



End-to-End Hybrid Refractive-Diffractive Lens Design with Differentiable Ray-Wave Model

XINGE YANG, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

MATHEUS SOUZA, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

KUNYI WANG, King Abdullah University of Science and Technology (KAUST), Canada and University of British Columbia, Canada

PRANEETH CHAKRAVARTHULA, University of North Carolina at Chapel Hill (UNC), United States of America

QIANG FU, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

WOLFGANG HEIDRICH, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

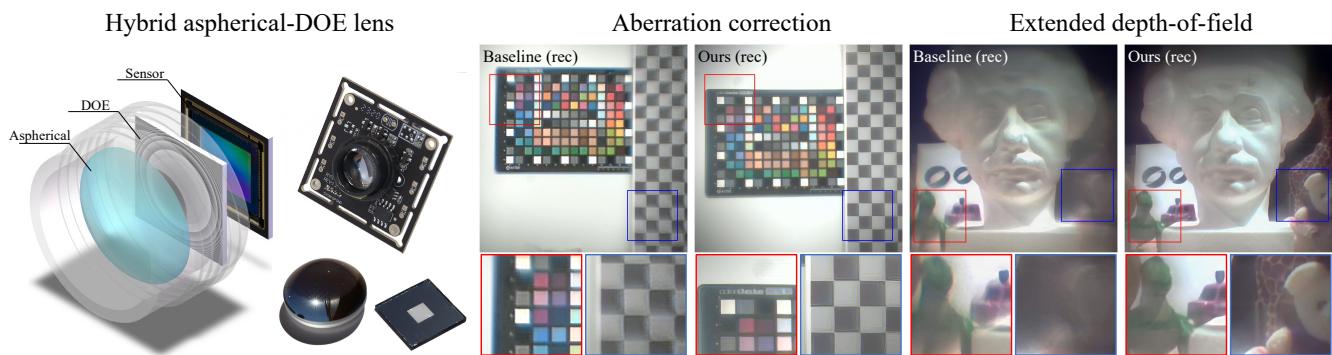


Fig. 1. We propose a differentiable ray-wave imaging model that accurately simulates both aberrations and phase modulation while enabling end-to-end optimization for the hybrid refractive-diffractive lens and the image reconstruction network. To experimentally demonstrate the effectiveness of the proposed model, we present a hybrid aspherical-DOE lens prototype (left) and investigate two applications: aberration correction (middle) and large field-of-view extended depth-of-field imaging (right).

Hybrid refractive-diffractive lenses combine the light efficiency of refractive lenses with the information encoding power of diffractive optical elements (DOE), showing great potential as the next generation of imaging systems. However, accurately simulating such hybrid designs is generally difficult, and in particular, there are no existing differentiable image formation models for hybrid lenses with sufficient accuracy.

In this work, we propose a new hybrid ray-tracing and wave-propagation (ray-wave) model for accurate simulation of both optical aberrations and diffractive phase modulation, where the DOE is placed between the last

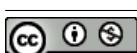
refractive surface and the image sensor, i.e. away from the Fourier plane that is often used as a DOE position. The proposed ray-wave model is fully differentiable, enabling gradient back-propagation for end-to-end co-design of refractive-diffractive lens optimization and the image reconstruction network. We validate the accuracy of the proposed model by comparing the simulated point spread functions (PSFs) with theoretical results, as well as simulation experiments that show our model to be more accurate than solutions implemented in commercial software packages like Zemax. We demonstrate the effectiveness of the proposed model through real-world experiments and show significant improvements in both aberration correction and extended depth-of-field (EDoF) imaging. We believe the proposed model will motivate further investigation into a wide range of applications in computational imaging, computational photography, and advanced optical design. Code will be released upon publication.

CCS Concepts: • Computing methodologies → Computational photography.

Additional Key Words and Phrases: Differentiable Optics, Computational Imaging, End-to-end Optical Design, Computational camera

ACM Reference Format:

Xinge Yang, Matheus Souza, Kunyi Wang, Praneeth Chakravarthula, Qiang Fu, and Wolfgang Heidrich. 2024. End-to-End Hybrid Refractive-Diffractive Lens Design with Differentiable Ray-Wave Model. In *SIGGRAPH Asia 2024 Conference Papers (SA Conference Papers '24)*, December 03–06, 2024, Tokyo, Japan. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3680528.3687640>



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

SA Conference Papers '24, December 03–06, 2024, Tokyo, Japan

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1131-2/24/12

<https://doi.org/10.1145/3680528.3687640>

1 Introduction

End-to-end optical design [Metzler et al. 2020; Shi et al. 2022; Sitzmann et al. 2018; Sun et al. 2020; Wetzstein et al. 2020; Wu et al. 2019] has demonstrated remarkable potential in scientific imaging [Baek et al. 2021; Dun et al. 2020; Jeon et al. 2019; Li et al. 2022; Metzler et al. 2020; Shi et al. 2022; Sun et al. 2020; Tseng et al. 2021a], computer vision [Côté et al. 2023; Ikoma et al. 2021; Tseng et al. 2021b; Wu et al. 2019; Yang et al. 2023b], and advanced optical system design [Yang et al. 2023a; Zheng et al. 2023b], surpassing classical lens design approaches. An end-to-end lens system comprises an optical encoder, such as a refractive lens, the diffractive optical element (DOE), and/or metasurface [Chen et al. 2016; Tseng et al. 2021a], which captures information from the real world, and a neural network decoder that reconstructs the final output, which can be an image or other visual representation. The optics and the network are jointly optimized using gradient backpropagation to find the optimal computational optical system for a given task. Among existing end-to-end computational lenses, hybrid refractive-diffractive lenses [Flores et al. 2004; Wang et al. 2016] combine the advantages of both refractive and diffractive optics, showcasing the great potential for next-generation optical imaging systems.

However, no existing differentiable image formation model provides sufficient accuracy for hybrid refractive-diffractive lenses, posing a significant challenge for the end-to-end design of such systems. The commonly used paraxial wave optics model [Goodman 2005] idealizes the refractive lens as a thin phase plate and neglects optical aberrations. The ray tracing model, employed in commercial optical design software such as Zemax [Zemax LLC 2023] and in recent works [Shih and Renshaw 2024; Zhang et al. 2024], simplifies the diffractive element as local gratings and introduces an approximation for diffraction simulation [Fischer et al. 2000]. However, both of these simplified optical models fail to accurately simulate aberrations and diffraction simultaneously, and they rely on strong assumptions to function. It may be possible to exploit the generalized pupil function to include aberrations of the refractive lenses using Zernike polynomials [Goodman 2005; Wyant and Creath 1992]. However, for each field of view and each wavelength, the Zernike coefficients must be calculated individually. In addition, full inverse design with pupil functions is not possible since there is no mapping back from pupil functions to lens geometry.

In this paper, we propose a differentiable ray-tracing and wave-propagation (ray-wave) model for accurate simulation of hybrid lens systems. The ray-wave model can simulate both refractive optical aberrations and diffractive phase modulation in a differentiable manner, enabling end-to-end co-design of refractive lenses, DOEs, and neural networks. Specifically, we study a hybrid lens configuration where the DOE is placed between the refractive lenses and the image sensor. The proposed ray-wave model first performs coherent ray tracing [Chen et al. 2021; Mout et al. 2018] to accurately simulate the aberrated amplitude and phase map, followed by scalar Rayleigh-Sommerfeld diffraction [Goodman 2005] to incorporate diffractive phase modulation. Free-space propagation to the image plane is then performed for point spread function (PSF) calculation. In contrast to existing imaging models, our ray-wave model combines the advantages of both ray tracing and wave optics, providing

accurate simulation for optical aberrations and diffractive phase modulation.

First, we validate the proposed ray-wave model by comparing the PSF simulation results with both theoretical predictions and the ray tracing model. The ray-wave model provides accurate simulation results and is capable of handling discontinuous diffractive phase maps, whereas the ray tracing model fails to accurately capture real diffractive phenomena. Next, we design a compound hybrid refractive-diffractive lens integrated with an image reconstruction network for computational imaging. We then compare its performance with both a paraxial achromat design and a design generated using Zemax optical software. The proposed end-to-end hybrid lens design demonstrates superior image reconstruction quality compared to existing methods.

Furthermore, we demonstrate a hybrid aspherical-DOE lens prototype (Fig. 1) featuring a large field-of-view (FoV), compact form factor, and high image quality. To validate the effectiveness of our model, we investigate two real-world applications. First, we perform an end-to-end design of a DOE for computational aberration correction. Experimental results show that our hybrid lens successfully mitigates optical aberrations of the refractive component, producing high-quality images. Second, we design a DOE for large FoV extended depth-of-field (EDoF) imaging. Unlike existing DOE designs that idealize the refractive lens as a thin phase plate, our design explicitly accounts for lens aberrations. As a result, our DOE design significantly improves the reconstructed image quality, especially in off-axis regions where aberrations are more pronounced.

