Wirtinger Holography for Near-Eye Displays

PRANEETH CHAKRAVARTHULA, University of North Carolina at Chapel Hill
YIFAN PENG, Stanford University
JOEL KOLLIN, Microsoft Research
HENRY FUCHS, University of North Carolina at Chapel Hill
FELIX HEIDE, Princeton University

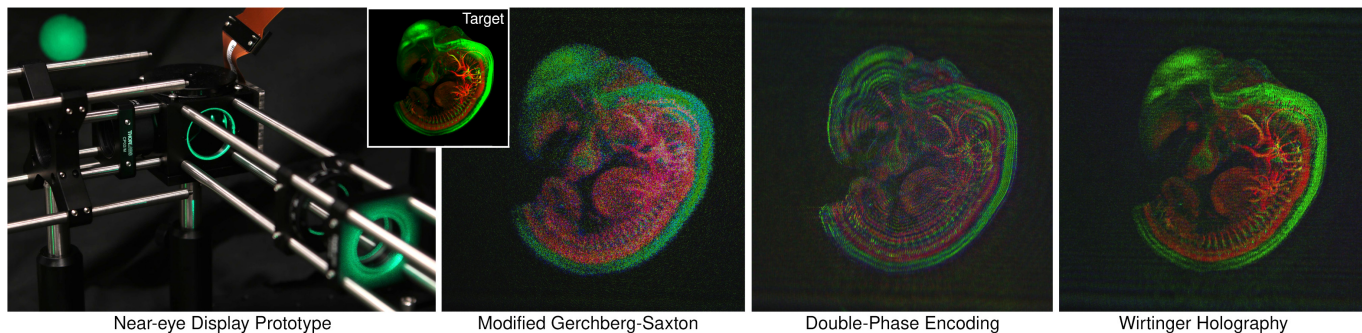| Near-eye Display Prototype | Modified Gerchberg-Saxton | Double-Phase Encoding | Wirtinger Holography |

Fig. 1. Wirtinger Holography on a Display Prototype. To compute phase-only modulation patterns, we depart from existing iterative projection algorithms, such as error-reduction Gerchberg-Saxton methods [Gerchberg 1972; Peng et al. 2017], and heuristic encoding approximations, such as the double phase encoding method [Hsueh and Sawchuk 1978]. Instead, we revisit the use of formal optimization using complex Wirtinger derivatives for the underlying phase retrieval problem. We built a near-eye display prototype using a phase-only spatial light modulator (SLM) and off-the-shelf optics (left). Compared to holographic reconstructions at a set focal distance from the existing methods [Peng et al. 2017] (center left) and [Maimone et al. 2017] (center right), the proposed Wirtinger holography substantially reduces reconstruction artefacts on our prototype, while achieving an order of magnitude reduced error in simulation. Mouse embryo image by Miltenyi Biotec.

Near-eye displays using holographic projection are emerging as an exciting display approach for virtual and augmented reality at high-resolution without complex optical setups — shifting optical complexity to computation. While precise phase modulation hardware is becoming available, phase retrieval algorithms are still in their infancy, and holographic display approaches resort to heuristic encoding methods or iterative methods relying on various relaxations.

In this work, we depart from such existing approximations and solve the phase retrieval problem for a hologram of a scene at a single depth at a given time by revisiting complex Wirtinger derivatives, also extending our framework to render 3D volumetric scenes. Using Wirtinger derivatives allows us to pose the phase retrieval problem as a quadratic problem which can be minimized with first-order optimization methods. The proposed Wirtinger

Authors' addresses: Praneeth Chakravarthula, cpk@cs.unc.edu, University of North Carolina at Chapel Hill, Chapel Hill, NC; Yifan Peng, evanpeng@stanford.edu, Stanford University, Stanford, CA; Joel Kollin, jkollin@microsoft.com, Microsoft Research, Redmond, WA; Henry Fuchs, fuchs@cs.unc.edu, University of North Carolina at Chapel Hill, Chapel Hill, NC; Felix Heide, fheide@cs.princeton.edu, Princeton University, Princeton, NJ.

Holography is flexible and facilitates the use of different loss functions, including learned perceptual losses parametrized by deep neural networks, as well as stochastic optimization methods. We validate this framework by demonstrating holographic reconstructions with an order of magnitude lower error, both in simulation and on an experimental hardware prototype.

CCS Concepts: • **Hardware → Emerging optical and photonic technologies**; • **Applied computing → Physics**; • **Human-centered computing → Displays and imagers**; • **Computing methodologies → Computer graphics**.

Additional Key Words and Phrases: computer generated holography, near-eye display, augmented reality, virtual reality, vergence-accommodation conflict, computational displays

1 INTRODUCTION

Over the last few years, near-eye display approaches have been emerging at a rapid pace, ultimately promising practical and comfortable virtual reality (VR) and augmented reality (AR) in the future. Head-mounted displays for stationary virtual reality use-cases have become ubiquitous consumer products, having solved resolution, field of view (FOV), tracking and latency limitations of the past

decades of research. Today's AR devices partially inherit these advances but are still limited to large form factors, and are unable to allow for occlusions and continuous focus cues needed to avoid the vergence-accommodation conflict. Approaches using conventional optics to address these individual problems are often restricted to large setups, possibly involving actuation [Chakravarthula et al. 2018]. Near-eye displays using light field modulation [Huang et al. 2015a; Lanman and Luebke 2013] offer an alternative approach with impressive results but only provide sparse ray-space sampling at acceptable spatial resolutions — the spatio-angular resolution trade-off is fundamentally limited by diffraction.

In theory, digital holography offers an elegant way to solve the complex optical design challenges for near-eye displays. Instead of physically controlling the emitted wavefront of light with cascades of (variable) refractive optics and conventional displays, holographic displays shift this process to computation. Given a scene intensity image and associated depth map, holographic display methods solve for the states of a phase (and amplitude) modulator to encode the corresponding phase in an incident source wavefront. Assuming for a moment that the encoding process is perfect, this approach allows for high-resolution and continuous focus cues. Building on precise phase modulation hardware and a large body of existing work in holography, the recent seminal works of Maimone et al. [2017] and Shi et al. [2017] demonstrate impressive high-FOV holographic near-eye displays in a light-weight wearable form-factor. Although these approaches offer insight into the promise of holographic near-eye displays, they rely on heuristic holographic phase encoding schemes that severely restrict the achievable image quality and future research. Akin to early image processing methods, it is challenging to solve for different loss functions, or research additional modulation schemes or hybrid systems.

In this work, we deviate from heuristic solutions to holographic phase retrieval. Instead, we rely on formal optimization enabled by complex Wirtinger derivatives. With a differentiable forward model and its (Wirtinger) gradient in hand, we formulate a quadratic loss function that is solved via a first-order optimization method. Specifically, we demonstrate that the problem is not required to be relaxed, as in recent lifting methods [Candes et al. 2015], but can be directly solved with standard methods. We show that the phase retrieval problem for digital fresnel holography can be solved efficiently using a quasi-Newton method or stochastic gradient descent. We achieve an order of magnitude lower reconstruction error, i.e. 10 dB PSNR improvement. We validate the proposed method in simulation and using an experimental prototype, demonstrating that this improvement eliminates severe artefacts present in existing methods. We validate the flexibility of the proposed phase retrieval method by modifying the objective with a learned perceptual loss, which can be optimized using vanilla stochastic gradient descent (SGD) methods. Moreover, we show, in simulation, that the proposed framework allows to incorporate additional optics or modulation in a plug-and-play fashion.

Specifically, we make the following contributions in this work:

- We introduce complex Wirtinger derivatives for holographic display formation, that allow us to directly solve holographic phase retrieval problems as formal optimization problems using first-order optimization algorithms.

- We validate the proposed differentiable framework by solving phase retrieval of holographic projections with constant focus over the image as a least-squares problem using off-the-shelf quasi-Newton optimization. The resulting holographic reconstructions have an order of magnitude lower error than previous methods.

- We demonstrate the flexibility of the Wirtinger framework by solving for a perceptual loss function parameterized by a convolutional deep neural network. Moreover, in simulation, we show that the method facilitates alternative optical setups such as cascaded modulation on two sequential phase SLMs.

- We assess the proposed framework experimentally with a prototype near-eye holographic display setup. The proposed method reduces severe artefacts of existing holographic imaging approaches.

Limitations. Although the proposed approach provides unprecedented holographic phase retrieval quality, while being flexible as a formal optimization framework, we do not achieve real-time frame rates on consumer hardware. The runtime is comparable to that of the modified Gerchberg-Saxton method from [Peng et al. 2017], at about 30 sec for a full HD hologram on a consumer laptop computer with partial GPU acceleration. However, akin to ray tracing, which recently has been enabled at real-time rates using dedicated hardware, we hope that a low-level implementation on next-generation GPU hardware, or dedicated hardware, could overcome this limitation in the future. In the meantime, it is practical to use a computationally cheap initial iterate combined with Wirtinger refinement. While optimizing for 3D holograms (defined as a hologram which addresses points at different depths simultaneously) is not within the scope of this work, we anticipate the extension of Wirtinger holography to 3D holograms by us or others in the future. In addition, it is possible to extend the technique described in this paper to render high-quality 3D volumetric scenes by dynamically adjusting the focal distance to the foveal region of interest, as demonstrated earlier by [Maimone et al. 2017].

