Research Article | Vol. 32, No. 16 / 29 Jul 2024 / Optics Express 28402
Optics EXPRESS

Hybrid design scheme for enabling large-aperture diffractive achromat imaging

JIAN ZHANG,[1,2,3,†] ZIYU ZHAO,[1,2,3,†] ZEYING FAN,[1,2,3] HONGFEI JIAO,[1,2,3] ZHANSHAN WANG,[1,2,3,4] XINBIN CHENG,[1,2,3,4] YIFAN PENG,[5] AND XIONG DUN[1,2,3,*]

[1]*Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China*
[2]*MOE Key Laboratory of Advanced Micro-Structured Materials, Shanghai 200092, China*
[3]*Shanghai Frontiers Science Center of Digital Optics, Shanghai 200092, China*
[4]*Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092, China*
[5]*Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China*
[†]The authors contributed equally to this work.
[*]*dunx@tongji.edu.cn*

Abstract: Diffractive achromats (DAs) combined with image processing algorithms offer a promising lens solution for high-performance ultra-thin imagers. However, the design of large-aperture DAs that align seamlessly with image processing algorithms remains challenging. Existing sequential methods, which prioritize focusing efficiency in DAs before selecting an algorithm, may not achieve a satisfactory match due to an ambiguous relationship between efficiency and final imaging quality. Conversely, image-quality-oriented end-to-end design often entails high computational complexity for both front-end optics and back-end algorithms, impeding the development of large-aperture designs. To address these issues, we present a hybrid design scheme that begins with end-to-end optimization of the DA with the simplest image processing algorithm, i.e., Wiener filter, significantly reducing the back-end complexity. Subsequently, we apply complex algorithm fine-tuning to further enhance image quality. We validate this hybrid design scheme through extensive investigations on several DA imagers. Our results demonstrate a reduction in memory requirement by approximately 50% while maintaining a high imaging quality with a reasonably large aperture. As a case in point, we simulated a DA imager with a 25 mm diameter aperture. Furthermore, our hybrid design scheme provides two crucial insights. Firstly, we find no strong linear correlation between focusing efficiency and imaging quality, which challenges the conventional understanding. Secondly, we establish a prediction formula for imaging quality, benefiting from the hybrid design scheme.

1. Introduction

The demand for high-performance optical imagers that are lightweight has been steadily increasing in academia and industry, particularly for applications such as biomedical imaging [1], machine vision [2], and consumer electronics [3]. However, traditional refractive optics encounter substantial obstacles in developing lightweight systems due to their bulky nature. Diffractive lenses [4–7] and metalenses [8,9] have emerged as promising alternatives, utilizing planar micro/nanostructures to simplify the complexity of optics and have shown great potential for lightweight optical imaging devices. However, their practical implementation is still limited by inherent chromatic aberrations [10,11], caused by phase wrapping, where wavelengths other than the designed wavelength deviate from the ideal phase distribution [12]. Although a number of studies have demonstrated achromatic diffractive lenses [13] and achromatic metalenses [14–16] (collectively referred to as diffractive achromats) through well-designed structures, achieving

#525558 https://doi.org/10.1364/OE.525558
Journal © 2024 Received 10 Apr 2024; revised 6 Jul 2024; accepted 15 Jul 2024; published 23 Jul 2024

Optics EXPRESS

achromaticity comes at a cost, as it introduces additional wavefront aberrations. Therefore, subsequent image processing procedures, that is, *computational imaging*, are often employed to correct the residual aberrations and improve the imaging quality. Diffractive achromats (DAs) combined with image processing algorithms are promising in achieving high-performance, lightweight DA imagers. Some pioneering work [17–21] has recently designed high-performance broadband computational imaging systems using this insight.

Currently, there are two primary types of DA imager designs. The first type is to optimize the structural parameters of the DAs with focusing efficiency as the optimization goal [5–7,22,23] and subsequently design an image processing algorithm to improve the imaging quality, as shown in Fig. 1(a). This method is often known as sequential design, which is simple to perform and does not require a large computational memory. However, the sequentially designed DA imagers often exhibit mediocre image quality due to the unclear relationship between focusing efficiency and final imaging quality. The second type is the end-to-end design approach [4,9,24–27] that jointly designs DAs and image processing algorithms with the final imaging quality as the optimization goal by simulating the entire imaging pipeline. This approach has been relatively successful in matching DA and image processing algorithms. However, complex front-end optics and numerous back-end algorithms have significant computational complexities, requiring enormous computing resources [28], as illustrated in Fig. 1(b). Although Dun et al. developed a concentric ring decomposition method [4] that reduces the front-end optics calculation, optimizing large aperture DAs remains challenging because of the complexity of the back-end algorithms.

