch $8\mu m$ to a lens of focal length $100mm$ achieves a large eyebox of size about $7mm$.

Note that bending the incident beam locally can also be achieved using alternate designs such as a prismlet array – tiling of micro-prisms each with a different tilt. However, the proposed design of overlapping arrays of lenslets ensures smooth transitions at the boundaries of local lenslets displayed on the SLM and ray biasing without chromatic artifacts.

Hologram Formation Model. As described above, we use a combination of SLM and a coplanar convex lens to achieve ray biasing and generate a virtual image one focal length away from the lens-SLM plane. We now use this display configuration and algorithmically incorporate the conjugate concave lenslet kernel of the same focal length as illustrated in Figure 5 to derive the hologram formation model. Note that this places the final image at optical infinity as the convex lens and the conjugate lens kernel effectively compensate each other. See Supplementary Material for additional details. Specifically, for any given plane wave illumination, the conjugate lens phase on the SLM deflects the light in a diverging manner to form a virtual image. This diverging light is collected by the convex lens and bent to form a parallel beam. Like a physical diverging concave lens, the deflection of light near the center of the hologram is smaller whereas the deflection is larger at the edges. However, note that the SLM can only support a maximum deflection equal to that of the maximum diffraction angle θ. Therefore, we confine the conjugate lens phase corresponding to a given target image pixel to this smaller diffraction support as illustrated in Figure 5(top). A local support region for a given image pixel can be thought of as a *central lens cut* from a larger concave lens. However, note that different *lens cuts* corresponding to different object points can overlap on the hologram plane.

We will now discuss the general holographic wave propagation model between the SLM plane and the image plane. For a target wave field at the image plane that we want the eye to see, $U_{\text{target}} = A_{\text{target}} e^{j\phi}$, where A_{target} is the target image amplitude and ϕ its corresponding phase, typically called "image plane phase" or "object phase". We can propagate U_{target} to the SLM plane by computing the continuous sum of overlapping concave *lens cut* phase functions weighted by the target pixel amplitudes. This formation of the complex wavefront U_{SLM} at the SLM-plane can be mathematically represented as

$$U_{\text{SLM}}(\bar{x}) = \int_{\bar{t} \in L} A_{\text{target}}(\bar{x}) e^{j\phi(\bar{x}^2)} e^{-\frac{jk}{2f}(\bar{x}-\bar{t})} d\bar{t}, \tag{3}$$

where $\bar{x}$ is a target point and $\bar{t}$ is the offset spatial translation of the concave lens as illustrated by the yellow inset in Figure. 5(top), and L denotes the pixels within the diffraction angle support region that the SLM provides. Note that the above Eq. (3) is a convolution.

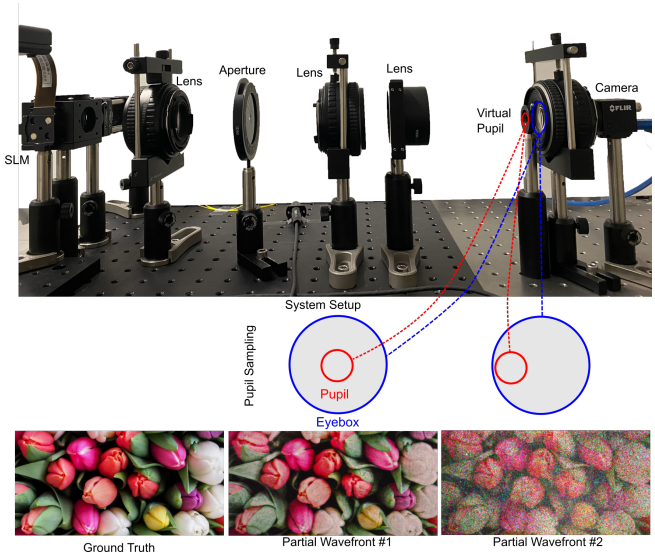


Fig. 7. *Prototype Holographic Display.* We built an experimental display to validate pupil-aware holographic phase retrieval. To mimic the pupil sampling of a human eye, we use an aperture on the eyebox plane, as can be seen in front of the camera, as a virtual pupil. The proposed pupil-aware holography method enables accurate holographic image reconstructions across the eyebox for diverse pupil states.

Hence, we can define the image formation forward model as

$$U_{\text{SLM}} = U_{\text{target}} * G, \tag{4}$$

where $*$ is the convolution operator, U_{SLM} is the complex hologram field, U_{target} is the target image field, and

$$G = e^{-j\frac{k}{2f}(x^2+y^2)} \tag{5}$$

is the field propagation kernel defined within the SLM diffraction angle support. Note that f here is the focal length of the large physical convex lens co-planar with the SLM and the kernel G is nothing but the conjugate lenslet displayed on the SLM. The convolution above can be computed in Fourier domain, thereby posing it as a band-limited angular spectrum propagation function [Goodman 2005]. Note that this image formation forward model *creates a coherent planar wavefront traveling towards the eyebox*, which is located at the focal plane of the ray-biasing lens.

Learning Pupil-aware Holograms. The object phase on the target image plane is responsible for the light distribution and hence energy distribution within the eyebox of the display. Note that although the eyebox supported by a display is larger, having the light energy concentrated in one region results in a smaller *effective eyebox*, see Figure 9. On the other hand, it also decides the quality of holographic images; a uniform object phase results in noise-free images with most energy concentrated into an extremely tiny effective eyebox whereas a completely random object phase results in extremely noisy images but with energy spread uniformly over the entire eyebox. As such, existing methods typically prefer a uniform phase on the target image plane for the quality of holographic projections. Although we can design a display that allows for a reasonably large eyebox, existing methods ignore that the *effective eyebox* where all the light energy is concentrated may still be very small as shown in Figure 9. As a result, the eye pupil movement outside of this *effective eyebox* results in losing a significant portion of the image, and sometimes the image itself.

In order to tackle these problems, we jointly optimize for the SLM phase, holographic image quality and the eyebox energy distribution to support diverse eye pupil states within the eyebox. We derive a differentiable pupil-sampled variant of the forward model from the previous section to propagate the wave field effectively between the hologram plane, image plane and the eyebox, and stochastically sample the eyebox to incorporate different pupil states. Using this differentiable forward model, we learn the SLM phase using stochastic gradient descent with Wirtinger gradients [Chakravarthula et al. 2019; Peng et al. 2020] in an iterative fashion. Integrated into the stochastic gradient descent scheme, we randomly sample different sizes and locations of the pupil within the eyebox, during the optimization. We initialize the optimizer with the hologram computed using the model discussed in the previous paragraph.

Specifically, for a given target image amplitude A_{target}, we optimize for the complex hologram field at the SLM $U_{\text{SLM}} = A_{\text{SLM}} e^{j\Phi_{\text{SLM}}}$ to produce high-fidelity reconstructions as well as energy distribution across the eyebox, and consequently indirectly optimize the phase distribution of the propagated wavefront on the image plane. To this end, we model the *reconstructed image at the image plane in two different ways.* We first propagate the incident wave from the SLM plane to the virtual image plane, filtering higher frequencies and orders within a 4F system, resulting in the fully-sampled wavefront simulated image on the target plane $U_{\text{target; full}}$ as

$$U_{\text{target; full}} = \text{Prop}_{\text{SLM}\to\text{target}}(U_{\text{SLM}}) = \mathcal{F}^{-1}\left(\mathcal{F}(U_{\text{SLM}}) \odot M_{\text{iris}}\right) * G^{\dagger}, \tag{6}$$

where $\text{Prop}_{X\to Y}(U)$ is the wave propagation operator, propagating a field U from the plane X to Y, M_{iris} is a binary mask used to filter higher orders within a 4F system and $G^{\dagger}$ is the complex conjugate of the kernel introduced in Eq. (5). We compute the convolution in Eq. (6) in the Fourier space, specifically as a modified angular spectrum propagation between the SLM and the target image planes.