2 RELATED WORK

Computational holograms are traditionally classified into Fourier holograms corresponding to far-field Fraunhofer diffraction, and Fresnel holograms which produce images in the near field, as determined by the Fresnel Number. The computational complexity between near-field and far-field holograms is substantial. While computing far-field holograms using 2D Fourier transforms can be accomplished using fast algorithms and optimized hardware, computing near-field Fresnel holograms admits no simple analytic solution. In fact, it is mathematically equivalent to inverting a generalized scalar diffraction integral [Underkoffler 1991]. A majority of work on this topic investigates numerical techniques for generating Fresnel fringe patterns. In this section, we briefly summarize prior art related to hologram computation and display.

2.1 Traditional Phase Retrieval Algorithms

A Fourier hologram produces a flat image at a far distance Fourier plane, and such a hologram is often computed using traditional phase retrieval techniques. Phase retrieval is the method of recovering an unknown signal from the measured magnitude of its Fourier transform. Since the phase is lost in the measurement of the signal, the inverse problem of recovering it is generally ill-posed. However, the phase can be perfectly recovered, in theory, by solving a set of non-linear equations if the measurements are sufficiently oversampled [Bates 1982]. Early methods of phase retrieval included error reduction methods using iterative optimization [Gerchberg 1972; Lesem et al. 1969], together with an assumption on a non-zero support of the real-valued signal, with applications in optics, crystallography, biology and physics. Extension 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]. Phase-retrieval methods using first-order non-linear optimization have been explored in the past to characterize complex optical systems [Fienup 1993; Gonsalves 1976; Lane 1991], eventually also sparking recent work on using alternative direction methods for phase retrieval [Marchesini et al. 2016; Wen et al. 2012], non-convex optimization [Zhang et al. 2017], and methods overcoming the non-convex nature of the phase retrieval problem by lifting, i.e. relaxation, to a semidefinite [Candes et al. 2013] or linear program [Bahmani and Romberg 2017; Goldstein and Studer 2018].

2.2 Computational Fresnel holograms

Fresnel hologram computation can be categorized into two classes. 1) Geometry-based techniques, which model three-dimensional scene geometry as a collection of emitters; either point-emitters (point-source method) or polygonal tiles (polygon-based method). The collective interference of these emitters with a reference wave is computed at a set of discretized locations throughout the combined field to generate a hologram of the scene [Benton and Bove Jr 2008]. 2) Image-based techniques, which leverage the advantage of computer graphics rendering techniques along with wave propagation. Next, we review geometry and image-based methods.

Point-source methods. Waters et al. [1966] were the first to propose using a collection of points rather than finite sized objects to model a scene for holographic image generation. By using a look-up table of precomputed elemental fringes, Lucente et al. [1993] sped up the point-source hologram computation to under one second. Recent point-source based CGH computation methods leverage the parallelization of a computer graphics card [Chen and Wilkinson 2009; Masuda et al. 2006; Petz and Magnor 2003]. Most recent work of Maimone et al. [2017] present point-source based CGH computation for holographic near-eye displays for both virtual and augmented reality. All of these methods have in common that the optical transmission process of different primitives is modeled to be independent, and hence, it is challenging to accurately represent view-dependent mutual-occlusion and shading. Moreover, a continuous parallax demands a very dense set of point-sources which requires a large compute budget.

Polygonal methods. Polygonal methods for computing CGH also have existed for decades [Leseberg and Frère 1988]. The basic idea of this approach to CGH is to represent a 3D object as a collection of tilted and shifted planes. The diffraction patterns from a tilted input plane can be computed by the fast Fourier transform (FFT) algorithm and an additional coordinate transformation. This analysis can also be done in the spatial frequency domain by using the translation and rotation transformations of the angular spectra permitting the use of FFT algorithms [Tommasi and Bianco 1993]. Moreover, a property function can be defined for each planar input surface to provide texture [Matsushima 2005] and shading [Ahrenberg et al. 2008; Matsushima 2005]. Researchers also have explored geometric facet selection methods by ray tracing [Kim et al. 2008], silhouette methods [Matsushima and Nakahara 2009; Matsushima et al. 2014] and inverse orthographic projection techniques [Jia et al. 2014] to provide occlusion culling effects.

Image-based methods. Image-based holography techniques can be broadly categorized into light field and layer-based methods. Light field holograms, also known as holographic stereograms, partition a hologram spatially into elementary hologram patches, each producing local ray distributions (images) that together reconstruct multiple views supporting intra-ocular occlusions [Lucente and Galyean 1995; Smithwick et al. 2010; Yamaguchi et al. 1993]. Holographic stereograms can be paired with the point-source method to enhance the image fidelity to provide accommodation and occlusion cues [Shi et al. 2017; Zhang et al. 2015]. Layer-based methods, in contrast, compute the hologram by slicing objects at multiple depths and superimposing the wavefronts from each slice on the hologram plane [Bayraktar and Özcan 2010; Zhao et al. 2015]. Layer-based and light field methods can both be combined to produce view-dependent occlusion effects [Chen and Chu 2015; Zhang et al. 2016]. While prior work on holographic phase retrieval addressed computing holograms for both 2D and 3D scenes, the focus of this paper lies on rendering high-quality images of scenes at a single depth plane. The method can be extended to volumetric scenes as discussed in Section 6.2.

3 COMPUTATIONAL DISPLAY HOLOGRAPHY

Computational display holography (CDH) simulates the physical processes of an optical hologram recording, and reconstruction, using numerical methods. The computed holographic image is brought to life by a holographic display, typically consisting of an illuminating source and a phase-modifying (sometimes also amplitude-modifying) spatial light modulator (SLM). The phase of the SLM describes the delay of the incident wave phase introduced by the SLM element. Note that the geometrical (ray) optics model, commonly used in computer graphics, models light as rays of photon travel instead of waves. Although being an approximation to physical optics, i.e. ignoring diffraction, ray optics still can provide an intuition: the perpendiculars to the waves can be thought of as rays, and, vice versa, phase intuitively describes the relative delay of photons traveling along these rays. We refer the reader to [Goodman 2005] for a detailed review.
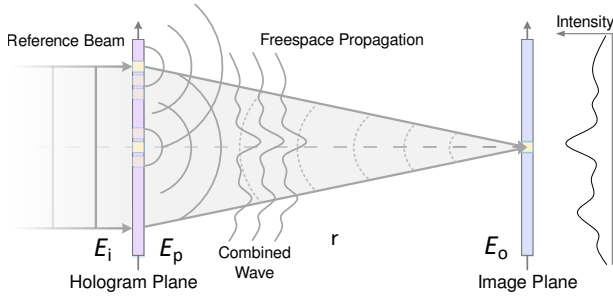
Fig. 2. Schematic of scalar diffraction model, that the incident wave (E_i) first gets modulated by the phase delays on the hologram, where each pixel then emits a sub-wave (E_p) that propagates in free space to the image plane. The complex amplitude of one point on the image plane (E_O) is the integration of sub-waves propagated from all pixels on the hologram plane. Note, r here represents the Euclidean distance. As such, given a target propagated distance and a principle wavelength, a sharp point spread function on the image plane can be obtained under the point-source propagation model, corresponding to the phase distribution of a lens on the hologram plane.

In the following, we review Fresnel holography, which can produce near-field imagery for near-eye display applications. Specifically, we review propagation for scenes modeled as point source emitters and a convolutional forward model, which both provide intuition for the image formation process. The proposed method is not limited to either of these propagation models but, indeed, flexible to support a variety of different propagation models, including Fourier-based propagation and the angular spectrum propagation, all of which we review in the supplemental material.

3.1 Point-source Propagation

Computing a hologram requires knowledge of both the illuminating reference wave and the object wave. Given a 3D model of the scene, the object wave can be modeled as a superposition of waves from spherical point emitters at the location of object points and propagating towards the hologram plane where the SLM is located, see Fig. 2. The object wave E_O reaching a pixel at location $\vec{p}$ on the SLM at the hologram plane, from a point source located at an object point at $\vec{o}$, can be expressed as

$$E_O(\vec{p}) = a_o \mathbf{exp}^{j\left(\phi_o + \frac{2\pi r(\vec{p}, \vec{o})}{\lambda}\right)}, \qquad (1)$$

where λ is the illumination reference wavelength, a_o is the amplitude of wave at the object point $\vec{o}$, ϕ_o is the (initial) phase associated with each diffuse object point and $r(\vec{p}, \vec{o})$ is the Euclidean distance between the object point and the SLM pixel.

