

基于计算成像的端到端延展景深衍射元件的设计

吉佳瑞, 谢洪波, 杨磊*

天津大学精密仪器与光电子工程学院光电信息技术教育部重点实验室, 天津 300072

摘要 提出一种基于计算成像理论的端到端衍射元件设计方法, 通过全局性优化方案将光学设计和图像复原作为整体, 从而降低前端光学系统的成像质量要求, 并利用图像复原算法去除残余像差以简化系统。所提设计方法涵盖光场传播、探测器去噪和图像后处理等关键环节的模型建立与联合优化。该设计方案可用于景深延展的轻薄型衍射元件的设计, 且所适用的大景深的简单光学系统具有较高的成像质量。

关键词 成像系统; 端到端; 景深延展; 图像复原; 衍射元件

中图分类号 O436

文献标志码 A

DOI: 10.3788/AOS231275

1 引言

学术界和工业界一直致力于实现小型化和轻量化的光学成像结构^[1-3]。然而, 传统光学设计为获得更多信息, 通常需要设计复杂结构或采用特殊材料和特殊面形等^[4-7], 导致结构的体积和制作成本显著增加。为解决这一问题, 计算成像技术引起了研究人员的兴趣^[8-10], 该技术基于全局性优化思想, 将光学设计和图像复原作为整体, 从而降低前端光学系统设计的成像质量要求, 并利用图像复原算法去除残余像差以简化系统。这种方法将光学设计的难度转移到算法上, 能够显著减少光学系统的复杂性, 同时保证成像质量。

2012 年, Schuler 等^[11]设计出仅包含一片球面镜头的单透镜成像系统, 并提出一种交替去除马赛克的非盲图像复原算法。为解决单透镜成像色差较大的问题, 2013 年 Heide 等^[12]提出一种基于交叉通道先验的图像复原算法, 且通过自制的单透镜相机获得较高质量的图像。此后, 包含光学系统与图像处理的协同设计方法受到了国内外研究者的关注^[13-15]。为了进一步弥合光学设计和算法开发之间的截断, 数据驱动的端到端 (E2E) 单透镜的优化模型被提出^[10, 16-19], 该方法凭借较高的设计灵活性和准确性展示了其在光学成像系统设计中的巨大潜力。

计算成像技术在景深延展衍射元件设计中发挥了关键作用^[20]。基于波前编码的景深延展技术通过衍射元件 (例如立方相位板) 操控光场的传输状态^[21-22], 并借助后端图像算法来获得清晰的全焦图像。值得注意的是, 虽然立方相位板方法包含了计算成像的思想, 但由于点扩散函数 (PSF) 的深度一致性是前端光学设计

过程的优化目标, 也是光学系统的强制评价标准, 因此该方法并未建立全局优化框架, 而是将实现光学系统和图像处理算法的局部最优作为目标。端到端的衍射元件大景深模型利用现代深度学习框架来集成光学元件与图像恢复算法, 以实现成像系统的全局最优, 显著降低设计复杂度^[16]。已有的端到端系统大多使用单个衍射光学元件 (DOE), 并要求该元件同时承担聚焦和扩展景深的功能, 但受到色差等因素的影响, 其实现成像质量的突破较为困难。

本文提出一种端到端框架, 整合光学结构和后端图像算法, 用于设计具备景深延展功能的单个 DOE。所提出的端到端设计范式整合 PSF 设计、成像模型和深度图像恢复网络, 能够打破前端光学设计和后端图像处理之间的数据流动壁垒, 省去光学系统的强制评价指标, 以图像质量为最终评价标准, 在给定景深范围和 PSF 的一致性之间找到最佳组合。所提出的端到端设计方法还包括透镜 (组) 的设计, 从而将衍射元件的设计难度降低到仅涉及景深扩展功能。所提系统在光学结构和后端处理算法方面取得平衡, 为小型化、轻量化和低成本的光学系统设计提供技术依据。

2 光学成像系统的端到端模型

所提端到端模型如图 1 所示, 其框架包含光学成像和图像恢复两部分。光学成像模型生成 PSF 来模拟探测器采集的模糊图像, 随后经过图像恢复模型得到清晰图像。所提模型在优化过程中, 系统的光学成像和图像恢复模型产生的损失函数将同时驱动图像恢复算法和光学元件进行更新。本节对该正向模型的各

收稿日期: 2023-07-17; 修回日期: 2023-08-18; 录用日期: 2023-09-18; 网络首发日期: 2023-09-22

通信作者: *yanglei@tju.edu.cn

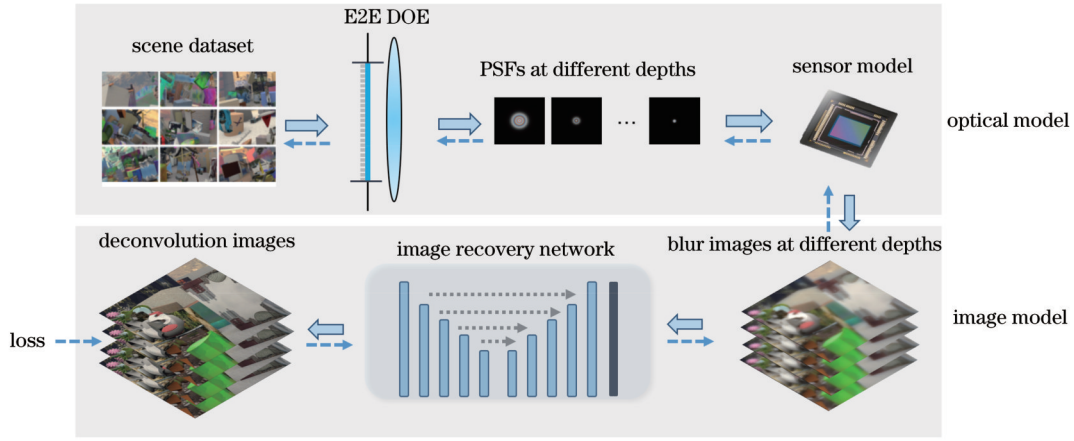


图1 光学成像系统的端到端模型框架

Fig. 1 End-to-end model framework of an optical imaging system

部分进行描述,并以一种适用于后续端到端优化的方式进行直接参数化。

2.1 光学模型

所提模型中假设光是空间不相干的,基于傅里叶

光学理论^[23],PSF被视为点光源在光学系统传播过程中在探测器上形成的图像。因此,光学成像模型包括3个部分:点光源发出光线、光学系统传播和探测器接收,如图2所示。

