Neural Holography

Yifan Peng
Stanford University

Suyeon Choi
Stanford University

Nitish Padmanaban
Stanford University

Jonghyun Kim
NVIDIA, Stanford University

Gordon Wetzstein
Stanford University

ABSTRACT

Holographic displays promise unprecedented capabilities for direct-view displays as well as virtual and augmented reality (VR/AR) applications. However, one of the biggest challenges for computer-generated holography (CGH) is the fundamental tradeoff between algorithm runtime and achieved image quality, which has prevented high-quality holographic image synthesis at fast speeds. Moreover, the image quality achieved by most holographic displays is low, due to the mismatch between physical light transport of the display and its simulated model. Here, we develop an algorithmic CGH framework that achieves unprecedented image fidelity and real-time framerates. Our framework comprises several parts, including a novel camera-in-the-loop optimization strategy that allows us to either optimize a hologram directly or train an interpretable model of the physical light transport and a neural network architecture that represents the first CGH algorithm capable of generating full-color holographic images at 1080p resolution in real time.

CCS CONCEPTS

• **Hardware → Displays and imagers**.

KEYWORDS

computational displays, holography, virtual and augmented reality

ACM Reference Format:
Yifan Peng, Suyeon Choi, Nitish Padmanaban, Jonghyun Kim, and Gordon Wetzstein. 2020. Neural Holography. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Emerging Technologies (SIGGRAPH '20 Emerging Technologies), August 17, 2020.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388534.3407295

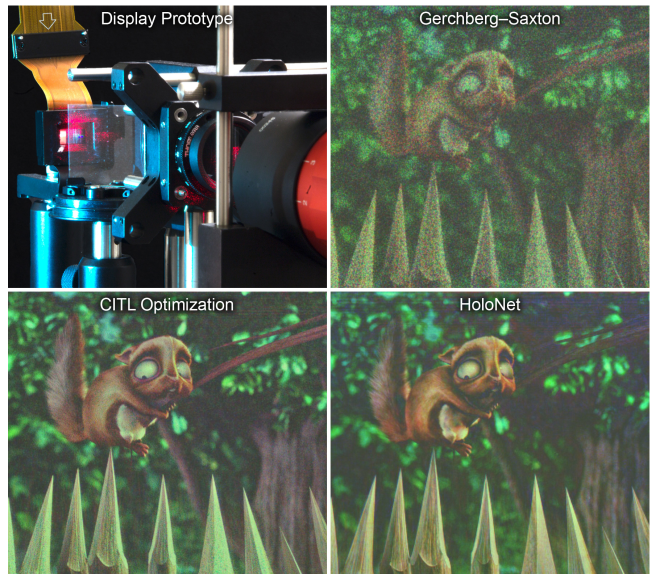

Figure 1: Conventional computer-generated holography algorithms, such as Gerchberg–Saxton, suffer from artifacts. We introduce camera-in-the-loop (CITL) optimization strategies and a neural network, HoloNet, that achieve unprecedented holographic image quality and real-time framerates.

1 INTRODUCTION

Computer-generated holography has recently experienced a renaissance in the computational optics and computer graphics communities. For direct-view displays, holography enables glasses-free 3D display modes and in VR/AR systems, holography has the potential to optimize some of the biggest remaining challenges, such as dynamic image and eyebox steering capabilities, focus cues, vision correction, device form factors, as well as image resolution and brightness. However, the challenge of robustly achieving high image fidelity with experimental holography setups while simultaneously achieving real-time performance remains unsolved. This challenge presents a major roadblock for making holographic displays a practical (near-eye) display technology.

Since the invention of the holographic principle in the late 1940s, much progress has been made. The laser enabled the first optical holograms, and digital computers and spatial light modulators (SLMs) enabled holographic video based on computer-generated holography (CGH). While early efforts aimed for holographic television [1], more recent work has focused on holographic near-eye displays for VR/AR applications [2, 5–7, 9, 10]. Over the last few decades, significant research efforts have been dedicated to advancing CGH algorithms, where a wave field is digitally propagated from a target plane or volume to the source or SLM plane, for example using the angular spectrum method [4]. To enforce the phase-only constraints imposed by current SLMs, direct methods use phase coding [7] to approximate the complex-valued wave field on the SLM with a phase-only field. To achieve the same goal, iterative methods use optimization approaches based on phase retrieval [3]. Typically, direct methods are faster than iterative approaches but offer lower image quality (see Fig. 2).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGGRAPH '20 Emerging Technologies, August 17, 2020, Virtual Event, USA
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7967-0/20/08.
https://doi.org/10.1145/3388534.3407295

2 RESULTS

We have developed an algorithmic CGH framework based on variants of stochastic gradient descent (SGD) to address these long-standing challenges. Using the proposed framework, we design a novel camera-in-the-loop (citl) optimization strategy that allows us to iteratively optimize a hologram using a hybrid physical–digital light transport model. To this end, we show phase patterns on a phase-only spatial light modulator (SLM), capture the result with a camera, calculate the error for some target image, and backpropagate the error into the phase pattern to iteratively improve the phase pattern. We show that this CGH approach achieves the best image fidelity to date, because it directly evaluates the error between target image and synthesized hologram using the physically observed holographic image (see Fig. 1).