Equation 1 describes the phase patterns corresponding to each object point on the hologram plane. Under a paraxial approximation these phase patterns act as Fresnel lenses, also see supplemental material. Owing to the optical path difference, the spatial frequencies of these phase patterns increase from the center towards the edges, where the light is deflected maximally. However, the maximum deflection angle supported by an SLM is limited, hence restricting the lens phase patterns to a smaller region. Therefore, the overall hologram can be thought of as a superposition of local lens phase patterns, or phase Fresnel lenses, corresponding to individual target

object points

$$H(\vec{p}) = \sum_{o \in s_p} a_o \mathbf{exp}^{j\left(\phi_o + \frac{2\pi r(\vec{p}, \vec{o})}{\lambda}\right)} = \sum_{o \in s_p} a_o L_o(\vec{p}), \qquad (2)$$

where s_p is the set of object points whose phase Fresnel lenses are defined at a pixel $\vec{p}$, and L is the lens phase function of the phase Fresnel lens for a given object point.

3.2 Convolutional Propagation

The point-wise integration described in the previous paragraphs requires intensive computation when all object points of a full three-dimensional scene are considered. If we make the assumption that the scene is located on a single depth plane, this integration is equivalent to the convolution of the target image with a complex valued lens phase function L [Maimone et al. 2017]

$$H = A * L, \qquad (3)$$

where $*$ is the complex-valued convolution operator, A represents the complex amplitude of the underlying image, and L is a spatially invariant lens phase function which is the same for all object points $\vec{o}$. Although the focus, i.e. depth, over the image is constant, the focus may be changed globally either on a per-frame basis or in a gaze-contingent manner by employing eye tracking [Maimone et al. 2017].

For a given hologram H from Equation 3, an approximate holographic image on the image plane, after propagating the modulated wave field, can be formulated as a convolution of the hologram with the complex conjugate of the lens function

$$I = |H * \overline{L}|^2 = |\mathcal{F}^{-1}(\mathcal{F}[H] \circ \mathcal{F}[\overline{L}])|^2, \qquad (4)$$

where $| \cdot |^2$ denotes the element-wise absolute value squared operation, $\overline{L}$ is the conjugate lens phase function, $\mathcal{F}$ is the Fourier transform operator, $\circ$ is the Hadamard element-wise multiplication and I is the intensity of the image.

4 WIRTINGER HOLOGRAPHY

In this section, we present a framework for computing high-quality phase-only holograms for near-eye display applications. We start by formulating phase hologram computation as a general optimization problem independently of the propagation model, and only assuming that this forward model is *differentiable*. Once cast as an optimization problem, we show how first-order optimization techniques can be applied to solve this complex non-convex problem.

For the sake of brevity, the penalty functions and corresponding gradients in this section are defined for matrices in *vectorized* form. For the ease of notation, we use matrices and their vectorized versions interchangeably. Note that, in practice, we do not explicitly form million-dimensional vectors for optimizing holograms, but instead perform all operations on tensors for efficient memory usage and computation.

4.1 Synthesizing Optimal Phase Holograms

Section 3 discussed the computation of Fresnel holograms, and the reconstruction of holographic images by propagating the field from the hologram plane to the image plane using different propagation models. If z is the complex wave field at the image plane, the error

between the reconstructed image $|z|^2$ produced by the hologram H and the target image I can be computed as

$$Err = f(|z|^2, I) = f(z), \tag{5}$$

for a general penalty function f. Note that the field z is a function of the hologram H, and is the output of the propagation function $\mathcal{P}$, as discussed in Section 3, that is

$$z = \mathcal{P}(H). \tag{6}$$

The phase-only hologram H has a constant amplitude across the hologram plane with each of its entries of the form

$$H(\Phi) = c\mathbf{exp}^{j\Phi} = c(h_i) = c\mathbf{exp}^{j\phi_i}, \tag{7}$$

where c is the (constant) amplitude, and ϕ_i is the phase associated with the i-th pixel on the SLM located at the hologram plane. Without loss of generality, the value of c may be set to unity for all practical purposes. We are interested in computing the phase-only hologram to display on a phase modulating SLM which can be formulated as solution to the following non-convex error minimization problem

$$\Phi_{opt} = \underset{\Phi}{\text{minimize}} \quad \underbrace{f(|\mathcal{P}(H(\Phi))|^2, I)}_{Err(\Phi)} + \gamma||\nabla\Phi||^2 \quad , \tag{8}$$

where $||\nabla\Phi||^2$ is an optional regularizer on the phase patterns, which may, for example, enforce constraints on the phase variations to account for SLM limitations, e.g., limited modulation range of frequency. We show that this non-convex holographic phase retrieval problem can be solved using first-order optimization algorithms. In the following section, we discuss calculating the Wirtinger gradient for the objective function before describing how we apply first-order optimization methods.

4.2 Computing the Complex Gradient

Chain rule. The gradient of the objective from Equation 8 can be calculated from Equations 5, 6 and 7, using chain rule as follows

$$\frac{d(Err)}{d\Phi} = \underbrace{\frac{df}{dz}}_{I} \underbrace{\frac{dz}{dH}}_{II} \underbrace{\frac{dH}{d\Phi}}_{III} . \tag{9}$$

Non-Holomorphic Objective. We assume that our objective is a real-valued non-convex function of complex variables, that is $f : \mathbb{C}^n \mapsto \mathbb{R}$, where n is the number of elements in z (equal to that in H). A complex function that is complex differentiable at every point in its domain is called a holomorphic function. However, a real-valued function of a complex argument cannot be holomorphic unless it is a constant — its derivative is not defined in its domain. The derivative of such a holomorphic real-valued function is always zero, refer to supplemental material. Since our objective function is not a constant, the derivatives of any order are not defined in its domain. We note that [Fienup 1993] arrived at a relatively similar analytic expression for the gradient of an l_2-penalty, see discussion in the supplemental material.

Next, we introduce Wirtinger derivatives [Remmert 2012] as a gradient approximation that is defined. Specifically, in Section 4.3, we discuss computing Part *I* of Equation 9 which is the derivative of the scalar error with respect to the complex wave field on the image plane. Part *II* of the chain rule expression in Equation 9 depends on the propagation model. We derive this derivative for the convolutional propagation model in Section 4.4, to which we apply Part *III*. We derive Wirtinger derivatives for other propagation models in the supplemental material.

4.3 Wirtinger Derivatives

To overcome the undefined nature of the gradient of Part *I*, we use the following approximation for partials of a scalar function of a complex vector $f : \mathbb{C}^n \mapsto \mathbb{R}$

$$d(Err) = df(z) = Re\langle \nabla f, dz \rangle, \tag{10}$$

where Re denotes the real part of a complex number and $\langle ., . \rangle$ denotes the inner product of two vectors. This definition, although not the exact gradient, is consistent with the main properties of gradients used in first-order optimization methods: 1) the gradient defines the direction of maximal rate of change of the function, and 2) the gradient being zero is a necessary and sufficient condition to determine a stationary point of a function.

To obtain ∇f, we compute the Wirtinger derivatives of the function f with respect to the complex argument z. Assume $f(z)$ to be a function of two independent vectors z and $\bar{z}$, its Wirtinger derivative can be defined as

$$\begin{aligned} df &= (\nabla_z f)^T dz + (\nabla_{\bar{z}} f)^T d\bar{z} \\ &= (\nabla_z f)^T dz + \overline{(\nabla_z f)^T dz} \\ &= 2Re\left((\nabla_z f)^T dz \right) \\ &= Re\langle 2\nabla_{\bar{z}} f, dz \rangle \end{aligned} . \tag{11}$$

We simplify the computation of $\nabla f(z)$ as a scaled gradient of f with respect to only $\bar{z}$

$$\nabla f(z) = 2\nabla_{\bar{z}} f. \tag{12}$$

Equation 12 is valid for any general scalar function of complex variables. However, following Equation 5, we further decompose Equation 12 as

$$\nabla_{\bar{z}} f = \underbrace{\frac{df}{d(|z|^2)}}_{A} \circ 2\underbrace{\nabla_{\bar{z}}(|z|^2)}_{B}. \tag{13}$$

Note that $|z|^2$ is the intensity of the reconstructed image on the image plane. The derivative of a scalar loss with respect to the reconstructed holographic image (Part *A*) is defined for any differentiable penalty function — even for a neural-network based loss function, where one can obtain gradients using back propagation. To plug a custom loss function, standard or neural network learned, into our framework only requires a Hadamard product of the gradient of the error function with respect to the predicted image (Part *A*) and the Wirtinger derivative of field intensity on the image plane (Part *B*), which is given by the scaled value of the field itself as

$$2\nabla_{\bar{z}}(|z|^2) = 2\nabla_{\bar{z}}(z\bar{z}) = 2z. \tag{14}$$

For example, for an l_2-penalty, i.e., setting $f(\cdot) = \|\cdot\|^2$, the value of ∇f can be obtained as follows

$$\nabla f = 2\nabla_{\bar{z}}f \quad = \quad \overbrace{\frac{d\left(\frac{1}{2}\||z|^2 - I\|^2\right)}{d(|z|^2)}}^{A} \circ \overbrace{2\nabla_{\bar{z}}(|z|^2)}^{B} \quad (15)$$

$$= \quad 2\left(|z|^2 - I\right) \circ 2z.$$

Note: In the following sections, vec denotes the vectorization operator for a multi-dimensional signal, $\mathcal{F}[.]$ denotes a Discrete Fourier Transform (DFT) operator applied to multi-dimensional signal, and F represents the corresponding matrix applied to the vectorized signal, in the sense that

$$\mathcal{F}[H] = F\,vec(H). \quad (16)$$

The operator $\circ$ denotes the element-wise (Hadamard) product, and $F^\dagger$ is the Hermitian (conjugate transpose) of F. In the following, we use the fact that F is unitary, and therefore $F^{-1} = F^\dagger$. We introduce the DFT matrix F here for notational brevity. In practice, the DFT is performed by the Fast Fourier Transform (FFT) algorithm and the matrix F is never explicitly constructed. Note also that H is a function of Φ, the phase that we are interested in, and FH actually means $F\,vec(H)$, as shorthand notation. All multiplications, unless specifically stated, are element-wise multiplications.

