Research Article Vol. 33, No. 9 / 5 May 2025 / Optics Express 19399
Optics EXPRESS

View synthesis for 3D computer-generated holograms using deep neural fields

KENNETH CHEN,[1,*] ANZHOU WEN,[2,3] YUNXIANG ZHANG,[1] PRANEETH CHAKRAVARTHULA,[2,3] AND QI SUN[1]

[1]New York University, Tandon School of Engineering, New York, NY 11201, USA
[2]The University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC 27599, USA
[3]The University of North Carolina at Chapel Hill, Department of Applied Physical Sciences, Chapel Hill, NC 27599, USA
[*]kennychen@nyu.edu

Abstract: Computer-generated holography (CGH) simulates the propagation and interference of complex light waves, allowing it to reconstruct realistic images captured from a specific viewpoint by solving the corresponding Maxwell equations. However, in applications such as virtual and augmented reality, viewers should freely observe holograms from arbitrary viewpoints, much as how we naturally see the physical world. In this work, we train a neural network to generate holograms at any view in a scene. Our result is the Neural Holographic Field: the first artificial-neural-network-based representation for light wave propagation in free space and transform sparse 2D photos into holograms that are not only 3D but also freely viewable from any perspective. We demonstrate by visualizing various smartphone-captured scenes from arbitrary six-degree-of-freedom viewpoints on a prototype holographic display. To this end, we encode the measured light intensity from photos into a neural network representation of underlying wavefields. Our method implicitly learns the amplitude and phase surrogates of the underlying incoherent light waves under coherent light display conditions. During playback, the learned model predicts the underlying continuous complex wavefront propagating to arbitrary views to generate holograms.

1. Introduction

The ultimate goal of displays is to visually teleport an environment – often from the camera captures of their physical counterparts – with the identical viewing experience as if we were in the real world. Holography is a promising technology that can optically record and replay the wavefront from a 3D environment [1]. Computer-generated holography (CGH) numerically simulates this wave propagation and recreates the underlying 3D scene when the recorded pattern is displayed on a spatial light modulator (SLM) illuminated with coherent light. CGH has been shown as a promising technology for virtual/augmented reality (VR/AR) [2–4] and biological discoveries [5]. Although accurately computing the interference patterns is challenging, recent advances in deep learning-based methods [6,7] have demonstrated unprecedented success in computing holograms with low power, high quality, and in real-time [8–12], making CGH promising for telepresence.

Computer-generated hologram phases are typically computed from individual frames of a given 2D image [13,14], 4D light fields [15,16], depth maps [8,17,18], or multi-plane focal stacks [19,20]. This means that existing CGH methods can only reconstruct holograms for limited views that have already been captured, but fail in accounting for arbitrary and novel viewing directions. Consequently, 3D CGHs so far are largely confined to computer-simulated virtual imagery, such as hand-crafted polygonal meshes [21–23], and not yet facilitating the free-viewing telepresence experience of physical environments.

#559364 https://doi.org/10.1364/OE.559364
Journal © 2025 Received 13 Feb 2025; revised 2 Apr 2025; accepted 3 Apr 2025; published 25 Apr 2025

Optics EXPRESS

Three-dimensional imaging techniques have evolved over the years to acquire geometric light propagation with photorealism [24], even with non-line-of-sight [25–27] and distant [28] objects. However, existing methods only solve Maxwell's equations under several approximations for tractable computational complexity, and therefore cannot accurately represent complex phenomena seen from arbitrary views, such as diffraction around partial occlusions [8,15]. This forms a significant impediment to achieving true 3D telepresence: how to represent wave propagation in all directions.

In this work, we bridge the gap between image-based 3D CGH generation and sparse real-world camera photos to achieve free-view, 6 degree of freedom (DOF), and continuous visual reproduction of 3D holograms. To this end, we devise a physics-based implicit hologram representation with a deep neural network, dubbed neural holographic fields (NHFs). We first generate a dataset of per-view light fields using a neural view synthesis technique, which are then transformed to holograms using a stereogram approach. Then, we train a neural network to predict the stereograms from single RGB views, for real-time inference. The NHFs impose the underlying omnidirectional wave propagation in a physical scene with sparse amplitude-only photos. This way, it presents an implicit proxy of the real-world scene and captures complex photorealistic effects. Moreover, our framework builds the real-world scene representation from just a handful of images captured using common consumer smartphones. This not only eliminates the need for creating photorealistic computer models with trillions of polygons and using sophisticated capture equipment, but also for the first time introduces a practical and comprehensive pipeline for camera-display dual telepresence using CGH. See Visualization 1 for an animation example.