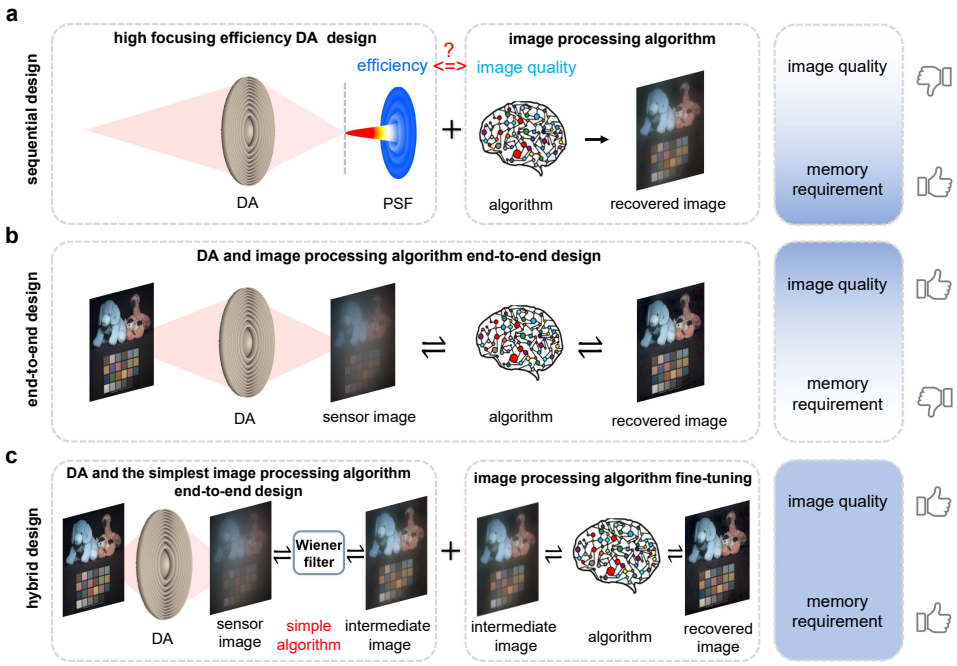

Fig. 1. Comparison of different diffractive achromat (DA) design strategy. (a) The sequential design optimizes the DAs with focusing efficiency, then followed by the design of an image processing algorithm. (b) The end-to-end design optimizes the DAs and image processing algorithms together. (c) The proposed hybrid design scheme first end-to-end design DAs and the simplest image processing algorithm, that is, the Wiener filter, then followed by a complex algorithm fine-tuning to further enhance image quality.

In this study, we present a hybrid design scheme, shown in Fig. 1(c), that first jointly optimizes DA and the simplest image algorithm, namely the Wiener filter, and subsequently fine-tunes

Optics EXPRESS

the image algorithm by cascading deep learning after the Wiener filter, while maintaining the DA constant. By decomposing the computational complexity into two distinct processes, our strategy reduces the memory requirement and enables the optimization of large-aperture DAs. We validate the effectiveness of our approach by optimizing and analyzing several DA imagers with varying parameters. Our hybrid design scheme attains high imaging performance while maintaining minimal memory requirement. To demonstrate the superiority of the proposed scheme, we present a DA imager design with an ultra large-aperture diameter of 25 mm, which, to the best of our knowledge, represents the largest DA imager in the visible spectrum employing an end-to-end design architecture. In addition, based on this hybrid design strategy, we uncover two critical insights. First, we find that a robust linear correlation between focusing efficiency and image quality is absent, implying that relying solely on optimizing DAs for focusing efficiency may be insufficiently reliable. Second, we present a formula to predict the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of DA imagers, providing a valuable reference for the design of desired DAs.

2. Hybrid design scheme for DA imagers

2.1. End-to-end design of the DA and Wiener filtering

In the first stage of our hybrid design scheme, the DA and Wiener filter are jointly learned by constructing a differentiable pipeline. This pipeline consists of a forward imaging model, a Wiener filter image recovery model [29], and a loss function.

The forward imaging model incorporates the calculation of the point spread function (PSF) and the simulation of sensor imaging. We assume that the DA exhibits rotational symmetry and consists of concentric rings with a ring interval (feature size) d and varying heights h_m, where $m = 1, 2, \ldots, N$, $N = D/2d$, and D is the aperture. The complex transmittance function of the DA can be expressed using a matrix:

$$\mathbf{t}(r,\lambda) = \left[\begin{array}{cccc} e^{ik(n(\lambda)-1)h_1} & e^{ik(n(\lambda)-1)h_2} & \cdots & e^{ik(n(\lambda)-1)h_N} \end{array} \right], \tag{1}$$

where $k = 2\pi/\lambda$ is the wave number, $n(\lambda)$ is the refractive index of DA material. Based on the concentric ring decomposition [4], the rotationally symmetric PSF model of DA is as follows:

$$\mathbf{PSF}(\rho,\lambda) = |\mathbf{t}(r,\lambda)\mathbf{H}(r,\rho)|^2, \tag{2}$$

where ρ represents the polar coordinates in the image plane, and $\mathbf{H}(r,\rho)$ is a matrix formed by the Rayleigh-Sommerfeld diffraction field of each concentric ring. This matrix is calculated using the angle spectrum method [30,31] in advance. The $\mathbf{H}(r,\rho)$ used in our work differs from Ref. [4], as we consider a more rigorous situation.

