Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP) © Optica Publishing Group 2023

Learned Large Field-of-View Imaging with Efficient Shift-Variant Wave Optics Modeling

Haoyu Wei,[1] Xin Liu,[1,2] Edmund Y. Lam,[1,*] Yifan Peng[1,*]

[1] *Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China*

[2] *College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China*

[*] *elam@eee.hku.hk, evanpeng@hku.hk*

Abstract: Efficiently modeling shift-variant point spread functions is essential in designing large field-of-view imaging systems. We present an end-to-end optimized large field-of-view imaging framework that models off-axis diffraction at ultra-fast speed.

1. Introduction

In the era of computational imaging, enabling a large field-of-view (FoV) imaging modality requires the joint optimization of the optics, such as diffractive optical elements (DOE), and the reconstruction algorithms [3]. A prerequisite for such a scheme is an accurate, efficient, and differentiable simulation of the forward model at large incident angles. However, existing solutions are either inefficient or non-differentiable [3, 5]. For instance, under the non-paraxial assumption, the sampling strategy should be carefully optimized to prevent error propagation or excessive waste of resources induced by under-or-over-sampling. Hereby, we propose a large-FoV imaging model assisted with accurate and efficient numerical simulation. By utilizing the shifting properties of the Fourier transform, we greatly relieve the sampling requirement hence shortening the computation time. We demonstrate faithful reconstructions at around 35° FoV by jointly optimizing a DOE encoded in the aperture plane and a reconstruction network.

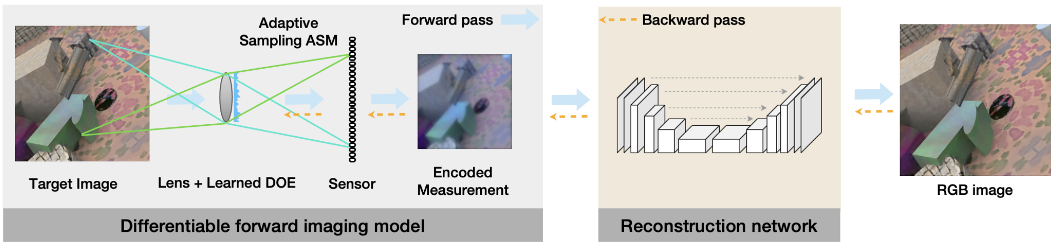

Fig. 1. End-to-end optimization framework for large-FoV imaging. A target image is convolved with shift-variant PSFs to produce an encoded image, which is decoded using a reconstruction network.

2. Method

We design an end-to-end optimized large FoV imaging system composed of a camera with a learned DOE and a reconstruction network, as illustrated in Fig. 1. In the forward model, an off-axis spherical wave originating from the object space is modulated by the camera lens and the DOE, then a point spread function (PSF) is calculated using the angular spectrum method (ASM) [1]. We divide the object space into several sub-regions, within which the PSFs are assumed shift-invariant. Therefore, the captured image is generated by convolving each PSF kernel with the corresponding object space. In the reconstruction process, the captured image is pre-inverted and an RGB image is recovered by the reconstruction network that includes a UNet structure [4].

Our adaptive sampling angular spectrum method fully optimizes both spatial and frequency sampling requirements using the shifting property of the Fourier transform, enabling accurate and efficient large-FoV imaging modeling. In the spatial domain, we identify an asymmetry in the angular spectrum of an input field, primarily caused by its off-axis components. This asymmetry causes an increase in the sampling rate as the spectrum is located in high-frequency regions, according to the Nyquist sampling theorem [1]. Suppose an oblique wave with phase $\phi_u(\xi, \eta)$ arrives right after the aperture, with the center of its angular spectrum at (f_{Xc}, f_{Yc}). We compensate a linear term $\phi_{\text{comp}} = -2\pi(f_{Xc}\xi + f_{Yc}\eta)$ to it such that the resulting field has a frequency center at zero. Then

