Learned Propagation Model with Complex Convolutions for Holographic Systems

Manu Gopakumar
Stanford University
manugopa@stanford.edu

Bryan Chiang
Stanford University
bhchiang@stanford.edu

1. Motivation

In recent years, advancements in virtual reality displays have greatly reduced visual discrepancies between the actual world and displayed scenes. Higher spatial resolution has allowed displays to more accurately depict fine details, and higher frame rate has produced smoother viewing experiences. However, recreating the visual focus cues found in the real world has remained challenging since this requires both spatial and angular control of the light emitted by a display.

Prior work has used different trade-offs to achieve the desired angular resolution. Various light field display implementations have tackled this issue by trading away spatial resolution for angular resolution [5]. Recent multifocal display prototypes have created focus visual cues by presenting objects at different depths in subsequent frames presenting a trade-off between frames per scene and depth resolution [3]. Varifocal display prototypes have presented scenes at a single depth which is adjusted based on which part of the scene the viewer is looking at [1]. This avoids any of the trade-offs of the other approaches but requires very fast and accurate eye-tracking.

These limitations motivate yet another approach for reproducing visual focus cues, holography. By directly simulating the light wavefront which contains both spatial and angular information, holographic displays aim to sidestep the difficulties faced by the discussed prior approaches. While the theoretical potential for the holographic displays is great, physical implementations have struggled to approach their simulated potential. This mismatch occurs because the images produced by holographic displays are very sensitive to aberrations which current physical setups have failed to eliminate completely. This motivates the need for better propagation models which can accurately simulate the effect of these aberrations.

2. Discussion of Related Work

The original implementations of computer-generated holography modeled holographic display systems as phase modulation of constant uniform illumination which is incident on a spatial light modulator (SLM). This constant amplitude wavefront with modulated phase is then propagated from the SLM to the viewer with the angular spectrum method (ASM) [6]. Unfortunately, this simple propagation model fails to account for many of the effects present in a physical system.

Recent work has sought to model these unaccounted for effects by introducing additional learned terms to the simulated model [7]. These terms explicitly model the effect of content independent source and target field variations, Zernike aberrations along the optical propagation, phase non-linearities, and content-dependent undiffracted light. While these physically inspired parameters provided interpretability for the learned operations, they lack the flexibility to fully model the aberrations of the physical system.

Other recent work has used a more flexible model to learn the propagation of the light through the aberrated physical system [2]. An image-to-image translation UNet is used to learn the transformation from images predicted by the original ASM propagation model and the captured images. This UNet provides more flexibilty for the learned propagation operator, but it is limited because it is trying to learn the aberrations as a function of just the amplitude of the ASM predicted wavefront.

3. Overview

In the proposed approach, a UNet will be directly applied to the complex ASM predicted wavefront in order to learn the aberrations of the physical system. Since this UNet learns the aberration operator on the wavefront, it has more information than prior approaches which operated solely on the amplitude of the ASM predicted wavefront. This should better model the physical aberrations since they are a function of the complex wavefront not just the amplitude of the wavefront at the target plane. Applying a UNet on the complex amplitude does require some modifications of conventional UNet architectures. The conv and convtranspose layers of Pytorch do not support complex convolution, which motivates the use of JAX which has more developed support for complex valued operations. The normalization layers and activation functions of conventional UNet architectures are also designed for real values, so operators designed for complex values can be substituted into this complex UNet [4]. With these changes, a more accurate and efficient forward model should hopefully be learned.

3.1. Milestones and Timeline

Milestones:

- Implement ASM propagation with JAX
- Implement UNet in JAX with complex convolution layers
- Construct and train propagation model using ASM propagation and UNet with complex convolution
- Experiment with different complex activation functions and normalization layers
- Compare proposed model results with UNets with real convolution
- Test optimized phase patterns with best complex UNet configuration
- Create presentation
- Write report

3.2. Timeline

- Week 1 (by Feb 19, 2021)

  Implement ASM propagation with JAX

  Implement a UNet in JAX with complex convolution between the layers

- Week 2 (by Feb 26, 2021)

  Construct and train propagation model using ASM Propagation and UNet with complex convolution

  Experiment with different complex activation functions and normalization layers

- Week 3 (by Mar 12, 2021)

  Compare proposed model results with UNets with real convolution

  Test optimized phase patterns with best complex UNet configuration

- Week 4 (by Mar 19, 2021)

  Create presentation

  Write report

References

[1] K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)*, 36(6):1–13, 2017.

[2] P. Chakravarthula, E. Tseng, T. Srivastava, H. Fuchs, and F. Heide. Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. *ACM Transactions on Graphics (TOG)*, 39(6):1–18, 2020.

[3] J.-H. Chang, B. V. K. Vijaya Kumar, and A. C. Sankaranarayanan. Towards multifocal displays with dense focal stacks. *SIGGRAPH Asia / ACM Trans. Graphics*, 37(6):1–13, 2018.

[4] A. Hirose. *Complex-valued neural networks: theories and applications*, volume 5. World Scientific, 2003.

[5] H. Hua and B. Javidi. A 3d integral imaging optical see-through head-mounted display. *Optics express*, 22(11):13484–13491, 2014.

[6] 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.

[7] Y. Peng, S. Choi, N. Padmanaban, J. Kim, and G. Wetzstein. Neural holography. In *ACM SIGGRAPH 2020 Emerging Technologies*, pages 1–2. 2020.