Time-multiplexed Neural Holography: 
A Flexible Framework for Holographic Near-eye Displays with 
Fast Heavily-quantized Spatial Light Modulators

Suyeon Choi∗
suyeon@stanford.edu
Stanford University
USA
Manu Gopakumar∗
manugopa@stanford.edu
Stanford University
USA
Yifan Peng
evanpeng@stanford.edu
Stanford University
USA
Jonghyun Kim
jonghyunk@nvidia.com
NVIDIA and Stanford University
USA
Matthew O’Toole
mpotoole@cmu.edu
Carnegie Mellon University
USA
Gordon Wetzstein
gordon.wetzstein@stanford.edu
Stanford University
USA

RGBD Supervision
Ours, 8 frames
Focal Stack Supervision
Ours, 8 frames
Light Field Supervision
OLAS, Padmanaban et al.
Ours, 8 frames

Figure 1: Computer-generated holography (CGH) results captured with a display prototype that uses a fast, low-precision (i.e., 4 bit) phase spatial light modulator (SLM). When supervised with 2.5D RGBD images, our approach (2nd column) provides a better image quality than the state-of-the-art neural 3D holography algorithm [Choi et al. 2021a] (1st column) using this low-precision SLM. Our CGH framework is flexible in not only enabling 2.5D but also 3D focal stack and 4D light field supervision. The former approach (3rd column) results in the best in-focus (red boxes) and out-of-focus (white boxes) image quality among 2.5D and 3D CGH algorithms. Our 4D light field–supervised approach (5th column) outperforms the recently proposed OLAS method [Padmanaban et al. 2019] (4th column) by a large margin and utilizes the space–bandwidth product more effectively, as shown by the simulated light fields in the lower right images.

ABSTRACT
Holographic near-eye displays offer unprecedented capabilities for virtual and augmented reality systems, including perceptually important focus cues. Although artificial intelligence–driven algorithms for computer-generated holography (CGH) have recently made much progress in improving the image quality and synthesis efficiency of holograms, these algorithms are not directly applicable to emerging phase-only spatial light modulators (SLM) that are extremely fast but offer phase control with very limited precision. The speed of these SLMs offers time multiplexing capabilities, essentially enabling partially-coherent holographic display modes. Here we report advances in camera-calibrated wave propagation models for these types of holographic near-eye displays and we develop a CGH framework that robustly optimizes the heavily quantized phase patterns of fast SLMs. Our framework is flexible in supporting runtime supervision with different types of content, including 2D and 2.5D RGBD images, 3D focal stacks, and 4D light fields. Using our framework, we demonstrate state-of-the-art results for all of these scenarios in simulation and experiment.

CCS CONCEPTS
• Hardware → Emerging technologies; • Computing methodologies → Computer graphics.
KEYWORDS
computational displays, holography, virtual reality

1 INTRODUCTION
Holographic near-eye displays for virtual and augmented reality (VR/AR) applications offer many benefits to wearable computing systems over conventional microdisplays. These include high peak brightness, power efficiency, support of perceptually important focus cues and vision-correcting capabilities [Kim et al. 2021], as well as thin device form factors [Kim et al. 2022; Maimone and Wang 2020]. Yet, the image quality achieved by computer-generated holography (CGH) lags far behind that of conventional displays, requiring further advancements in the algorithms driving holographic displays.

Recently, artificial intelligence (AI) methods have enabled significant improvements in image quality [Chakravarthula et al. 2020; Choi et al. 2021a; Peng et al. 2020] and speed [Horisaki et al. 2018; Peng et al. 2020; Shi et al. 2021] of holographic displays. These algorithms, however, are primarily applicable to slow liquid crystal–based (LC) spatial light modulators (SLMs) that offer control of the phase of a coherent light source at high precision. Emerging micro-electromechanical (MEMS) phase SLMs [Bartlett et al. 2019] offer potential benefits over LC-based systems in being more light efficient, significantly faster, better suited to operate across a wide range of wavelengths, and more stable for varying temperatures. Indeed, MEMS-based amplitude SLMs are one of the most popular technology choices for many display applications, including projectors, so MEMS-based phase SLMs may also become increasingly important for holography applications. Unfortunately, the algorithms developed for high-precision LC-based phase SLMs suffer from a degradation in image quality and fail to fully utilize time-multiplexing when used with the high framerate, heavily quantized phase control that MEMS-based SLMs offer. For example, DLP’s phase SLM by Texas Instruments only offers up to 4 bits of precision or, similarly, 16 unevenly distributed discrete levels of phase control at frame rates of 1440 Hz [Bartlett et al. 2019; Ketchum and Blanche 2021].