The two-dimensional PSF is obtained by rotating the $\mathbf{PSF}(\rho,\lambda)$ along the optical axis. Similar to the Ref. [4], we use a relatively small sensor size to save computational memory. However, we not only incorporate energy regularization (see loss function) but also modify the $\mathrm{PSF}_{sensor}(x,y,\lambda)$ on the sensor to account for the possibility of energy outside the sensor. This additional energy is collected and added as a uniformly distributed background $\mathrm{PSF}_{bg}(x,y,\lambda)$.

For brevity, we assume the DA imager is shift invariant [31–33], which is valid for within a diagonal field of view (FOV) of 30°. See Supplement 1 for detailed analysis. Larger FOV scenarios require consideration of shift variant, which involves advanced off-axis modeling [34] and can be explored in future work. Therefore, the images $y(x,y)$ acquired by the DA and sensor can be expressed as follows:

$$y(x,y) = \int_{\lambda_{\min}}^{\lambda_{\max}} \left(\left(\mathrm{PSF}_{sensor}(x,y,\lambda) + \mathrm{PSF}_{bg}(x,y,\lambda) \right) * I(x,y,\lambda) \right) R_c(\lambda) d\lambda, \tag{3}$$

Optics EXPRESS

where $R_c(\lambda)$ is the spectral response curve of sensor, $*$ represents convolution, and $I(x, y, \lambda)$ is the scene. By modeling the sensor noise, the sensor images $b(x, y)$ can be written as:

$$b(x, y) = \eta_g(y(x, y), \sigma_g) + \eta_p(y(x, y), a_p), \tag{4}$$

where $\eta_g(y(x, y), \sigma_g) \sim \mathcal{N}(y(x, y), \sigma_g^2)$ is the Gaussian noise component and $\eta_p(y(x, y), a_p) \sim \mathcal{P}(y(x, y)/a_p)$ is the Poisson noise component.

The Wiener filter image recovery model is represented as:

$$\tilde{I}_c(x, y) = F^{-1}\left(\frac{H(u, v)^* ((1 + \gamma)B(u, v))}{|H(u, v)|^2 + \gamma}\right), \tag{5}$$

where $F^{-1}(\cdot)$ is the inverse Fourier transform, (u, v) represents frequency-space coordinates, $*$ represents a conjugate operator, γ is a learnable parameter, and $H(u, v)$ and $B(u, v)$ are the Fourier transforms of PSF and $b(x, y)$, respectively. The sensor images are fed into the Wiener filter image recovery model to recover the scene image.

DA and Wiener filtering optimizations are achieved by minimizing the following loss function:

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + \alpha\mathcal{R}_{\text{PSF}}, \tag{6}$$

where $\mathcal{L}_{\text{recon}}$ represents the fidelity, namely the ℓ_2 norm, α is the regularization weight, and $\mathcal{R}_{\text{PSF}}$ represents energy regularization [4], which limits the amount of diffraction energy falling on the sensor as much as possible, allowing us to use a relatively small sensor size for saving computational memory.

2.2. Fine-tuning with a neural network complex algorithm

In the second stage of the hybrid design scheme, we fine-tune the image processing algorithm with a more complex neural network algorithm. We maintain the DA optimized in the first stage, and connect the Res-Unet network [35] after the Wiener filtering to further denoise the intermediate image, which is already deconvolved using the Wiener filter. Specifically, we begin by extracting the DA height distribution from the optimized result in the first stage. We then utilize the forward model of the DA to process all the hyperspectral images in the dataset, resulting in the generation of image pairs comprising sensor images and RGB ground truth. These image pairs are then employed for retraining the imaging processing algorithm. During the retraining process, we optimize the Res-Unet network by connecting it in series after the Wiener filter, and jointly optimizing them by minimizing the loss function. The loss function is adjusted to exclusively include the first fidelity term $\mathcal{L}_{\text{recon}}$ in Eq. (6). Supplement 1 provides details regarding the Res-Unet network architecture and configuration.

Our hybrid design scheme significantly reduces memory requirement. Figure 2 compares the memory requirement of the hybrid design scheme and the vanilla end-to-end design method (refer to Supplement 1 for the evaluation method). It can be seen that when designing DA with an aperture greater than 20 mm using the vanilla end-to-end design method, the memory requirement reaches 12 GB, exceeding the memory limitations of many consumer-level GPUs, such as the NVIDIA 3060Ti GPU.

By contrast, in the first stage of the hybrid design scheme, although the memory consumption also increases with aperture size, using the simplest image processing algorithm reduces memory requirement to 6.5 GB for a DA with 20 mm aperture, a saving of approximately 50%. In the second stage, the memory requirement is independent of the aperture as the front-end optics remains fixed. In addition, through this decomposition trick, our scheme reduces the number of optimization variables, potentially improving the convergence. Refer to the analysis in Section 3.1 for details.