The main contributions of this paper can be summarized as:

- We present a differentiable ray-wave model for hybrid refractive-diffractive lenses. The proposed model can accurately simulate both optical aberrations and diffractive phase modulation. It facilitates end-to-end optimization of refractive lenses, DOEs, and image reconstruction networks.
- We demonstrate a hybrid aspherical-DOE lens prototype. Two real-world applications are investigated to validate the effectiveness of our model: computational aberration correction and large FoV EDoF imaging. Our hybrid lens achieves high-quality imaging performance with a large FoV and compact form factor.

2 Related Works

2.1 End-to-End Optical Design

Classical optical design [Fischer et al. 2000; Kidger 2001; Kingslake 2012] optimizes the lens system for the best optical performance independently of the image processing algorithm. However, with the advancement of deep learning and computer vision, the intermediate raw captures of cameras can be post-processed by neural networks, eliminating the need for them to be the design objective. Following this idea, end-to-end optical design [Sitzmann et al. 2018; Sun et al. 2020; Tseng et al. 2021a; Wetzstein et al. 2020; Wu et al. 2019], jointly optimizes the optical system (typically containing both refractive and diffractive elements) and the image processing network for the best final output. In an end-to-end optical design pipeline, a differentiable image formation model is employed to simulate the raw captures, which are then fed into the image processing

network. The optical system and the network can be jointly optimized using gradient backpropagation and deep learning algorithms. Recent research on end-to-end optical design has demonstrated outstanding performance in various applications, e.g., hyperspectral imaging [Baek et al. 2021; Dun et al. 2020; Jeon et al. 2019; Li et al. 2022], high-dynamic-range imaging [Metzler et al. 2020; Sun et al. 2020], seeing through obstructions [Shi et al. 2022], depth estimation [Chang and Wetzstein 2019; Ikoma et al. 2021; Wu et al. 2019], object detection [Côté et al. 2023; Tseng et al. 2021b], and compact optical systems [Chakravarthula et al. 2023; Tseng et al. 2021a].

An accurate and differentiable image formation model is crucial for end-to-end optical design. Existing end-to-end design works typically use paraxial optical models [Goodman 2005; Shi et al. 2022; Sitzmann et al. 2018; Sun et al. 2020], which idealize the refractive lens as a thin phase plate and neglect optical aberrations. However, this paraxial approximation is inaccurate and cannot optimize the refractive lens, limiting the designed optical systems to small fields of view and low aberration performance. Differentiable ray tracing [Côté et al. 2023; Sun et al. 2021; Wang et al. 2022; Yang et al. 2023a] is another widely used approach for the simulation of refractive lenses and can be extended to simulate diffractive surfaces by approximating them as local gratings [Fischer et al. 2000]. Ray tracing for diffractive surfaces has been employed in both commercial optical design software, such as Zemax [Zemax LLC 2023], and recent research works [Shih and Renshaw 2024; Zhang et al. 2024; Zhu et al. 2023]. However, while ray tracing can approximate the light propagation direction after diffractive surfaces, it cannot accurately represent real diffraction phenomena. Additionally, it requires the phase map of the diffractive surface to change slowly, whereas in end-to-end optical design, the DOE often has rapidly changing phase patterns. In this work, we propose a hybrid ray tracing and wave propagation (ray-wave) model for the accurate simulation of optical aberrations and phase modulation in hybrid refractive-diffractive lenses. This model combines ray tracing for aberration simulation with wave propagation for diffraction simulation.

2.2 Hybrid Refractive-Diffractive Lens

Hybrid refractive-diffractive lenses [Flores et al. 2004; Stone and George 1988; Wang et al. 2016] have showcased the great potential for next-generation optical imaging systems with a compact form factor, powerful information encoding capability, and outstanding light efficiency. Existing works on hybrid lenses can be summarized into two groups. In classical optical design, the DOE is primarily utilized for chromatic aberration correction with reduced physical size, given the reversed direction of dispersion in diffractive and refractive optics [Chen et al. 2018; Flores et al. 2004; Stone and George 1988; Wang et al. 2016]. For instance, Canon has revealed several patents for compact hybrid refractive-diffractive lenses [Minami and Yamada 2011]. Although the simulation and design process of these products are not publicly known, they use slowly varying phase patterns for which ray-tracing approaches work well. In more recent end-to-end optical design works [Shi et al. 2022; Sun et al. 2020; Wu et al. 2019], where the DOE is typically designed with more complex phase patterns, sometimes even discontinuous with

pixel-wise optimization, the ray-tracing models fail. Although existing end-to-end optical design works have demonstrated remarkable performance in numerous applications, the idealization of the DOE as a thin phase plate under the paraxial approximation is inaccurate. This limitation restricts the designed DOE to function only for low aberrations and small FoV when combined with refractive lenses. Additionally, the paraxial optics model can not optimize the refractive lens. So far, there is no existing work that can accurately simulate and optimize refractive lenses with complex and discontinuous diffractive surfaces.

3 Methods

3.1 Differentiable Ray-Wave Imaging Model

Considering a hybrid refractive-diffractive lens with the DOE at the last surface, where diffraction mainly occurs on the DOE surface (see e.g., Fig. 2). We propose a differentiable ray-wave method to accurately model the optical aberrations and diffractive phase modulation. The ray-wave model consists of two parts: coherent ray tracing to calculate the aberrated complex wave field at the DOE plane, followed by DOE phase modulation and free-space wave propagation to the sensor plane, where the PSF is calculated.

3.1.1 Coherent Ray Tracing. Monte Carlo ray tracing [Lafontaine 1996; Li et al. 2018; Wang et al. 2022] is a widely used method to simulate light propagation by modeling optical light as a group of rays. In the coherent ray tracing stage, we sequentially calculate the ray intersection \mathcal{I} and refraction \mathcal{R} at each lens surface S for each ray, recording the position \mathbf{o} , direction \mathbf{d} , and phase ϕ . The ray tracing process through a refractive lens can be represented as:

$$\begin{cases} \mathcal{I}_n(S_n) : (\mathbf{o}^{n-1}, \mathbf{d}^{n-1}, \phi^{n-1}, \lambda) \mapsto (\mathbf{o}^n, \mathbf{d}^{n-1}, \phi^n, \lambda), \\ \mathcal{R}_n(S_n) : (\mathbf{o}^n, \mathbf{d}^{n-1}, \phi^n, \lambda) \mapsto (\mathbf{o}^n, \mathbf{d}^n, \phi^n, \lambda), \\ (\mathbf{o}, \mathbf{d}, \phi, \lambda) = (\mathcal{R}_N \mathcal{I}_N \dots \mathcal{R}_1 \mathcal{I}_1)(\mathbf{o}^0, \mathbf{d}^0, \phi^0, \lambda), \end{cases} \quad (1)$$

where λ is the wavelength of the light, and N is the number of lens surfaces. The phase change ϕ is calculated by the optical path difference. We assume the point light source is coherent, and all rays have the same initial phase. The intersection \mathcal{I} can be solved by Newton's method and the refraction \mathcal{R} can be solved by Snell's law, which have been described in detail in existing works [Chen et al. 2021; Wang et al. 2022; Yang et al. 2023a]. Double-precision arithmetic is used to address the precision problem during the phase calculation, as the wavelength of optical rays is much smaller than the physical size of the optical lens.