The focus of our work is to extend AI-driven CGH algorithms to operate with emerging fast but heavily quantized phase SLMs. This is a non-trivial task, because quantization is non-differentiable, so the standard machine learning toolset does not directly apply in these settings. Moreover, most of the degrees of freedom of a holographic display stem from their ability to create constructive and destructive interference, which can only be achieved instantaneously in time but not between time-multiplexed frames. It is thus not clear whether the partially-coherent holographic display mode enabled by the fast SLM speed is actually beneficial when combined with a limited precision of phase control or how it affects image quality. We propose an algorithmic CGH framework that robustly optimizes holograms in these mathematically challenging scenarios and explore the aforementioned tradeoff, demonstrating significant benefits in image quality and space–bandwidth utilization [Yoo et al. 2021] of higher-speed phase SLMs. Moreover, we develop a learned propagation model that is more flexible than previously proposed alternatives in allowing us to calibrate it using 3D multiplane supervision but leverage a variety of target content, including 2D images, 2.5D RGBD images, 3D focal stacks, and 4D light fields, for supervision during runtime.

Specifically, our contributions include the following:

- a new variant of a camera-calibrated wave propagation model for holographic displays, which is flexible in enabling runtime supervision by 2D, 2.5D, 3D, or 4D content;
- a framework for robust CGH optimization with fast but heavily quantized phase-only SLMs;
- experimental demonstration of improved image quality and better utilization of the SLM’s space–bandwidth product enabled by our framework.

Source code for this paper is available at computationalimaging.org.

2 RELATED WORK
Many aspects of holographic displays, including optics, SLMs, and algorithms, have advanced considerably over the last few years. Detailed discussions of many of these advancements can be found in the survey papers by Yaras [2010], Park [2017], and Chang et al. [2020]. A recent roadmap article by Javidi et al. [2021] also outlines current and future research efforts of digital holography in non-display areas, including 3D imaging and microscopy.

Our work primarily focuses on advancing the algorithms driving holographic near-eye displays. In a nutshell, the CGH problem comprises several parts. First, the target content is specified in some format that needs to be converted to a complex-valued wavefield, such as point clouds [Fienup 1982; Gerchberg 1972; Maimone et al. 2017; Shi et al. 2017, 2021], polygons [Chen and Wilkinson 2009; Matsushima and Nakahara 2009], light rays [Wakunami et al. 2013; Zhang et al. 2011], image layers [Chen et al. 2021; Chen and Chu 2015; Zhang et al. 2017], or light fields [Benton 1983; Kang et al. 2008; Lucente and Galvean 1995; Padmanaban et al. 2019; Ziegler et al. 2007]. Second, this wavefield needs to be encoded by a phase-only SLM, which can be achieved by fast, direct phase coding approaches [Hsueh and Sawchuk 1978; Lee 1970; Maimone et al. 2017] or slow, iterative solvers, such as classic Gerchberg–Saxton-type algorithms [Fienup 1982; Gerchberg 1972] or variants of stochastic gradient descent [Chakravarthula et al. 2019; Peng et al. 2020].

Yet, the simulated wave propagation models used by most of these CGH algorithms do not always model the physical optics faithfully, thereby degrading image quality. Moreover, the computational complexity of these algorithms often prevents them from being practical in the power-constrained settings of a wearable computing system. Emerging artificial intelligence–driven CGH approaches have focused on addressing these limitations. For example, surrogate gradient methods that use a camera in the loop (CITL)
for hologram optimization can significantly improve image quality [Choi et al. 2021b; Peng et al. 2021, 2020]. Alternatively, differentiable wave propagation models can be learned to calibrate for the gap between simulated models and physical optics [Chakravarthula et al. 2020; Choi et al. 2021a; Kavakli et al. 2022; Peng et al. 2020]. Moreover, neural networks can be trained to enable real-time CGH algorithms [Horisaki et al. 2021, 2018; Peng et al. 2020; Shi et al. 2021].

Note that our work is concurrently and independently developed from the very recent work by Lee et al. [2022]. Although both works share some similarity in applying constrained gradient descent methods to optimize binary or heavily-quantized phase holograms, our framework outperforms the counterpart with the use of a learned propagation model for better image quality, the ability to effectively handle SLMs with varied bit depths and non-linear quantizations, and compatibility with a wide range of supervision sources.

### 3 A FLEXIBLE FRAMEWORK FOR CGH

In Fresnel holography, a collimated coherent light beam illuminates an SLM with a source field \( u_{src} \), and the light reflected in response reproduces a target intensity distribution. To generate this hologram, a phase-only SLM imparts a spatially-varying delay \( \phi \) on the phase of the field. After propagating a distance \( z \) from the SLM, the resulting complex-valued field \( u_z \) is given by the following image formation model:

\[
\begin{align*}
\begin{aligned}
\phi (x, y, \lambda) &= f (u_{SLM}(x, y, \lambda), z), \\
u_{SLM}(x, y, \lambda) &= e^{i\phi (x, y, \lambda)} u_{src}(x, y, \lambda),
\end{aligned}
\end{align*}
\]

where \( \lambda \) is the wavelength of light, \( x, y \) are the transverse coordinates, and \( u_{SLM} \) is the modulated field at the SLM. The wave propagation operator \( f \) models free-space propagation between two parallel planes separated by a distance \( z \). For notational convenience, we will omit the dependence on \( x, y, \lambda \) and the source field \( u_{src} \). The intensity pattern generated by this display at distance \( z \) in front of the SLM when showing phase \( \phi \) is therefore \( |f(e^{i\phi}, z)|^2 \).