In addition, we also incorporate a forward model that includes the eye pupil. We propagate the complex wave field from the SLM to the eyebox, attenuate the eyebox with a stochastically sampled eye pupil mask M and then propagate the sampled eyebox field to the image plane to reconstruct the image as seen by the eye pupil. Note that in our display configuration, the eyebox plane lies at the Fourier plane of the ray-biasing lens co-planar with the SLM. Hence the wave propagation between the target image plane and the eyebox plane can be modeled by the far-field Fraunhofer propagation. Therefore, the pupil-sampled reconstructed image $U_{\text{target;pupil}}$, can then be expressed as

$$U_{\text{target;pupil}} = \mathcal{F}^{-1}\left(M \odot \mathcal{F}(U_{\text{SLM}}) \odot M_{\text{iris}}\right) * G^{\dagger}. \tag{7}$$

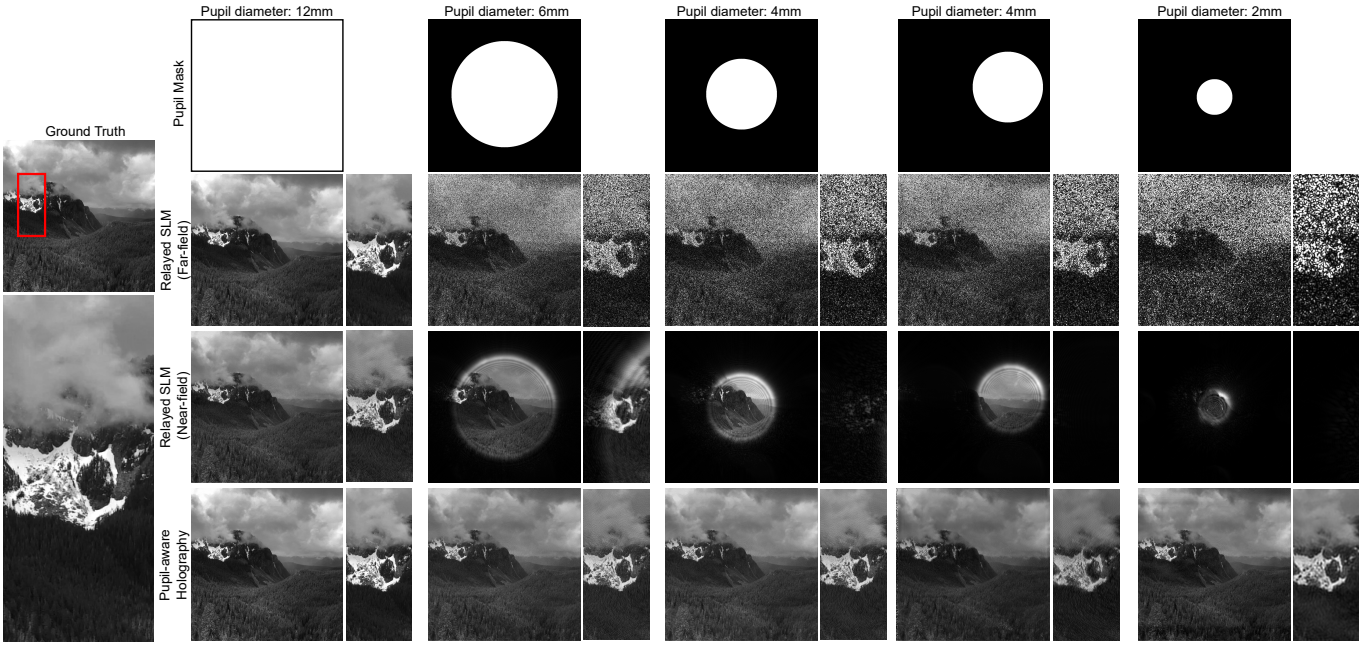


Fig. 8. *Evaluating Display Configurations (Simulation).* We study commonly used near- and far-field holographic display configurations and the proposed wide eyebox variant for evaluating our pupil-aware holography method. Note, that each row corresponds to a different display configuration and hence comes with a slightly modified hologram generation algorithm to account for setup changes. More details on the display setup configurations can be found in the Supplementary Material Figures S3, S4 and S5. See Figure 3 for SLM relayed and non-relayed setup schematics, and Figure 4 for a small pupil forming Maxwellian-style display. We test the configurations for five pupil masks with different sizes and locations shown on top, where the white mask represent the eye pupil. All existing methods, with varying pupil states, produce either speckle while truncating the spatial frequencies that are admitted into the pupil in the far-field configuration (second row), or truncate the image itself as shown in the near-field configuration (third row). The proposed variant of the wide eyebox display allows us to study pupil-aware holography that incorporates diverse pupil states .

The image formation between the SLM and target image planes ensures that the SLM phase produces the appropriate noise-free reconstructions when the entire wavefront is sampled. The image formation model between the target image plane and the stochastically sampled eyebox plane ensures that the image quality is maintained for a variety of pupil states across the eyebox. Therefore, we use both fully-sampled and pupil-sampled reconstructed images to find the optimal complex wave field at the SLM ($U_{\text{SLM}} = A_{\text{SLM}} e^{j\Phi_{\text{SLM}}}$) by solving the following optimization problem

$$A_{\text{SLM}}, \Phi_{\text{SLM}} = \underset{\{A'_{\text{SLM}}, \Phi'_{\text{SLM}}\}}{\arg\min} \; \mathcal{L}(|U_{\text{target;full}}|, A_{\text{ref}}) + \mathcal{L}(|U_{\text{target;pupil}}|, A_{\text{ref}}), \tag{8}$$

where A_{ref} is the desired reference image and $\mathcal{L}$ is a custom penalty function which we detail below. $U_{\text{target;full}}$ and $U_{\text{target;pupil}}$ are as defined in Eq. 6 and Eq. 7 respectively. Figure 6 provides an illustration of this method.

Within the proposed optimization, we employ stochastic sampling of eyebox to account for a variety of pupil states. This also allows for a *pupil-aware* distribution of energy within the eyebox. Optimizing for a complex SLM wave field allows for more degrees of freedom than a phase-only optimization, and offers an optically elegant way of removing noise from the images. For the noise removal, we filter higher frequencies corresponding to the complex wave field on the SLM during the optimization. This forces noise to the higher diffraction orders. To optically filter these higher diffraction orders carrying the noise, we encode the normalized complex SLM wave field (i.e., $Ae^{j\phi} \in \mathbb{C}(0 \leq A \leq 1)$) into a phase-only pattern by interleaving the amplitude and phase using a double phase amplitude coding method [Hsueh and Sawchuk 1978]

$$Ae^{j\phi} = 0.5e^{j(\phi - cos^{-1}A)} + 0.5e^{j(\phi + cos^{-1}A)}. \tag{9}$$

We use a checkerboard mask to select the interleaving phase values from the decomposed two phase-only holograms. The interleaving into the high frequency checkerboard pattern result in high diffraction orders carrying high frequency noise which we filter in the Fourier plane.