After tracing rays through the refractive lens and reaching the DOE surface, we can calculate the complex wave field by coherent superposition. The wave field before the DOE surface can be represented as

$$\mathbf{U}_{\text{DOE}^-} = \sum_{i=1}^{spp} u_i \exp(j\phi_i) \cos(\mathbf{d}_i, \mathbf{n}), \quad (2)$$

where u_i is the amplitude and ϕ_i is the phase of the i -th optical ray. \mathbf{n} is the normal vector of the DOE surface, and $\cos(\mathbf{d}_i, \mathbf{n})$ represents the obliquity factor in the Rayleigh-Sommerfeld theorem [Goodman 2005]. j is the imaginary unit, and spp is the number of optical

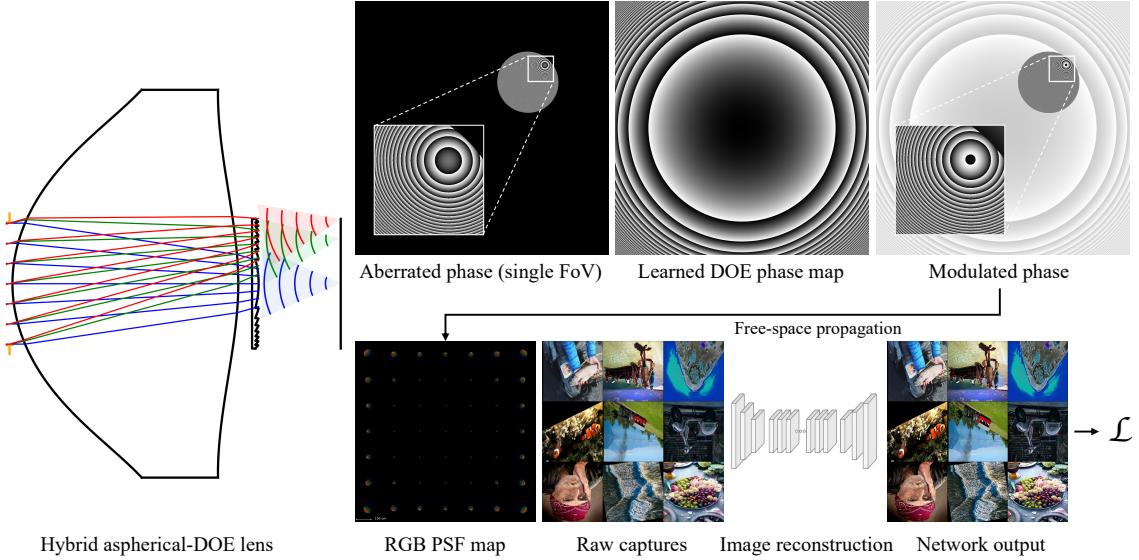


Fig. 2. A differentiable ray-wave model is proposed, enabling accurate simulation of both refractive aberrations and diffractive phase modulations. The ray-wave model initially calculates the complex wave field at the DOE plane using coherent ray-tracing. The aberrated wave field encodes the amplitude and phase aberrations introduced by the refractive lens. Subsequently, the wave field is modulated by the DOE phase profile and propagated to the sensor image plane for PSF calculation. The sensor-captured images are simulated with full FoV RGB PSFs and then fed into the downstream image reconstruction network. The pipeline is fully differentiable, enabling gradient backpropagation for end-to-end design of both the optics and the neural network.

rays sampled from each point source, which is set to 10^6 in our experiments.

Energy decay due to Fresnel transmission [Born and Wolf 2013] is often ignored at the geometric lens design stage, and dealt with in a separate design step for optical coatings [Baumeister 2004]. In order to allow direct comparisons with existing systems we also ignore the energy loss along a ray; therefore, the amplitude u_i of each ray is identical and set to 1. The complex amplitude of each optical ray is assigned to its surrounding 4 pixels with weights determined by the distance to each pixel point. This inverse bilinear interpolation enables gradient calculation when converting a group of rays to a wave field and also reduces the aliasing problem caused by sub-pixel phase shifts [Mout et al. 2018]. The wave field U_{DOE^-} records both the amplitude and phase aberrations introduced by the imperfections of the refractive lens. An example of the phase of a single-FoV wave field before the DOE surface is shown in Fig. 2, the majority of the phase is undefined (black) because light is converged to a small region by the refractive lens.

3.1.2 DOE Phase Modulation. The DOE surface introduces a phase modulation to the wave field passing through it, as illustrated in Fig. 2. Based on scalar diffraction theory [Goodman 2005] and using Kirchhoff boundary conditions [Born and Wolf 2013], the phase change introduced by the DOE can be expressed as

$$U_{\text{DOE}^+} = U_{\text{DOE}^-} \exp \left(j \frac{2\pi}{\lambda} (n_\lambda - 1) h(x, y) \right), \quad (3)$$

where $h(x, y)$ is the 2D height map of the DOE surface, and n_λ is the refractive index of the DOE substrate material at the wavelength λ .

For compatibility with Zemax, the DOE design is internally represented as the phase ϕ_0 at the nominal design wavelength λ_0 , and then remapped to the corresponding phases of other wavelengths during simulation. Given that

$$\phi_0 = \frac{2\pi}{\lambda_0} (n_0 - 1) h(x, y), \quad (4)$$

by substituting Eq. (4) into Eq. (3), the wave field after DOE modulation can be expressed as

$$U_{\text{DOE}^+} = U_{\text{DOE}^-} \exp \left(j \frac{n_\lambda - 1}{n_0 - 1} \frac{\lambda_0}{\lambda} \phi_0 \right). \quad (5)$$

In our experiments, λ_0 is set to $0.55 \mu\text{m}$ and ϕ_0 is the optimizable parameter. In Eq. (5), the DOE phase map ϕ_0 can be a discontinuous function to model the multi-level DOE features, eliminating the continuous condition in existing methods [Fischer et al. 2000; Zhu et al. 2023].

3.1.3 PSF Calculation. The modulated wave field then propagates to the sensor plane. This free-space propagation can be calculated using the angular spectrum method [Goodman 2005], i.e.,

$$U_{\text{Sensor}} = \mathcal{F}^{-1}(\mathcal{F}(U_{\text{DOE}^+})H), \quad (6)$$

where H is the transfer function, while \mathcal{F} and \mathcal{F}^{-1} respectively represent the Fourier transform and its inverse. The Nyquist sampling criterion [Goodman 2005; Mehrabkhani and Schneider 2017] requires a small sampling step of the wave field U_{DOE^+} for accurate calculation. Therefore, in Eq. (2), we choose a sampling resolution of an integer multiple (typically twice) of the DOE feature size and upsample the DOE phase map to the same resolution as the wave

field. Specifically, for the lens example in Fig. 2, the DOE is defined over a $3\text{ mm} \times 3\text{ mm}$ area with a $1\text{ }\mu\text{m}$ sampling resolution. Hence, we sample a $6,000 \times 6,000$ wave field in Eq. (2) and upsample the DOE phase map to the same resolution. During the free-space wave propagation, we also pad the wave field with zeros to twice the physical size and resolution to cover off-axis waves [Matsushima 2010], corresponding to a $12,000 \times 12,000$ resolution.

The PSF of the entire system is calculated by squaring the amplitude of the wave field at the sensor image plane, which can be represented as

$$\text{PSF} = |\mathbf{U}_{\text{Sensor}}|^2. \quad (7)$$

On the sensor image plane, the sampling resolution is determined by the camera sensor used, which is typically smaller than the wave field $\mathbf{U}_{\text{Sensor}}$. Therefore, an undersampling of the intensity distribution is performed. Subsequently, we crop the valid region, determined by the perspective relation of the refractive lens, to obtain a smaller PSF of the hybrid lens.

We encourage the readers to refer to existing works [Born and Wolf 2013; Chen et al. 2021; Goodman 2005; Mout et al. 2018; Wang et al. 2022; Yang et al. 2023a] for more details about coherent ray tracing and wave propagation methods. Our experimental code will also be released to assist in understanding the proposed method.

3.2 Mixed Precision End-to-End Optical Design

A mixed-precision training method is proposed to bridge the double-precision optical simulation with single-precision network training. Specifically, we observe that the precision challenge primarily arises from phase calculations during coherent ray tracing and wave propagation, while the intensity PSF is less sensitive to precision. To address this, we introduce a differentiable PyTorch autograd function that converts the double-precision PSF to single-precision after Eq. (7) for image simulation and network training. In the backpropagation phase, the single-precision PSF gradients from the network are converted back to double-precision and continue backpropagating through the optics.

We build the differentiable ray-wave model on top of an open-source ray tracer, DeepLens [Wang et al. 2022; Yang et al. 2023a], using the PyTorch framework. The memory consumption of the proposed differentiable ray-wave model is high due to million-scale ray tracing, high-resolution wave propagation, and downstream network processing. Specifically, for the hybrid aspherical-DOE lens (shown in Fig. 2) with given experimental settings (10^6 rays, $6,000 \times 6,000$ field), the single FoV RGB PSF calculation and backpropagation require approximately 35 GB of GPU memory, which gives a theoretical lower bound of GPU memory requirement. The memory consumption for multi-FoV PSF calculation and end-to-end lens design can be reduced by several strategies, such as multi-GPU parallelization, gradient checkpointing, patch backpropagation, and adjoint rendering [Nimier-David et al. 2020; Teh et al. 2022; Vicini et al. 2021; Yang et al. 2023a]. In our end-to-end hybrid lens design experiments, we calculate 10×10 RGB PSFs for accurate simulation of the hybrid lens system. We choose three wavelengths to represent the broadband for each color channel and randomly select one in each iteration to calculate the PSF during the training. The PSFs are

then convolved with the input image batch for output image simulation and downstream network training, which has been described in detail in existing end-to-end lens design works [Côté et al. 2023; Sitzmann et al. 2018; Tseng et al. 2021a]. See more experimental details in the following sections and Supplementary Materials.