When using low-bit SLMs for time-multiplexed holography, the effect of quantization is not negligible. To model a quantized phase-only SLM with \( M \times N \) pixels, where every pixel offers phase control with limited precision, we define a quantization operator \( q \):

\[
q : \mathbb{R}^{M \times N} \rightarrow Q^{M \times N}, \quad \phi \mapsto q(\phi) = \Pi Q(\phi),
\]

where \( \Pi \) is the projection operator that maps the continuous phase value to the closest discrete phase in the feasible set \( Q \) supported by the SLM.

Our framework approaches computer-generated holography with a differentiable camera-calibrated image formation model (Sec. 3.1), an optimization procedure designed for quantized SLMs (Sec. 3.2), and a family of loss functions supervised on either 2D, 2.5D, 3D, or 4D content to produce time-multiplexed holograms (Sec. 3.3). Figure 2 illustrates our model and optimization pipeline.

#### 3.1 Camera-calibrated Wave Propagation Model

Recent work on holographic displays has demonstrated that the naive application of simulated wave propagation models, like the angular spectrum method (ASM) [Goodman 2014], to holographic displays fails to account for the non-idealities of the physical optical system, such as phase distortions of the SLM, optical aberrations, and the limited diffraction efficiency of the SLM [Chakravarthula et al. 2020; Choi et al. 2021a; Peng et al. 2020]. This discrepancy between simulated and physical image formation adversely affects image quality, but can be overcome by learning to calibrate for the physical optics using a differentiable, neural network–parameterized propagation model.

Here, we propose a variant of the learned model recently proposed by Choi et al. [2021a]:

\[
\begin{align*}
\begin{aligned}
f_{\text{mod}}(u_{\text{SLM}}, z) &= \text{CNN}_{\text{input}} \left( \mathcal{P}_{\text{ASM}} \left( \text{CNN}_{\text{SLM}} (u_{\text{src}} e^{i\phi_{\text{src}} u_{\text{SLM}}}, z) \right) \right), \\
\mathcal{P}_{\text{ASM}} (u, z) &= \int \mathcal{F} (u) \cdot \mathcal{H} (f_x, f_y, \lambda, z) e^{i2\pi f_x x + f_y y} d f_x d f_y, \\
\mathcal{H} (f_x, f_y, \lambda, z) &= a_f e^{i \frac{2\pi}{\lambda} z \sqrt{1 - (f_x^2 + f_y^2 + \lambda^2 z)}} e^{i\phi_f},
\end{aligned}
\end{align*}
\]
Table 1: Comparison of different calibrated wave propagation models. All models are trained on 6 of the 7 planes. PSNR is evaluated for training and test sets as well as for the 7th held-out plane. The number of parameters of each model is also reported. Training details are listed in Supplement S2.4.

| Models                          | Params | Train | Test | Held-out |
|---------------------------------|--------|-------|------|----------|
| NH [Peng et al. 2020]           | 4.1M   | 26.7  | 27.1 | 26.3     |
| NH3D [Choi et al. 2021a]        | 68.5M  | 34.4  | 32.4 | 31.9     |
| Our model, CNNs only            | 6.2M   | 31.6  | 29.7 | 30.0     |
| + asrc                          | 7.2M   | 35.3  | 35.4 | 32.3     |
| + asrc + φsrc                   | 8.2M   | 36.2  | 36.3 | 33.0     |
| + asrc + φsrc + φy              | 12.3M  | 36.4  | 36.4 | 32.8     |
| + asrc + φsrc + φy + lut        | 12.3M  | 36.4  | 36.4 | 32.8     |
| + asrc + φsrc + φy + lut        | 16.4M  | 36.7  | 36.7 | 32.6     |

where CNN_{\text{src}} and CNN_{\text{target}} are convolutional neural networks that operate on the complex field at the SLM and target planes. The target plane is a distance \( z \) from the SLM. In addition, \( a_{\text{src}} \) and \( \phi_{\text{src}} \) are learned to account for content-independent spatial variations in amplitude and phase of the incident source field at the SLM plane while \( a_{\text{y}} \) and \( \phi_{\text{y}} \) are added to the ASM propagation to learn spatial variations in amplitude and phase in the Fourier plane similarly to the learned complex convolutional kernel presented by Kavaki et al. [2022].

Similar to Choi et al., we capture a training and a test set comprised of a large number of SLM phase patterns and corresponding amplitude images recorded at a set of distances \( \{ j \} \), \( j = 1 \ldots J \) with our prototype holographic display. Using a standard stochastic gradient descent–type solver, we then fit the parameters of the CNNs, \( \text{CNN}_{\text{src}} \), and \( \text{CNN}_{\text{target}} \), as well as \( a_{\text{src}}, \phi_{\text{src}}, \phi_{\text{y}}, \) and \( a_{\text{y}} \) to learn the calibrated wave propagation model. The model used in this framework builds upon the model from Choi et al. by using the terms \( \phi_{\text{src}}, \phi_{\text{y}}, \) and \( a_{\text{y}} \) to learn many of the content-independent non-idealities of the holographic system. The source terms can efficiently model the effects of non-ideal illumination at the SLM plane, and the Fourier plane terms can compactly account for the effects of non-ideal optical filtering. Together these terms enable the use of smaller convolutional neural networks to learn the content-dependent non-idealities, such as the spatially varying pixel response at the SLM.