Our framework also allows us to automatically calibrate a differentiable light transport model of the physical display. This calibration procedure builds on automatic differentiation and is akin to the training phase of a neural network, where many example images are presented on a physical setup and the error between captured result and target image is backpropagated into a differentiable proxy of the physical hardware system. This proxy models the intensity distribution of the laser source on the SLM, the nonlinear mapping from voltage to phase delay of each SLM pixel, and optical aberrations between the SLM and the target image plane. Unlike the global lookup tables used by commercial SLMs to model voltage-to-phase mapping, our differentiable model is capable of modeling a unique mapping function per SLM pixel and automatically calibrate them. This model-based approach requires a camera in the loop for the training stage. Once the proxy model is optimized, it can be directly used to synthesize new holograms of images that are not included in the training dataset without requiring a camera.

We have also developed a neural network architecture, HoloNet, that is trained with our citl-trained model, to enable full-color high-quality holographic images at 1080p resolution in real time. This network takes an sRGB image as input, which is first converted to amplitude values, then passed to a target-phase-generator subnetwork, which predicts a phase on the target plane. The field is propagated to the SLM plane and processed by a phase encoder, which produces a final phase-only representation to be displayed on the SLM. The quality of the results achieved by HoloNet approaches that of our iterative citl holography optimization (see Fig. 1).

Finally, we explore two approaches to extending the proposed algorithms to 3D holographic image presentation: a holographic varifocal and a multiplane display mode. Both of these display modes have the potential to mitigate the vergence–accommodation conflict as well as to optimize visual comfort and perceptual realism in VR/AR systems. As opposed to conventional varifocal [8] or multifocal displays, our holographic varifocal and multifocal display modes do not require mechanical actuation of any display components, focus-tunable optics, stacked waveguides or microdisplays, or high-speed SLMs. Our 3D holographic display modes either synthesize a single high-quality 2D hologram at varying distances to the SLM or they optimize 2D holograms at multiple distances to be displayed simultaneously by a single SLM.

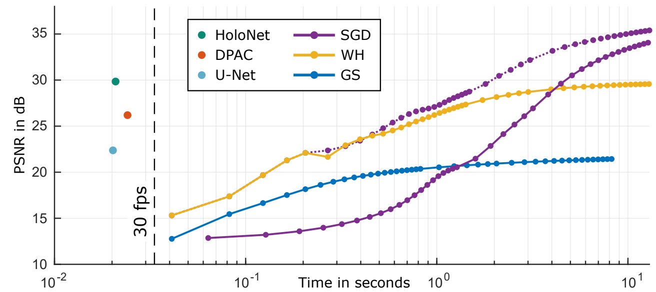

Figure 2: Direct CGH algorithms achieve real-time rates, but HoloNet is the only one to also achieve a peak signal-to-noise ratio (PSNR) of ≈30 dB. Iterative algorithms, such as Gerchberg–Saxton (GS) or Wirtinger Holography (WH), offer a slightly improved quality at the cost of extensive compute times. Our SGD algorithm achieves the best image quality among all CGH algorithms. PSNR values are averaged over 100 test images.

Please watch our video to see all of these results and visit our website at www.computationalimaging.org.

At the intersection of graphics and computational optics, advanced computer-generated holography algorithms are a key enabling technology for 3D virtual and augmented reality applications. With our work, we take first steps to combine classical CGH algorithms and optical systems with modern machine-learning techniques to address several long-standing challenges, such as speed and image quality. We believe that our work paves the way for a new era of neural holographic displays.

ACKNOWLEDGMENTS

We thank Julien Martel and our sponsors: Ford, NSF (1553333, 1839974), a Sloan Fellowship, and a PECASE by the ARO.

REFERENCES

[1] Stephen A. Benton and V. Michael Bove. 2008. *Holographic Imaging*. Wiley-Interscience.
[2] Praneeth Chakravarthula, Yifan Peng, Joel Kollin, Henry Fuchs, and Felix Heide. 2019. Wirtinger Holography for Near-eye Displays. *ACM Trans. Graph. (SIGGRAPH Asia)* 38, 6 (2019).
[3] James R Fienup. 1982. Phase retrieval algorithms: a comparison. *Applied optics* 21, 15 (1982), 2758–2769.
[4] Joseph W Goodman. 2005. *Introduction to Fourier optics*. Roberts and Company.
[5] Changwon Jang, Kiseung Bang, Gang Li, and Byoungho Lee. 2018. Holographic Near-eye Display with Expanded Eye-box. *ACM Trans. Graph. (SIGGRAPH Asia)* (2018).
[6] 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)* (2017).
[7] Andrew Maimone, Andreas Georgiou, and Joel S Kollin. 2017. Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph. (SIGGRAPH)* (2017).
[8] Nitish Padmanaban, Robert Konrad, Tal Stramer, Emily A. Cooper, and Gordon Wetzstein. 2017. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *PNAS* (2017).
[9] Nitish Padmanaban, Yifan Peng, and Gordon Wetzstein. 2019. Holographic Near-eye Displays Based on Overlap-add Stereograms. *ACM Trans. Graph. (SIGGRAPH Asia)* (2019).
[10] 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. (SIGGRAPH Asia)* (2017).