2. Generating holograms for novel views

As illustrated in Fig. 1(A), in our experiments, we freely captured unique images (4K resolution, later scaled down to 1280x720 for network training) that sufficiently cover all regions of a scene.

This acquisition step only takes several minutes, and we demonstrate that as few as 50 captured images can adequately represent a physical scene. However, more structured or densely-sampled captures may improve the results, as shown in the study in Fig. 3 (D) which compares image quality against the number of input images. The camera captures are then processed using a structure from motion (SfM) model [29,30] to estimate camera parameters.

We model and represent a target scene as a set of anisotropic 3D Gaussian primitives with learned position, covariance, opacity, and view-dependent color [31]. Given an input camera view, these Gaussians are projected to 2D and rendered into an RGB image I via alpha blending. The color of a single pixel in an image located at coordinates $\theta = (\theta_x, \theta_y)$ with camera location $\mathbf{t}$ and viewing direction $\mathbf{R}$ is computed through alpha blending as

$$
\begin{aligned}
I_{(\theta_x,\theta_y)} = f(\theta, \mathbf{R}, \mathbf{t}) &= \sum_i \mathbf{c}_i \alpha_i g(\theta, \mu_i, \Sigma_i) \prod_{j=1}^{i-1} (1 - \alpha_j g(\theta, \mu_i, \Sigma_i)) \\
g(\theta, \mu_i, \Sigma_i) &= e^{-\frac{1}{2}(\theta-\mu_i)^\intercal \Sigma_i^{-1} (\theta-\mu_i)}.
\end{aligned}
\tag{1}
$$

An image-based loss (L1 norm) between the synthesized and input images is computed, and the gradients are used for a backward pass that optimizes the spherical harmonics parameters, opacity α, positions μ, and 3D covariance Σ of the Gaussians. The learned model represents a holographic radiance field, a function that approximates the radiance emitted from a spatial location $\mathbf{t}$ and viewing direction $\mathbf{R}$ [32].

2.1. Per-view light field synthesis

The learned Gaussian-based omnidirectional radiance field encodes light wave propagation in the 3D environment. By querying the learned radiance field for any given camera view, we

Optics EXPRESS

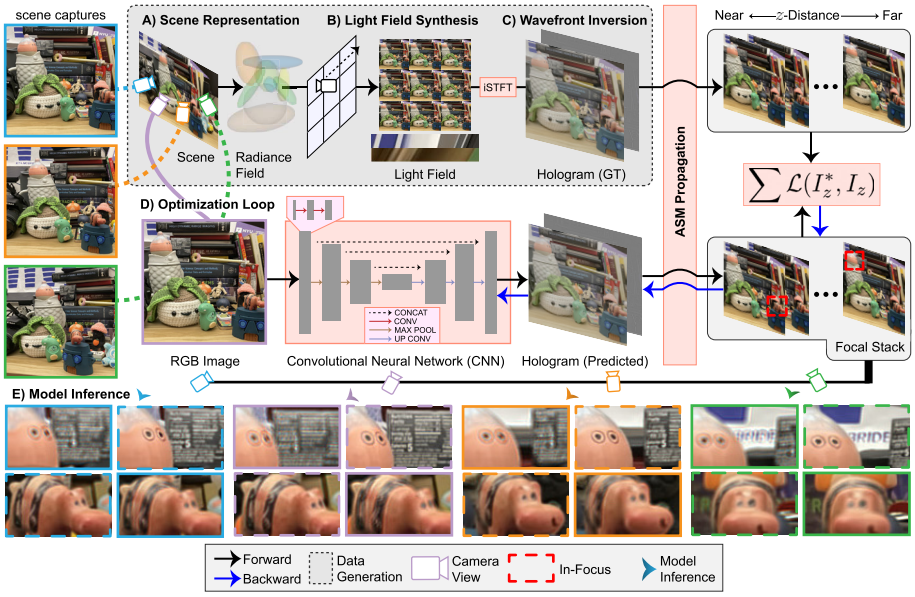

Fig. 1. *Learning and reconstructing free-view neural holographic fields.* (**A**) Camera captures are used to optimize a scene representation model. (**B**) Light field elemental views are rendered by iteratively evaluating the radiance field at uniformly spaced camera positions. Epipolar slices show that the scene representation model can synthesize new views with high image quality. (**C**) The light field is converted to a complex wavefront by computing the inverse of the short-time Fourier transform (iSTFT). (**D**) The model weights of a CNN are optimized by backpropagating the errors of a focal stack loss. The ASM is used to simulate wavefront propagation at different distances from a reference plane. (**E**) Insets of model predictions for four test views are shown, with near and far focus.