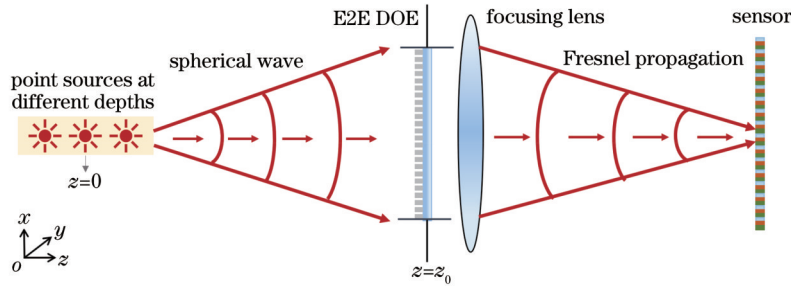


图2 端到端框架中的光学模型

Fig. 2 Optical model in the end-to-end framework

2.1.1 点光源

在空间中传播一定距离 z_0 的光波 U_0 表示为

$$U_0(x_0, y_0, z = z_0) = A_0(x_0, y_0) \exp[j\Phi_0(x_0, y_0, z_0)], \quad (1)$$

式中: A_0 为振幅; Φ_0 为相位。本实验模拟点光源的入射情况,假设点光源位于 $z=0$ 处,在传播 z_0 距离后到达光学元件时的相位 $\Phi_0 = k(x_0^2 + y_0^2 + z_0^2)$ 。

2.1.2 光学系统

1) 衍射元件

对于单个折射光学元件或DOE,其对光场的相位调制与元件的折射率 n 和厚度 h 有关,即

$$\Phi_{\text{DOE}} = k(n-1)h(x_0, y_0), \quad (2)$$

式中: $k = 2\pi/\lambda$ 为波数, λ 为波长; x_0, y_0 为光学元件相位的坐标。

2) 聚焦透镜

为降低单个DOE的成像压力,本文引入聚焦透镜,其相位调制表示为

$$\Phi_{\text{Lens}} = k\left(f - \sqrt{x_0^2 + y_0^2 + f^2}\right), \quad (3)$$

式中: f 为聚焦透镜的焦距。经过聚焦透镜后,光场的分布为

$$U_{\text{Lens}}(x_0, y_0, z) = P_1(x_0, y_0) \exp[j(\Phi_0 + \Phi_{\text{DOE}} + \Phi_{\text{Lens}})], \quad (4)$$

式中: $P_1(x_0, y_0)$ 为孔径函数。当光学元件的直径为 D 时,孔径函数的定义为

$$P_1(x_0, y_0) = \begin{cases} 1, & x^2 + y^2 \leq D \\ 0, & x^2 + y^2 > D \end{cases} \quad (5)$$

2.1.3 探测器接收

光场经光学元件调制,在自由空间中传播 L 距离后到达探测器,此时光场在探测器上的分布为

$$U_s(x_1, y_1, z) = \mathcal{F}^{-1}\left\{\mathcal{F}(U_s) \exp(jkL) \times \exp[-j\pi\lambda L(f_x^2 + f_y^2)]\right\}, \quad (6)$$

式中: $f_x = \Delta x_0/2$; $f_y = \Delta y_0/2$; \mathcal{F} 和 \mathcal{F}^{-1} 分别为傅里叶变换与傅里叶逆变换。式(6)使用能够模拟的近场和远场传播,适用于任意景深的成像模拟。

本文使用点光源在探测器上形成的图像表示

PSF, 因此需要计算其强度值:

$$I_{\text{PSF}}(x_1, y_1, z) \propto |U_s(x_1, y_1, z)|^2. \quad (7)$$

需要注意的是, PSF 是与波长相关的, 其强度由复值波场的模平方进行描述。

2.2 图像模型

1) 由 PSF 生成图像

在近轴近似条件下, 光学成像结果 I' 是清晰场景 I 和 PSF 的卷积。PSF 是与波长相关的函数, 因此, 当对全可见光波段进行建模时, 需要对不同的波长进行积分, 即

$$I'(x, y) = \int \eta(\lambda) I(x, y, \lambda) * I_{\text{PSF}}(x, y, \lambda) d\lambda, \quad (8)$$

式中: $\eta(\lambda)$ 表示传感器对每个颜色通道的波长灵敏度; * 表示卷积操作。

2) 探测器模拟

光学成像结果 I' 被集成到传感器像素上并受到噪声的影响, 形成传感器图像 I_s , 即

$$I_s(x, y) = s[I'(x, y) + \sigma], \quad (9)$$

式中: σ 为传感器噪声, 本文使用高斯噪声进行建模; $s(\cdot)$ 为相机响应函数, 其取值范围为 $[0, 1]$ 。

3) 图像复原

图像复原是一项具有挑战性的任务。传统的方法通常依赖先验知识来设计特定场景的特殊算法流程, 在处理复杂的图像退化时效果受限, 而近年发展的深度学习技术为图像复原提供了新的机遇。其中, 基于卷积神经网络的 U-Net 模型被广泛应用于图像复原领域^[24]。

所采用的 UNet 网络架构如图 3 所示, 该网络由编码器和解码器组成。输入图像首先经过编码器的 4 个下采样操作, 逐步降低空间分辨率并提取高维特征, 捕捉图像的全局和局部上下文信息; 其次, 经过 2 个卷积层, 以进一步提取输入图像的高维特征。解码器包含 4 个上采样操作和跳跃连接, 以逐步恢复图像的细节和空间分辨率。本实验设置的不同层级的特征通道数分别为 16、32、64、128、256。