Next, we derive the gradient for the convolutional propagation model from Sec. 3.2. Please refer to the supplemental material for the gradients for other propagation models.

4.4 Gradient for Convolutional Propagation

As discussed in Section 3.2, the image on the destination plane can be computed as $I' = |H*L|^2$. Expressing convolution as a multiplication in the Fourier domain, the wave field at the destination plane can be obtained as $z = \mathcal{F}^{-1}(\mathcal{F}[H] \circ \mathcal{F}[L])$. Deriving Part II from the partial computed in Equation 10 yields

$$\begin{aligned} d(Err(H)) &= Re\langle\nabla f, dz\rangle \\ &= Re\langle\nabla f, d(F^\dagger(FH)(FL))\rangle \\ &= Re\langle\nabla f, F^\dagger FLFdH\rangle \\ &= Re\langle F^\dagger(FL)^\dagger F\nabla f, dH\rangle \\ &= Re\langle F^\dagger(FL)^* F\nabla f, dH\rangle \end{aligned} \quad (17)$$

Finally, evaluating Part III with $H = \mathbf{exp}^{j\Phi}$, we derive the partial of the loss function with respect to the phase Φ as follows

$$\begin{aligned} d(Err(\Phi)) &= Re\langle F^\dagger(FL)^* F\nabla f, d(\mathbf{exp}^{j\phi})\rangle \\ &= Re\langle -j\mathbf{exp}^{-j\phi}F^\dagger(FL)^* F\nabla f, d(\phi)\rangle \end{aligned} \quad (18)$$

Since the phase Φ is real-valued, the above inner-product becomes the gradient, that is

$$\nabla Err(\Phi) = Re(-j\mathbf{exp}^{-j\phi}F^\dagger(FL)^* F\nabla f). \quad (19)$$

5 SETUP AND IMPLEMENTATION

We assess the holograms computed by the proposed method in simulation and experimentally using a hardware display prototype. We use several loss functions, including a neural network based learned perceptual error metric, and compute all the corresponding gradients according to Section 4.3.
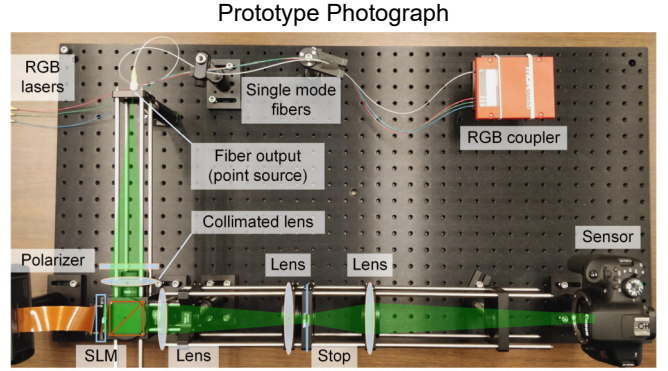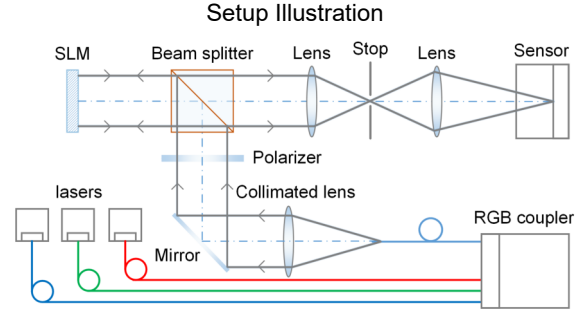
Fig. 3. Schematic of our prototype holographic near-eye display. We couple lasers of three different wavelengths into a single-mode fiber whose output end can be treated as a point source. The light emitted from the single-mode fiber is polarized and collimated before incident on the phase-only reflective SLM, where the holograms are displayed. The modulated wave is imaged using a camera at the eye position.

5.1 Hardware Prototype

The prototype display system for experimental validation is similar to the one demonstrated by Maimone et al. [2017] but differs in a variety of implementation details. We use plane wave illumination in the projection light engine. Our prototype display includes a HOLOEYE LETO-I liquid crystal on silicon (LCoS) reflective phase-only spatial light modulator with a resolution of 1920×1080. The pixel pitch of the SLM is 6.4 μm, resulting in the active area of 12.28 mm $\times$ 6.91 mm. We use a single optical fiber that is coupled to three laser diodes emitting at wavelengths 446 nm, 517 nm, and 636 nm, in combination with collimating optics, to illuminate the SLM. The laser power is controlled by a laser diode controller, and the output is linearly polarized as required by the SLM. The SLM has a refresh rate of 60 Hz, and is illuminated by the laser source in a color field sequential manner.

Figure 3 shows the optical design of the proposed prototype system. The light from the laser diodes is collimated using an achromatic doublet before being modulated by the LETO phase SLM, and then focused by a 150mm achromatic lens. An iris is placed at an intermediate image plane to discard unwanted higher diffraction orders, especially for double phase encoding holograms. Finally, a combination of 80 mm and 150 mm achromatic lenses is used to relay the image to the SLM sensor plane. An optional zero-order

stop can be placed immediately after the 80mm achromat. The optical relay lens system was optimized to have minimal aberrations, allowing a fair evaluation of different methods with well-corrected images. To assess the display's image quality, we capture images from the prototype holographic display using a Canon Rebel T6i DSLR camera body (without camera lens attached) with an output resolution of 6000×4000, and a pixel pitch of 3.72 μm.

Note that achieving higher image quality in our experimental setup requires an intermediate image plane where zero-order stop and other filter stops can be placed to filter out undiffracted and higher-order diffracted light. Using off-the-shelf optics to create a high-quality image relay resulted in a fairly long optical system. Miniaturizing our setup to an actual head-mounted display remains an exciting area of future research.

5.2 Implementation

We compute phase-only holograms on a PC with an Intel Xeon 2.4GHz CPU and 64GB of RAM, and an NVIDIA GeForce GTX 1080 GPU. For quadratic loss functions we use quasi-Newton optimization, and implement the proposed method in MATLAB using the L-BFGS [Liu and Nocedal 1989] quasi-Newton method. In particular, we use the minFunc optimization package [Schmidt 2005]. For loss functions requiring stochastic gradient optimization, such as neural-network based losses, we implement the proposed method in TensorFlow [Abadi et al. 2016] and use the Adam optimizer. Specifically, we use a learning rate of 0.001 and an exponential decay rate of 0.9 for the first moment estimate, and an exponential decay rate of 0.999 for the second moment estimate. We obtain the gradient for the loss function component (i.e Part I of Equation 9) from backpropagation in TensorFlow, especially for losses parameterized by convolutional neural networks. We then use this gradient to compute the complex Wirtinger gradient (i.e. Part II and Part III of Equation 9) and feed it back to the optimizer for computing optimal phase holograms. Full color holograms are generated by performing optimization sequentially for all three color channels.

To compare runtimes of our Wirtinger Holography algorithm against the state of the art modified-GS method [Peng et al. 2017] and double phase amplitude encoding of Fresnel holograms [Maimone et al. 2017], we measure computation time needed to generate a 1920×1080 hologram both on MATLAB running only on CPU, and on Tensorflow with native GPU support.

MATLAB. We compute phase holograms on a per-channel basis on the CPU, in MATLAB, for all three methods. The double phase encoding method, using only one convolution, is the fastest with only 7 sec on MATLAB when performed using FFTs for a large lens kernel size. The Modified GS algorithm takes about 95 sec for 60 iterations for generating a full HD image. Note that increasing the number of iterations also increases the quality of GS hologram. Without warmstarting and initializing with random phase, our MATLAB implementation takes about 70 sec to typically surpass a PSNR of 30dB in about 20 iterations.