Optics EXPRESS

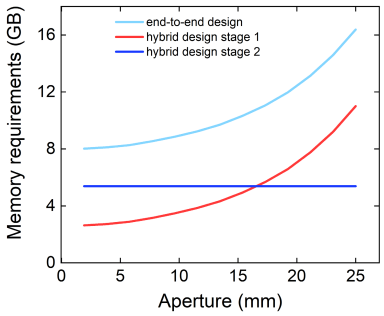

Fig. 2. Memory consumption comparison of the hybrid design scheme and end-to-end design method [4] under different aperture settings. Note that the memory requirement is also related to the feature size of DA. In this figure, the feature size is 2 μm. If it is 1 μm, the memory requirement of the hybrid design stage 1 and end-to-end design needs to be doubled.

3. Evaluation in simulation

3.1. Effectiveness validation of the hybrid design scheme

To validate the effectiveness of the hybrid design scheme in designing DAs, we optimized 21 DA imagers to compare the design results with the vanilla end-to-end design method [4]. These DA imagers spanned aperture sizes from 1 to 10 mm, F-numbers ranging from 1 to 5, and maximum heights from 2 to 16 μm. Specific parameters can be found in Supplement 1. The ring feature size was set to 1 μm for all DAs. The material for DAs was PMMA. The achromatic spectrum ranged from 400 to 700 nm and comprised 11 channels with intervals of 30 nm. The initial height map was the Fresnel lens distribution. The sensor pixel size was set at 6 μm with a sensor count of 512×512. The sensor noise, including Gaussian σ_g and Poisson a_p noises, was configured at 0.001 and 0.005, respectively. The starting learning rates for the DA and image processing algorithms were set to 10^{-6} and 10^{-4}, respectively, with the Adam optimizer. The learning rate decayed linearly during optimization, and the cut-off learning rate was set to 10^{-10}. The energy regularization weight α in the loss function was fixed at 0.005. The datasets comprise the Harvard [36], ICVL [37], CAVE [37], and NUS [38] datasets. 14 hyperspectral images are selected as the test set and the remaining 361 images as the training set.

Figure 3 compares the optimization performance of the hybrid design scheme and the vanilla end-to-end design method [4]. The imaging quality of the DA imagers designed by these two design methods has a good linear correlation. These results prove that directly using complex reconstruction algorithms, such as deep learning, to optimize DA imagers is not required. As an alternative, simple algorithms, such as the Wiener filter algorithm, can be used to optimize DAs, and then deep learning can be cascaded to fine-tune the image processing algorithms. Moreover, by using the red isoline to divide and compare the PSNR (SSIM) values between the two design methods, we find that most DAs designed using the hybrid method exhibit higher PSNR (SSIM) than that of the vanilla end-to-end method.

3.2. Application validation of a large-aperture DA imager

We designed a 25 mm diameter DA imager as a demonstration example using the proposed hybrid design scheme. For this aperture size DA imager, the vanilla end-to-end design method would require at least 16.5 GB memory according to Fig. 2, which exceeds the memory limitation of most consumer-level GPUs. However, using our scheme, the required memory is reduced to approximately 11 GB. The focal length of the DA was 180 mm. The maximum height was 4 μm, and the feature size was 2 μm. Other settings were the same as those mentioned in Section 3.1.

Optics EXPRESS

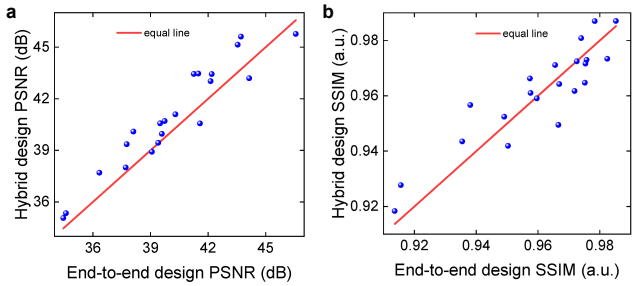

Fig. 3. Optimization performance validation of the hybrid design. (a) PSNR comparison of different DA imagers using the hybrid design scheme and end-to-end design method [4]. (b) SSIM comparison. The vertical and horizontal coordinate values of the blue dots indicate the PSNR (SSIM) of DA imagers designed using the two methods, respectively. The red isoline represents their equal value region.

