Optica Imaging Congress 2025 (3D, COSI, DH, IS, pcAOP, RadIT) © Optica Publishing Group 2025

Enabling Wide Field-of-View Imaging via Jointly Learned Off-aperture Encoding

Haoyu Wei,[1] Edmund Y. Lam,[1,2] and Yifan Peng[1,3]

[1] *Electrical and Electronic Engineering Department, The University of Hong Kong, Pok Fu Lam, Hong Kong*

[2] *elam@eee.hku.hk,* [3] *evanpeng@hku.hk*

Abstract: We propose an end-to-end off-aperture diffractive optical element (DOE) design, co-optimizing position and optical properties. This overcomes on-aperture DOE limitations in localized aberration correction, significantly improving wide field-of-view image quality. © 2025 The Author(s)

1. Introduction

Deep Optics, employing joint design of optics and algorithms, has shown promise across diverse imaging applications by enhancing encoding capabilities beyond traditional heuristic optical designs [4]. However, conventional designs predominantly places DOEs at or near the aperture plane for global wavefront modulation [7], which becomes a bottleneck for localized aberration correction, particularly in wide field-of-view (FoV) imaging. To overcome this limitation, we investigate a co-optimized off-aperture DOE placement, demonstrating its capacity for localized wavefront control and improved aberration correction. A lightweight image restoration network then decodes the sensor measurements and removes optical blur across different FoVs. The system is illustrated in Fig. 1. Our findings reveal that off-aperture DOEs significantly outperform on-aperture counterparts in compact wide-FoV imagers, achieving over 5dB PSNR improvement and showcasing the potential of our approach.

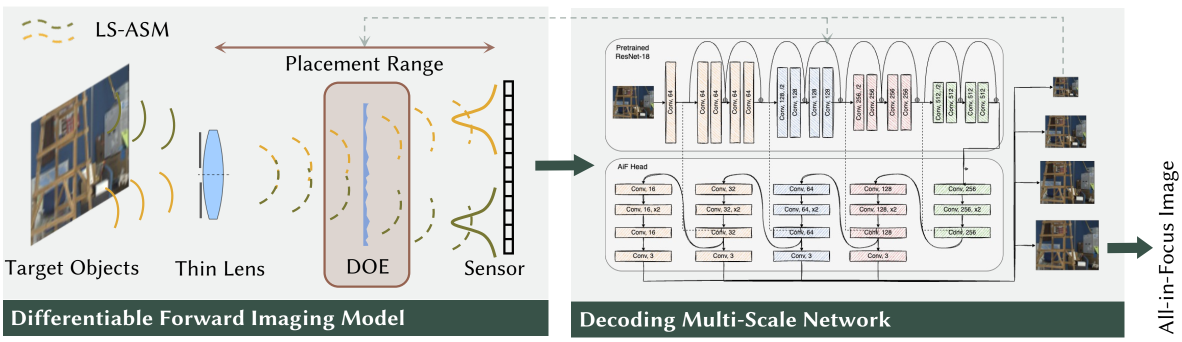

Fig. 1. Illustration of the end-to-end optimization procedure for compact wide-FoV imaging. A DOE and its location are co-optimized with a multi-scale decoding network.

2. Method

We design a fully end-to-end optimized wide-FoV imaging system by placing a DOE at an optimized location between the aperture and the sensor plane. The permissible DOE placement range, without increasing the camera's form factor, spans from the principal plane (or aperture, denoted as $l = 0.0$) to the sensor plane ($l = 1.0$). For ease of assembly, we heuristically restrict the optimization range to $l \in [0, 0.4]$ and enforce this constraint during training using a soft bounding penalty in the regularization. To model the two-step off-axis wave propagation efficiently, we employ the least sampling angular spectrum method [6], which enables fast, accurate simulations at large angles. This reduces computational costs by propagating only between local windows on the DOE and sensor, avoiding full-wave simulations. During optimization, we randomly sample two angles per forward pass to ensure stable convergence of the DOE phase profile. The encoded measurements are restored by a multi-scale image restoration network, including a pre-trained ResNet18 [3] as the feature extractor and an interconnected decoding head. The network outputs images at four resolutions for loss and gradient computation, ultimately reconstructing an all-in-focus, high-quality image.