Table 1 quantitatively assesses the effect of these physically-inspired parameters by evaluating the performance of different calibrated wave propagation models on a captured dataset. All models are trained over 6 intensity planes, corresponding to 0.0 D, 0.5 D, 1.0 D, 1.5 D, 2.5 D, and 3.0 D in the physical space. A 7th plane at 2.0 D is set as the held-out plane for evaluation. In this table, we also ablate the performance of an additional \( \text{lut} \) parameter to optionally learn the feasible set \( Q \) of quantized values supported by the SLM. We observe that our model (bottom row) significantly reduces the number of parameters when compared to the original NH3D model, while still producing the highest PSNR metrics on the test set and the held-out plane. Notably, the lagging performance of the NH model, which is purely composed of physically-inspired terms, illustrates the substantial benefit of incorporating the flexibility of CNNs in a calibrated propagation model. Further details on our model architecture and training are included in Supplement S2.4.

3.2 Optimizing Phase Patterns for Quantized SLMs

Emerging MEMS-based phase SLMs are fast but offer only a limited precision for controlling phase. DLP’s phase SLM by Texas Instruments (TI) [Bartlett et al. 2019], for example, runs at a maximum framerate of 1440 Hz grayscale but only offers 4 bits, or 16 discrete phase levels, at each of the frames. We therefore need to derive methods that allow us to optimize phase patterns for heavily quantized phase SLMs. The primary problem is that the quantization function \( q \) is not differentiable. To this end, we discuss and evaluate several strategies for dealing with \( q \) assuming some simple 2D loss function \( \mathcal{L} \left[ s \cdot \left| f_{\text{real}} \left( \text{iq}(\phi, 0) \right) \right| - a_{\text{target}} \right] \), where \( a_{\text{target}} \) is the desired 2D amplitude, and \( s \) is a scale parameter that is optimized along with \( \phi \).

The naive solution to dealing with \( q \) is to simply ignore it. Specifically, the phase pattern \( \phi \) can be optimized given a 2D target amplitude image \( a_{\text{target}} \) and quantized to the available precision after the optimization. This is the approach typically adopted by state-of-the-art CGH algorithms that work well for liquid-crystal–type phase SLMs, because these SLMs offer 8 bit or higher precision phase modulation. TI’s MEMS device enables time multiplexing but only offers 4 bits, which makes this approach impractical (see Fig. 3). Instead, the reference code supplied with the SLM implements a variant of projected gradient descent [Boyd et al. 2004], which projects the iteratively updated solution onto the feasible set of quantized values \( Q \). This approach is equivalent to a gradient descent–type update scheme that applies \( q \) after each iteration \( k \) as:

\[
\phi(k) \leftarrow \phi(k-1) - a \left( \frac{\partial \mathcal{L}}{\partial \phi} \right) \left( \mathcal{L} \left[ s \cdot \left| f_{\text{real}} \left( \text{iq}(\phi(k)) \right) \right| - a_{\text{target}} \right] \right)
\]

(4)

As an alternative solution to solving these types of problems, surrogate gradient methods are often used [Bengio et al. 2013; Zenke and Ganguli 2018]. Here, the forward pass is computed using the correct quantization function \( q \) but during the error backpropagation pass, the gradients of a differentiable proxy function \( \hat{q} \) are used. This enables improved optimization of phase patterns through a quantization layer with the minimal overhead of computing the proxy gradients:

\[
\phi(k) \leftarrow \phi(k-1) - a \left( \frac{\partial \mathcal{L}}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial \phi} \right) \left( \mathcal{L} \left[ s \cdot \left| f_{\text{real}} \left( \text{i\hat{q}}(\phi(k-1)) \right) \right| - a_{\text{target}} \right] \right)
\]

(5)

Perhaps the most common choice for \( \hat{q} \) is a sigmoid function, whose slope can be gradually annealed during training [Bengio et al. 2013; Chung et al. 2016; Zenke and Ganguli 2018].

We propose the use of a continuous relaxation of categorical variables using Gumbel-Softmax [Jang et al. 2016; Maddison et al. 2016] for optimizing heavily quantized phase values in CGH applications. This approach has several desirable properties. First, the Gumbel noise and categorical relaxation prevent the optimization from getting stuck in local minima, which is perhaps the primary benefit over other surrogate gradient methods. Second, annealing of the temperature parameter \( \tau \) of the softmax as well as the shape of the score function are directly supported. Formally, this approach
is written as:

\[ \tilde{q}(\phi) = \sum_{l=1}^{L} Q_l \cdot G_l(\text{score}(\phi, Q)), \]

\[ G_l(z) = \frac{\exp((z_l + g_l)/\tau)}{\sum_{l=1}^{L} \exp((z_l + g_l)/\tau)}, \]

\[ \text{score}_l(\phi, Q) = \sigma(w \cdot \delta(\phi, Q_l))\left(1 - \sigma(w \cdot \delta(\phi, Q_l))\right), \]

where \( g_l \sim \text{Gumbel}(0, 1) \) is the Gumbel noise for all of the \( l = 1, \ldots, L \) categories, i.e., quantized phase levels, \( \sigma \) is a sigmoid function, \( \delta \) is the signed angular difference, and \( w \) is a scale factor (see Jang et al. [2016] and the supplement for additional details).