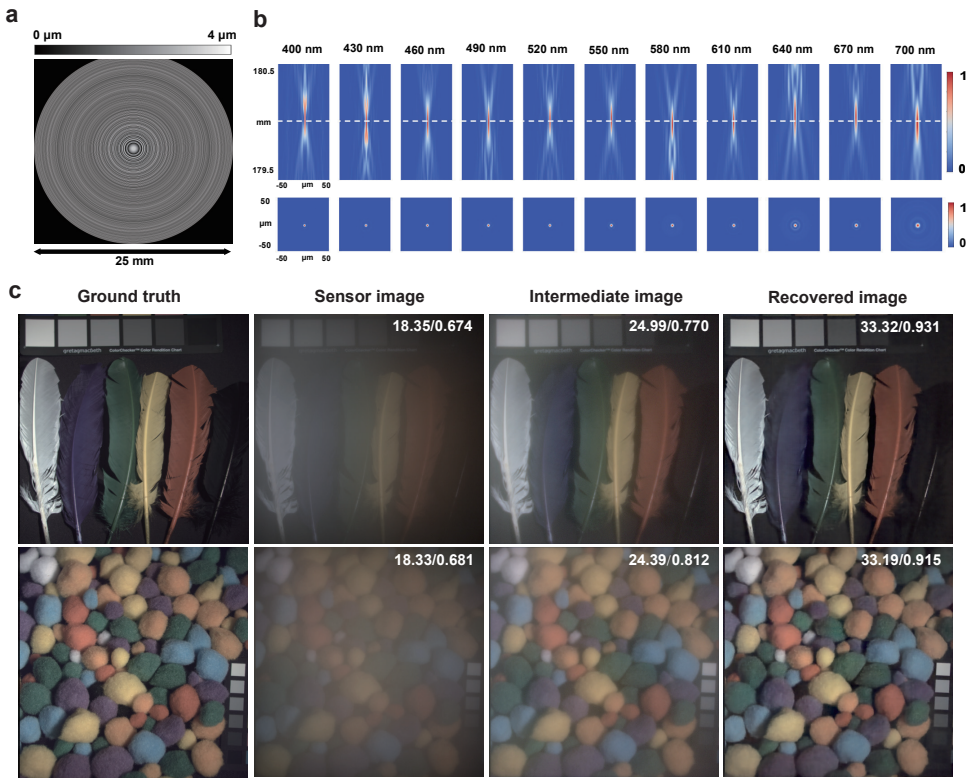

Fig. 4. Results of simulating a 25 mm diameter aperture DA imager. (a) Height map. (b) Light field intensities along the optical axis and at the focal plane. (c) Sensor images and recovery images. The inset values indicate the PSNR and SSIM.

Figure 4(a) shows the height map of the designed DA. We analyzed its achromatic characteristics based on two-dimensional angular spectrum diffraction. Figure 4(b) shows the focusing fields along the optical and focal planes. Across all 11 wavelengths between 400-700 nm, the light converged at a consistent point (z = 180 mm). Figure 4(c) shows the simulated imaging results of the DA imager. The initial restoration of image details through Wiener filtering achieved an average PSNR of 24.68 dB and SSIM of 0.780. Subsequently, the cascaded Res-Unet network

Optics EXPRESS

almost eliminated residual aberrations, resulting in final recovered images with an average PSNR of 33.52 dB and SSIM of 0.926. Thus, the proposed hybrid design scheme realizes an ultra-large-aperture DA design with high imaging performance.

4. Insights regarding DA imagers

4.1. Relationship between focusing efficiency and imaging quality

The relationship between a DA's focusing efficiency and its imaging quality has been rarely investigated in previous studies. By conducting statistical analysis, we explored the relationship between focusing efficiency and image quality, as well as between Strehl ratio and image quality. Specifically, we calculated the focusing efficiency, Strehl ratio and PSNR (SSIM) of the DAs, as detailed in Section 3.1, considering 21 parameter groups designed using the hybrid design scheme. The focusing efficiency is defined as the ratio of the power within three times the full width at half maximums ($3w$) of PSF of the DAs over the entire transmitted power on the exit pupil. To consider the size of the PSFs, based on the function from equation 5 of reference [6], the efficiency is divided by w^2, where w is normalized by the diameter of the diffraction-limited focus disc ($\lambda/2NA$, full width at half maximum, strictly speaking). The averaging focus efficiency of the DA is defined as $\langle Eff \rangle_\omega / w^2_{\max}$, where $w^2_{\max}$ is the maximum normalized diameter of PSF across 11 wavelengths. Strehl ratio is the ratio of the maximum value of PSF over the maximum value diffraction-limited PSF. In this study, Strehl ratio refers to the average Strehl ratio of the DA across 11 wavelengths.

We considered two imaging quality metrics (PSNR and SSIM) before and after Res-Unet fine-tuning (only Wiener filter deconvolution). Figures 5(a)-(b) and (e)-(f) respectively show the results before Res-Unet fine-tuning. To mitigate statistical errors resulting from local optima, we optimized the DA for each set of parameters ten times and plotted the error bars. It is evident that a straightforward linear relationship between either DA focusing efficiency or Strehl ratio and resulting image quality is absent. Despite certain DA imagers exhibiting notably high focusing efficiency and high Strehl ratio, the final imaging quality is not satisfactory. Figures 5(c)-(d) and (g)-(h) show the results after fine-tuning the DA using Res-Unet. The correlation between them became less apparent. These results lack error bars because we optimized the Res-Unet only once for each set of DA imagers due to the time-consuming nature of Res-Unet training. In addition, we plotted the MTF graphs for all DAs in Supplement 1 and examine the MTF value at the sensor cutoff frequency. We found it does not have a linear relationship with PSNR/SSIM either.