synthesize a local light field as an intermediate representation for 3D holograms. For a specific camera orientation and position in 3D space, we obtain images of the scene by evaluating the rendering function f for all pixel coordinates. This allows sampling at arbitrary precision, and enables accurate reconstruction of structured light fields by translating a virtual camera,

$$L\left(j,k,\theta_x,\theta_y\right) = f\left(\left(\theta_x,\theta_y\right),\mathbf{R},\mathbf{t}_{(j,k)}\right),$$
$$\mathbf{t}_{(j,k)} = \mathbf{t} + \mathbf{R}\begin{pmatrix} j\gamma_x \\ k\gamma_y \\ 0 \end{pmatrix}, \tag{2}$$

where j,k are x and y indices of light field elemental views, and γ_x, γ_y are constants which control the distance between elemental images in the x and y directions, respectively. The vector $\mathbf{t}_{(j,k)}$ denotes the translation along the camera's local z plane. Light field elemental views can be sampled by uniformly evaluating Eq. (2) at evenly spaced intervals. This formulation defines the light field central view at $(j,k) = (0,0)$, i.e., when camera translation is unchanged. Epipolar images from the synthesized light fields are extracted by sampling along horizontal and vertical lines, and are displayed in Fig. 1. These images show that the scene representation can interpolate between training images, even at small translation distances.

Optics EXPRESS

2.2. Wavefront inversion

At the core of our NHF representation is a method that computes complex holograms through the inversion of an input light field, similar to prior literature [33]. We first describe how a complex wavefront, $U = ae^{i\phi}$ where a is amplitude and ϕ phase, is related to a light field. Light field angular views can be recovered by computing local samples of a propagated wavefront through a 2D Short-Time Fourier Transform (STFT),

$$\begin{aligned} &S(x,y,f_x,f_y) \\ &\quad = \mathrm{STFT}_{(\theta_x,\theta_y)}\{U(x,y)\} \iint_{-\infty}^{\infty} U(x',y')w(x'-x,y'-y)e^{-2\pi i(f_x x'+f_y y')}dx'dy', \end{aligned} \tag{3}$$

in which a window function w (Hann window in our experiments) constrains the wavefront to a limited set of frequencies. We apply the STFT here because it computes a local frequency spectrum, where the application of a window function results in local light field angular views. Here, the variables x, y are spatial position, θ_x, θ_y are angles, and f_x, f_y are spatial frequencies in the x and y directions which are related to the sine of light field viewing angles as

$$f_x = \frac{\sin(\theta_x)}{\lambda}. \tag{4}$$

Here, λ is the wavelength of a monochromatic illumination source. In practice, three coherent monochromatic sources illuminate the SLM sequentially to display full-color images. The corresponding light field, L, can be computed as the square magnitude of the local STFT spectrum, as

$$L(x,y,\theta_x,\theta_y) = \left|S(x,y,f_x,f_y)\right|^2, \tag{5}$$

where the operator $|\cdot|^2$ is the element-wise squared absolute, or the real-valued intensity of the complex field. In order to recover the underlying complex wavefield from a source light field, we invert the STFT as follows,

$$\begin{aligned} &U(x,y) \\ &= \iint_{-\infty}^{\infty} w(x-x',y-y') \iint_{-\infty}^{\infty} L(x',y',\theta_x,\theta_y)e^{\left(2\pi(f_x x'+f_y y')+\Phi(x,y,f_x,f_y)\right)i}df_x df_y dx'dy' \\ &= \mathrm{STFT}^{-1}\left\{\sqrt{L(x,y,\theta_x,\theta_y)}e^{\Phi(x,y,f_x,f_y)i}\right\}. \end{aligned} \tag{6}$$

Notably, light field images typically do not contain phase information Φ, and heuristics are used to solve the inversion problem [21,34]. This bidirectional transform between light field and hologram can be generalized by the Wigner distribution function (WDF) [35,36], but is typically solved using stereogram approaches [34,37] due to the computational intractability of inverting the WDF.

An illustration of our data processing and hologram generation pipeline is shown in Fig. 1. In summary, the data generation pipeline creates paired training data consisting of RGB images corresponding to light field central views, and complex holograms generated from the light field via wavefront inversion. We pre-compute the data generation step and use it as a training dataset for an efficient deep learning-based optimization pipeline, which we describe next.