### 3.3 Runtime Supervision of Time-multiplexed Holograms

Fast MEMS-based phase SLMs can produce higher-quality holograms through time multiplexing, i.e., intensity averaging of multiple frames. Given our camera-calibrated wave propagation model (Sec. 3.1), we optimize for time-multiplexed holograms using different target content at runtime.

#### 2D Holography

In this case, we wish to synthesize a 2D intensity image at a distance \( z \) in front of the phase SLM. The distance can be fixed or dynamically varied in software to enable a varifocal holographic display mode. For this purpose, we specify the loss:

\[ L_{2D} = \mathcal{L}\left(\sum_{t=1}^{T} \frac{1}{T} \sum_{f=1}^{T} \left| f_{\text{model}}\left(e^{iq(\phi^{(t)})}, z\right) - a_{\text{target}} \right|^2 \right), \]

between the target amplitude image \( a_{\text{target}} \) and the simulated holographic image and solve for \( \phi \). We can easily formulate a time-multiplexed variant of the CGH problem using this loss function by summing over \( t = 1 \ldots T \) squared amplitudes, i.e., intensities, where \( T \) refers to the total number of time-multiplexed frames that can be displayed throughout the exposure time of the human eye. The simplest example of the loss function \( L \) is an \( l_2 \) loss although other loss functions, such as perceptually motivated image quality metrics, could be applied as well.

#### 2.5D Holography using RGBD Input

Using the multiplane loss function presented by Choi et al. [2021a], holograms can be synthesized to generate a 2D set of intensities at depths specified by a depth map. We refer the interested reader to Supplement S2.5 for the loss function and an additional discussion on utilizing time multiplexing to produce natural blur with 2.5D supervision.

#### 3D Multiplane Holography

True 3D holography can be achieved by optimizing a single SLR phase pattern \( \phi \) or a series of time-multiplexed patterns \( \phi^{(t)} \) for the target amplitude of a focal stack \( a_{\text{target}} \). The corresponding loss function in our framework looks very similar to that of the 2D hologram above, although it is evaluated over the set of focal slices \( \{j\} \):

\[ L_{3D} = \mathcal{L}\left(\sum_{t=1}^{T} \frac{1}{T} \sum_{f=1}^{T} \left| f_{\text{model}}\left(e^{iq(\phi^{(t)})}, z(j)\right) - a_{\text{target}} \right|^2 \right), \]

Effectively optimizing this focal stack loss using the full blur available within the diffraction angle of the SLM requires time multiplexing as illustrated in Supplement S2.6.

#### 4D Light Field Holography

Finally, we can also supervise our CGH framework using the amplitudes of a 4D target light field \( \mathbf{I}_{\text{target}} \). For this purpose, a differentiable hologram-to-light field transform is required, which can be calculated using the Short-time Fourier transform (STFT) [Padmanaban et al. 2019; Zhang and Levoy 2009]:

\[ L_{4D} = \mathcal{L}\left(\sum_{t=1}^{T} \frac{1}{T} \sum_{f=1}^{T} \left| \text{STFT}\left(f_{\text{model}}\left(e^{iq(\phi^{(t)})}, z\right)\right) - \mathbf{I}_{\text{target}} \right|^2 \right), \]

By utilizing time multiplexing, our optimized holograms can uniquely reproduce a set of light field views that fully covers the SLM’s space–bandwidth product as detailed in Supplement S2.7.
4 EXPERIMENTS

To evaluate our novel algorithms, we use a benchtop 3D holographic display prototype. This prototype includes a FISBA RGBBeam fiber-coupled module with red, green, and blue optically aligned laser diodes for illumination and a TI DLP6750Q1EVM phase SLM for high-speed quantized phase modulation. We capture the images produced by this prototype with a FLIR Grasshopper3 12.3 MP color USB3 sensor through a Canon EF 35mm lens with focus controlled by an Arduino microcontroller. Further details of the prototype are included in Supplement S1.

Comparing CGH Algorithms. We compare several CGH approaches for the task of optimizing phase patterns for a fast phase SLM with 4 bits, or 16 phase levels, in Fig. 3. The naive approach, which quantizes the phase after optimization, performs poorly, as measured by the peak signal-to-noise ratio (PSNR). The surrogate gradient (SG) method used with the gradients of sigmoid and those of the Gumbel-Softmax are significantly better than other methods, with Gumbel-Softmax outperforming all other methods by a large margin, especially for higher-speed SLMs. This experiment represents the TI SLM with 4 bits and up to 480 Hz color, i.e., 8 multiplexed frames each running at 60 Hz so a total of 480 Hz. We evaluate other bit depths in the supplement and show similar trends. Finally, Gumbel-Softmax can be used as part of an SG method (Eq. 5) using only its gradients \( \partial \omega / \partial \phi \) or it can be used to replace \( q \) by \( \tilde{q} \) also in the forward image formation. We found the former performs better in most settings, and therefore only report these results in the paper; see the supplement for evaluations of the latter approach.