TensorFlow. To compute phase patterns for all three colors using TensorFlow with GPU acceleration, modified GS takes about 40 sec for 60 iterations, double phase encoding method takes about 10 ms
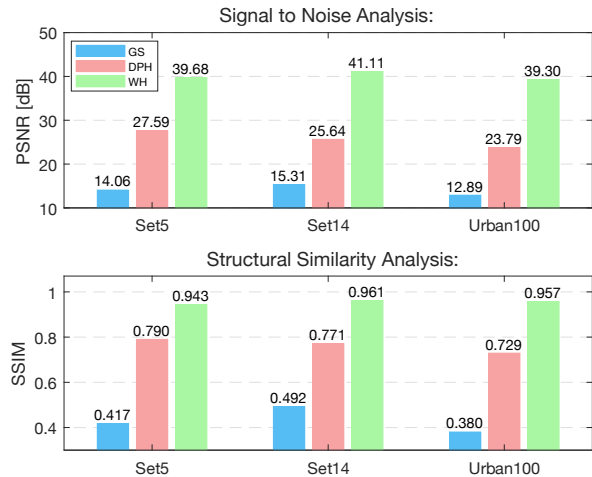

Fig. 4. Evaluation of reconstructed holographic images for three different datasets. The GS-based phase retrieval and our algorithm ran until the PSNR values were converged. The evaluation is done on grayscale images for a 532 nm green laser source.

while our method, initialized with random phase, takes about 30 sec before reaching a PSNR of over 30dB in about 200 iterations using the Adam optimizer with the parameters discussed above.

Note that the runtimes for Wirtinger Holography is comparable or less than that of modified GS algorithm, and achieve significant quality improvement in very few iterations.

6 ANALYSIS

In this section, we analyze the proposed method in simulation, and we validate its flexibility for different objective functions, forward models and setup configurations in simulation. First, we compare our approach against the state-of-the-art holographic phase retrieval methods from [Peng et al. 2017], which is a modified Gerchberg-Saxton method (GS), and the double phase holography method (DPH) proposed in [Maimone et al. 2017]. For quantitative comparisons of the reconstructed holographic images, we evaluate peak signal-to-noise ratio (PSNR) as a mean-squared error metric, and structural similarity Index (SSIM) for perceptual image quality.

Figure 4 lists PSNR and SSIM metric evaluation results for the proposed method, using an ℓ_2 loss and convolutional forward model, compared against modified GS phase retrieval, and DPH, over the following super-resolution image benchmark datasets: Set5 [Bevilacqua et al. 2012], Set14 [Zeyde et al. 2010] and Urban100 [Huang et al. 2015b]. Among these datasets, Set5 and Set14 consist of natural scenes and Urban100 contains challenging urban scenes images with details in different frequency bands. For a fair comparison, the intensity of the reconstructed images is adjusted to have the same mean value as their corresponding target images, as proposed in [Yoshikawa et al. 2016].

We find that our method produces superior reconstruction quality approaching 40 dB on an average. The proposed method outperforms double phase encoding holograms by *more than 10 dB* across all datasets and the modified GS method from [Peng et al. 2017] by
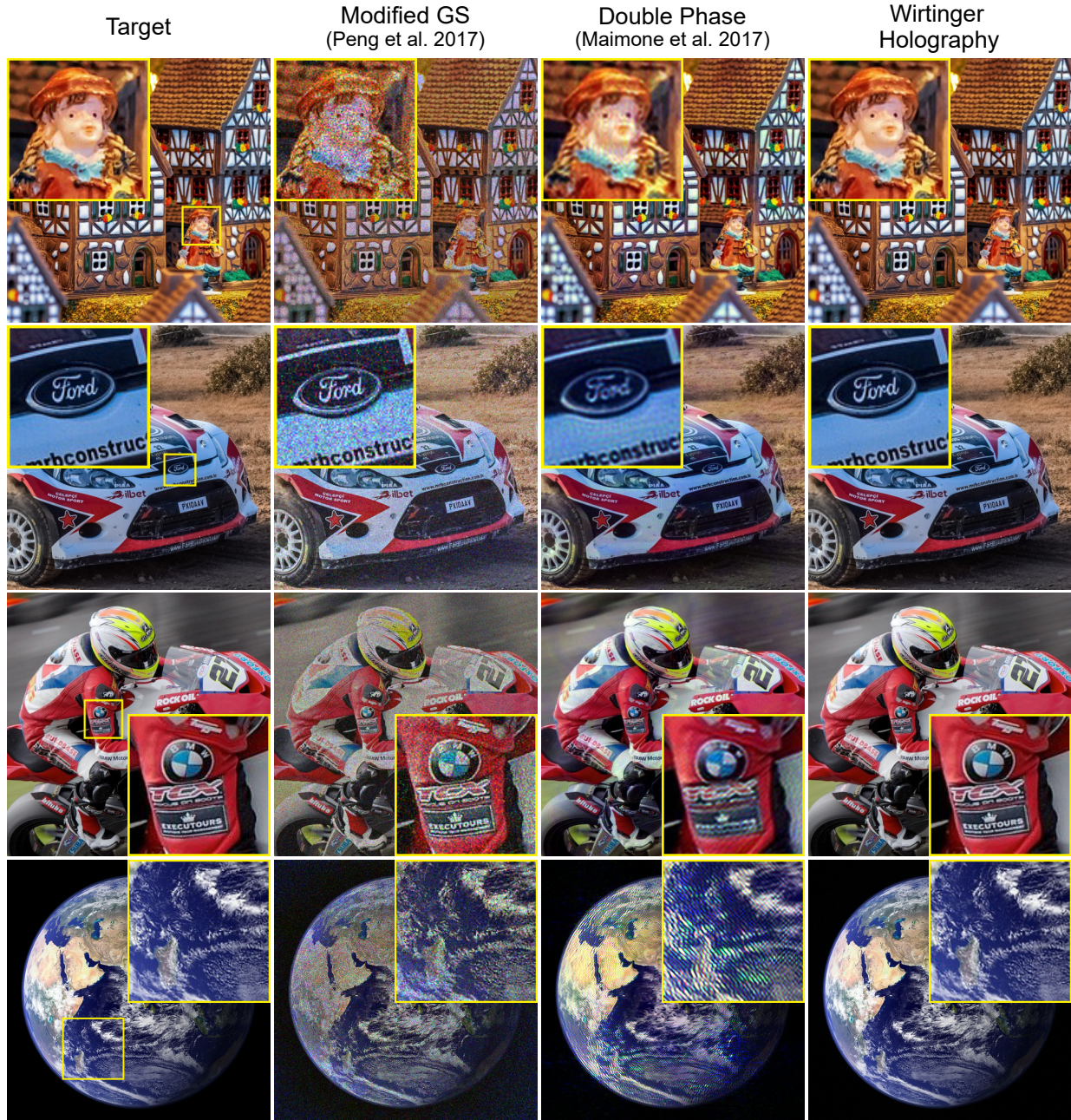
Fig. 5. Synthetic results reconstructed from holograms generated using different algorithms. From left to right, we show the reference, reconstructed results of modified GS iterative algorithm, double phase amplitude encoding direct propagation method, and the proposed Wirtinger holography, respectively. Note, for this comparison we set the propagated distance 200 mm, hologram pixel pitch 6.4 μm, and three principal wavelengths 638 nm, 520 nm, 450 nm for assessment. Image credits: Chritmas motif by Couleur, Rally by Dimitris Vestsikas, Motorbike by Steve Sewell and Earth by WikiImages.

more than 20 dB on average. Also note that the SSIM measurements validate that our method produces reconstructions with perceivably fewer errors as compared to double-phase holograms. Figure 5 shows qualitative comparisons for selected image patches. Please see additional examples in the supplemental material.

Resolution. The double phase encoding method from Maimone et al. [2017] produces a good approximation of the target image. However, encoding a single complex value into two phase values results in loss of resolution in holographic images. In contrast, our method produces high quality reconstructions without significant

loss of resolution. For example, notice the enhancement of text in Figure 5 (rows 2 and 3), and the reproduction of details (3rd row).

Ringing. Ringing is a ripple artefact that occurs near sharp edges. Although ringing artefacts from double phase encoding are subtle for smooth images, they are prominent in images with high frequency details. These artefacts are particularly apparent as they can differ in the color channels, resulting in chromatic ringing. The last row of Figure 5 shows an extreme case where ringing significantly degrades the holographic image quality produced by a double phase hologram. In contrast, the proposed method produces a faithful reconstruction of the image.

Noise and Contrast. Compared to the modified GS method from Peng et al. [Peng et al. 2017], the proposed method suffers from significantly less reconstruction noise, an artefact type typical to Gerchberg-Saxton-style methods. Moreover, Wirtinger Holography allows for improved contrast and reduced chromatic artefacts due to the improved quality across all color channels.