3. Learning neural holographic fields

3.1. NHF neural network

Generating the holograms for each view is impractical, especially when rendering light fields with high angular or spatial resolution. On average, this process can take more than 3 minutes

Optics EXPRESS

to generate a monocolor hologram for a single view, and can consume significant memory for wavefront inversion. We train a neural network which learns complex holograms from light field central views to bypass this compute-heavy phase hologram generation step. Our neural network-based hologram generation takes a single RGB image (rendered view from the view synthesis algorithm at the captured camera positions) , and outputs the corresponding six-dimensional hologram amplitude and phase (ie three channels each for RGB color). This is different from traditional 3D hologram generation, which require 3D supervision (e.g. depth maps [38,39], focal stacks [40,41]) to produce accurate focusing effects. The NHF model is a fully convolutional neural network (CNN) with an encoder-decoder structure, and is trained on the simulated data. Skip connections are used to improve the ability of later upsampling layers by reintroducing features lost during downsampling in earlier layers of the network. A simplified schematic of our neural network model as well as its inputs and outputs is shown in Fig. 1.

3.2. Image formation

Images generated by a holographic display are simulated by computing the propagation of a source wavefront, U_s, created by a coherent beam by using the angular spectrum method (ASM) [42]. The modulation of a wavefront by a spatial light modulator (SLM) to a target plane at distance z from the SLM plane can be computed as follows,

$$\begin{aligned} \mathcal{P}(U_s;z) &= \mathcal{F}^{-1}\Big(\mathcal{F}\big\{U_s(u,v;0)\big\} \circ H(f_x,f_y;z)\Big) \\ H(f_x,f_y;z) &= \begin{cases} e^{i2\pi z\sqrt{\frac{1}{\lambda^2}-f_x^2-f_y^2}}, & \text{if } \sqrt{f_x^2+f_y^2}<\frac{1}{\lambda} \\ 0, & \text{otherwise} \end{cases}, \end{aligned} \tag{7}$$

where z is the propagation distance, H is the ASM propagation kernel, and $\mathcal{F}(\cdot)$, $\mathcal{F}^{-1}(\cdot)$ are the 2D Fourier transform and its inverse, which decomposes the wavefront into its plane wave components. Notably, the ASM propagation module is a differentiable operation, and physical wavefront propagation can be simulated to reconstruct holographic images in a learning-based pipeline.

3.3. Training procedure

We use stochastic gradient descent to train the parameters, Θ, of our CNN to produce holograms by minimizing a task-specific loss function,

$$\Theta^* = arg\,min_\Theta \sum_{z\in[z_{\text{near}},z_{\text{far}}]} \mathcal{L}\Big(\mathcal{P}\Big(\mathcal{M}_\Theta\big\{L(0,0)\big\};z\Big),\mathcal{P}\Big(U;z\Big)\Big) \tag{8}$$

where $\mathcal{M}_\Theta$ is our neural network model with trainable parameters Θ. The final trained parameters Θ^* of the network are used to output three-dimensional hologram amplitude and phase a, ϕ (a six-dimensional image) which minimizes this loss function to approximate a ground truth wavefront. Our loss function samples the volume within near and far planes $[z_{\text{near}}, z_{\text{far}}]$ at multiple depths, z, by propagating the wave field using the ASM and consists of both a photometric and perceptual loss,

$$\begin{aligned} &\mathcal{L}\Big(\mathcal{P}\Big(\mathcal{M}_\Theta\big\{L(0,0)\big\};z\Big),\mathcal{P}\Big(U;z\Big)\Big) \\ &= \underbrace{\lambda_{\text{L2}}\sum\Big(\big|\mathcal{P}(\mathcal{M}_\Theta\big\{L(0,0)\big\};z)\big|^2-\big|\mathcal{P}(U;z)\big|^2\Big)^2}_{\text{Photometric Operator}} + \underbrace{\lambda_{\text{vgg}}\mathcal{L}_{\text{vgg}}\Big(\big|\mathcal{P}(\mathcal{M}_\Theta\big\{L(0,0)\big\};z)\big|^2,\big|\mathcal{P}(U;z)\big|^2\Big)}_{\text{Perceptual Loss}} \end{aligned} \tag{9}$$