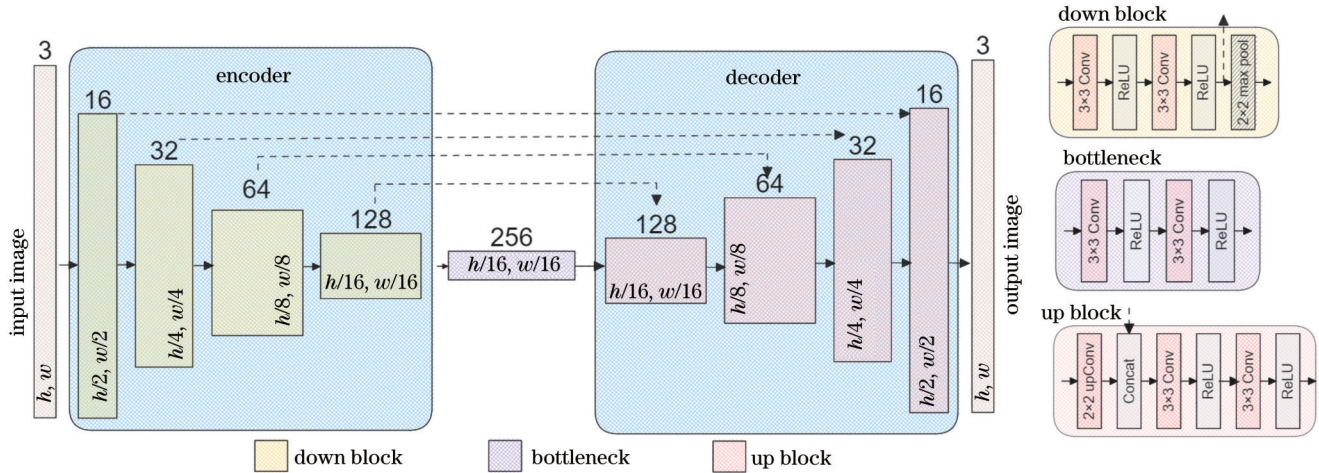


图 3 UNet 网络架构

Fig. 3 UNet network architecture

3 大景深端到端模型的建立

3.1 优化约束

将 2.1 节描述模型的每个阶段表示为可微分的模块后, 将模型有效集成到现代深度学习工具的工作流程中。使用 TensorFlow 的框架实现一个优化 DOE 的随机梯度方法。光学高度图 h 是优化变量, 光学元件尺寸、传感器像素尺寸和传感器噪声水平被设置为超参数。

在 RGB 图像数据集上优化模型时, 光学元件和重建方法中都使用图像的 L_1 损失进行变量优化, 以减小输出图像与真实图像之间的像素级绝对差异。 L_1 损失的定义为

$$L_1 = \frac{1}{N} \sum_{i=1}^N |y_i - x_i|, \quad (10)$$

式中: N 为像素数量; y_i 为输出图像中的像素值; x_i 为

清晰图像相应位置的像素值。

3.2 扩展景深的光学约束

对于端到端框架, 光学约束是为承担不同成像任务的光学系统而设计的, 其目的是确保光学任务和系统模型之间的一致性, 对光学元件的优化方向进行引导, 使整个系统更快优化。

扩展景深成像的思想是设计近似深度不变的 PSF, 并将探测器记录的图像应用反卷积以获得全聚焦图像。因此, 在所提模型中引入对 PSF 相似性的约束:

$$\mathcal{R}(f_{\text{PSF}, s}) = \sum_{m=1}^{L-1} \sum_{n=m+1}^L (p_m - p_n)^2, \quad (11)$$

式中: p_m 和 p_n 分别表示第 m 个和第 n 个深度处 PSF 的对应位置处的像素值; $f_{\text{PSF}, s}$ 为第 m 个和第 n 个深度处 PSF 的逐像素差异。因此, 本文网络的损失函数为

$$L = L_1 + \mathcal{R}(f_{\text{PSF}, s}). \quad (12)$$

此外,在建立扩展景深成像系统时,为实现可见光波段的成像模拟,将光谱离散为 486 nm、588 nm 和 656 nm 3 个波长。设定聚焦透镜的焦距为 50 mm,光学元件到探测器的传播距离为 52.63 mm。在选定景深范围时,采用 7 个离焦量(-30、-20、-10、0、10、20、30)交替优化的方法。离焦量是一个无量纲的物理量,表达式为

$$\phi_{\text{df}} = \frac{\pi D^2}{4\lambda} \left(\frac{1}{l} - \frac{1}{l_0} \right), \quad (13)$$

式中: D 为透镜的入瞳直径,本文的设计值为 5 mm; l 为物距; l_0 为对距离,本文的设计值为 1 m。

3.3 DOE 的设计约束

DOE 是平面透镜,它依靠由微小特征引起的相位延迟以期望的方式进行光波衍射。在进行优化设计时,其高度图 h 是优化的变量,然而, h 是二维高度图,直接进行优化将耗费大量运算资源,导致不稳定性增强。因此,本文分别设计两种方案对扩展景深的光学元件进行约束,这两种方案分别是三次方相位板和旋转对称 DOE。

1) 三次方相位板

三次方相位板的相位是连续的,其数学表达式为

$$\Phi(x, y) = a(x^3 + y^3). \quad (14)$$

三次方相位板被广泛应用于景深延展场景中,在式(14)中,只需优化变量 a ,计算量显著降低。

2) 旋转对称 DOE

三次方相位板虽然计算简单,但其优化的自由度较低。为了提升优化的灵活度,设计了一种旋转对称的相位板。该相位板将沿径向的离散一维数据作为优化变量,并通过旋转获得二维高度图 h ,即

$$h(x, y) = f_{\text{rotate}}[h(r)], \quad (15)$$

式中: f_{rotate} 表示旋转函数。本文通过极坐标的方法实现旋转函数,即为 $h(r)$ 赋予旋转角 θ ,将其旋转 360° : $x = r \cos \theta, y = r \sin \theta, \theta \in [0, 360^\circ)$,可在提高优化灵活度的同时限制内存的使用。