Learning Physical Filters. We visualize in Figure 4 the performance of our learned model in accurately approximating the optical filter, which is an iris in the physical display system. As expected, values outside the filters are all zeros. The shape of blade edges is robustly learned with our model and scales with wavelength as expected. The variance of diameter size also aligns with the variance of wavelength. Refer to Figure S7 in the supplement for visualization of the full model.

Assessing 2D Holography. We present in Figure 5 experimental results of 2D holographic display assessing different CGH algorithms and different multiplexing schemes. In this experiment, we compare SGD algorithms using the ASM with Naive quantization, our model with Naive quantization, and our model with Gumbel-Softmax (GS). We observe two insights. First, the use of our calibrated wave propagation model corrects for most artifacts present in the physical display. Second, applying the GS operation leads to better performance in such heavily-quantized optimization problems. Refer also to Figures S8–9, as well as Tables S1 and S2 in the supplementary document for both quantitative and qualitative assessments of other examples.

Assessing 3D Holography. We present in Figure 6 experimental results of 3D holographic display assessing different CGH algorithms. In this experiment, we compare SGD algorithms with the prior state-of-the-art NH3D model and Naive quantization using RGBD input [Choi et al. 2021a] with 1 frame and 8 multiplexed frames, respectively, our model with Gumbel-Softmax (GS), and our model with GS using focal stack supervision. PSNR metrics are provided in the caption. Using only a single frame results in speckly in-focus content (shown with red squares in Figure 6). Even with multiple frames, RGBD supervision produces speckle in the unconstrained out-of-focus regions. However, with our focal stack supervision and time multiplexing, we observe natural out-of-focus blur, while still preserving sharpness for the in-focus content. For example, the branch at the intermediate depth is sharp, and the sky in the background is smooth. In the supplement, we show extensive evaluations and ablations of 3D multiplane CGH methods for more 3D scenes (Figures S3–4 and S10–16).

Assessing 4D Light Field Holography. We present in Figure 7 experimental results of 4D light field–supervised holographic display, assessing different CGH algorithms. In this experiment, we compare the OLAS [Padmanaban et al. 2019] algorithm, our approach using light field–supervision with the ASM and naive quantization...
Figure 6: Comparison of 3D CGH algorithms using experimentally captured data. Here, we compare SGD algorithms with the prior state-of-the-art NH3D model and Naive quantization using RGBD input [Choi et al. 2021a] with 1 frame and 8 multiplexed frames, respectively, our model with Gumbel-Softmax (GS), and our model with GS using focal stack supervision. The corresponding PSNR metrics are 24.3 dB, 25.8 dB, and 26.7 dB with respect to the RGBD all-in-focus targets (left 3 columns), and 26.9 dB with respect to the focal stack (right column). For close-ups, red squares indicate where the camera is focused at three distances (from top to bottom: far, intermediate, and near).

Figure 7: Comparison of 4D light field–supervised CGH algorithms using experimentally captured data. Here, we compare the OLAS algorithm [Padmanaban et al. 2019] (1st column) without time multiplexing, and three variants of our approach: ASM-Naive without time multiplexing (2nd column) and with 8 multiplexed frames (3rd column) and Model-GS with 8 multiplexed frames (4th column). For close-ups, red squares indicate where the camera is focused at two distances (top: far, bottom: near). Since OLAS deterministically computes a single phase pattern for a target light field, there would be no variation between time-multiplexed frames.

(ASM-Naive), and our approach with the camera-calibrated wave propagation model and Gumbel-Softmax (Model-GS) to account for the low bit depth of the SLM. The OLAS algorithm requires light field and depth maps for each light field view as input and it does not support time multiplexing. Both variants of our method do not require depth maps and jointly optimize 8 time-multiplexed frames using SGD. For each example scene, we show close-ups of
content at two distances (far, near). We observe that our framework exhibits the best image quality for both in-focus (red squares) and out-of-focus regions (white squares). Refer also to Figures S5 and S17 in the supplementary document for additional simulation and experimental results.

5 DISCUSSION

In summary, we present a new framework for computer-generated holography. This framework includes a camera-calibrated wave propagation model that combines parts of the recently proposed model in a novel way to achieve a better performance with fewer model parameters. We explore surrogate gradient methods for optimizing the heavily quantized SLM patterns of emerging MEMS-based phase SLMs and show the Gumbel-Softmax algorithm to outperform other approaches. Our framework is flexible in supporting 2D, 2.5D, 3D, and 4D supervision at runtime and we show state-of-the-art results in all of these scenarios with our near-eye holographic display prototypes.