We solve the above optimization problem using first-order iterative stochastic gradient descent methods as both the forward and backward image formations are differentiable with respect to the complex wave field at the SLM. For the image loss function $\mathcal{L}$, we use a weighted combination of ℓ_2 penalty $\mathcal{L}_{\ell_2}$, SSIM $\mathcal{L}_{\text{SSIM}}$, perceptual penalty based on VGG-19 $\mathcal{L}_{\text{PERC}}$ [Johnson et al. 2016], and Watson FFT $\mathcal{L}_{\text{WFFT}}$ [Czolbe et al. 2020], that is:

$$\mathcal{L} = \lambda_{\ell_2}\mathcal{L}_{\ell_2} + \lambda_{\text{SSIM}}\mathcal{L}_{\text{SSIM}} + \lambda_{\text{PERC}}\mathcal{L}_{\text{PERC}} + \lambda_{\text{WFFT}}\mathcal{L}_{\text{WFFT}}. \tag{10}$$

We use a least-square penalty $\mathcal{L}_{\ell_2}$ for per-pixel accuracy in the reconstruction and $\mathcal{L}_{\text{SSIM}}$ as a hand-crafted perceptual loss. The second perceptual penalty $\mathcal{L}_{\text{PERC}}$ compares the image features from

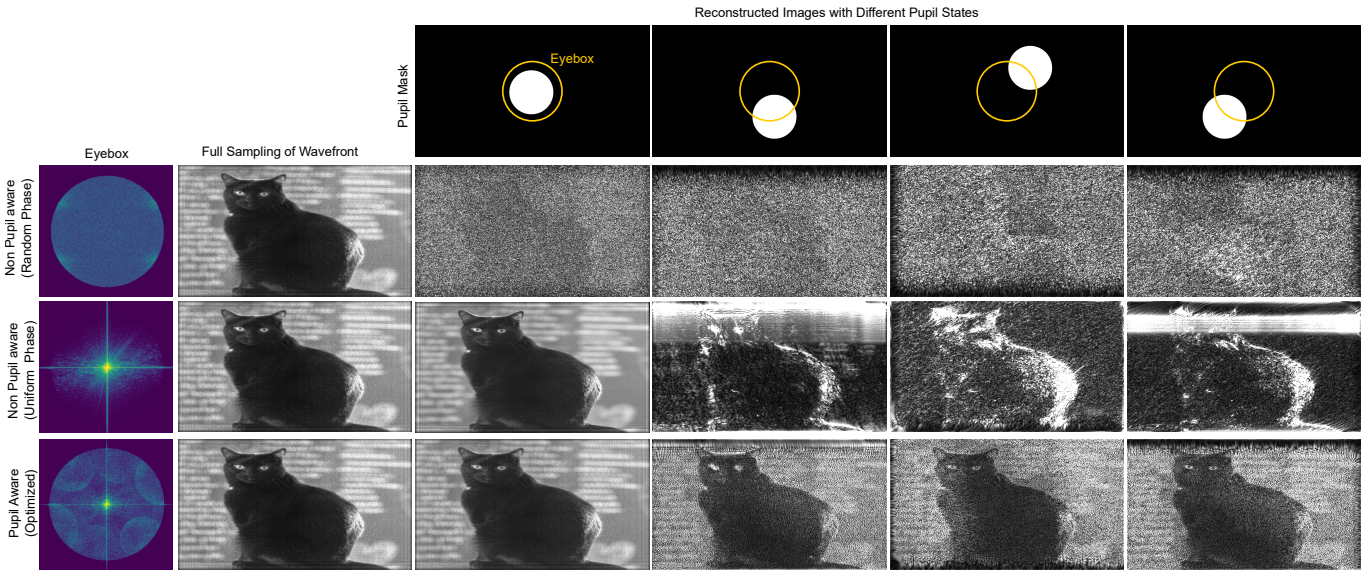


Fig. 9. *Pupil-aware vs non-pupil-aware holography on the same display setup (Simulation).* Traditional non-pupil-aware holographic phase retrieval assumes an unknown phase on the image plane, which is typically either random (top row) or smooth (middle row). While a random object phase on the image plane generates a uniform eyebox energy, it results in the reconstructed image corrupted with extreme noise (top row). On the other hand, a uniform object phase results in energy concentrated within the eyebox, which results in noise-free reconstructions at the center of eyebox and loss of image as the pupil moves away from it, resulting in a tiny effective eyebox (middle row). The proposed pupil-aware holography results in high-fidelity reconstructions by distributing the energy across the eyebox (bottom row).

activation layers in a pre-trained VGG-19 neural network, that is,

$$\mathcal{L}_{\text{PERC}} = \sum_l v_l ||\phi_l(x) - \phi_l(y)||_1, \tag{11}$$

where ϕ_l is the output of the l-th layer of the pre-trained VGG-19 network and v_l are the corresponding penalty-balancing weights. Specifically, we use the outputs of ReLU activations just before the first two maxpool layers, i.e., relu1_2 and relu2_2. The loss term $\mathcal{L}_{\text{PERC}}$ therefore helps recover finer details. However, note that the VGG-19 network is optimized for classification and detection tasks, and is "robust" to the perceptual influence of artifacts such as noise. Therefore, we further augment the reconstruction quality by adopting the Watson FFT error function which is crafted specifically for human visual system, based on Watson's visual perception model [Watson 1993], see Supplemental Document. We find the combination of the above losses helps steer the optimization towards holograms that maintain high quality across the eyebox.

5 IMPLEMENTATION

5.1 Software

We tested our pupil-aware holography concept in simulation implemented using PyTorch running on an NVIDIA P100 GPU. PyTorch now provides complex Wirtinger gradients within its auto-differentiation modules making implementation of the optimization scheme with state-of-the-art first order optimizers straightforward. We notice that different optimizers result in slightly different reconstruction quality. We use the Adam optimizer for solving Eq. (8) with a learning rate of 0.01 and other default parameters. We assume an SLM of 1080×1920 pixel resolution with a pixel pitch of $8\mu m$ to match our physical hardware setup. All the input images and the holographic phase output was maintained at 1080×1920 resolution. The simulated Fourier space filter and the eye pupil were chosen to match the physical size of the prototype display. Specifically the Fourier space filter dictates the size of the eyebox in our display prototype and is maintained at $6mm$.

We assume the least knowledge of the eye position and hence the pupil is equally probable across the eyebox. Therefore, we sample the size and position of the pupil mask from a uniform random distribution. With position randomly sampled across the eyebox, the pupil size varied between 2 mm and 6 mm modeling an average human eye. Modeling the pupil rotation and translation effects within a 3D eyebox volume to produce view dependent effects is an interesting future direction. The near-fied images were placed 100 mm from the SLM whereas the far-field images are generated at optical infinity from the SLM. We initialize the optimization with the non-pupil-sampling hologram modeled by Eq. (4), and run the optimizer for 500 iterations until convergence. The overall optimization takes about 2 seconds.