Optics EXPRESS

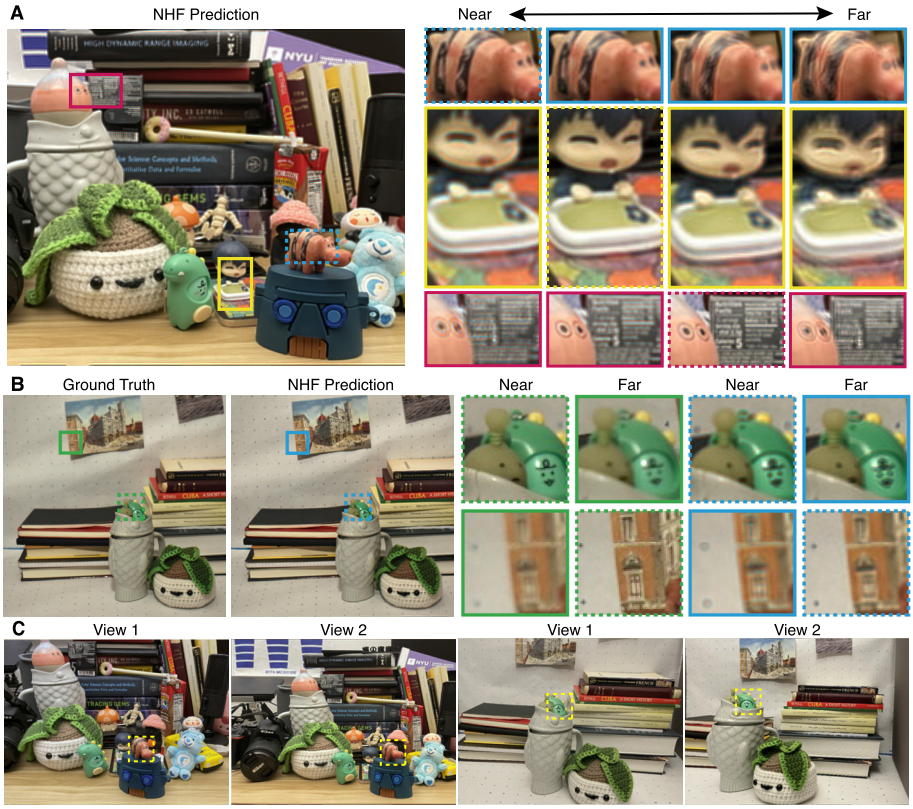

Fig. 2. *Simulated results*. (**A**) NHF-predicted results with all-in-focus DSLR test image as input, as well as focal stacks with zoomed insets. (**B**) Another example scene, including insets with near and far focus. (**C**) New scene views predicted by NHF with near focus. In all examples, dashed boxes represent in-focus regions. Ground truth images were captured using a different camera than used for training of the NHF model, so color correction was applied to displayed images for better comparison.

The terms $\lambda_{\text{vgg}}, \lambda_{l_2}$ are weights on each loss term, which are experimentally tuned. Please refer to the Supplement 1 for more discussion and an ablation study on these terms. Here, we assume $L(0,0)$ is the central light field elemental image, taken at camera parameters $(\mathbf{R}, \mathbf{t})$. The first term in the loss is the photometric operator, or the mean square error loss, which intends to minimize the distance of individual pixel intensities between predicted and ground truth focal stack images. The second term is a VGG loss, which is a perceptually-inspired image distance metric [43]. VGG loss computes the difference between predicted and target images projected to a pre-trained CNN backbone's feature space. This loss is used to optimize for higher-order perceptual artifacts, rather than pixel-wise error, which is the purpose of the photometric error term. After computing the loss between simulated holographic reconstructions predicted by the CNN and the target images, the errors are backpropagated to the neural network weights which are updated using the computed gradients. See the Supplement 1 for pseudocode of our pipeline.

3.4. *Performance evaluation*

The NHF CNN can converge within approximately 40 epochs, and only requires a small dataset (as generated by the procedure in the previous section) to generalize to new views. Due to this, the depth and width of our network need not be large to generalize to unseen views, as long as images