4 Model Accuracy

The ray-tracing and wave propagation described above are individually well understood and validated, however, their combination to model hybrid systems is new and must be evaluated. Unfortunately, there is no “ground truth” reference for complex hybrid systems that could be compared against. We therefore turn to a simplified optical system that can be simulated with our hybrid model, but also entirely with scalar diffraction theory as a ground truth solution. The optical system (Fig. 3) consists of an ideal (i.e., un-aberrated) thin lens and a DOE. The thin lens model can be easily simulated with both ray-tracing and scalar diffraction theory [Goodman 2005], allowing us to compare our hybrid model to the scalar diffraction solution as well as a hybrid simulation provided by the commercial Zemax lens design package. Zemax uses a ray-tracing model with an additional ray-bending term based on the grating equation to approximate diffraction [Yu et al. 2011; Zemax LLC 2023].

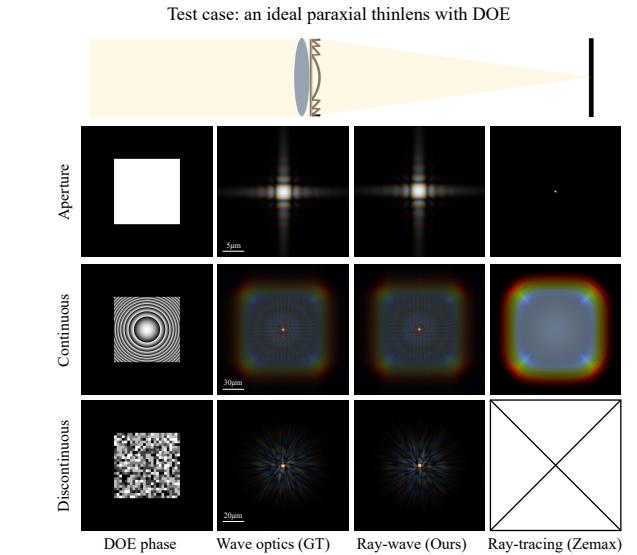


Fig. 3. PSF simulation of different optical models. An ideal paraxial thin lens with a DOE is used for testing. For this specific case, wave optics gives accurate simulation results [Goodman 2005], therefore serving as the ground truth. Different DOE phase maps are studied, our ray-wave model gives identical results with the ground truth, while the ray tracing model fails to address the real diffractive phenomena and can not function for discontinuous phase maps.

The specifications of the test are as follows. The lens consists of a thin lens ($f = 100\text{ mm}$) and a DOE attached to it. Three different DOE phase maps are studied: (1) a constant phase with a square aperture to simulate regular aperture diffraction; (2) a continuous phase map to simulate an ideal Kinoform diffractive surface [Jordan et al. 1970];

and (3) a discontinuous phase map to simulate both the multi-level fabrication process and discontinuous DOEs in existing works [Shi et al. 2022; Sun et al. 2020]. The simulation results indicate that our ray-wave model yields virtually identical results to the ground truth, while the ray tracing model fails. Specifically, the edge diffraction effect of the aperture mask is entirely ignored by the ray tracing model, and the central bright spot of the continuous phase map is caused by the interference of the diffracted light, which is also not captured by the ray tracing model. Moreover, for multi-level discontinuous phase maps, the ray tracing model cannot function because it requires a continuous condition to calculate the gradient of the local phase [Fischer et al. 2000]. More PSF simulation results are presented in the Supplementary Material.

A comprehensive comparison with other existing simulation methods for refraction and diffraction simulation is presented in Table 1. Specifically, we primarily consider imaging models that are or can be easily designed to be differentiable. Chen et al. [Chen et al. 2021] proposed a ray tracing model to simulate the aperture diffraction in a refractive lens; however, this exit-pupil diffraction method cannot function for diffractive optical elements. Zhu et al. [Zhu et al. 2023] introduced a wave-ray model, which converts the wavefront after phase modulation into a group of optical rays for ray tracing. Similar to Zemax [Zemax LLC 2023], this model also relies on the gradient calculation of the phase map [Yu et al. 2011], making it inaccurate and unable to function for binary and discontinuous phase maps. We compare the performance of these methods in the Supplementary Material.

Table 1. Comparison of different hybrid refractive-diffractive lens simulation models. A detailed explanation is provided in the supplementary material.

	Paraxial optics	Zemax [2023]	Chen et al. [2021]	Zhu et al. [2023]	Ours
Optical model	Wave	Ray	Ray	Wave-ray	Ray-wave
Accuracy	✗	✗	✓	✗	✓
Optical aberration	✗	✓	✓	✓	✓
Edge diffraction	✓	✗	✓	✓	✓
Phase modulation	✓	✓	✗	✓	✓
Discontinuous phase	✓	✗	✗	✗	✓
Differentiable	✓	✓	✗	✓	✓
End-to-end design	✓	✗	✗	✓	✓

5 End-to-end hybrid lens design

We first conduct an end-to-end hybrid lens design by simulation. A compound refractive lens, a DOE, and an image reconstruction network are jointly optimized. The refractive lens has a total length of 75 mm, a focal length of 47 mm, a sensor size of 8 mm × 8 mm, a diagonal FoV of 14° at F/4. The refractive lens exhibits strong chromatic aberrations, being a good candidate for DOE aberration correction. The DOE is placed between the refractive lens and the image sensor, with a physical size of 8 mm × 8 mm. The designed DOE phase map ϕ_0 is parameterized as $\phi_0 = \sum_{i=1}^k \alpha_i \rho^{2i}$, where ρ is the normalized radial distance from the center of the DOE, and α_i is the i -th even-polynomial coefficient. In our experiments, we set $k = 4$. Both the aberrated wave field U_{DOE} and the DOE phase map have a resolution of 5,000 × 5,000. During the end-to-end lens design process, the refractive lens, DOE, and an imaging

reconstruction network are simultaneously optimized to achieve the best image output. The sensor has a resolution of 2,000 × 2,000, corresponding to a pixel pitch of 4 μm. We sample 10 × 10 RGB PSFs for full-resolution image simulation. For image reconstruction, we adopt NAFNet [Chen et al. 2022], which demonstrates outstanding image deblurring performance and fast convergence. The image reconstruction loss is used for end-to-end lens design to learn the optimal optics-network joint system, which can be represented as:

$$\mathcal{L} = \mathcal{L}_{\text{pixel}}(\mathcal{N}(\text{PSF} * \mathbf{I}), \mathbf{I}) + \alpha \mathcal{L}_{\text{percep}}(\mathcal{N}(\text{PSF} * \mathbf{I}), \mathbf{I}), \quad (8)$$

where \mathbf{I} is the input object image, which is also used as the ground truth since we aim to optimize for the best imaging performance. \mathcal{N} represents the reconstruction network, $\mathcal{L}_{\text{pixel}}$ is the pixel-wise loss, $\mathcal{L}_{\text{percep}}$ is the perceptual loss, and α is the weight of the perceptual loss. In our experiments, we use mean squared error loss as the pixel-wise loss and VGG loss [Johnson et al. 2016] as the perceptual loss with $\alpha = 0.1$. DIV2K dataset [Agustsson and Timofte 2017] is used for training and testing. The end-to-end training runs for 50 epochs, then we fix the hybrid lens and fine-tune the image reconstruction network with sensor noise for 50 epochs.

Paraxial wave optics and ray tracing models are used for comparison: (1) The paraxial wave model (Fig. 4) represents the compound refractive lens as a phase plate. We pre-calculate the on-axis chromatic aberrations of the refractive lens and design the DOE for achromatic purposes [Chen et al. 2018; Wang et al. 2016]. Specifically, we use the Fresnel phase DOE and analytically solve it to reduce the chromatic aberration at different wavelengths. (2) For the ray tracing model (Fig. 4), we load the hybrid lens into Zemax and jointly optimize the refractive lens with the DOE, using the “Binary2” surface with the same polynomial order as ours, operated by an experienced optical designer. We then load the optimized hybrid lens into our code and continue optimizing the optics to close the gap between our simulation and Zemax, with RMS spot size used as the optimization objective.