4.2. Prediction formula for DA imaging quality

DA imagers with varying apertures, focal lengths, and maximum heights exhibit varying levels of image quality. To find the underlying patterns, we analyzed the results obtained from Wiener filtering. Given the small number of optimizable variables in Wiener filtering, the image quality trend is more apparent than that of the Res-Unet. For an in-depth discussion on the effects of F number, maximum height, and aperture on imaging quality, refer to Figure S5 in Supplement 1. Overall, we observe that smaller apertures, larger F-numbers, and higher maximum heights tend to lead to DA imagers with higher PSNR and SSIM values.

Expanding on our findings, we can summarize a prediction formula for the image quality of DAs when Wiener filtering is employed as the joint design algorithm:

$$\mathrm{PSNR}(D,F,h) = a_1 \frac{F^{a4}\left(a_5 + h^{a6}\right)}{a_2 + D^{a3}}, \quad \mathrm{SSIM}(D,F,h) = b_1 \frac{F^{b4}\left(b_5 + h^{b6}\right)}{b_2 + D^{b3}}, \tag{7}$$

where D, h, and F are the aperture, maximum height, and F number of DA, respectively. The fitting coefficients a_i and $b_i (i = 1, 2, \ldots, 6)$ are listed in Table 1. The fitting function is 4-dimensional, and visualizing the fitting error is inconvenient. Thus, we fixed maximum height

Optics EXPRESS

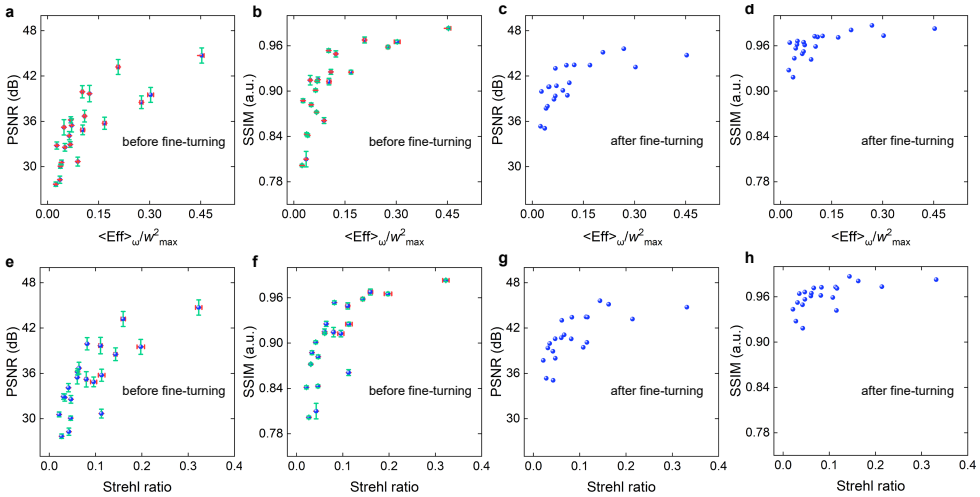

Fig. 5. Distribution of PSNR and SSIM versus focusing efficiency and Strehl ratio. (a)-(b) Scatter plot showing focusing efficiency versus PSNR (SSIM) using Wiener filtering recovery. Red error bars indicate variance in focusing efficiency, and green error bars indicate variance in PSNR (SSIM). (c)-(d) Scatter plot showing corresponding statistical results after Res-Unet fine-tuning. (e)-(f) Scatter plot showing Strehl ratio versus PSNR (SSIM) using Wiener filtering recovery. (g)-(h) Scatter plot showing Strehl ratio versus PSNR (SSIM) after Res-Unet fine-tuning.

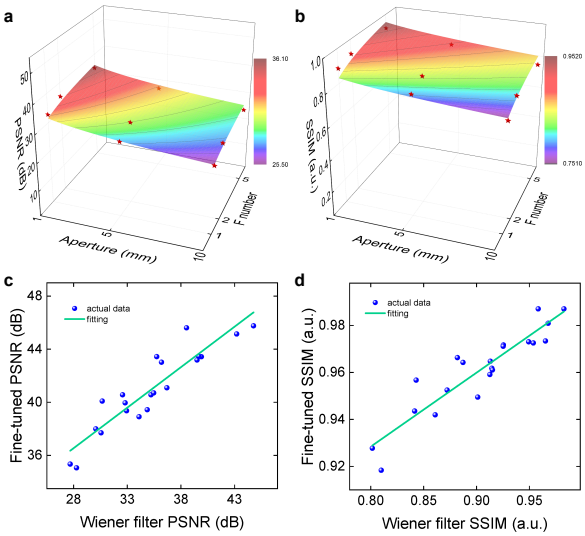

Fig. 6. Fitting errors and the image quality improvement correspondence. (a)-(b) PSNR (SSIM) fitting errors of the prediction formula. The red stars and colorful surfaces represent the actual and fitting data, respectively. (c)-(d) The correspondence between the PSNR (SSIM) recovered by Wiener filtering and that fine-tuned after Res-Unet.

dimension and changed the other three dimensions for display. Figures 6(a) and (b) show the fitting errors when the maximum height with 2 μm. See Supplement 1 for fitting errors when fixing other maximum heights. The maximum PSNR and SSIM errors are within 1.84 dB and 0.055, respectively. Note that although the prediction formula was obtained using 1–10 mm

Optics EXPRESS

apertures, 1–5 F-numbers, and 2–16 μm maximum heights, it remains valid outside these ranges. For example, the PSNR and SSIM prediction errors are only 2.12 dB and 0.049, respectively, for the 25 mm DA described in Section 3.2.