Optics EXPRESS

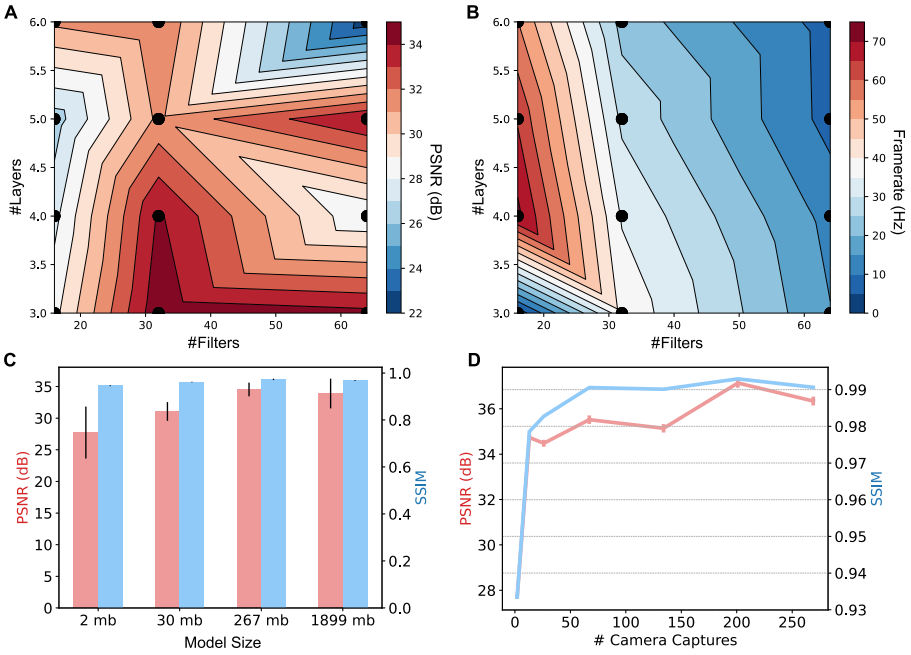

Fig. 3. *Performance evaluation*. (**A**) PSNR (dB) as a function of different neural network parameters. In this figure, the number of convolutional filters (or kernels) and the depth of the network are compared. Black scatter points are parameters sampled for these experiments. (**B**) The run-time for our model is also measured. (**C**) Additional PSNR (red, left *y*-axis) and SSIM (blue, right *y*-axis) numbers are reported for different model sizes, by varying the number of network layers. (**D**) Comparison of number of input views (between 2 and 269 training images) and model performance. Error bars represent one standard deviation.

are in the domain of the original scene images. Figure 2 shows simulated results generated by our NHF pipeline. We evaluate the performance of our neural network by comparing qualitative results to ground truth focal stack images as captured by a NIKON D3500 DSLR. The simulated images not only exhibit high image quality, but also accurate refocusing cues.

We conducted an analysis of the running time of our network as a function of several parameters, such as the number of model layers and filtering dimensions. When implemented on a GPU, the NHF CNN can achieve a forward evaluation framerate of >70Hz for images of resolution 1024×576, compared to the approximately 3.5 minute run time of the wavefront inversion pipeline. A quantitative evaluation of the image quality (PSNR, in units of dB) of images synthesized by our model across different parameters are provided in Fig. 3 (A) and (B). Additionally, we test with different model sizes, as well as with different numbers of training images (between 2-269 images), as shown in Fig. 3 (C) and (D). We find that increasing the model size (number of parameters) beyond a certain point provides negligible performance increase, and may even degrade performance. We also found that performance saturates at around 100 input training views. See the Supplement 1 for additional training details, pseudocode, more results, etc, and Visualization 1 for an animation of our model's view interpolation ability.

4. Display prototype

Our holographic display prototype setup uses a Fisba Ready fiber-coupled laser and phase only LCoS SLM (HOLOEYE PLUTO-2.1-VIS-014) with resolution of 1920 × 1080 and a pixel pitch of 8 µm. The laser provides wavelengths of 450 nm, 520 nm and 660 nm with per-diode

Optics EXPRESS

Fig. 4. *Experimental results.* Left column are simulated results predicted by NHF, and the right are experimental results captured from our holographic display. Dashed boxes represent regions which are in focus.

power control. The fiber-coupled laser is mounted using a Thorlabs FC/APC fiber adapter plate with Kinematic Mount. In our implementation, images are captured using a FLIR color camera for each wavelength and then concatenated channel-wise during post-processing. Other components include one non-polarizing cube beamsplitter with 50 : 50 energy split ratio (Thorlabs BS031), linear polarizers (Thorlabs LPNIRB100), one-axis motorized translation state (Thorlabs PT1/M-Z9) and relay imaging lenses, as shown in Supplement 1 Figure 8. The holograms predicted by NHF are sent to the display after double phase encoding [17,44], which allows us to display the predicted 6-channel complex holograms on our phase-only SLM. Figure 4 shows experimentally captured holographic images, compared to ground truth and simulated results. Supplement 1 Figure 9 shows monochromatic results. Please refer to the Supplement 1 for additional experimental results.