Three final designs are presented in Fig. 4. The orange dashed lines indicate the original refractive lens, and the DOE phase map is shown on the right. The log-scale PSFs at 4 FoVs from on-axis to full-FoV are calculated using our proposed model, which has been proven to yield accurate simulations in the previous section. The paraxial wave optics model fails to consider the off-axis aberrations, thus leaving significant chromatic aberrations at large FoVs. For both ray optics and our ray-wave model, the PSFs at all FoVs are optimized to be small. The RMS spot size is calculated with the ray tracing model and listed in Table 2. The Zemax design performs the best in terms of RMS spot size, which is the loss it was designed for.

We further evaluate the three lenses based on high-level imaging quality. For each lens, we simulate sensor-captured images using our proposed ray-wave model and train an image reconstruction network for each different lens to achieve the best output quality. In this network fine-tuning stage, 40×40 PSFs are used for accurate image simulation, and Gaussian noise is added to simulate sensor noise. As presented in Table 2, PSNR, SSIM, and 1-LPIPS metrics are calculated for the evaluation of both simulated images (“raw”) and reconstructed images (“rec”). Our end-to-end designed hybrid lens using the ray-wave model successfully outperforms the other two designs in terms of image quality, since we directly optimize the final

Table 2. Performance comparison of different hybrid lens designs by simulation. The PSNR, SSIM, and 1-LPIPS matrix are calculated for both simulated (“raw”) and reconstructed (“rec”) images.

Lens design	RMS radius (on-axis/off-axis/avg) ↓	PSNR/SSIM/1-LPIPS (raw) ↑	PSNR/SSIM/1-LPIPS (rec) ↑
Paraxial wave optics	16.1/55.0/32.0	17.7/0.520/0.443	25.041/0.688/0.569
Ray tracing (Zemax)	7.6/19.6/10.1	24.8/0.728/0.706	36.462/0.963/0.924
Ray-wave model (ours)	10.7/28.4/15.4	27.1/0.795/0.713	39.9/0.982/0.963

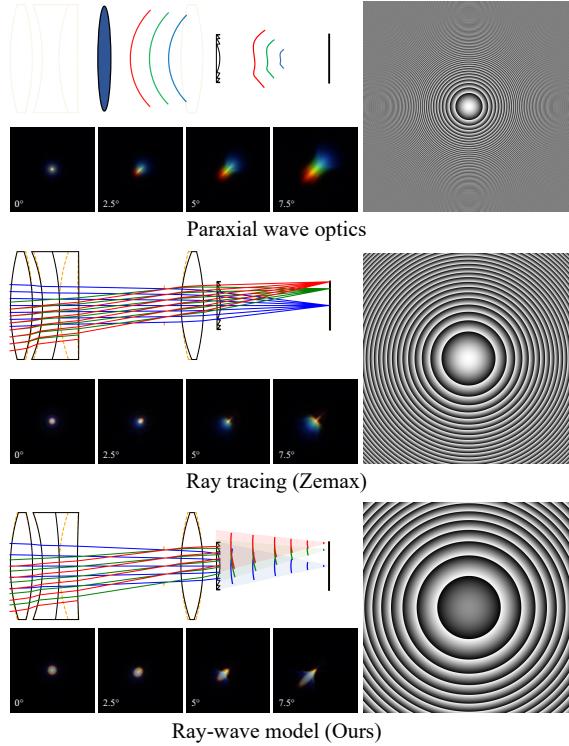


Fig. 4. Hybrid lens designs using different optical models and optimization methods. Top: achromatic DOE design using the paraxial optical model, minimizing paraxial chromatic aberration. Middle: hybrid lens design using a ray-tracing model, optimizing the RMS spot size in Zemax. Bottom: hybrid lens design using the proposed ray-wave model, optimizing the final image quality. PSFs at different FoVs are calculated with the proposed ray-wave model, shown in the log scale.

image, whereas the other two models cannot perform end-to-end design. Example “raw” and “rec” images are provided in Fig. 5, with zoomed image patches. More results and evaluations are provided in the Supplementary Material.

6 Hybrid aspheric-DOE lens prototype

To evaluate the proposed method with real-world experiments, we designed and built a large FoV compact aspherical-DOE hybrid lens prototype. Since we lack the fabrication capabilities for refractive lenses, we utilize an off-the-shelf aspherical lens and use our hybrid framework to only optimize the DOE, taking into account the real aberrations of the refractive lens. An aspherical lens with known surface data (Optolife Optics, China.) with a focal length

of 7.5 mm is chosen for evaluation, as shown in Fig. 1. The image sensor (OmniVision OV2710) has a pixel pitch of $3 \mu\text{m} \times 3 \mu\text{m}$, and the center $1,000 \times 1,000$ imaging region is used for the experiments. The diagonal FoV of the refractive lens is 35° at an effective F/2.2, leading to significant off-axis optical aberrations. Various fabrication techniques can be employed to fabricate DOEs, including lithography+etching [Fu et al. 2022; Shi et al. 2022; Zheng et al. 2023a], lithography+deposition [Amata et al. 2023], and lithography+nanoimprint [Fu et al. 2021]. The DOE has a feature size of $1 \mu\text{m} \times 1 \mu\text{m}$ and a physical size of $3 \text{ mm} \times 3 \text{ mm}$. To best preserve the edge features and micro-profiles, we choose to fabricate the DOE with high spatial resolution in a 16-level structure. The lens and DOE are assembled with a 3D-printed mount for the final prototype.

Two applications are demonstrated to evaluate the performance of our hybrid aspherical-DOE lens: (1) computational aberration correction and (2) large FoV and EDoF imaging. Specifically, existing diffractive EDoF works [Li et al. 2023; Pinilla et al. 2022; Seong et al. 2023] idealize the refractive lens as a paraxial thin lens, ignoring aberrations and functioning only for a small FoV. Our proposed model accurately simulates optical aberrations, enabling large FoV EDoF imaging. Moreover, since we do not optimize the refractive lens part, the aberrated wavefield at the DOE surface can be pre-calculated and stored, making the memory and time consumption for the single FoV PSF simulation the same as the paraxial optical model. All experimental settings are kept consistent (see Supplementary Material for more details).

6.1 Computational Aberration Correction

An aberration correction DOE is end-to-end designed with the reconstruction network. The same DOE parameterization is adopted for computational aberration correction. The optimized DOE phase map and a microscope image of the fabricated DOE are presented in Fig. 6. For comparison, the original aspherical lens with a blank DOE (to maintain the same aperture) is selected as the baseline experiment. Quantitative evaluation scores are measured on the simulated dataset, as presented in Table 3. In Fig. 7, both the raw captured (“raw”) and reconstructed (“rec”) images in the real world are shown for the hybrid lens and the refractive lens. The results demonstrate that the hybrid lens can successfully correct the refractive aberrations and achieve better image quality, especially for the off-axis image regions where aberrations are more significant. More results and experimental details are provided in the Supplementary Material. Moreover, inserting the DOE into an existing refractive lens at the back does not increase the form factor, showing great potential for cameras with highly constrained physical size, such as cellphone cameras.

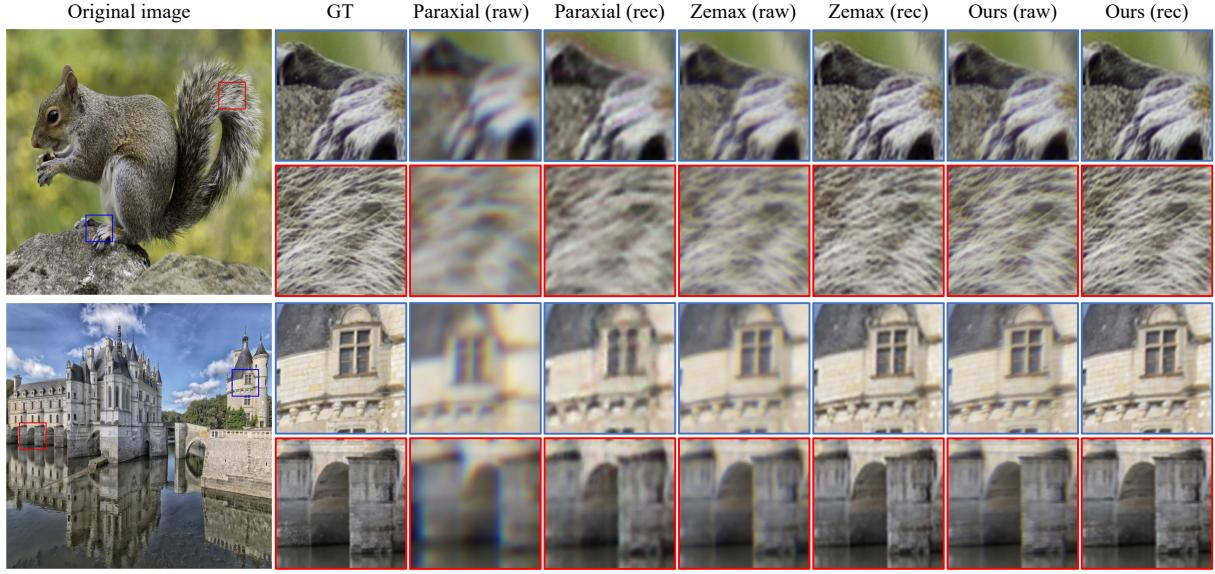


Fig. 5. Comparison of image quality between different hybrid lens design approaches. The paraxial wave optics model fails to correct aberrations across the full FoV, resulting in degraded image quality in both raw simulations and network reconstructions. The ray tracing method in Zemax corrects aberrations but uses the RMS spot size as the objective, precluding end-to-end lens design. Our ray-wave model accurately simulates full-FoV PSFs and is fully differentiable, enabling end-to-end co-design of the hybrid lens and the network, thus achieving the best reconstruction quality. Raw simulated images and corresponding reconstructed images are shown for each design approach, with zoomed-in patches highlighting the differences in image quality across the FoV.