5.2 Hardware Prototype

To validate our simulation, we built a prototype holographic display shown in Figure 7 with a HOLOEYE PLUTO liquid crystal on silicon (LCoS) reflective phase-only spatial light modulator with a resolution of 1920×1080 and a pixel pitch of $8\mu m$. This SLM is illuminated by a collimated and linearly polarized beam in a color sequential manner from optical fibers emitting at a wavelengths of $636nm$,

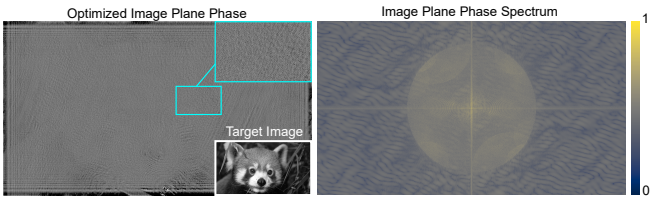


Fig. 10. *Optimized Object Phase and its Spectrum.* The optimized object phase resulting from the proposed method contains some of the target image structure overlaid with high frequencies. A log spectrum of this phase image shows the presence of low frequencies (similar to constant object phase) and high frequencies (similar to random object phase). The circular shape is from Fourier filtering of higher orders within a 4F system, which is also the size of the eyebox. This object phase results in a pupil-aware eyebox energy distribution.

520*nm* and 450*nm* and controlled using a ThorLabs KLD101 Kinesis K-Cube laser diode driver. The illuminated beam that is modulated by the phase-only SLM is focused by a Pentax 645N 75mm on an intermediate plane where an iris is placed to discard higher diffraction orders and conjugate images. We then relay the SLM with another Pentax 645N 75mm lens. Note that both the 75mm Pentax lenses form a 4F system with unit magnification, relaying the filtered SLM onto a virtual SLM plane as shown in Figure 7. We then place a Thorlabs AC-508-100-A-ML 100*mm* focal length achromatic doublet coplanar with the virtual SLM to create ray-biasing as described in Section 4. This configuration created an eyebox of 6*mm* where we place an iris on a motorized translation stage for sampling the eyebox. The sampled eyebox field is then measured by a Point Grey FLIR machine vision camera with a focusing lens.

We also built prototype displays for the relayed-SLM configuration showing both near-field and far-field holograms. As comparable holographic setups have been described in detail in a large body of existing work, we refer to the Supplemental Document for more details on the hardware configuration of these setups.

6 ASSESSMENT AND ANALYSIS

6.1 Synthetic Evaluation

Next, we investigate the pupil dependence of image quality in simulation for different near- and far-field display configurations (see Figure 3) as described in Section 3. We then evaluate the effectiveness of pupil-aware holographic phase retrieval on the proposed wide eyebox display configuration (see Figure 5) and compare it to non-pupil-aware holographic reconstructions on the *same* display. For these simulations, all the input images and output holograms used a resolution of 1920×1080 and an SLM pixel pitch of $8\mu m$ to match our hardware. Finally, we discuss extending our pupil-aware holography method to larger étendue displays, validating it on 16× and 64× larger étendue.

Evaluating Different Display Configurations. Figure 8 compares the fidelity of the reconstructed holograms for three different display configurations in diverse pupil states. Specifically, we show five different pupil states including the 12 mm virtual pupil diameter that samples the entire wavefront. While humans have pupils that span a

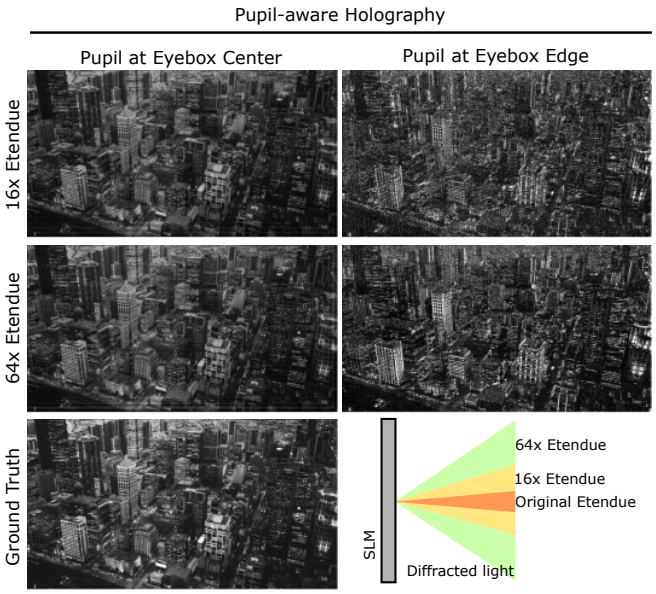


Fig. 11. *Pupil-aware Holography on Large Étendue Displays (Simulation).* We apply pupil-aware holographic phase retrieval on 16× and 64× large étendue display compared to our prototype hardware display. The results here demonstrate that our method scales to future large étendue displays

diameter of around 4 mm in normal conditions, we note that a very large synthetic pupil diameter of 12 mm is a commonly used setting in various conventional holographic displays [Chakravarthula et al. 2020; Peng et al. 2020]. In this unrealistically large pupil state, the entire wavefront impinging the pupil is collected by the eye lens and focused to the retina. Another way to achieve this is to use an eyepiece to relay the wavefront into the eye pupil, but this reduces the size of the eyebox to typically less than a millimeter, and hence the image can completely disappear even for minute eye pupil movements, see Figure 4. Collecting the full wavefront ensures high-fidelity reconstruction of holographic images across all tested configurations as shown in the first column of Figure 8. We next evaluate the configurations on more realistic pupil parameters in the range from 2 mm to 6 mm in different locations. This results in partial sampling of the modulated wavefront as shown in the pupil masks at the top row of Figure 8. For the far-field relayed SLM configuration, the pupil diameter mainly determines the degree of artifacts such as the speckle noise, while the location of the pupil does not affect the overall performance. This is due to the random phase distribution typical to the far field holograms which results in a relatively uniform but random energy distributed within the eyebox. In contrast, pupil sampling acts as a cropping operation for near-field holograms with additional possible diffraction artifacts due to the size of the pupil and the propagation distance of the projected light. The qualitative difference between the far-field and the near-field setups manifests in whether the image degradation appears globally or locally on the retina plane. In particular, displaying a near-field hologram, which is generally preferred due to their high-fidelity reconstructions relative to the far-field, mandates that