Optics EXPRESS

5. Conclusion

Existing CGH can only be generated from pre-captured imagery, such as RGB-D maps, to synthesize accurate 3D holograms. This view-dependent generation framework limits its potential in free-viewing and interactive scenarios, such as VR/AR. To overcome these barriers, we introduce the first free-viewing neural encoding for CGH–the neural holographic field (NHF). Using just a few freely captured photos from a smartphone as heuristic knowledge of a physical environment, NHF employs a neural network to generate corresponding 3D and colored holograms for any unseen camera view in real-time. We envision NHF paving the way for practical applications of CGH in VR/AR with the aid of AI. In the future, we plan to explore solutions with generative models [45] to create realistic, free-viewing holograms without relying on camera captures. Furthermore, NHF has shown promising results for inferring 3D holograms at arbitrary novel viewpoints in real-time, making it well-suited for holographic displays with small baseline parallax, such as wearable AR/VR systems. Meanwhile, an exciting future direction is to extend this approach to large-scale multiview holographic displays.

Funding. National Science Foundation (2225861, 2232817, 2107454).

Disclosures. The authors declare no competing interests.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. P.-A. Blanche, A. Bablumian, R. Voorakaranam, *et al.*, "Holographic three-dimensional telepresence using large-area photorefractive polymer," Nature **468**(7320), 80–83 (2010).
2. M. Gopakumar, G.-Y. Lee, S. Choi, *et al.*, "Full-colour 3d holographic augmented-reality displays with metasurface waveguides," Nature **629**(8013), 791–797 (2024).
3. K. Wakunami, P.-Y. Hsieh, R. Oi, *et al.*, "Projection-type see-through holographic three-dimensional display," Nat. Commun. **7**(1), 12954 (2016).
4. J. Yang, L. S. Li, Q. He, *et al.*, "An ultrahigh-fidelity 3d holographic display using scattering to homogenize the angular spectrum," Sci. Adv. **9**(41), eadi9987 (2023).
5. D. R. Smith, D. G. Winters, and R. A. Bartels, "Submillisecond second harmonic holographic imaging of biological specimens in three dimensions," Proc. Natl. Acad. Sci. **110**(46), 18391–18396 (2013).
6. C. Zhang, X. Liu, L. Wang, *et al.*, "Lensless efficient snapshot hyperspectral imaging using dynamic phase modulation," Photon. Res. **13**(2), 511–526 (2025).
7. C. Zhang, W. Liu, J. Li, *et al.*, "Tunable optimally-coded snapshot hyperspectral imaging for scene adaptation," Laser Photonics Reviews n/a, 2401921.
8. L. Shi, B. Li, C. Kim, *et al.*, "Towards real-time photorealistic 3d holography with deep neural networks," Nature **591**(7849), 234–239 (2021).
9. Y. Wang, P. Chakravarthula, Q. Sun, *et al.*, "Joint neural phase retrieval and compression for energy-and computation-efficient holography on the edge," ACM Transactions on Graphics (TOG) **41**, 1–16 (2022).
10. C. Chang, K. Bang, G. Wetzstein, *et al.*, "Toward the next-generation vr/ar optics: a review of holographic near-eye displays from a human-centric perspective," Optica **7**(11), 1563–1578 (2020).
11. M. H. Eybposh, N. W. Caira, M. Atisa, *et al.*, "Deepcgh: 3d computer-generated holography using deep learning," Opt. Express **28**(18), 26636–26650 (2020).
12. P. Chakravarthula, E. Tseng, T. Srivastava, *et al.*, "Learned hardware-in-the-loop phase retrieval for holographic near-eye displays," ACM Transactions on Graphics (TOG) **39**(6), 1–18 (2020).
13. Y. Peng, S. Choi, J. Kim, *et al.*, "Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration," Sci. Adv. **7**(46), eabg5040 (2021).
14. P. Chakravarthula, Y. Peng, J. Kollin, *et al.*, "Wirtinger holography for near-eye displays," ACM Transactions on Graphics (TOG) **38**(6), 1–13 (2019).
15. P. Chakravarthula, E. Tseng, H. Fuchs, *et al.*, "Hogel-free holography," ACM Transactions on Graphics **41**, 1–16 (2022).
16. S. Choi, M. Gopakumar, Y. Peng, *et al.*, "Time-multiplexed neural holography: a flexible framework for holographic near-eye displays with fast heavily-quantized spatial light modulators," in *ACM SIGGRAPH 2022 Conference Proceedings*, (2022), pp. 1–9.
17. L. Shi, B. Li, and W. Matusik, "End-to-end learning of 3d phase-only holograms for holographic display," Light: Science Applications **11**(1), 247 (2022).