6.0.1 Alternative Penalty Functions. To validate the flexibility of the proposed framework, we show results for three different penalty functions for the proposed setup and the convolutional forward model. We evaluate the following penalty functions (see Sec. 4.1): 1) ℓ_2-norm least-squares objective 2) *MS-SSIM* [Wang et al. 2003] error minimization as a hand-crafted perceptual accuracy 3) the *learned perceptual similarity metric* [Zhang et al. 2018] as a deep neural network perceptual quality metric.

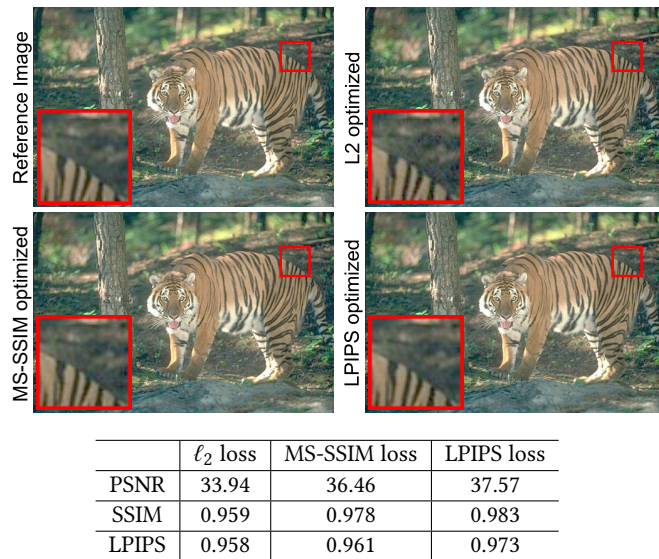|	ℓ_2 loss	MS-SSIM loss	LPIPS loss
PSNR	33.94	36.46	37.57
SSIM	0.959	0.978	0.983
LPIPS	0.958	0.961	0.973

Fig. 6. Quantitative and qualitative assessments of applying different penalty functions for the proposed Wirtinger holography method. Please zoom into the electronic version of this document for better viewing.

Figure 6 shows evaluations on a custom dataset of 20 images randomly picked from Set5, Set14, Urban100 and BSDS100 [Martin et al. 2001] datasets to cover images with a range of spatial frequencies. Optimizing for hand-crafted or learned perceptual loss

metrics improves the respective metric value by a significant margin over a least-squares loss function. The margin between the learned LPIPS loss and the least-squares loss is the largest with about 3.5 dB or 0.02 in SSIM. Note that the LPIPS loss function offers the best performance in all three image quality metrics, illustrating the promise of our flexible framework for future research on formal optimization for phase retrieval. The qualitative examples on the top in Fig. 6 show perceptually strong differences especially in uniform and smooth regions, where localized errors average out in the least-squares estimate but are penalized by the learned perceptual metric.

6.1 Generalization to Alternative Optical Designs

To demonstrate the ability of the proposed method to generalize to other optical setups, we also simulate a stacked setup of two cascaded SLMs, illustrated in Figure 7. For conventional transmissive displays, stacking two displays with a lateral offset of half a pixel has allowed researchers to synthesize images with greater spatial resolution than that afforded by any single SLM [Heide et al. 2014]. We apply this idea to holographic display generation. Instead of calculating a spatially superresolved hologram and decomposing it into two stacked layers using complex matrix factorization [Peng et al. 2017], our framework facilitates end-to-end optimization of cascaded holograms.

A cascade of two phase-only modulators combine additively in phase, or multiplicatively in their complex wave fields. Optically cascading a first SLM (A) with phase ϕ_A of M pixels, and a second SLM (B) with phase ϕ_B of M pixels, results in a complex wave field that can be expressed as:

$$H = \mathbf{exp}^{j\Gamma_{AB}\Phi_A} \circ \mathbf{exp}^{j\Gamma_{BA}\Phi_B} = \mathbf{exp}^{j(\Gamma_{AB}\Phi_A + \Gamma_{BA}\Phi_B)}, \qquad (20)$$

where Γ_{AB} and Γ_{BA} are sparse transformation matrices of size $4M \times M$ that model the relative pixel shifts of the two phase-SLMs with respect to a 2× higher-resolution virtual SLM with $4M$ pixels. Specifically, Γ_{AB} maps one pixel from SLM A to four pixels from the virtual SLM, and similarly Γ_{BA} represents the mapping for SLM B. For a half-pixel shift between both SLMs, these two matrices are binary matrices. The above equation can be easily plugged into our forward model, and the individual SLM phase patterns can be computed as

$$\Phi_{opt,A}, \Phi_{opt,B} = \underset{\Phi_A,\Phi_B}{\text{minimize}} \quad f(|\mathcal{P}(\mathbf{exp}^{j(\Gamma_{AB}\Phi_A + \Gamma_{BA}\Phi_B)})|^2, I)$$

$$(21)$$

We simulate two reflective mode phase-only SLMs with the specifications of the HOLOEYE LETO used in our hardware prototype, but with half the resolution (i.e. 960×540) in both dimensions to generate a hologram of four times the size of any SLM (i.e. 1920 × 1080). Figure 7 shows simulated reconstruction of images produced by a dual-SLM cascaded holographic display. Notice that the cascaded hologram spatially superresolves the image by a factor of four, i.e. twice as many pixels along each axis. The cascaded hologram offers fine detail with high contrast. While this approach effectively promises next generation results with currently available SLM hardware, we note that, in practice, alignment and calibration of the
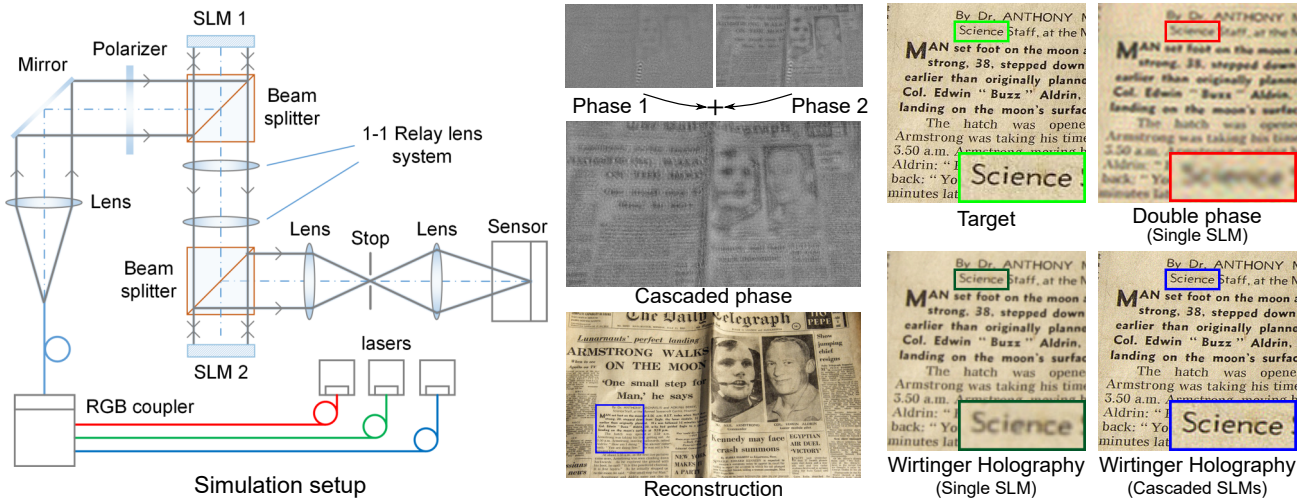
Fig. 7. Schematic of setup for realizing a cascaded SLM holographic display (left), where two SLMs are relayed by optics to perform the phase modulation defined by complex matrix multiplication. The two phase holograms superresolves a high resolution phase hologram that corresponds to reconstruct a high quality image (center). Synthetic comparison results produced by different solutions are shown (right). Image credits Sandid/Pixabay.

overlap might be challenging without micron-accurate mechanical alignment methods, and hence require significant engineering effort to enable efficient manufacturing processes.

6.2 3D Holograms

Although the focus of the proposed method is on improving the quality of 2D holographic images, our optimization framework can be extended to 3D volumetric scenes. One can slice the 3D scene into multiple depth planes and superpose all complex holograms corresponding to each depth plane, thereby forming a true 3D hologram. However, this method is not optimal and causes slow down of a factor proportional to the number of depth planes, and also will not support smooth depth changes. One option is to generate 2D holograms approximating spatially variant focus by fast depth switching. With gaze tracking, we can measure the depth at which the user is looking and compute the complex 2D hologram by changing the focus of the scene to that depth, providing precise focus in the fovea region [Maimone et al. 2017]. In other words, we can render a dense focal-stack in lieu of computing the full 3D hologram, thereby addressing the vergence-accommodation conflict, see Figure 2 of supplemental material. Since the eye can focus only at a single depth any given time, presenting depth of field blur for points close to each other within the fovea region can be addressed in the image space [Cholewiak et al. 2017]. Please refer to Section 3 of the supplemental material for additional details.