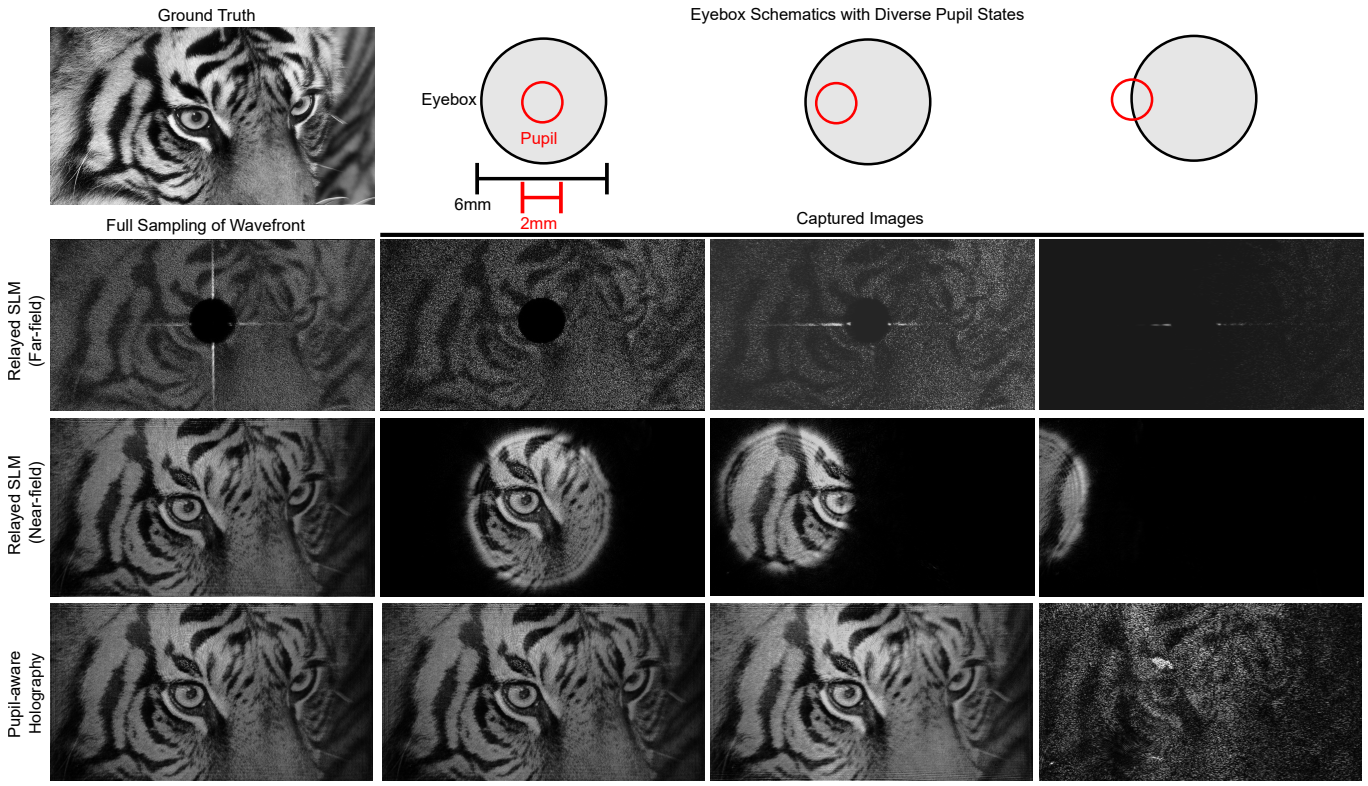


Fig. 12. *Experimental Evaluation of Various Display Configurations* We validate the choice of our wide eyebox variant of holographic display compared to common existing display configurations by capturing the reconstructed holograms on a hardware prototype at three different pupil states shown in the top row. Please see Supplementary Material (Figures S3,S4 and S5) for details on the hardware display configurations. While the far- and near-field relayed SLM configurations allow for accurate reconstructions when the wavefront is fully sampled, they suffer from speckle and cropping artifacts due to partial pupil sampling (second and third row). In contrast, the proposed display configuration allows for evaluating pupil-aware holography for diverse pupil states across the eyebox. All holograms are computed using stochastic gradient descent iterative optimization.

the eyebox is always smaller than the size of the eye pupil for high-fidelity reconstructions. The proposed method provides accurate holographic reconstructions across diverse pupil states including both pupil diameter and locations as shown in Figure 8. This is because we model the image formation at the SLM plane as a near-field propagation thus ensuring high-fidelity reconstructions, the final image received by the eye is that of a far field due to the lens placed co-planar with the SLM thus ensuring a sufficiently large effective eyebox.

Evaluating Pupil-aware Phase Retrieval on the Same Display. We next analyze the effectiveness of the proposed pupil-aware holography optimization on the proposed display configuration, compared to conventional phase retrieval methods where the unknown object phase is chosen to be either random or uniform [Maimone et al. 2017; Shi et al. 2021]. Note that the choice of the object phase on the target image plane effects both the reconstruction quality of holographic projections as well as the energy distribution within the eyebox. Figure 9 shows the simulated energy distribution at the eyebox plane for three different experiments on the same display configuration. The first column shows the eyebox energy distribution for random and uniform object phase, and for pupil-aware optimized eyebox. A completely random object phase results in a uniform distribution of light in the eyebox, but at the cost of image quality for different pupil states, see Figure 9 (top row). On the other hand, although the size of the eyebox supported by the display is larger, a hologram computed using a uniform phase on the target plane results in energy concentrated in the center, see Figure 9 (middle row). Therefore, an image is visible only if the eye pupil samples the center of the eyebox where the maximum intensity is found. Any deviation from the center results in a loss of image. This reduces the eyebox effectively to a very small region in the center. However, with our pupil-aware optimization, we notice that the energy in the eyebox is distributed similar to that of a constant phase in the center and approaches pseudo random phase towards the edge of the eyebox (Figure 9 (bottom row)), thus finding a tradeoff between both extremes. This energy distribution leads to higher fidelity reconstructions across diverse pupil states. Pupil-aware hologram computation also results in high image fidelity across the eyebox. As also reported in Table 1, while non-pupil-awareness result in rapid degradation of images, pupil-aware holograms maintain image fidelity even at the eyebox extremes.

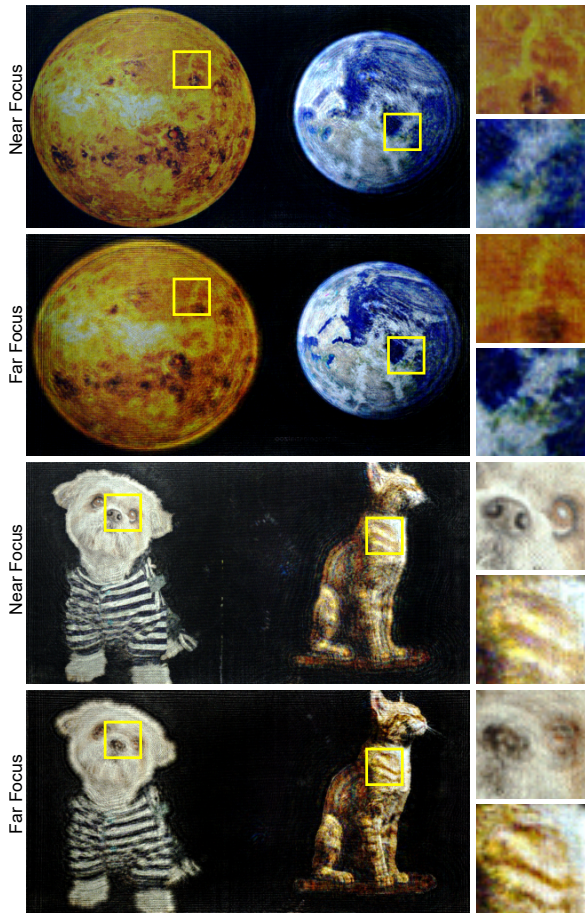


Fig. 13. *Experimental 3D Holographic Display.* We demonstrate that the proposed method can also be extended to 3D holography. To this end, we show multiplane 3D holographic projections on an experimental propotoype. That is, using the proposed method we optimize single SLM pattern to simultaneously project imagery at both near and far distances. The images shown here are measured only by changing the camera focus, and the corresponding in-focus and out-of-focus imagery can be observed in the insets.

Table 1. Quantitative results for synthetic evaluation of non-pupil-aware and pupil-aware holograms. We evaluate the generated holographic reconstructions for each method across the center of the eyebox and the far eccentricity edge of the eyebox over a dataset of randomly selected 20 images and we report the average quantitative PSNR metric scores.