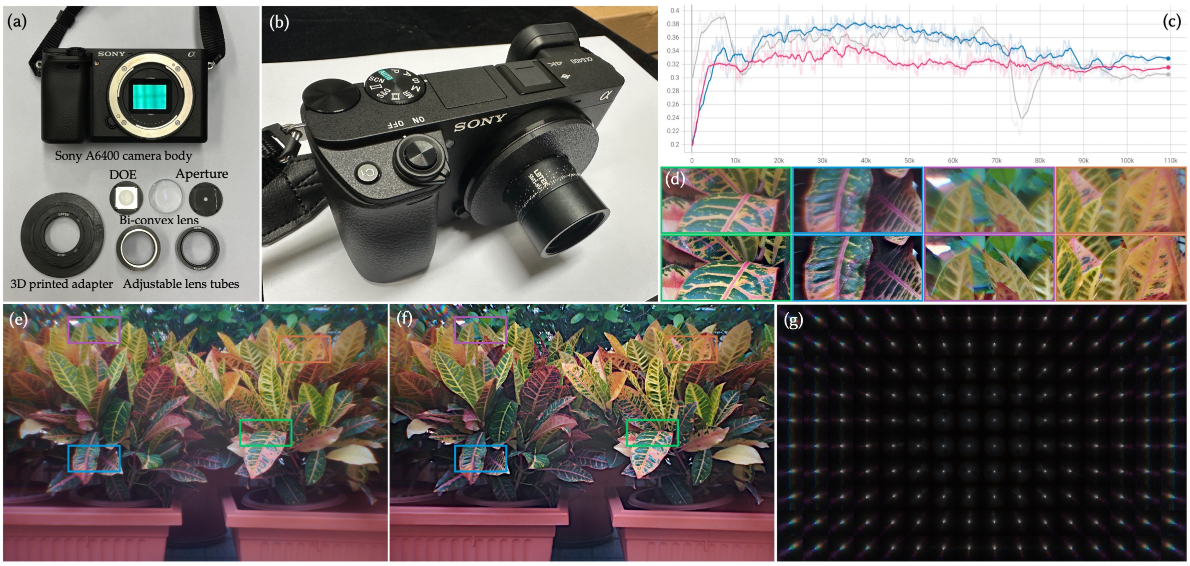

Fig. 2. (a, b) The fabricated prototype. (c) Optimization curves of DOE placement, converging to approximately $l = 0.3$. (d) Zoom-in patches for comparison. (e, f) Real captured outdoor raw measurement and reconstruction. (g) Optimized point spread functions (PSFs) across FoVs.

3. Results

We design a compact wide-FoV imaging system, comprising a single thin refractive lens paired with a rotationally symmetric off-aperture DOE, optimized for a 45° FoV. Our prototype (Fig. 2) features a 35mm effective focal length with an f/12 aperture, focused at 1.4m. We trained our multiscale network for 50 epochs on a diverse dataset of 26,281 images from three datasets, with separate test subsets for evaluation. The results are presented in Fig. 2. Through multiple optimization runs, the DOE positions consistently converged to approximately $l = 0.3$, which guided our fabrication. The PSFs across the FoV confirm localized phase control, achieving focused light at all incident angles. Visual and quantitative evaluations (Table 1) demonstrate the effectiveness of our off-aperture DOE approach over on-aperture systems while maintaining compact form factor.

Table 1. Imaging performance in PSNR↑ (dB) and SSIM↑ with on- or off-aperture DOEs.

DOE location / Time (sec)	On-aperture	Optimized off-aperture
Sceneflow [5]	30.65 / 0.964	**35.63 / 0.977**
Dualpixel [2]	27.83 / 0.922	**31.70 / 0.946**
InStereo2K [1]	30.97 / 0.966	**35.39 / 0.974**

Funding

This work was partially supported by the Research Grants Council of Hong Kong (GRF 17208023).

References

1. Wei Bao, Wei Wang, Yuhua Xu, Yulan Guo, Siyu Hong, and Xiaohu Zhang. Instereo2k: a large real dataset for stereo matching in indoor scenes. *Science China Information Sciences*, 63:1–11, 2020.
2. Rahul Garg, Neal Wadhwa, Sameer Ansari, and Jonathan T Barron. Learning single camera depth estimation using dual-pixels. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7628–7637, 2019.
3. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
4. Daniel S Jeon, Seung-Hwan Baek, Shinyoung Yi, Qiang Fu, Xiong Dun, Wolfgang Heidrich, and Min H Kim. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Transactions on Graphics*, 2019.
5. 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/CVF Conference on Computer Vision and Pattern Recognition*, pages 4040–4048, 2016.
6. Haoyu Wei, Xin Liu, Xiang Hao, Edmund Y Lam, and Yifan Peng. Modeling off-axis diffraction with the least-sampling angular spectrum method. *Optica*, 10(7):959–962, 2023.
7. Haoyu Wei, Xin Liu, Edmund Y Lam, and Yifan Peng. Learned large field-of-view imaging with efficient shift-variant wave optics modeling. In *Computational Optical Sensing and Imaging*, pages CTu3B–3. Optica Publishing Group, 2023.