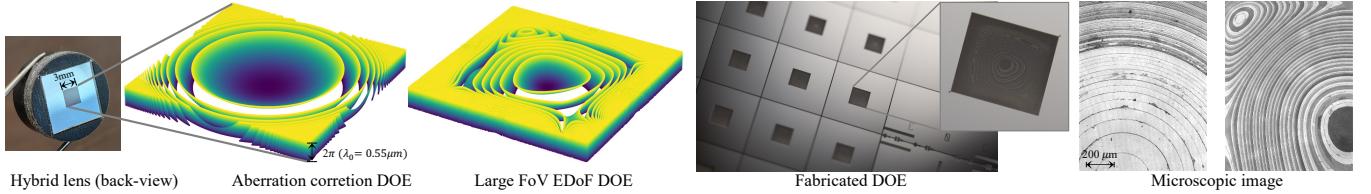


Fig. 6. Designed DOE phase map (left) and fabricated samples (right) for computational aberration correction and large FoV EDoF imaging. Microscopic images are taken by a Nikon objective, 5X/0.13 TI WD 9.3 427041.

Table 3. Performance comparison between refractive and hybrid lens designs on the simulated dataset.

Lens	PSNR/SSIM/1-LPIPS (raw)	PSNR/SSIM/1-LPIPS (rec)
Refractive lens	24.4/0.702/0.677	28.8/0.835/0.746
Hybrid lens (ours)	24.7/0.733/0.707	32.5/0.931/0.880

6.2 Aberration-Aware Large Field-of-View Extended-Depth-of-Field Imaging

An EDoF DOE is end-to-end designed with the proposed ray-wave model to image clearly from 20 cm to 10 m with a diagonal FOV of 35°, named “large FoV DOE”. The DOE is parameterized as $\phi_0 = \sum_{i=2}^k \beta_i \rho^i$. We set $k = 7$ in our experiments. The odd polynomials provide the EDoF capability, while the even polynomials correct optical aberrations. With accurate simulation of optical aberrations, the designed DOE can find the best phase map in the presence of optical aberrations, thus improving image quality and practical

performance in the real world. The optimized DOE phase map and a microscope image of the fabricated DOE are in Fig. 6.

Table 4. Performance comparison of EDoF DOE designed with different imaging models.

DOE	depth	PSNR/SSIM/1-LPIPS (raw)	PSNR/SSIM/1-LPIPS (rec)
Baseline	20 cm	18.7/0.497/0.531	16.1/0.341/0.509
	30 cm	19.0/0.539/0.473	18.7/0.520/0.437
	10 m	18.9/0.523/0.517	18.6/0.502/0.493
Ours	20 cm	21.5/0.575/0.587	27.5/0.821/0.782
	30 cm	22.4/0.659/0.635	28.9/0.869/0.842
	10 m	21.7/0.598/0.524	27.4/0.818/0.787

For comparison, a DOE is end-to-end designed using the paraxial optical model, referred to as the “paraxial DOE,” as presented in existing works [Pinilla et al. 2022; Sitzmann et al. 2018]. Subsequently, we evaluate the simulated images of the two DOEs using the proposed ray-wave model and reconstruct the simulated images

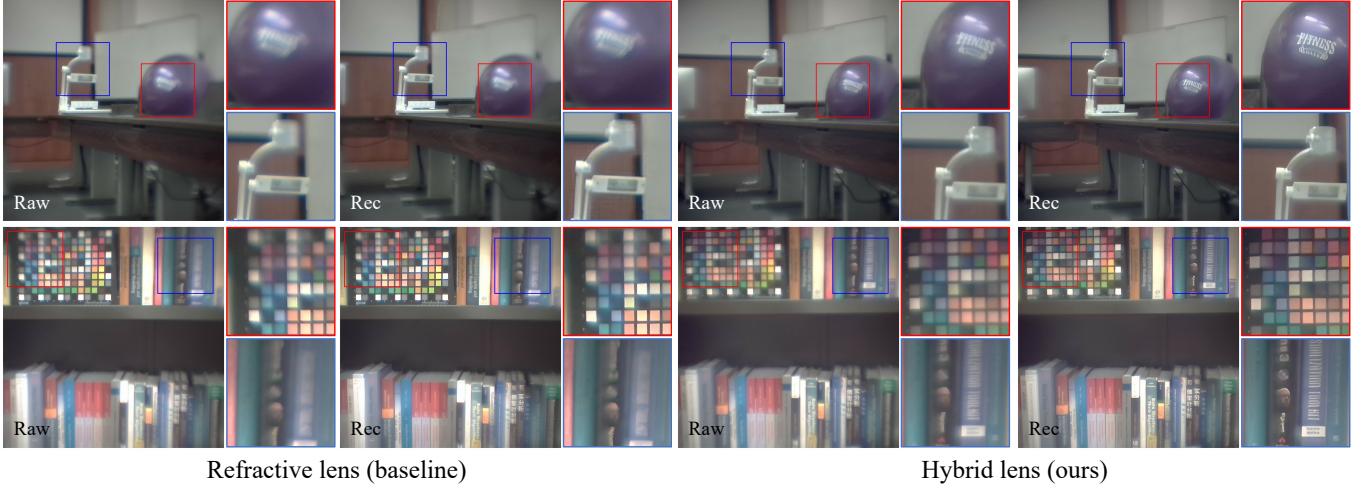


Fig. 7. Real-world image quality comparison between the refractive lens (baseline) and the hybrid lens (ours). Our designed and fabricated DOE corrects the aberrations of the refractive lens and achieves better image quality, especially for the off-axis image regions where aberrations are more significant.

with their respective reconstruction networks. Quantitative results measured on the simulated dataset are presented in Table 4, with qualitative results on the simulated dataset shown in Fig. 8. We then fabricated both DOEs for real-world experiments, and the results are depicted in Fig. 8. The results demonstrate that the hybrid lens can image clearly over a large depth-of-field and wide field-of-view (FoV). Notably, for off-axis image regions, the DOE designed with our ray-wave model can maintain excellent image quality, while the DOE designed with the paraxial optical model fails to image clearly. Additional results and experimental details are provided in the supplementary material.

7 Conclusion and Discussion

In this paper, we propose a differentiable ray-wave model for hybrid refractive-diffractive optical systems. The proposed model can accurately simulate both optical aberrations and diffractive phase modulation while also enabling gradient backpropagation for end-to-end design of hybrid lenses and neural networks. We validate the simulation accuracy by employing the ray-wave model to jointly design a hybrid lens consisting of a compound refractive lens, a DOE, and an image reconstruction network. We compare our results with existing methods and commercial software, such as Zemax, demonstrating the effectiveness of our approach. Furthermore, we demonstrate the practical applicability of our proposed model through an aspherical-DOE lens prototype for aberration correction and large field-of-view EDof imaging. Both simulations and real experiments show that the proposed prototype achieves high-quality imaging performance, outperforming existing methods.

The proposed model paves the way towards high-quality imaging systems with challenging optical requirements, such as compact cellphone lenses [Reshef et al. 2021; Williams et al. 2023] with large aperture and wide field of view, clip-in diffractive filters for commercial interchangeable-lens cameras [Totori 2023], and augmented reality wave-guide [Gopakumar et al. 2024; Tseng et al. 2024] for

near-display applications, to name just a few. Placing the DOE away from the Fourier plane is beneficial to shrink its size while maintaining a large entrance pupil of the system. In addition, optimizing the DOE position to the sensor adds further flexibility to the optimization. Last but not least, one limitation of placing the DOE as the last optical component in the hybrid system is owing to the current difficulties in reverse conversion from wave to rays, requiring further improvements in the model.