7 ASSESSMENT

Figure 8 and Figure 9 show experimentally acquired results from the prototype system presented in Section 5. We validate the proposed method by comparing different holographic phase retrieval methods in full color and for individual single-wavelength channels.

The results in Figure 8 demonstrate that the proposed approach suppresses most of the severe artefacts present in existing methods.

Specifically, reconstruction noise, as in the flat areas for the GS holograms, and ringing, as for the periphery in the TV line chart, limit the achieved resolution of previous methods. GS phase retrieval preserves high light efficiency, while its iterative projection of a random phase update on the image plane end inevitably causes severe reconstruction noise. The DPH method leads to the unwanted effect of a noticeable portion of light escaping the designated area. Moreover, it creates multiple higher-order twin images and suffers from severe ringing around high-contrast edges and in the periphery, which are particularly visible in the binary image in the second row of Figure 8. The proposed Wirtinger holography method recovers fine detail while suppressing severe ringing and reconstruction noise, resulting in improved resolution validated in the TV line chart areas in the top row of Figure 8

The full color results in Figure 9 confirm the trend of the single channel reconstructions. The proposed approach reduces ringing and reconstruction noise present in existing methods. As a result, fine details, as for the tiles on the roof of the houses or the microtubules in Figure 9, are displayed more faithfully with reduced chromatic artefacts. While this validates the proposed method, the margin compared to previous methods is not on par with the simulation results. This is due to residual artefacts resulting from a number of limitations of our specific low-cost prototype, which we note are not fundamental to the proposed method. As our prototype SLM is unsupported due to its age, we were only able to rely on the available look-up-tables, which did not match the wavelengths of our laser configuration. This results in inaccurate phase delays which severely affects methods relying on accurate forward models, such as the GS and proposed method. Moreover, the absence of a zero-order stop affects all methods uniformly in the proposed setup. Refer to the supplement for additional analysis. We note that even with these prototype limitations the proposed method outperforms existing holographic phase retrieval methods.
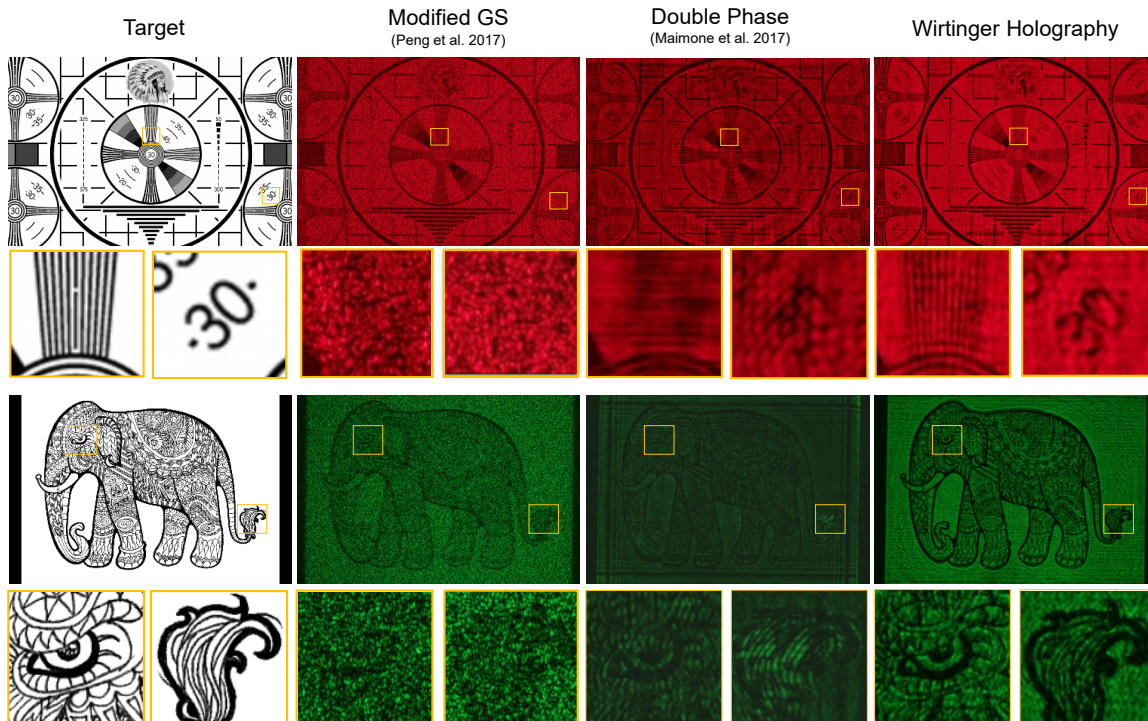
Fig. 8. Experimental results for existing holographic phase retrieval methods on our prototype display. We show here single-color reconstructions for the green and red wavelength channels of our system. The phase-only holograms used in this prototype are provided in the supplemental material. The images in each row are captured with same camera settings, using ISO 100 and same exposure time 10 ms. Please zoom into the electronic version of this document for better viewing.

8   DISCUSSION AND CONCLUSION

We have presented a phase retrieval method for holographic near-eye displays, which departs from existing heuristic algorithms, and instead relies on formal optimization using complex Wirtinger derivatives. This framework has allowed us to formulate holographic phase retrieval as a quadratic optimization problem that can be solved using first-order optimization methods. We have validated that the resulting holograms have an order of magnitude lower error, i.e. more than 10 dB PSNR improvement, when compared to recent state-of-the-art approaches. We have assessed the proposed method in simulation and using an experimental prototype system. Wirtinger holography eliminates the severe artefacts present in existing methods, and, in particular, enables us to minimize ringing and chromatic artefacts.

Although the proposed method provides unprecedented accuracy, its runtime is only comparable to that of the modified Gerchberg-Saxton method [Peng et al. 2017], with about 30 sec for a full HD hologram on a consumer laptop computer. However, note that we implemented our framework using high-level languages, without manual low-level code optimization, to allow for rapid adoption by researchers. Moreover, we do not use dedicated hardware for our holographic display algorithm. In the future, we envision efficient implementations or dedicated display hardware to overcome these limitations and achieve real-time frame rates, similar to how recent dedicated hardware made real-time ray tracing possible.

In addition to reducing error by an order of magnitude when compared to existing methods, we have validated the flexibility of the proposed system. Specifically, we have shown that Wirtinger holography facilitates the use of almost arbitrary penalty functions in a plug-and-play fashion, even including learned perceptual losses parametrized by deep neural networks, the use of different first-order optimization methods, including stochastic optimization methods, and that the proposed framework can be applied to different optical setup configurations. As such, Wirtinger holography paves the road towards high-quality artefact-free near-eye holographic displays of the future, and enabling future research towards this vision.
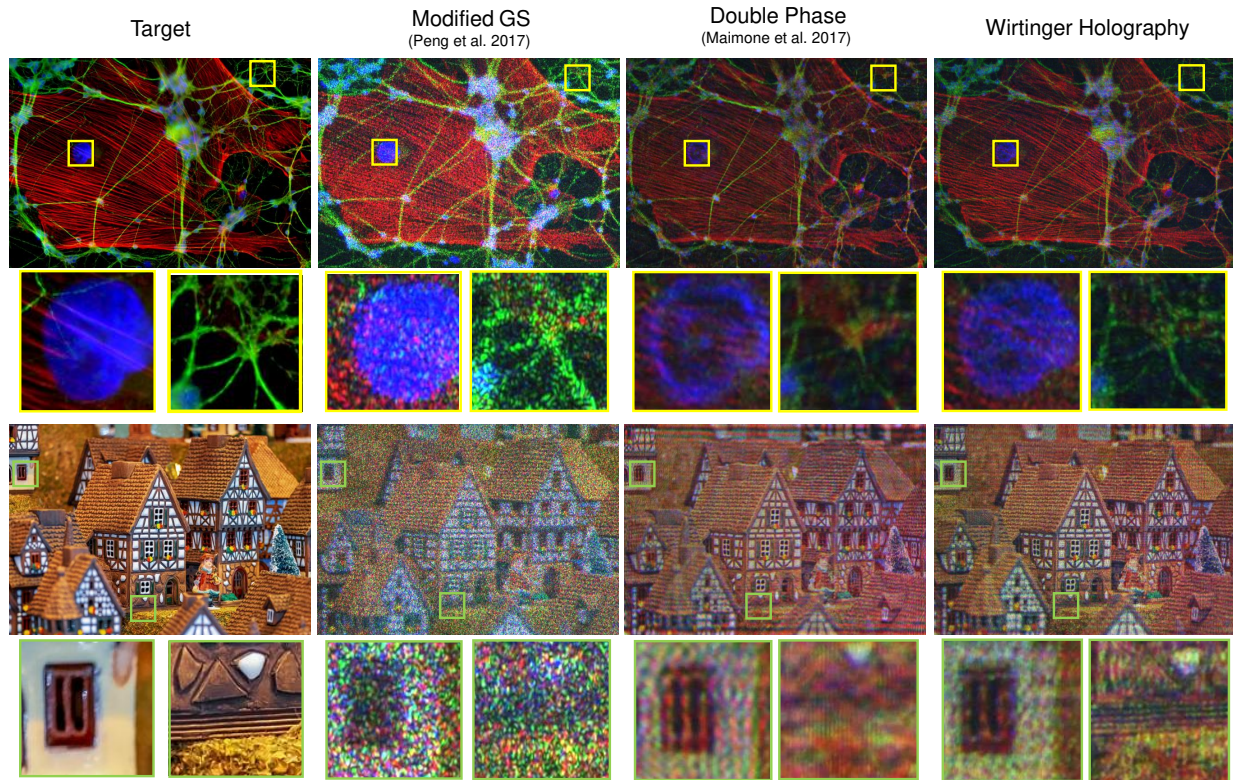
Fig. 9. Experimental results for existing holographic phase retrieval methods on our prototype display. We present RGB color images with each color channel captured sequentially. The phase-only holograms used in this prototype are provided in the supplemental material. The images in each row are captured with same camera settings, using ISO 100 and exposure time 10 ms. We have tuned the output power of three lasers before acquisition to approximately white-balance the illumination. Please zoom into the electronic version of this document for better viewing. Actin cell image by Jan Schmoranzer.