	Eyebox center	Eyebox eccentricity
Non pupil-aware	32.12 dB	4.98 dB
Pupil-awareness	29.8 dB	16.03 dB

Analysis of Optimized Object Phase. As discussed in Section 4, the object phase on the image plane effects the energy distribution on the eyebox plane and hence the effective size of the eyebox. On the other hand, it also effects the reconstructed image quality. A random object phase results in energy distributed across the eyebox but noisy images. A uniform object phase results in visibly less image artifacts but a small effective eyebox. Contrary to this, holograms learned using the proposed method result in high image quality as well as energy distributed across the eyebox. To understand this, we study the optimized target phase on the image plane as well as its frequency spectrum. An example is presented in Figure 10. We notice that the optimized object phase resulting from the proposed method perhaps surprisingly contains some of the structure of the image amplitude. We observe an impulse response in the center resembling that of uniform object phase. Most high frequency components of the target phase are to be limited to a disc of the size of the filter we employ in the Fourier plane to filter any higher orders. Repeating circular patterns appear within this region which we hypothesize results in meaningful images when the eye pupil diverges significantly from the eyebox. A similar distribution can be observed in the eyebox as well, see Figure 9. The higher frequencies outside of the central disc are filtered in the Fourier plane and hence might be carrying high frequency noise that is eventually filtered in the Fourier plane (optically with an iris in experimental hardware display).

Pupil-aware Holography on Large Étendue Displays. Today's SLM pixel technology does not support large étendue displays natively. Hence, we validate that pupil-aware holography expands to such future large étendue displays in simulation. The simulations in the previous paragraphs match our hardware which uses an SLM of 1920×1080 pixel count and a pitch of $8\mu m$. Next, we simulate a $16\times$ larger étendue display with a pixel count of 7680×4320 and pixel pitch of $2\mu m$, and a $64\times$ larger étendue display with a pixel count of 15360×8640 and a pixel pitch of $1\mu m$. The phase-only hologram SLM patterns are computed using the proposed method and report results in Figure 11. The results from smaller étendue displays transfer to this setting and we validate that optimizing holograms in the pupil-aware fashion results in high-fidelity reconstructions. The proposed method scales well to larger étendue displays as it only considers the pupil sampling and the energy distribution within the eyebox, independent of FoV of the display. We refer to the Supplementary Material for additional discussion on étendue and the eyebox considerations.

6.2 Experimental Evaluation

We validate the proposed pupil-aware holography method on a prototype display and report experimental findings. Specifically, we first compare different display configurations with our wide eyebox display configuration (also assessed in simulation in the previous section) to validate the choice of our display setup. We then validate our pupil-aware holograms by comparing to non-pupil-aware holograms on the *same experimental wide eyebox display configuration.*

Validating the Wide Eyebox Display Configuration. We next validate the proposed pupil-aware display configuration in conjunction with the proposes phase retrieval method. To this end, we display the SLM phase patterns for a given target image and measure the raw images with a 2 mm-diameter pupil (iris) sampling the 7mm eyebox at different locations as shown in the top row of Figure 12. The top row of the figure indicates the pupil sampling of the eyebox.

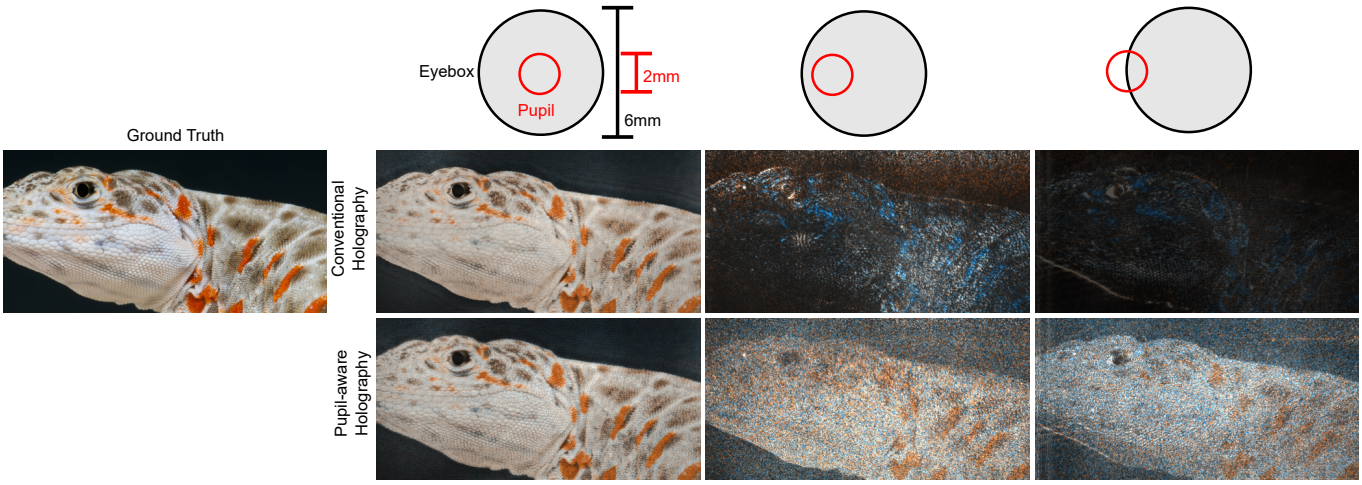


Fig. 14. *Experimental Evaluation of Pupil-aware Holography*. We validate the proposed pupil-aware holography on a wide eyebox hardware display setup, with the eyebox sampled by different pupil states (top). In contrast to conventional holographic methods that do not consider pupil sampling (middle), learning pupil-aware holograms lead to higher-fidelity reconstructions across the eyebox (bottom) only from partial wavefront sampling.

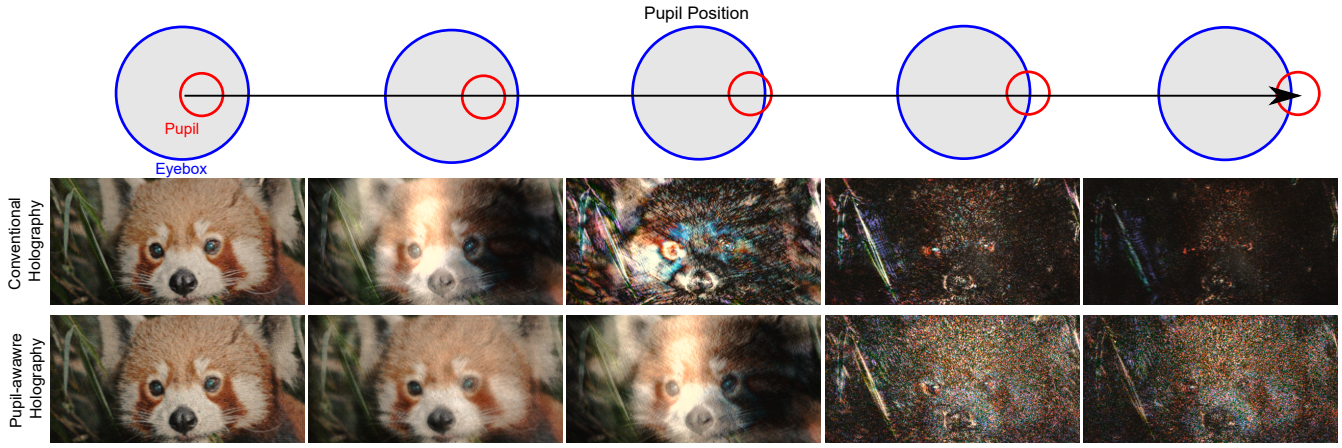


Fig. 15. *Experimental Evaluation of Densely-sampled Pupil States*. We densely sample the eyebox (top) to study the image quality as the pupil traverses the eyebox. Conventional holographic methods produce high-fidelity imagery at the center of the eyebox by concentrating the light energy distribution at the center. As a result, conventional non-pupil-aware methods results in complete loss of imagery as the pupil moves from the center (middle). Pupil-aware holography maintains the imagery throughout the eyebox as shown in the bottom row.