在优化过程中,高度图的特征尺寸为 $6 \mu\text{m}$,并将 20 nm 范围内的随机均匀噪声添加到高度图中,以增加光学元件对制造缺陷的鲁棒性。

4 端到端设计的实验

4.1 光学模型验证

为验证端到端网络中的光学模型,在 Zemax 软件中建立一个近轴系统,其入瞳直径为 5 mm,焦距为 50 mm,观察 Zemax 近轴系统与端到端系统中的 PSF,如图 4 所示。

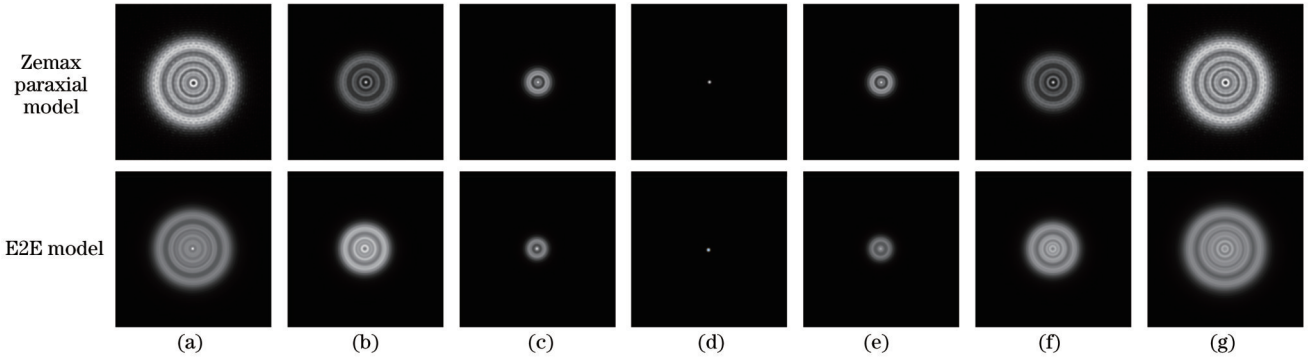


图 4 Zemax 近轴系统与端到端模型在不同离焦量处的 PSF 对比。(a) $\phi_{\text{df}} = -30$; (b) $\phi_{\text{df}} = -20$; (c) $\phi_{\text{df}} = -10$; (d) $\phi_{\text{df}} = 0$; (e) $\phi_{\text{df}} = 10$; (f) $\phi_{\text{df}} = 20$; (g) $\phi_{\text{df}} = 30$

Fig. 4 Comparison of PSF between Zemax paraxial model and end-to-end model at different defocus levels. (a) $\phi_{\text{df}} = -30$; (b) $\phi_{\text{df}} = -20$; (c) $\phi_{\text{df}} = -10$; (d) $\phi_{\text{df}} = 0$; (e) $\phi_{\text{df}} = 10$; (f) $\phi_{\text{df}} = 20$; (g) $\phi_{\text{df}} = 30$

从图 4 可以看到:当离焦量为 0,即点源位于对焦距离处时,两种成像模型在探测器上的成像均为一个点,可视为没有引入离焦像差;随着离焦量绝对值的增加,点源在探测器上呈现出菲涅耳衍射现象,衍射图像的直径增加,离焦像差也增大。通过对比端到端光学模型和 Zemax 常规光学模型的结果,发现它们的变化趋势一致,验证了端到端模型的有效性。

4.2 扩展景深成像实验

使用包含 21818 幅训练图像和 4248 幅测试图像的 FlyingThings 3D 数据集^[25]以及包含 2506 幅训练图像和 684 幅测试图像的 DualPixel 数据集^[26]进行交替训练,分别得到光学元件的设计以及最终成像结果。

图 5 展示了通过所提出的端到端设计方法获得的三次方相位板和旋转对称 DOE 的矢高图。

为对比衍射元件对景深延展的作用,首先将不含景深延展元件的 Zemax 近轴光学模型与 UNet 图像恢复方法直接结合,并作为原始系统。然后,将原始系统与包含衍射元件的端到端延展景深成像系统进行对比。为全面展示不同系统在不同场景中的景深延展效果,图 6(a)~(c)展示了来自 SceneFlow 测试集的图像复原效果,图 6(e)~(g)展示了来自 DualPixel 测试集的图像复原效果。如图 6(d)所示,在离焦量为 0 的被摄物体上,原始系统获得清晰图像;随着离焦量的增加,图像逐渐变模糊。相比之

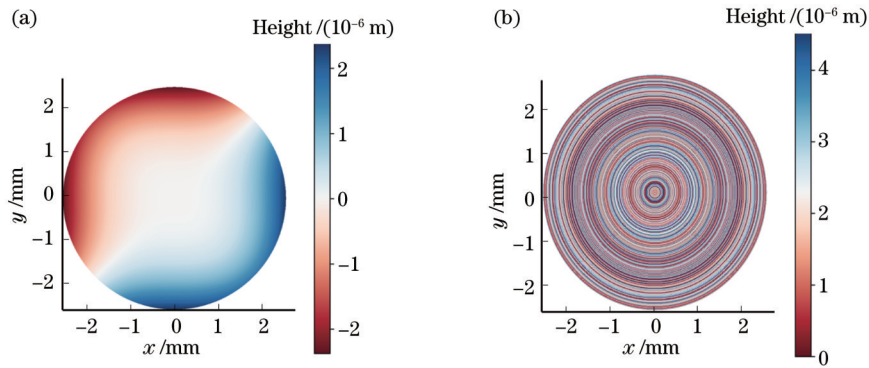


图 5 两种方案的景深延展衍射元件的矢高图。(a)三次方相位板;(b)旋转对称 DOE

Fig. 5 Height maps of depth-of-field extended diffractive element for two schemes. (a) Cubic phase plate; (b) rotationally symmetric DOE