Table 1. Fitting coefficient of the prediction formula for DA image quality

Fitting coefficient	1	2	3	4	5	6
a_i	28.5181	11.2420	0.5875	0.0840	12.0416	0.5632
b_i	0.6724	25.9461	0.5875	0.0840	31.8462	0.5632

The use of Res-Unet fine-tuning further enhances the PSNR and SSIM values. Figures 6(c) and (d) show the improvement correspondence in PSNR (SSIM) from the Wiener filtering to the Res-Unet fine-tuning. Using the prediction formula and the improvement correspondence shown in Fig. 6, one can readily specify the PSNR (SSIM) of the DA with arbitrary parameters set by the proposed hybrid design scheme.

5. Conclusion

In this study, we present a hybrid design scheme aimed at addressing the challenges associated with optimizing large-aperture DAs due to the optical and algorithmic complexities inherent in traditional end-to-end design methods. Our approach employs Wiener filter to initially optimize the DA, followed by fine-tuning using Res-Unet. This strategy significantly reduces memory requirement and facilitates large aperture DA optimization while achieving superior imaging performance compared with traditional end-to-end design method. Our hybrid design method is a versatile paradigm, we found that it is more pronounced when using more advanced and complex networks [9,19]. See Supplement 1 for detailed analyse and comparison.

In addition, we found that DAs exhibiting higher focusing efficiency do not consistently exhibit enhanced imaging quality. As a result, relying solely on optimizing DAs for focusing efficiency proves insufficiently reliable in all cases. It is imperative to adopt a methodology that jointly optimizes both the DA and downstream algorithms, including even the most rudimentary, to achieve high-performance imaging. We found that a DA characterized by a high maximum height, large F-number, and small aperture is more likely to deliver exceptional imaging quality. Our proposed prediction formula for the DA image quality offers a valuable reference for optical designers in selecting DA parameters. We envision our work enable large-aperture, lightweight, and high-performance optical imagers for applications in biomedical research and consumer electronics.

Funding. National Natural Science Foundation of China (62192774, 62105243, 61925504, 6201101335, 62020106009, 62192770, 62192772); Science and Technology Commission of Shanghai Municipality (17JC1400800, 20JC1414600, 21JC1406100); Shanghai Municipal Science and Technology Major Project (2021SHZDZX0100); Fundamental Research Funds for the Central Universities; the Special Development Funds for Major Projects of Shanghai Zhangjiang National Independent Innovation Demonstration Zone (Grant No. ZJ2021-ZD-008); Research Grants Council of Hong Kong (GRF 17201822).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. M. H. Tran and B. Fei, "Compact and ultracompact spectral imagers: technology and applications in biomedical imaging," J. Biomed. Opt. **28**(04), 040901 (2023).
2. D. L. Marks, D. S. Kittle, H. S. Son, *et al.*, "Gigapixel imaging with the aware multiscale camera," Opt. Photonics News **23**(12), 31 (2012).