Limitations and Future Work. Image quality could be further improved by increasing the precision and framerate of the employed phase SLMs and, importantly, by improving their diffraction efficiency. In Figure S6 of our supplement, we explore the simulated image quality with varying levels of time multiplexing and bit depth, but analytically deriving this landscape remains an interesting direction for future work to explore. Our algorithms do not run in real time, but require on the order of tens of seconds to a few minutes to compute a hologram. Neural networks could be employed to speed up the computation, as recently demonstrated by Horisaki et al. [2018], Peng et al. [2020], and Shi et al. [2021]. Due to their limited space–bandwidth product, holographic near-eye displays only provide a limited eye box, which could be addressed by dynamically steering it using eye tracking [Jang et al. 2017]. The depth of field of 3D-synthesized holograms in AR scenarios should match that of the user’s eye, which requires tracking their pupil diameter. Finally, we demonstrated our results on barchen prototype displays, which will have to be miniaturized into the impressive device form factors presented by Maimone et al. [2017] and Wang and Maimone [2020].

Conclusion. The algorithmic advances presented in this work help make holographic near-eye displays a practical technology for next-generation VR/AR systems.

ACKNOWLEDGMENTS

We thank Cindy Nguyen for helpful discussions. This project was in part supported by a Kwanjeong Scholarship, a Stanford SGF, Intel, NSF (award 1839974), a PECASE by the ARO (W911NF-19-1-0120), and Sony.

REFERENCES

Terry A. Bartlett, William C. McDonald, and James N. Hall. 2019. Adapting Texas Instruments DLP technology to demonstrate a phase spatial light modulator. In SPIE OPTO. Proceedings Volume 10932, Emerging Digital Micromirror Device Based Systems and Applications XI. 109520S.

Yoshua Bengio, Nicholas Léonard, and Aaron Courville. 2013. Estimating or propagating gradients through stochastic neurons for conditional computation. arXiv preprint arXiv:1308.3432 (2013).

Stephen A. Benton. 1983. Survey Of Holographic Stereograms. In Proc. SPIE, Vol. 0367.

Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. 2004. Convex optimization. Cambridge university press.

Praneeth Chakravartula, Yifan Peng, Joel Kollin, Henry Fuchs, and Felix Heide. 2019. Writinger Holography for Near-Eye Displays. ACM Trans. Graph. 38, 6 (2019).

Praneeth Chakravartula, Ehsan Tseng, Tarun Srivastava, Henry Fuchs, and Felix Heide. 2020. Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. ACM Trans. on Graph. (TOG) 39, 6 (2020), 1–18.

Chenliang Chang, Kiseung Bang, Gordon Wetzstein, Byoungho Lee, and Liang Gao. 2020. Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective. Optica 7, 11 (2020), 1563–1578.

Chun Chen, Byoungho Lee, Nan-Nan Li, Minseo Chae, Di Wang, Qiong-Hua Wang, and Byoungho Lee. 2021. Multi-depth hologram generation using stochastic gradient descent algorithm with complex loss function. Opt. Express 29, 10 (2021), 15089–15100.

Jhen-Si Chen and Daping Chu. 2015. Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications. Opt. Express 23, 14 (2015), 18143–18155.

Rick H-Y Chen and Timothy D Wilkinson. 2009. Computer generated hologram with geometric occlusion using GPU-accelerated depth buffer rasterization for three-dimensional display. Applied optics 48, 21 (2009), 4246–4255.

Suyeon Choi, Manu Gopakumar, Yifan Peng, Jonghyun Kim, and Gordon Wetzstein. 2021a. Neural 3D Holography: Learning Accurate Wave Propagation Models for 3D Holographic Virtual and Augmented Reality Displays. ACM Trans. Graph. (SIGGRAPH Asia) (2021).

Suyeon Choi, Jonghyun Kim, Yifan Peng, and Gordon Wetzstein. 2021b. Optimizing image quality for holographic near-eye displays with michelson holography. Optica 8, 2 (2021), 143–146.

Junyoung Chung, Sungjin Ahn, and Yoshua Bengio. 2016. Hierarchical multiscale recurrent neural networks. arXiv preprint arXiv:1609.01704 (2016).

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 phase from image and diffraction plane pictures. Optics 35 (1972), 237–246.

Joseph W. Goodman. 2014. Holography Viewed from the Perspective of the Light Field Camera. In Fringe 2013, Wolfgang Osten (Ed.). Springer Berlin Heidelberg, 3–15.

Ryoichi Horisaki, Yohi Nishizuki, Katsujiya Kitagawa, Mamoru Sato, and Jun Tanida. 2021. Three-dimensional deeply generated holography. Appl. Opt. 60, 4 (2021), A323–A328.

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

Chang-Kai Hsu and Alexander A. Sawchuk. 1978. Computer-generated double-phase holograms. Applied optics 17, 24 (1978), 3874–3883.

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 Trans. Graph. (SIGGRAPH Asia) 36, 6 (2017).

Eric Jang, Shixiang Gu, and Ben Poole. 2016. Categorical reparameterization with gumbel-softmax. arXiv preprint arXiv:1611.01444 (2016).