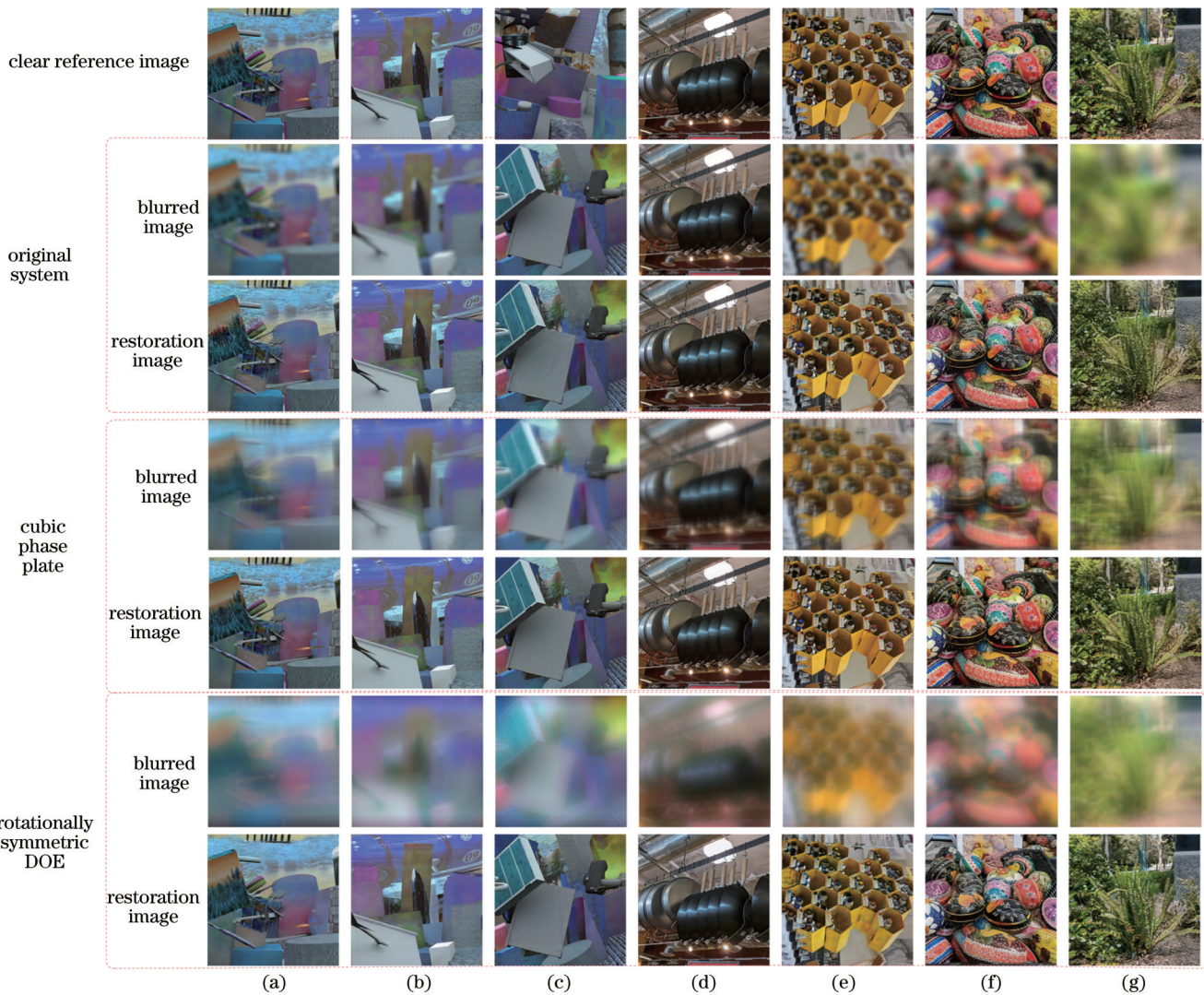


图 6 不同景深延展方法在不同离焦量的成像质量对比。(a) $\phi_{\text{def}} = -30$; (b) $\phi_{\text{def}} = -20$; (c) $\phi_{\text{def}} = -10$; (d) $\phi_{\text{def}} = 0$; (e) $\phi_{\text{def}} = 10$; (f) $\phi_{\text{def}} = 20$; (g) $\phi_{\text{def}} = 30$

Fig. 6 Comparison of image quality for different depth-of-field extension methods at various defocus amounts. (a) $\phi_{\text{def}} = -30$; (b) $\phi_{\text{def}} = -20$; (c) $\phi_{\text{def}} = -10$; (d) $\phi_{\text{def}} = 0$; (e) $\phi_{\text{def}} = 10$; (f) $\phi_{\text{def}} = 20$; (g) $\phi_{\text{def}} = 30$

下,所提出的景深延展系统尽管在对焦距离为 0 处仍然存在离焦模糊,但随着离焦量的增大,探测器上的模糊程度变化较小。观察复原原图像及其评价指标可以发现,当离焦量增大时,景深延展系统的复原图像

质量能够保持较高水准,而原始系统的图像质量明显下降。

为评估所提系统的景深延展能力,计算不同方法在测试数据集上的平均峰值信噪比(PSNR)和结构相

似性指数 (SSIM), 结果如图 7 所示。随着离焦量增加, SSIM 和 PSNR 整体下降, 说明增大离焦量会降低图像质量。采用旋转对称 DOE 和三次方相位板技术的图像相对于原始系统的图像在多数情况下具有更高的评价价值。具体而言, 三次方相位板在较小的离焦范

围内成像质量变化较平稳, 但在较大离焦量处性能下降, 尤其是在真实图像组成的 DualPixel 数据集上, 三次方相位板的成像效果不如旋转对称 DOE。相比之下, 旋转对称 DOE 在对焦位置及较大离焦量下均能够保持较高的成像质量。