The subsequent rows show the measured images for far- and near-field configurations where the SLM is relayed onto the pupil, and our wide eyebox display configuration. Please see Supplementaty Material for more details on the display configurations.

In the first column, we show the measured images for the three different configurations where the entire wavefront is sampled. We see that sampling the entire wavefront from the SLM results in high fidelity reconstructions. Note that measured images from the relayed SLM with far-field holograms have a dark region in the center due to a mask we used to block the high DC intensity of the unmodulated light from the SLM, see Supplementary Material for more details. When the entire wavefront is sampled, the pupil-relayed near-field configuration (third row) and our wide eyebox configuration (fourth row) demonstrate comparable reconstructions. However, matching the simulations, partial sampling of the wavefront due to the pupil causes severe degradation in image quality in both near- and far-field relayed-SLM configurations (second and third row). Specifically, while the far-field holograms suffer from reduced brightness and increased speckle artifacts, the near-field holograms suffer from the image being cropped and diffraction artifacts at the edges of the pupil. As the pupil moves towards the edge of the eyebox, these artifacts increase further until the entire image is lost in either cases, see last column of Figure 12. The proposed wide eyebox display configuration with pupil-aware holographic phase retrieval providing accurate reconstructions at different pupil states within the eyebox, as shown in the last row of Figure 12.

Note that, at the edge of the eyebox, as opposed to losing the image entirely, pupil-aware holograms demonstrate reconstructions that look similar to the far-field holograms in relayed-SLM configuration, see last row of fourth column and second row of Figure 12. As discussed in Section 6.1, the center of the eyebox of our optimized holograms achieves an energy distribution similar to that of uniform phase holograms resulting in noise-free reconstructions within the eyebox. As far eccentricities of the eyebox approach the distribution similar to a (pseudo) random phase, the reconstructions look akin to a random object phase hologram but still with visible image features, even when the pupil only partially overlaps with the eyebox. We do not evaluate the small eyebox Maxwellian-style display configurations [Maimone et al. 2017] here as the eyebox is typically less than a millimeter, as discussed in Section 3. Overall, we find that our wide eyebox display configuration is best suited for evaluating the pupil-aware holographic phase retrieval method.

Validating Pupil-aware Holography. We now validate the proposed pupil-aware holographic phase retrieval on the same wide eyebox display configuration described in Section 4 and evaluated above. To this end, we display phase patterns computed with and without pupil-awareness, using the proposed optimization method in both cases (with and without pupil sampling), on an experimental prototype display and report results in Figure 14. We also report a dense set of measured images with finer pupil sampling in Figure 15. With changes in the pupil diameter, the brightness of the images also change. Therefore we maintain uniform laser power for different pupil configurations for a fair comparison of the image quality. See Supplementary Material for additional results.

Figure 14 demonstrates the image fidelity achieved by our pupil-aware holography method over the eyebox. As described in the synthetic evaluations in Section 6.1, the energy distribution achieved within the eyebox of our pupil-aware holograms approach that of noise-free uniform target phase holograms in the center (where the eye pupil is expected to sample the most) and that of a pseudo-random target phase holograms at higher eccentricities of the eyebox. The relative pupil position within the eyebox is indicated on the top. Holograms computed for maximum image quality but without pupil-awareness result in high-fidelity reconstructions when the pupil samples the center of the eyebox where maximum energy is concentrated. However, as the pupil moves away from the eyebox, the image is lost, see top row of Figure 14. In contrast, our pupil-aware holographic phase retrieval distributes the energy throughout the eyebox to maintain the image quality as can be seen in bottom row of Figure 14. As the pupil is likely to sample the center of the eyebox the most, majority of the SLM bandwidth is used to redirect the light to produce noise-free reconstructions in lower eccentricities of the eyebox. However, at higher eccentricities approaching the edge of the eyebox, image reconstructions that approach random phase holograms are produced. This is in accordance with the eyebox and object phase spectrum analysis reported in Section 6.1. To better visualize the degradation of image quality, we report a finer pupil sampling of the eyebox and the corresponding holographic reconstructions in Figure 15. It can be observed that the image quality of pupil-aware holograms is maintained at eccentricities where the images of non-pupil-aware holograms are completely lost. See Supplementary Material for additional results.

3D Pupil-Aware Holography. Although studying eyebox dependence of holographic image quality is an unexplored area independently if 2D or 3D holograms are considered, the proposed pupil-aware holography also extends to 3D holographic displays and future very large étendue displays. To validate extension to 3D holographic displays, we demonstrate a multiplane display application in Figure 13 where the holograms are computed using our pupil-aware holographic phase retrieval. For data capture, the focus is changed between near and far objects by only changing the camera focus ring. For example, in the first column, the features on the dog such as the nose and the stripes on the shirt are sharply visible when the camera is focused to near distance whereas the stripes on the skin of the cat are blurred. When the camera is focused at far distance, the cat comes to focus whereas the dog appears blurred. Similar trend can be observed in the second column of Figure 13 where the features on the Earth and Venus go in and out of focus at near and far distances. Some of these features are zoomed and shown in the insets for accessibility.

7 CONCLUSION

We propose a holographic display method that is pupil-aware. Most existing holographic methods 1) do not consider the light distribution within the eyebox and 2) the wavefront sampling by the eye-pupil. With research on wide étendue displays progressing, an issue becomes apparent: the hologram wavefront is often cropped by the viewer pupil and this cropping can have a severe effect on the perceived images, an effect which was never simulated nor studied in the past. This is primarily because the étendue that today's displays are restricted by is too small to observe this issue. We find that large étendue comes at the cost of a high degree of speckle, or drastic loss of overall intensity, and, as such, even with an ideal 1 billion pixel SLM, existing holographic displays are impractical as they are all fundamentally subject to pupil sampling of the wavefront at the eyebox. We investigate and explain such pupil effects when using only partially sampled part of the eyebox. To generate pupil-aware holographic projections, we propose a content-driven display algorithm that generates high quality images across the eyebox regardless of the eye pupil size and location. We rely on a differentiable pupil-aware image formation model and a corresponding per-image optimization method that incorporate sampling over diverse pupil states, and, together, ensure image fidelity and energy distribution across the eyebox when solving for SLM patterns. We validate this method in simulation and with an experimental prototype system, where we achieve high image fidelity over the full eyebox extent. As such, we make a first step towards pupil-invariant large étendue displays of the future, which may make holography practical across application domains. Immediate next steps in this direction could be the joint optimization of custom étendue-expanding elements with pupil invariance – potentially paving the way towards fully pupil-invariant holography in the future.