Bab jamidi, Artur Carnicer, Aran Anan, George Barbastathis, Wen Chen, Pietro Ferraro, J. W. Goodman, Ryoichi Horisaki, Kedar Khare, Malgorzata Kujawińska, Rainer A. Leitgeb, Pierre Marquet, Takanori Nomura, Aydogan Ozcan, YongKeun Park, Giancarlo Pedrini, Pascal Picart, Joseph Rosen, Genaro Saavedra, Natal T. Shaked, Adrian Stern, Enrique Tajahuerce, Lei Tian, Gordon Wetzstein, and Masahiro Yamaguchi. 2021. Roadmap on digital holography. Opt. Express 29, 22 (2021).

Hoonjong Kang, Takeshi Yamaguchi, and Hiroshi Yoshikawa. 2008. Accurate phase-added stereogram to improve the coherent stereogram. Applied optics. 47, 19 (2008).

Koray Kavakli, Hakum Urey, and Kian Akshait. 2022. Learned holographic light transport. Appl. Opt. 61, 5 (2022), B50–B55.

Remington S Ketchum and Pierre-Alexandre Blanche. 2021. Diffraction efficiency characteristics for MEMS-based phase-only spatial light modulator with nonlinear phase distribution. In Photonics, Vol. 8. Multidisciplinary Digital Publishing Institute, 62.

Dongyoo Kim, Seung-Woo Nam, Kiseung Bang, Byoungho Lee, Seungjae Lee, Youngmo Jeong, Jong-Mo Seo, and Byoungho Lee. 2021. Vision-correcting holographic display: evaluation of aberration correcting hologram. Biomed. Opt. Express 12, 8 (2021), 5179–5195.

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.

Byoungho Lee, Dongyoo Kim, Seungjae Lee, Chun Chen, and Byoungho Lee. 2022. High-contrast, speckle-free, true 3D holography via binary CGH optimization. arXiv preprint arXiv:2201.04035 (2022).

Wai Hon Lee. 1970. Sampled Fourier transform hologram generated by computer. Applied Optics 9, 3 (1970), 639–643.

Mark Lucente and Tinsley A Galyean. 1995. Rendering interactive holographic images. In ACM SIGGRAPH 387–394.

Chris J Maddison, Andryi Mnih, and Yee Whye Teh. 2016. The concrete distribution: A continuous relaxation of discrete random variables. arXiv preprint arXiv:1611.00712 (2016).
Andrew Maimone, Andreas Georgiou, and Joel S Kollin. 2017. Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.* (SIGGRAPH) 36, 4 (2017), 85.

Andrew Maimone and Junren Wang. 2020. Holographic Optics for Thin and Lightweight Virtual Reality. *ACM Trans. Graph.* (SIGGRAPH) 39, 4 (2020).

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.

Nitish Padmanaban, Yifan Peng, and Gordon Wetzstein. 2019. Holographic Near-eye Displays Based on Overlap-add Stereograms. *ACM Trans. Graph.* 38, 6 (2019).

Jae-Hyeung Park. 2017. Recent progress in computer-generated holography for three-dimensional scenes. *Journal of Information Display* 18, 1 (2017), 1–12.

Yifan Peng, Suyeon Choi, Jonghyun Kim, and Gordon Wetzstein. 2021. Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration. *Science Advances* (2021).

Yifan Peng, Suyeon Choi, Nitish Padmanaban, and Gordon Wetzstein. 2020. Neural holography with camera-in-the-loop training. *ACM Trans. Graph.* 39, 6 (2020), 1–14.

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 Trans. Graph.* 36, 6 (2017).

Liang Shi, BeiChen Li, Changkil Kim, Petr Kelihhofer, and Wojciech Matusik. 2021. Towards real-time photorealistic 3D holography with deep neural networks. *Nature* 591, 7849 (2021), 234–239.

Koki Wakunami, Hiroaki Yamashita, and Masahiro Yamaguchi. 2013. Occlusion culling for computer generated hologram based on ray-wavefront conversion. *Optics express* 21, 19 (2013), 21811–21822.

Fahri Yaras, Hoonjung Kang, and Levent Onural. 2010. State of the Art in Holographic Displays: A Survey. *Journal of Display Technology* 6, 10 (2010), 443–454.

Donghoon Yoo, Youngjin Jo, Seung-Woo Nam, Chun Chen, and ByoungHo Lee. 2021. Optimization of computer-generated holograms featuring phase randomness control. *Opt. Lett.* 46, 19 (2021), 4769–4772.

Friedemann Zenke and Surya Ganguli. 2018. Superspike: Supervised learning in multilayer spiking neural networks. *Neural computation* 30, 6 (2018), 1514–1541.

Hao Zhang, Liangcai Cao, and Guofan Jin. 2017. Computer generated hologram with occlusion effect using layer-based processing. *Applied optics* 56, 13 (2017).

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

Zhengyun Zhang and M. Levoy. 2009. Wigner distributions and how they relate to the light field. In *Proc. ICCP.* IEEE, 1–10.

Remo Ziegler, Simon Bucheli, Lukas Ahrenberg, Marcus Magnor, and Markus Gross. 2007. A Bidirectional Light Field-Hologram Transform. In *Computer Graphics Forum (Eurographics),* Vol. 26. 435–446.