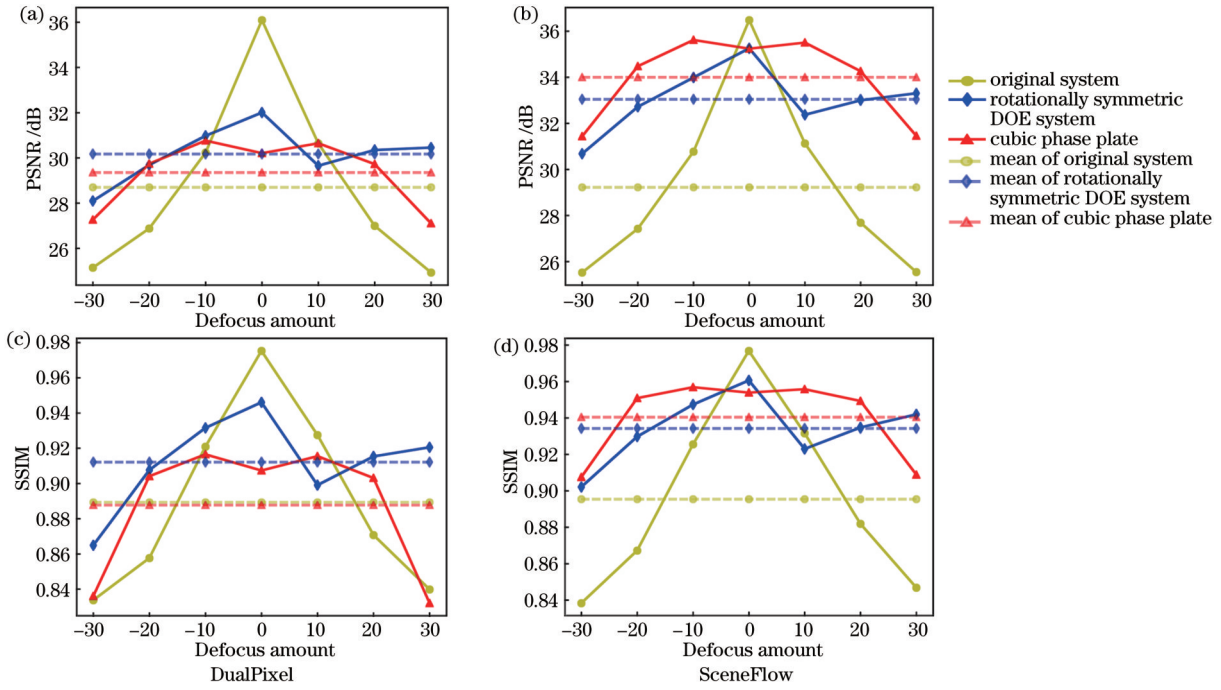


图 7 不同景深延展方法在测试集上的 PSNR 和 SSIM 随离焦量的变化。(a) (b) PSNR; (c) (d) SSIM

Fig. 7 Variation of PSNR and SSIM for different depth-of-field extension methods on the test set. (a) (b) PSNR; (c) (d) SSIM

为验证所提设计方法在不同景深处的鲁棒性, 对离焦范围内非设计值位置, 即离焦量为 -26、-16、

-6、6、16、26 的位置进行测试, 结果如图 8 所示。可以看到, 景深非设计值位置的评价结果趋势与设计值位

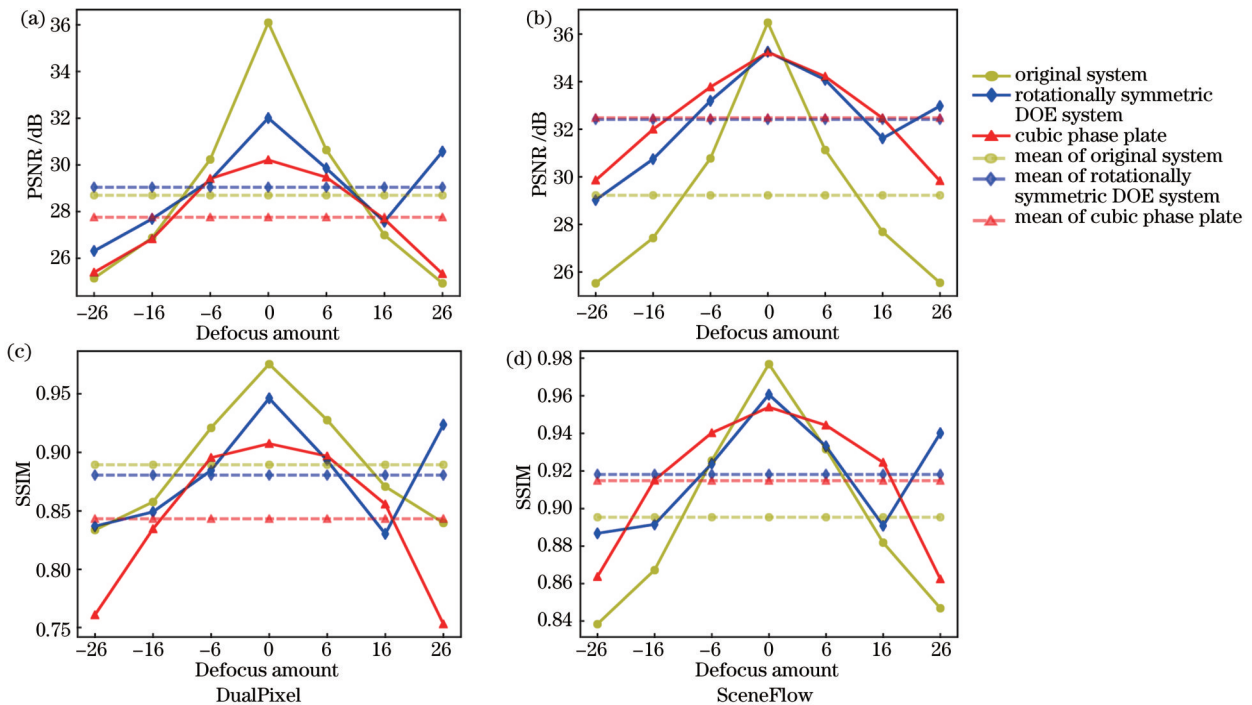


图 8 不同景深延展方法在测试集上非设计值处的 PSNR 和 SSIM 随离焦量的变化。(a) (b) PSNR; (c) (d) SSIM

Fig. 8 Variation of PSNR and SSIM for different depth-of-field extension methods at non-designed values on the test set. (a) (b) PSNR; (c) (d) SSIM