References

- Eirikur Agustsson and Radu Timofte. 2017. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 126–135.
- Hadi Amata, Qiang Fu, and Wolfgang Heidrich. 2023. Additive fabrication of SiO₂-based micro-optics with lag-free depth and reduced roughness. *Opt. Express* 31, 25 (2023), 41533–41545.
- Seung-Hwan Baek, Hayato Ikoma, Daniel S Jeon, Yuqi Li, Wolfgang Heidrich, Gordon Wetzstein, and Min H Kim. 2021. Single-shot hyperspectral-depth imaging with learned diffractive optics. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2651–2660.
- Philip Baumeister. 2004. *Optical coating technology*. Vol. 137. SPIE press.
- Max Born and Emil Wolf. 2013. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier.
- Praneeth Chakravarthula, Jipeng Sun, Xiao Li, Chenyang Lei, Gene Chou, Mario Bujelic, Johannes Froesch, Arka Majumdar, and Felix Heide. 2023. Thin On-Sensor Nanophotonic Array Cameras. *ACM Transactions on Graphics (TOG)* 42, 6 (2023), 1–18.
- Julie Chang and Gordon Wetzstein. 2019. Deep optics for monocular depth estimation and 3d object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 10193–10202.
- Hou-Tong Chen, Antoinette J Taylor, and Nanfang Yu. 2016. A review of metasurfaces: physics and applications. *Reports on progress in physics* 79, 7 (2016), 076401.
- Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. 2022. Simple baselines for image restoration. In *European conference on computer vision*. Springer, 17–33.
- Shiqi Chen, Huajun Feng, Dexin Pan, Zhihai Xu, Qi Li, and Yueteng Chen. 2021. Optical aberrations correction in postprocessing using imaging simulation. *ACM Transactions on Graphics (TOG)* 40, 5 (2021), 1–15.
- Wei Ting Chen, Alexander Y Zhu, Jared Sisler, Yao-Wei Huang, Keroles MA Yousef, Eric Lee, Cheng-Wei Qiu, and Federico Capasso. 2018. Broadband achromatic metasurface-refractive optics. *Nano letters* 18, 12 (2018), 7801–7808.
- Geoffroi Côté, Fahim Mannan, Simon Thibault, Jean-François Lalonde, and Felix Heide. 2023. The differentiable lens: Compound lens search over glass surfaces and materials for object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision*

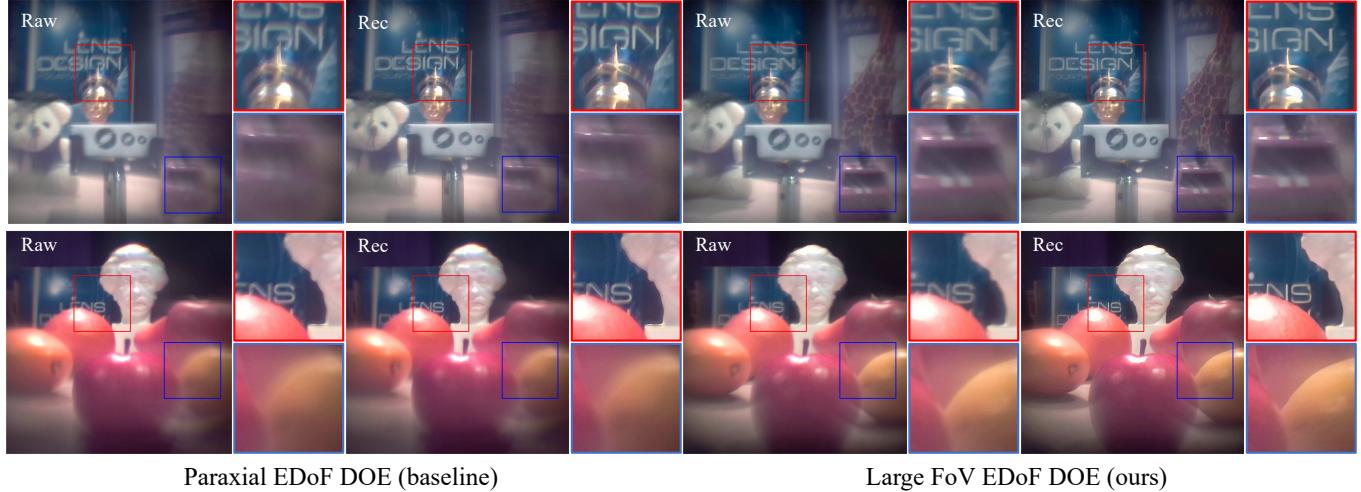


Fig. 8. Real-world image quality comparison of extended depth-of-field imaging between the DOEs design with the paraxial optical model (baseline) and the proposed ray-wave model (ours). Hybrid lens with our designed and fabricated DOE achieves better image quality over a large FoV and a large depth of field, particularly in the off-axis regions which are neglected during the design using the paraxial optical model.