REFERENCES

Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. 2016. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16).* 265–283.

Lukas Ahrenberg, Philip Benzie, Marcus Magnor, and John Watson. 2008. Computer generated holograms from three dimensional meshes using an analytic light transport model. *Applied optics* 47, 10 (2008), 1567–1574.

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.

R Bates. 1982. Fourier phase problems are uniquely solvable in mute than one dimension. I: Underlying theory. *Optik (Stuttgart)* 61 (1982), 247–262.

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.

Muharrem Bayraktar and Meriç Özcan. 2010. Method to calculate the far field of three-dimensional objects for computer-generated holography. *Applied optics* 49, 24 (2010), 4647–4654.

Stephen A Benton and V Michael Bove Jr. 2008. *Holographic imaging.* John Wiley & Sons.

Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. 2012. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. (2012).

Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. 2015. Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory* 61, 4 (2015), 1985–2007.

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.

J-S Chen and DP Chu. 2015. Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications. *Optics express* 23, 14 (2015), 18143–18155.

Rick H-Y Chen and Timothy D Wilkinson. 2009. Computer generated hologram from point cloud using graphics processor. *Applied optics* 48, 36 (2009), 6841–6850.

Steven A Cholewiak, Gordon D Love, Pratul P Srinivasan, Ren Ng, and Martin S Banks. 2017. ChromaBlur: Rendering chromatic eye aberration improves accommodation and realism. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 210.

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

James R Fienup. 1993. Phase-retrieval algorithms for a complicated optical system. *Applied optics* 32, 10 (1993), 1737–1746.

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).

RA Gonsalves. 1976. Phase retrieval from modulus data. *JOSA* 66, 9 (1976), 961–964.

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

Felix Heide, Douglas Lanman, Dikpal Reddy, Jan Kautz, Kari Pulli, and David Luebke. 2014. Cascaded displays: spatiotemporal superresolution using offset pixel layers. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 60.

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

Fu-Chung Huang, Kevin Chen, and Gordon Wetzstein. 2015a. 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), 60.

Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. 2015b. Single Image Super-Resolution From Transformed Self-Exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 5197–5206.

Jia Jia, Juan Liu, Guofan Jin, and Yongtian Wang. 2014. Fast and effective occlusion culling for 3D holographic displays by inverse orthographic projection with low angular sampling. *Applied optics* 53, 27 (2014), 6287–6293.

Hwi Kim, Joonku Hahn, and Byoungho Lee. 2008. Mathematical modeling of triangle-mesh-modeled three-dimensional surface objects for digital holography. *Applied optics* 47, 19 (2008), D117–D127.

RG Lane. 1991. Phase retrieval using conjugate gradient minimization. *journal of Modern Optics* 38, 9 (1991), 1797–1813.

Douglas Lanman and David Luebke. 2013. Near-eye light field displays. *ACM Transactions on Graphics (TOG)* 32, 6 (2013), 220.

Detlef Leseberg and Christian Frère. 1988. Computer-generated holograms of 3-D objects composed of tilted planar segments. *Applied optics* 27, 14 (1988), 3020–3024.

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.

Dong C Liu and Jorge Nocedal. 1989. On the limited memory BFGS method for large scale optimization. *Mathematical programming* 45, 1-3 (1989), 503–528.

Mark Lucente and Tinsley A Galyean. 1995. Rendering interactive holographic images. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques.* ACM, 387–394.

Mark E Lucente. 1993. Interactive computation of holograms using a look-up table. *Journal of Electronic Imaging* 2, 1 (1993), 28–34.

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.

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.

D. Martin, C. Fowlkes, D. Tal, and J. Malik. 2001. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In *Proc. 8th Int'l Conf. Computer Vision*, Vol. 2. 416–423.

Nobuyuki Masuda, Tomoyoshi Ito, Takashi Tanaka, Atsushi Shiraki, and Takashige Sugie. 2006. Computer generated holography using a graphics processing unit. *Optics Express* 14, 2 (2006), 603–608.

Kyoji Matsushima. 2005. Computer-generated holograms for three-dimensional surface objects with shade and texture. *Applied optics* 44, 22 (2005), 4607–4614.

Kyoji Matsushima and Sumio Nakahara. 2009. Extremely high-definition full-parallax computer-generated hologram created by the polygon-based method. *Applied optics* 48, 34 (2009), H54–H63.

Kyoji Matsushima, Masaki Nakamura, and Sumio Nakahara. 2014. Silhouette method for hidden surface removal in computer holography and its acceleration using the switch-back technique. *Optics express* 22, 20 (2014), 24450–24465.

Yifan Peng, Xiong Dun, Qilin Sun, and Wolfgang Heidrich. 2017. Mix-and-match holography. *ACM Transactions on Graphics* (2017).

Christoph Petz and Marcus Magnor. 2003. Fast hologram synthesis for 3D geometry models using graphics hardware. In *Proc. SPIE*, Vol. 5005. 266–275.

Reinhold Remmert. 2012. *Theory of complex functions.* Vol. 122. Springer Science & Business Media.

Mark Schmidt. 2005. minFunc: unconstrained differentiable multivariate optimization in Matlab. *Software available at http://www. cs. ubc. ca/˜ schmidtm/Software/minFunc. htm* (2005).

Liang Shi, Fu-Chung Huang, Ward Lopes, Wojciech Matusik, and David Luebke. 2017. Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 236.

Quinn YJ Smithwick, James Barabas, Daniel E Smalley, and V Michael Bove. 2010. Inter-active holographic stereograms with accommodation cues. In *Practical Holography XXIV: Materials and Applications*, Vol. 7619. International Society for Optics and Photonics, 761903.

Tullio Tommasi and Bruno Bianco. 1993. Computer-generated holograms of tilted planes by a spatial frequency approach. *JOSA A* 10, 2 (1993), 299–305.

John Stephen Underkoffler. 1991. *Toward accurate computation of optically reconstructed holograms.* Ph.D. Dissertation. Massachusetts Institute of Technology.

Zhou Wang, Eero P Simoncelli, and Alan C Bovik. 2003. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, Vol. 2. Ieee, 1398–1402.

James P Waters. 1966. Holographic image synthesis utilizing theoretical methods. *Applied physics letters* 9, 11 (1966), 405–407.

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.

Masahiro Yamaguchi, Hideshi Hoshino, Toshio Honda, and Nagaaki Ohyama. 1993. Phase-added stereogram: calculation of hologram using computer graphics technique. In *Proc. SPIE*, Vol. 1914. 25–31.

Hiroshi Yoshikawa, Takeshi Yamaguchi, and Hiroki Uetake. 2016. Image quality evaluation and control of computer-generated holograms. In *Practical Holography XXX: Materials and Applications*, Vol. 9771. International Society for Optics and Photonics, 97710N.

Roman Zeyde, Michael Elad, and Matan Protter. 2010. On single image scale-up using sparse-representations. In *International conference on curves and surfaces.* Springer, 711–730.

Hao Zhang, Yan Zhao, Liangcai Cao, and Guofan Jin. 2015. Fully computed holographic stereogram based algorithm for computer-generated holograms with accurate depth cues. *Optics express* 23, 4 (2015), 3901–3913.

Hao Zhang, Yan Zhao, Liangcai Cao, and Guofan Jin. 2016. Layered holographic stereogram based on inverse Fresnel diffraction. *Applied optics* 55, 3 (2016), A154–A159.

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.

Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. *arXiv preprint* (2018).

Yan Zhao, Liangcai Cao, Hao Zhang, Dezhao Kong, and Guofan Jin. 2015. Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method. *Optics express* 23, 20 (2015), 25440–25449.