置类似,能够体现出本文算法的鲁棒性。此外,非设计值位置的评价参数均值略低于设计值处,这一现象源于深度学习网络的数据依赖性。

实验结果显示,将 UNet 图像恢复能力与传统成像系统相结合,能够实现一定程度的景深延展效果。通过端到端网络优化光学系统和图像恢复算法,光学设计系统和图像处理算法之间的障碍得以打破,使得光学-图像的整体优化成为可能。所提方法不仅能够优化经典相位板的设计,还可以自行设计约束,如所提出的旋转对称相位板,从而突破传统工具对光学设计的限制。

5 结 论

提出一种基于计算成像理论的端到端光学设计方法,为光学设计和图像复原过程建立完整模型,并通过全局优化链路来改善 DOE 的设计流程和性能。所提方法可用于大景深成像系统的设计,验证了其应用于景深延展的有效性。实验结果显示,端到端优化的旋转对称 DOE 和三次方相位板技术能够有效改善离焦量处于 $[-30, 30]$ 范围内的图像质量。所提方法突破了传统光学设计中需要使用复杂结构、特殊材料和特殊面形的限制,为实现小型化、轻量化和低成本的光学系统提供了新的可能性。此外,提出新颖的旋转对称 DOE 设计来扩大光学设计自由度。本研究对于计算光学成像领域的进一步发展具有一定的理论和实际意义。

参 考 文 献

- [1] Flores A, Wang M R, Yang J J. Achromatic hybrid refractive-diffractive lens with extended depth of focus[J]. *Applied Optics*, 2004, 43(30): 5618-5630.
- [2] Marks D L, Kittle D S, Son H S, et al. Gigapixel imaging with the AWARE multiscale camera[J]. *Optics and Photonics News*, 2012, 23(12): 31.
- [3] Bian Y X, Jiang Y N, Huang Y R, et al. Smart-phone phase contrast microscope with a singlet lens and deep learning[J]. *Optics & Laser Technology*, 2021, 139: 106900.
- [4] Mallon J, Whelan P F. Calibration and removal of lateral chromatic aberration in images[J]. *Pattern Recognition Letters*, 2007, 28(1): 125-135.
- [5] Cui J L, Huang W. Optical aberration correction for simple lenses via sparse representation[J]. *Optics Communications*, 2018, 412: 201-213.
- [6] 张金凯, 陈晓波, 邵海燕, 等. 基于混合镜头模型的变焦镜头设计[J]. *光学精密工程*, 2018, 26(1): 38-47.
Zhang J K, Chen X B, Shao H Y, et al. Zoom lens design using mixed lens modules[J]. *Optics and Precision Engineering*, 2018, 26(1): 38-47.
- [7] 王翼昂, 胡洋, 朴明旭, 等. 含单层衍射元件的可见宽波段计算成像系统设计[J]. *光学学报*, 2023, 43(5): 0522001.
Wang Y A, Hu Y, Piao M X, et al. Design of visible broadband computational imaging system with single-layer diffractive element[J]. *Acta Optica Sinica*, 2023, 43(5): 0522001.
- [8] Li W L, Yin X Q, Liu Y, et al. Computational imaging through chromatic aberration corrected simple lenses[J]. *Journal of Modern Optics*, 2017, 64(20): 2211-2220.
- [9] Qiang W, Lee S. Optimal design of chip-on-the-tip endoscope with a large depth of field[C]//*Frontiers in Optics 2015*, October 18-22, 2015, San Jose, California. Washington, DC: OSA, 2015: JTU4A.21.
- [10] Metzler C A, Ikoma H, Peng Y F, et al. Deep optics for single-shot high-dynamic-range imaging[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1372-1382.
- [11] Schuler C J, Hirsch M, Harmeling S, et al. Non-stationary correction of optical aberrations[C]//2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE Press, 2012: 659-666.
- [12] Heide F, Rouf M, Hullin M B, et al. High-quality computational imaging through simple lenses[J]. *ACM Transactions on Graphics*, 2013, 32(5): 149.
- [13] Li W L, Liu Y, Yin X Q, et al. A computational photography algorithm for quality enhancement of single lens imaging deblurring[J]. *Optik*, 2015, 126(21): 2788-2792.
- [14] 谭政, 相里斌, 吕群波, 等. 基于像差选择性校正的光学-数字联合设计[J]. *光子学报*, 2018, 47(5): 0511001.
Tan Z, Xiangli B, Lü Q B, et al. Optics/digital processing co-design method based on aberration optional-correcting[J]. *Acta Photonica Sinica*, 2018, 47(5): 0511001.
- [15] Liu Y K, Zhang C Y, Kou T D, et al. End-to-end computational optics with a singlet lens for large depth-of-field imaging[J]. *Optics Express*, 2021, 29(18): 28530-28548.
- [16] Sitzmann V, Diamond S, Peng Y F, et al. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging[J]. *ACM Transactions on Graphics*, 2018, 37(4): 114.
- [17] Chang J L, Wetzstein G. Deep optics for monocular depth estimation and 3D object detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 10192-10201.
- [18] Sun Q L, Zhang J, Dun X, et al. End-to-end learned, optically coded super-resolution SPAD camera[J]. *ACM Transactions on Graphics*, 2014, 39(2): 9.
- [19] 南亦澜, 申俊飞, 张启灿. 基于多光谱融合的端到端计算成像[J]. *激光与光电子学进展*, 2022, 59(20): 2011010.
Nan Y L, Shen J F, Zhang Q C. End-to-end computational imaging based on multispectral fusion[J]. *Laser & Optoelectronics Progress*, 2022, 59(20): 2011010.
- [20] 刘江, 苗二龙, 王学亮, 等. 利用泽尼克位相型光瞳滤波器实现景深延拓及超分辨[J]. *光学学报*, 2015, 35(12): 1211002.
Liu J, Miao E L, Wang X L, et al. Depth of field extending and super-resolving with phase pupil filter of Zernike polynomials[J]. *Acta Optica Sinica*, 2015, 35(12): 1211002.
- [21] Dowski E R, Cathey W T. Extended depth of field through wave-front coding[J]. *Applied Optics*, 1995, 34(11): 1859-1866.
- [22] 邵加强, 苏宙平. 基于机器学习设计连续相位分布的衍射光学元件[J]. *光学学报*, 2023, 43(3): 0323001.
Shao J Q, Su Z P. Design of diffractive optical elements with continuous phase distribution based on machine learning[J]. *Acta Optica Sinica*, 2023, 43(3): 0323001.
- [23] Goodman J W. *Introduction to Fourier optics*[M]. 2nd ed. New York: McGraw-Hill, 1996.
- [24] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention - MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [25] Mayer N, Ilg E, Häusser P, et al. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016,