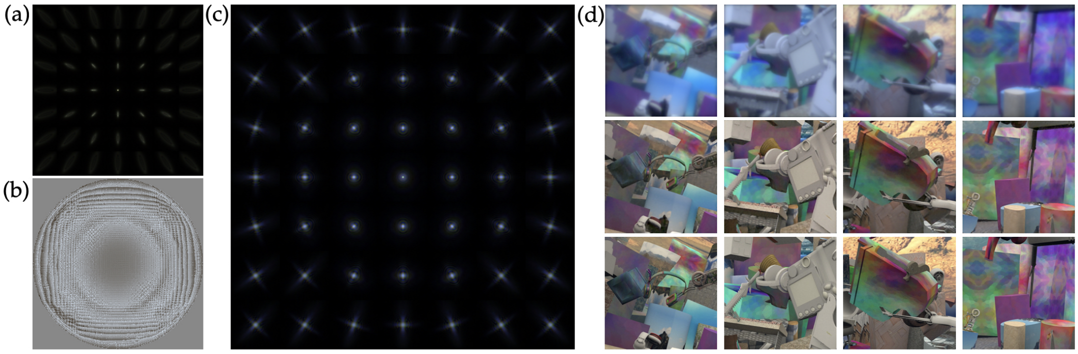

Fig. 2. Full FoV imaging results. (a) Initial PSFs without DOE at different angles of incidence. (b) Optimized DOE. (c) Optimized PSFs. (d) The measurements (first row), reconstructions (second row), and ground truth images (third row).

the spatial sampling rate is reduced by exactly $(2|f_{Xc}|, 2|f_{Xc}|)$. In the frequency domain, we similarly shift the observation window from (x_c, y_c) toward the origin by applying a linear phase term $2\pi(f_X x_c + f_Y y_c)$ to the transfer function, where (f_X, f_Y) are the frequency coordinates, hence reducing the frequency sampling rate.

3. Experiment and Results

In the forward imaging model, we compute an off-axis complex PSF for a thin-lens RGB camera with a focal length of 35 mm and f-number of 16. The PSF is formed by focusing an off-axis spherical wave that originates from 1.7 m away from the aperture. The proposed model is trained end-to-end with the Scene Flow dataset [2], and tested with images of 448×448 pixels. The object space is divided into 7×7 sub-regions and four groups of incident angles, including 0, 0.5, 0.707 of the FoV, and the full FoV. The optimized PSFs and DOE are presented in Fig. 2. It is seen that the DOE is prone to produce shift-invariant PSFs with more concentrated energy to better reconstruct an all-in-focus image. Visual results of the captured measurements, reconstructions, and ground truth images are demonstrated in Fig. 2 (d). The recovered images still preserve fine details in regions of large FoV. We further compute the runtime of a single forward iteration using our model to capture one image and compare our adaptive sampling scheme with the state-of-the-art sampling method Shift-BEASM [5]. The experiments are run on a workstation with an AMD Ryzen Threadripper Pro 3955WX CPU. As in Table 1, our method outperforms Shift-BEASM by 176×. In addition, the NUFFT implementation used in Shift-BEASM does not allow for GPU computation, whereas we use a matrix-based PyTorch implementation that easily fits into a learning-based framework, further accelerating the speed. A prototype for the proposed camera is under development.

Table 1. Computation runtime using our approach versus Shift-BEASM.

Time (sec) \ Angle (°)	0	17.34	24.52	34.68
Shift-BEASM	16.25	16.83	15.86	15.90
Ours	0.11	0.13	0.09	0.09

Funding

This work is supported by Research Grants Council of Hong Kong (ECS 27212822, GRF 17201822).

References

1. Joseph W Goodman. *Introduction to Fourier Optics*. W.H. Freeman and Company, 4th edition, 2017.
2. Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4040–4048, June 2016.
3. Yifan Peng, Qilin Sun, Xiong Dun, Gordon Wetzstein, Wolfgang Heidrich, and Felix Heide. Learned large field-of-view imaging with thin-plate optics. *ACM Trans. Graph.*, 38(6):219–1, November 2019.
4. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, October 2015.
5. Wenhui Zhang, Hao Zhang, Kyoji Matsushima, and Guofan Jin. Shifted band-extended angular spectrum method for off-axis diffraction calculation. *Optics Express*, 29(7):10089–10103, March 2021.