ACKNOWLEDGMENTS

Praneeth Chakravarthula was partly supported by NSF grants 1840131 and 1405847. Felix Heide was supported by an NSF CAREER Award (2047359), a Sony Young Faculty Award, a Project X Innovation Award, and an Amazon Science Research Award.

REFERENCES

Kaan Akşit, Ward Lopes, Jonghyun Kim, Peter Shirley, and David Luebke. 2017. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 1–13.

A Terry Bahill, Michael R Clark, and Lawrence Stark. 1975. The main sequence, a tool for studying human eye movements. *Mathematical biosciences* 24, 3-4 (1975), 191–204.

S. Bahmani and J. Romberg. 2017. Phase Retrieval Meets Statistical Learning Theory: A Flexible Convex Relaxation. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (Proceedings of Machine Learning Research)*, Aarti Singh and Jerry Zhu (Eds.), Vol. 54. PMLR, Fort Lauderdale, FL, USA, 252–260.

Heinz H Bauschke, Patrick L Combettes, and D Russell Luke. 2003. Hybrid projection–reflection method for phase retrieval. *JOSA A* 20, 6 (2003), 1025–1034.

Edward Buckley, Adrian Cable, Nic Lawrence, and Tim Wilkinson. 2006. Viewing angle enhancement for two-and three-dimensional holographic displays with random superresolution phase masks. *Applied optics* 45, 28 (2006), 7334–7341.

Emmanuel J Candes, Thomas Strohmer, and Vladislav Voroninski. 2013. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics* 66, 8 (2013), 1241–1274.

Praneeth Chakravarthula, David Dunn, Kaan Akşit, and Henry Fuchs. 2018. Focusar: Auto-focus augmented reality eyeglasses for both real world and virtual imagery. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2906–2916.

Praneeth Chakravarthula, Yifan Peng, Joel Kollin, Henry Fuchs, and Felix Heide. 2019. Wirtinger holography for near-eye displays. *ACM Transactions on Graphics (TOG)* 38, 6 (2019), 213.

Praneeth Chakravarthula, Ethan Tseng, Henry Fuchs, and Felix Heide. 2022. Hogel-free Holography. *ACM Transactions on Graphics (TOG)* (2022).

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), 186.

Praneeth Chakravarthula, Zhan Zhang, Okan Tursun, Piotr Didyk, Qi Sun, and Henry Fuchs. 2021. Gaze-contingent retinal speckle suppression for perceptually-matched foveated holographic displays. *IEEE Transactions on Visualization and Computer Graphics* 27, 11 (2021), 4194–4203.

Suyeon Choi, Manu Gopakumar, Yifan Peng, Jonghyun Kim, and Gordon Wetzstein. 2021. 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.

Steffen Czolbe, Oswin Krause, Ingemar Cox, and Christian Igel. 2020. A Loss Function for Generative Neural Networks Based on Watson's Perceptual Model. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Inc., 2051–2061.

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 transactions on visualization and computer graphics* 23, 4 (2017), 1322–1331.

M Hossein Eybposh, Nicholas W Caira, Mathew Atisa, Praneeth Chakravarthula, and Nicolas C Pégard. 2020. DeepCGH: 3D computer-generated holography using deep learning. *Optics Express* 28, 18 (2020), 26636–26650.

James R Fienup. 1982. Phase retrieval algorithms: a comparison. *Applied optics* 21, 15 (1982), 2758–2769.

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

Tom Goldstein and Christoph Studer. 2018. PhaseMax: Convex phase retrieval via basis pursuit. *IEEE Transactions on Information Theory* (2018).

Joseph W Goodman. 2005. *Introduction to Fourier optics*. Roberts and Company Publishers.

Ryoichi Horisaki, Ryosuke Takagi, and Jun Tanida. 2018. Deep-learning-generated holography. *Applied optics* 57, 14 (2018), 3859–3863.

CK Hsueh and AA Sawchuk. 1978. Computer-generated double-phase holograms. *Applied optics* 17, 24 (1978), 3874–3883.

Changwon Jang, Kiseung Bang, Gang Li, and Byoungho Lee. 2019. Holographic near-eye display with expanded eye-box. *ACM Transactions on Graphics (TOG)* 37, 6 (2019), 195.

Changwon Jang, Kiseung Bang, Seokil Moon, Jonghyun Kim, Seungjae Lee, and Byoungho 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), 190.

Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In *European Conference on Computer Vision (ECCV)*.

Jonghyun Kim, Manu Gopakumar, Suyeon Choi, Yifan Peng, Ward Lopes, and Gordon Wetzstein. 2022. Holographic Glasses for Virtual Reality. In *Proceedings of the ACM SIGGRAPH*. 1–8.

Grace Kuo, Laura Waller, Ren Ng, and Andrew Maimone. 2020. High resolution étendue expansion for holographic displays. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 66.

LB Lesem, PM Hirsch, and JA Jordan. 1969. The kinoform: a new wavefront reconstruction device. *IBM Journal of Research and Development* 13, 2 (1969), 150–155.

Conny Lu, Praneeth Chakravarthula, Yujie Tao, Steven Chen, and Henry Fuchs. 2020. Improved vergence and accommodation via Purkinje Image tracking with multiple cameras for AR glasses. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 320–331.

D Russell Luke. 2004. Relaxed averaged alternating reflections for diffraction imaging. *Inverse problems* 21, 1 (2004), 37.

Andrew Maimone, Andreas Georgiou, and Joel S Kollin. 2017. Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 85.

Andrew Maimone and Junren Wang. 2020. Holographic optics for thin and lightweight virtual reality. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 67–1.

Stefano Marchesini, Yu-Chao Tu, and Hau-tieng Wu. 2016. Alternating projection, ptychographic imaging and phase synchronization. *Applied and Computational Harmonic Analysis* 41, 3 (2016), 815–851.

Jae-Hyeung Park and Mehdi Askari. 2019. Non-hogel-based computer generated hologram from light field using complex field recovery technique from Wigner distribution function. *Optics express* 27, 3 (2019), 2562–2574.

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), 185.

Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. 2018. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light: Science & Applications* 7, 2 (2018), 17141.

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.

Koki Wakunami, Po-Yuan Hsieh, Ryutaro Oi, Takanori Senoh, Hisayuki Sasaki, Yasuyuki Ichihashi, Makoto Okui, Yi-Pai Huang, and Kenji Yamamoto. 2016. Projection-type see-through holographic three-dimensional display. *Nature Communications* 7 (2016).

Andrew B Watson. 1993. DCT quantization matrices visually optimized for individual images. In *Human vision, visual processing, and digital display IV*, Vol. 1913. International Society for Optics and Photonics, 202–216.

Zaiwen Wen, Chao Yang, Xin Liu, and Stefano Marchesini. 2012. Alternating direction methods for classical and ptychographic phase retrieval. *Inverse Problems* 28, 11 (2012), 115010.

Xinxing Xia, Yunqing Guan, Andrei State, Praneeth Chakravarthula, Tat-Jen Cham, and Henry Fuchs. 2020. Towards Eyeglass-style Holographic Near-eye Displays with Statically Expanded Eyebox. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 312–319.

Jingzhao Zhang, Nicolas Pégard, Jingshan Zhong, Hillel Adesnik, and Laura Waller. 2017. 3D computer-generated holography by non-convex optimization. *Optica* 4, 10 (2017), 1306–1313.