Las Vegas, NV, USA. New York: IEEE Press, 2016: 4040-4048.

[26] Abuolaim A, Brown M S. Defocus deblurring using dual-pixel

data[M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision - ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12355: 111-126.

Design of End-to-End Depth-of-Field Extension Diffractive Optical Elements Based on Computational Imaging

Ji Jiarui, Xie Hongbo, Yang Lei*

Key Laboratory of Optoelectronics Information Technology, Ministry of Education, College of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

Abstract

Objective This study aims to develop compact and lightweight imaging optical structures, transcending the challenges posed by intricate architectures, specialized materials, and unique surface configurations prevalent in traditional optical design paradigms. In response to these challenges, we introduce computational imaging techniques, seamlessly integrating the realms of optical design and image restoration. This integration alleviates the intricacies associated with front-end optical system design while concurrently streamlining the process through the application of image restoration algorithms. By transposing the complexities of optical design into the algorithmic realm, we endeavor to reduce optical system complexity while preserving image quality.

Methods We propose an end-to-end (E2E) framework to facilitate the creation of the diffractive optical element (DOE) capable of extending the depth of field (DOF). This framework integrates point spread function (PSF) design, imaging models, and deep image restoration networks through the utilization of modern deep learning tools. As a significant departure from traditional practices, this framework eradicates the traditional segregation between front-end optical design and back-end image processing stages. This method uses image quality as the final evaluation criterion to find the optimal balance between the consistency of a given DOF range and PSF. Moreover, the holistic E2E approach introduced by this method encompasses the intricate task of designing lenses (or lens groups) to accomplish the focus function. This strategic integration effectively simplifies the design complexities intrinsic to DOE, steering the design focus exclusively toward extending the DOF. Specifically, we employ the phase coefficients of cubic phase plates and the one-dimensional height map of rotationally symmetric DOEs to facilitate the dimension reduction of optical design parameters. Network constraints encompass the L1 constraint as the loss function for the image, alongside the inclusion of PSF consistency at varying depths as a specialized constraint for large DOF design. The amalgamation of these constraints gives rise to the loss function for the E2E network, propelling the designed network toward optimization updates. To enhance the network's generalization capabilities, the proposed method undergoes alternating training on two datasets: the FlyingThings 3D dataset, containing 21818 training images and 4248 test images, and the DualPixel dataset, featuring 2506 training images and 684 test images. This dual dataset training regimen yields designs for optical components and culminates in the final imaging outcomes.

Results and Discussions The efficacy of the proposed large DOF optical model is robustly validated through comparative analysis with Zemax results, visually depicted in Fig. 4. Subsequently, the E2E approach is efficaciously applied to the design of large DOF imaging systems, encompassing the design of both rotationally symmetric DOEs and cubic phase plate, depicted through their respective height maps in Fig. 5. To provide a comprehensive portrayal of the DOF extension effects within varying scenes, Fig. 6 presents the imaging quality of diverse DOF extension methods across different defocus levels, incorporating images from both the FlyingThings 3D and DualPixel datasets. Additionally, Fig. 7 effectively captures the graphs detailing variations in PSNR and SSIM concerning distinct DOF extension methods over the test dataset, showcasing the relative stability of imaging quality changes within a smaller defocus range for cubic phase plates. However, they experience performance degradation under more pronounced defocus levels, particularly evident in the real-image context of the DualPixel dataset, where rotationally symmetric DOEs outperform cubic phase plates. In contrast, the rotationally symmetric DOEs consistently maintain high image quality both at the focal point and under larger defocus levels. Furthermore, we underscore the robustness of the design method by meticulously validating its performance using non-design values within defocus ranges, as exhaustively detailed in Table 1. Empirical evidence

derived from these experiments unequivocally demonstrates that E2E-optimized rotationally symmetric DOEs and cubic phase plates effectively elevate image quality within the defocus range of $[-30, 30]$.

Conclusions In summary, we introduce an E2E optical design method based on computational imaging. The overall design workflow and performance of DOE are successfully enhanced by constructing a comprehensive model that integrates two different domains, optical design, and image restoration and applying the idea of global optimization with image quality as the final evaluation criterion. The method reduces the requirement for imaging quality of the front-end optical system and eliminates residual aberrations using image restoration algorithms, thus realizing a compromise between optical design and image algorithms. The method covers key aspects of optical field propagation, detector noise, and image post-processing. By building the corresponding models and jointly optimizing the optical models and image algorithms with the modern deep learning models, we successfully design lightweight and thin DOEs suitable for extended DOF and achieve high-quality imaging in a simple optical system with significant DOEs. In conclusion, this research advances an E2E optical design method rooted in computational imaging, enhancing the design of DOEs for extended DOF. The implications of this work extend to the broader field of computational optical imaging, holding both theoretical and practical significance.

Key words imaging systems; end-to-end; depth-of-field extension; image restoration; diffractive optical elements