- and Pattern Recognition. 20803–20812.
- Xiong Dun, Hayato Ikoma, Gordon Wetzstein, Zhanshan Wang, Xinbin Cheng, and Yifan Peng. 2020. Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging. *Optica* 7, 8 (2020), 913–922.
- Robert Edward Fischer, Biljana Tadic-Galeb, Paul R Yoder, Ranko Galeb, Bernard C Kress, Stephen C McClain, Tom Baur, Richard Plympton, Bob Wiederhold, and Bob Grant Alastair J. 2000. *Optical system design*. Vol. 599. Citeseer.
- Angel Flores, Michael R Wang, and Jame J Yang. 2004. Achromatic hybrid refractive-diffractive lens with extended depth of focus. *Applied optics* 43, 30 (2004), 5618–5630.
- Qiang Fu, Hadi Amata, and Wolfgang Heidrich. 2021. Etch-free additive lithographic fabrication methods for reflective and transmissive micro-optics. *Opt. Express* 29, 22 (2021), 36886–36899.
- Qiang Fu, Dong-Ming Yan, and Wolfgang Heidrich. 2022. Diffractive lensless imaging with optimized Voronoi-Fresnel phase. *Opt. Express* 30, 25 (2022), 45807–45823.
- Joseph W Goodman. 2005. *Introduction to Fourier optics*. Roberts and Company publishers.
- Manu Gopakumar, Gun-Yeal Lee, Suyeon Choi, Brian Chao, Yifan Peng, Jonghyun Kim, and Gordon Wetzstein. 2024. Full-colour 3D holographic augmented-reality displays with metasurface waveguides. *Nature* (2024), 1–7.
- Hayato Ikoma, Cindy M Nguyen, Christopher A Metzler, Yifan Peng, and Gordon Wetzstein. 2021. Depth from defocus with learned optics for imaging and occlusion-aware depth estimation. In *2021 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–12.
- Daniel S Jeon, Seung-Hwan Baek, Shinyoung Yi, Qiang Fu, Xiong Dun, Wolfgang Heidrich, and Min H Kim. 2019. Compact snapshot hyperspectral imaging with diffracted rotation. (2019).
- Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2016. 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*. Springer, 694–711.
- JA Jordan, PM Hirsch, LB Lesem, and DL Van Rooy. 1970. Kinoform lenses. *Applied Optics* 9, 8 (1970), 1883–1887.
- Michael J Kidger. 2001. *Fundamental optical design*. SPIE Optical Engineering Press.
- Rudolf Kingslake. 2012. *Lens design fundamentals*. Elsevier.
- Eric Lafontaine. 1996. Mathematical models and Monte Carlo algorithms for physically based rendering. *Department of Computer Science, Faculty of Engineering, Katholieke Universiteit Leuven* 20, 74–79 (1996), 4.
- Lingen Li, Lizhi Wang, Weitao Song, Lei Zhang, Zhiwei Xiong, and Hua Huang. 2022. Quantization-aware deep optics for diffractive snapshot hyperspectral imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19780–19789.
- Tzu-Mao Li, Miika Aittala, Frédo Durand, and Jaakko Lehtinen. 2018. Differentiable monte carlo ray tracing through edge sampling. *ACM Transactions on Graphics (TOG)* 37, 6 (2018), 1–11.
- Yuqi Li, Qiang Fu, and Wolfgang Heidrich. 2023. Extended depth-of-field projector using learned diffractive optics. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE, 449–459.
- Kyoji Matsushima. 2010. Shifted angular spectrum method for off-axis numerical propagation. *Optics Express* 18, 17 (2010), 18453–18463.
- Soheil Mehrabkhani and Thomas Schneider. 2017. Is the Rayleigh-Sommerfeld diffraction always an exact reference for high speed diffraction algorithms? *Optics express* 25, 24 (2017), 30229–30240.
- Christopher A Metzler, Hayato Ikoma, Yifan Peng, and Gordon Wetzstein. 2020. Deep optics for single-shot high-dynamic-range imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1375–1385.
- Kazuhiro Minami and Kazuhiro Yamada. 2011. Diffraction optical element and optical instrument. JP2011221510A.
- Marco Mout, Andreas Flesch, Michael Wick, Florian Bochiort, Joerg Petschelt, and Paul Urbach. 2018. Ray-based method for simulating cascaded diffraction in high-numerical-aperture systems. *JOSA A* 35, 8 (2018), 1356–1367.
- Merlin Nimier-David, Sébastien Speirer, Benoit Ruiz, and Wenzel Jakob. 2020. Radiative backpropagation: An adjoint method for lightning-fast differentiable rendering. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 146–1.
- Samuel Pinilla, Seyyed Reza Miri Rostami, Igor Shevkunov, Vladimir Katkovnik, and Karen Egiazarian. 2022. Hybrid diffractive optics design via hardware-in-the-loop methodology for achromatic extended-depth-of-field imaging. *Optics Express* 30, 18 (2022), 32633–32649.
- Orad Reshet, Michael P DelMastro, Katherine KM Bearne, Ali H Alhulaymi, Lambert Giner, Robert W Boyd, and Jeff S Lundeen. 2021. An optic to replace space and its application towards ultra-thin imaging systems. *Nature communications* 12, 1 (2021), 3512.
- Baekchun Seong, Woovin Kim, Younghun Kim, Kyung-A Hyun, Hyo-Il Jung, Jong-Seok Lee, Jeonghoon Yoo, and Chulmin Joo. 2023. E2E-BPF microscope: extended depth-of-field microscopy using learning-based implementation of binary phase filter and image deconvolution. *Light: Science & Applications* 12, 1 (2023), 269.
- Zheng Shi, Yuval Bahat, Seung-Hwan Baek, Qiang Fu, Hadi Amata, Xiao Li, Praneeth Chakravarthula, Wolfgang Heidrich, and Felix Heide. 2022. Seeing through obstructions with diffractive cloaking. *ACM Transactions on Graphics (TOG)* 41, 4 (2022), 1–15.
- Ko-Han Shih and C Kyle Renshaw. 2024. Hybrid meta/refractive lens design with an inverse design using physical optics. *Applied Optics* 63, 15 (2024), 4032–4043.
- Vincent Sitzmann, Steven Diamond, Yifan Peng, Xiong Dun, Stephen Boyd, Wolfgang Heidrich, Felix Heide, and Gordon Wetzstein. 2018. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–13.
- Thomas Stone and Nicholas George. 1988. Hybrid diffractive-refractive lenses and achromats. *Applied optics* 27, 14 (1988), 2960–2971.
- Qilin Sun, Ethan Tseng, Qiang Fu, Wolfgang Heidrich, and Felix Heide. 2020. Learning rank-1 diffractive optics for single-shot high dynamic range imaging. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 1386–1396.
- Qilin Sun, Congli Wang, Fu Qiang, Dun Xiong, and Heidrich Wolfgang. 2021. End-to-end complex lens design with differentiable ray tracing. *ACM Trans. Graph* 40, 4

- (2021), 1–13.
- Arjun Teh, Matthew O'Toole, and Ioannis Gkioulekas. 2022. Adjoint nonlinear ray tracing. *ACM Transactions on Graphics (TOG)* 41, 4 (2022), 1–13.
- Yuki Totori. 2023. Image capturing device. JP2023128253A.
- Ethan Tseng, Shane Colburn, James Whitehead, Luocheng Huang, Seung-Hwan Baek, Arka Majumdar, and Felix Heide. 2021a. Neural nano-optics for high-quality thin lens imaging. *Nature communications* 12, 1 (2021), 6493.
- Ethan Tseng, Grace Kuo, Seung-Hwan Baek, Nathan Matsuda, Andrew Maimone, Florian Schiessl, Praneeth Chakravarthula, Qiang Fu, Wolfgang Heidrich, Douglas Lanman, et al. 2024. Neural étendue expander for ultra-wide-angle high-fidelity holographic display. *Nature communications* 15, 1 (2024), 2907.
- Ethan Tseng, Ali Mosteh, Fahim Mannan, Karl St-Arnaud, Avinash Sharma, Yifan Peng, Alexander Braun, Derek Nowrouzezahrai, Jean-Francois Lalonde, and Felix Heide. 2021b. Differentiable compound optics and processing pipeline optimization for end-to-end camera design. *ACM Transactions on Graphics (TOG)* 40, 2 (2021), 1–19.
- Delio Vicini, Sébastien Speierer, and Wenzel Jakob. 2021. Path replay backpropagation: Differentiating light paths using constant memory and linear time. *ACM Transactions on Graphics (TOG)* 40, 4 (2021), 1–14.
- Congli Wang, Ni Chen, and Wolfgang Heidrich. 2022. dO: A differentiable engine for deep lens design of computational imaging systems. *IEEE Transactions on Computational Imaging* 8 (2022), 905–916.
- Peng Wang, Nabil Mohammad, and Rajesh Menon. 2016. Chromatic-aberration-corrected diffractive lenses for ultra-broadband focusing. *Scientific reports* 6, 1 (2016), 21545.
- Gordon Wetzstein, Aydogan Ozcan, Sylvain Gigan, Shanhui Fan, Dirk Englund, Marin Soljačić, Cornelia Denz, David AB Miller, and Demetri Psaltis. 2020. Inference in artificial intelligence with deep optics and photonics. *Nature* 588, 7836 (2020), 39–47.
- George M Williams, Charles Dupuy, Jeremy Brown, Samuel Grimm, Hooman Akhavan, and J Paul Harmon. 2023. Three-dimensional gradient index microlens arrays for light-field and holographic imaging and displays. *Applied Optics* 62, 14 (2023), 3710–3723.
- Yicheng Wu, Vivek Boominathan, Huajin Chen, Aswin Sankaranarayanan, and Ashok Veeraraghavan. 2019. Phasacam3d—learning phase masks for passive single view depth estimation. In *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–12.
- James C Wyant and Katherine Creath. 1992. Basic wavefront aberration theory for optical metrology. *Applied optics and optical engineering* 11, part 2 (1992), 28–39.
- Xinge Yang, Qiang Fu, and Wolfgang Heidrich. 2023a. Curriculum learning for ab initio deep learned refractive optics. *arXiv preprint arXiv:2302.01089* (2023).
- Xinge Yang, Qiang Fu, Yunfeng Nie, and Wolfgang Heidrich. 2023b. Image Quality Is Not All You Want: Task-Driven Lens Design for Image Classification. *arXiv preprint arXiv:2305.17185* (2023).
- Nanfang Yu, Patrice Genevet, Mikhail A Kats, Francesco Aieta, Jean-Philippe Tetienne, Federico Capasso, and Zeno Gaburro. 2011. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *science* 334, 6054 (2011), 333–337.
- Zemax LLC. 2023. *Zemax OpticStudio User Manual*. Zemax LLC, Kirkland, WA, USA. Available from Zemax website: <https://www.zemax.com>.
- Qiangbo Zhang, Zeqing Yu, Mengguang Wang, Yiyang Liu, Changwei Zhang, Chang Wang, and Zhenrong Zheng. 2024. Centimeter-Scale Achromatic Hybrid Metalens Design: A New Paradigm Based on Differentiable Ray Tracing in the Visible Spectrum. *arXiv preprint arXiv:2404.03173* (2024).
- Cheng Zheng, Guangyuan Zhao, and Peter So. 2023b. Close the Design-to-Manufacturing Gap in Computational Optics with a Real2Sim Learned Two-Photon Neural Lithography Simulator. In *SIGGRAPH Asia 2023 Conference Papers*. 1–9.
- Yidan Zheng, Qiang Fu, Hadi Amata, Praneeth Chakravarthula, Felix Heide, and Wolfgang Heidrich. 2023a. Hexagonal diffractive optical elements. *Opt. Express* 31, 26 (2023), 43864–43876.
- Ziwei Zhu, Zhaocheng Liu, and Changxi Zheng. 2023. Metalens enhanced ray optics: an end-to-end wave-ray co-optimization framework. *Optics Express* 31, 16 (2023), 26054–26068.