Optics EXPRESS

3. N. Saha, M. S. Ifthekhar, N. T. Le, *et al.*, "Survey on optical camera communications: challenges and opportunities," IET Optoelectron. **9**(5), 172–183 (2015).
4. X. Dun, H. Ikoma, G. Wetzstein, *et al.*, "Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging," Optica **7**(8), 913–922 (2020).
5. S. Banerji, M. Meem, A. Majumder, *et al.*, "Imaging with flat optics: metalenses or diffractive lenses?" Optica **6**(6), 805–810 (2019).
6. X. Xiao, Y. Zhao, X. Ye, *et al.*, "Large-scale achromatic flat lens by light frequency-domain coherence optimization," Light: Sci. Appl. **11**(1), 323 (2022).
7. E. Bayati, R. Pestourie, S. Colburn, *et al.*, "Inverse designed extended depth of focus meta-optics for broadband imaging in the visible," Nanophotonics **11**(11), 2531–2540 (2022).
8. J. Zhang, X. Dun, J. Zhu, *et al.*, "Large numerical aperture metalens with high modulation transfer function," ACS Photonics **10**(5), 1389–1396 (2023).
9. E. Tseng, S. Colburn, J. Whitehead, *et al.*, "Neural nano-optics for high-quality thin lens imaging," Nat. Commun. **12**(1), 6493 (2021).
10. F. Presutti and F. Monticone, "Focusing on bandwidth: achromatic metalens limits," Optica **7**(6), 624–631 (2020).
11. J. Engelberg and U. Levy, "Achromatic flat lens performance limits," Optica **8**(6), 834–845 (2021).
12. L. Huang, S. Colburn, A. Zhan, *et al.*, "Full-color metaoptical imaging in visible light," Adv. Photonics Res. **3**(5), 2100265 (2022).
13. D. W. Sweeney and G. E. Sommargren, "Harmonic diffractive lenses," Appl. Opt. **34**(14), 2469–2475 (1995).
14. S. Wang, P. C. Wu, V.-C. Su, *et al.*, "A broadband achromatic metalens in the visible," Nat. Nanotechnol. **13**(3), 227–232 (2018).
15. Z. Zhang, L. Xie, J. Zhang, *et al.*, "Enhancing photon throughput of miniaturized passive depth-detection cameras via broadband dispersion-engineered metalenses," ACS Photonics **10**(10), 3789–3796 (2023).
16. S. Shrestha, A. C. Overvig, M. Lu, *et al.*, "Broadband achromatic dielectric metalenses," Light: Sci. Appl. **7**(1), 85 (2018).
17. J. E. Fröch, P. K. Chakravarthula, J. Sun, *et al.*, "Beating bandwidth limits for large aperture broadband nano-optics," arXiv, (2024).
18. S. Pinilla, J. E. Fröch, S. R. Miri Rostami, *et al.*, "Miniature color camera via flat hybrid meta-optics," Sci. Adv. **9**(21), eadg7297 (2023).
19. R. Maman, E. Mualem, N. Mazurski, *et al.*, "Achromatic imaging systems with flat lenses enabled by deep learning," ACS Photonics **10**(12), 4494–4500 (2023).
20. Y. Dong, B. Zheng, H. Li, *et al.*, "Achromatic single metalens imaging via deep neural network," ACS Photonics **11**(4), 1645–1656 (2024).
21. P. Chakravarthula, J. Sun, X. Li, *et al.*, "Thin on-sensor nanophotonic array cameras," ACM Transactions on Graphics (TOG) **42**(6), 1–18 (2023).
22. N. Mohammad, M. Meem, B. Shen, *et al.*, "Broadband imaging with one planar diffractive lens," Sci. Rep. **8**(1), 2799 (2018).
23. P. Wang, N. Mohammad, and R. Menon, "Chromatic-aberration-corrected diffractive lenses for ultra-broadband focusing," Sci. Rep. **6**(1), 21545 (2016).
24. V. Sitzmann, S. Diamond, Y. Peng, *et al.*, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," ACM Trans. Graph. **37**(4), 1–13 (2018).
25. Q. Sun, E. Tseng, Q. Fu, *et al.*, "Learning rank-1 diffractive optics for single-shot high dynamic range imaging," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2020), pp. 1386–1396.
26. Q. Sun, J. Zhang, X. Dun, *et al.*, "End-to-end learned, optically coded super-resolution spad camera," ACM Trans. Graph. **39**(6), 1–12 (2020).
27. X. Liu, L. Li, X. Liu, *et al.*, "Investigating deep optics model representation in affecting resolved all-in-focus image quality and depth estimation fidelity," Opt. Express **30**(20), 36973–36984 (2022).
28. G. Côté, S. Thibault, J.-F. Lalonde, *et al.*, "Joint optical design for computer vision tasks: challenges and solutions," in *International Optical Design Conference 2023*, vol. 12798 (SPIE, 2023), pp. 51–55.
29. A. Murli, L. D. Amore, and V. De Simone, "The wiener filter and regularization methods for image restoration problems," in *Proceedings 10th International Conference on Image Analysis and Processing*, (Ieee, 1999), pp. 394–399.
30. K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," Opt. Express **17**(22), 19662–19673 (2009).
31. E. Huggins, "Introduction to fourier optics," Phys. Teach. **45**(6), 364–368 (2007).
32. K. Yanny, K. Monakhova, R. W. Shuai, *et al.*, "Deep learning for fast spatially varying deconvolution," Optica **9**(1), 96–99 (2022).
33. K. Monakhova, J. Yurtsever, G. Kuo, *et al.*, "Learned reconstructions for practical mask-based lensless imaging," Opt. Express **27**(20), 28075–28090 (2019).
34. H. Wei, X. Liu, X. Hao, *et al.*, "Modeling off-axis diffraction with the least-sampling angular spectrum method," Optica **10**(7), 959–962 (2023).
35. K. He, X. Zhang, S. Ren, *et al.*, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016), pp. 770–778.

Optics EXPRESS

36. A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *CVPR 2011*, (IEEE, 2011), pp. 193–200.
37. F. Yasuma, T. Mitsunaga, D. Iso, *et al.*, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," IEEE Trans. on Image Process. **19**(9), 2241–2253 (2010).
38. R. M. Nguyen, D. K. Prasad, and M. S. Brown, "Training-based spectral reconstruction from a single rgb image," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VII 13*, (Springer, 2014), pp. 186–201.