Optics EXPRESS

18. D. Hiyama, T. Shimobaba, T. Kakue, *et al.*, "Acceleration of color computer-generated hologram from rgb–d images using color space conversion," Opt. Commun. **340**, 121–125 (2015).
19. G. Kuo, F. Schiffers, D. Lanman, *et al.*, "Multisource holography," ACM Transactions on Graphics (TOG) **42**(6), 1–14 (2023).
20. H. Zhang, L. Cao, and G. Jin, "Computer-generated hologram with occlusion effect using layer-based processing," Appl. Opt. **56**(13), F138–F143 (2017).
21. G. Makey, Ö. Yavuz, D. K. Kesim, *et al.*, "Breaking crosstalk limits to dynamic holography using orthogonality of high-dimensional random vectors," Nat. Photonics **13**(4), 251–256 (2019).
22. K. Matsushima and S. Nakahara, "Extremely high-definition full-parallax computer-generated hologram created by the polygon-based method," Appl. Opt. **48**(34), H54–H63 (2009).
23. T. Ichikawa, K. Yamaguchi, and Y. Sakamoto, "Realistic expression for full-parallax computer-generated holograms with the ray-tracing method," Appl. Opt. **52**(1), A201–A209 (2013).
24. D. Stellinga, D. B. Phillips, S. P. Mekhail, *et al.*, "Time-of-flight 3d imaging through multimode optical fibers," Science **374**(6573), 1395–1399 (2021).
25. X. Liu, I. Guillén, M. La Manna, *et al.*, "Non-line-of-sight imaging using phasor-field virtual wave optics," Nature **572**(7771), 620–623 (2019).
26. R. Cao, F. De Goumoëns, B. Blochet, *et al.*, "High-resolution non-line-of-sight imaging employing active focusing," Nat. Photonics **16**(6), 462–468 (2022).
27. M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," Nature **555**(7696), 338–341 (2018).
28. C. Wu, J. Liu, X. Huang, *et al.*, "Non–line-of-sight imaging over 1.43 km," Proc. Natl. Acad. Sci. **118**(10), e2024468118 (2021).
29. J. L. Schönberger, E. Zheng, M. Pollefeys, *et al.*, "Pixelwise view selection for unstructured multi-view stereo," in *European Conference on Computer Vision (ECCV)*, (2016).
30. J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016).
31. B. Kerbl, G. Kopanas, T. Leimkühler, *et al.*, "3d gaussian splatting for real-time radiance field rendering," ACM Transactions on Graphics **42**(4), 1–14 (2023).
32. B. Mildenhall, P. P. Srinivasan, M. Tancik, *et al.*, "Nerf: Representing scenes as neural radiance fields for view synthesis," Commun. ACM **65**(1), 99–106 (2021).
33. L. Huang, H. Chen, T. Liu, *et al.*, "Self-supervised learning of hologram reconstruction using physics consistency," Nature Machine Intelligence **5**(8), 895–907 (2023).
34. N. Padmanaban, Y. Peng, and G. Wetzstein, "Holographic near-eye displays based on overlap-add stereograms," ACM Transactions on Graphics (TOG) **38**(6), 1–13 (2019).
35. Z. Zhang and M. Levoy, "Wigner distributions and how they relate to the light field," in *2009 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2009), pp. 1–10.
36. F. Schiffers, P. Chakravarthula, N. Matsuda, *et al.*, "Stochastic light field holography," in *2023 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2023), pp. 1–12.
37. H. Zhang, Y. Zhao, L. Cao, *et al.*, "Fully computed holographic stereogram based algorithm for computer-generated holograms with accurate depth cues," Opt. Express **23**(4), 3901–3913 (2015).
38. D. Yang, W. Seo, H. Yu, *et al.*, "Diffraction-engineered holography: Beyond the depth representation limit of holographic displays," Nat. Commun. **13**(1), 6012 (2022).
39. Y. Qin, W.-Y. Chen, M. O'Toole, *et al.*, "Split-lohmann multifocal displays," ACM Transactions on Graphics (TOG) **42**(4), 1–18 (2023).
40. P. Chakravarthula, S.-H. Baek, F. Schiffers, *et al.*, "Pupil-aware holography," ACM Transactions on Graphics **41**(6), 1–15 (2022).
41. S. Choi, M. Gopakumar, Y. Peng, *et al.*, "Neural 3d holography: Learning accurate wave propagation models for 3d holographic virtual and augmented reality displays," ACM Transactions on Graphics (TOG) **40**(6), 1–12 (2021).
42. K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," Opt. Express **17**(22), 19662–19673 (2009).
43. J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands*, October 11-14, 2016, Proceedings, Part II 14, (Springer, 2016), pp. 694–711.
44. A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," ACM Transactions on Graphics (Tog) **36**(4), 1–16 (2017).
45. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," Advances in neural information processing systems **33**, 6840–6851 (2020).