Global Search Optics: Automatically Exploring Optimal Solutions to Compact Computational Imaging Systems

Yao Gao[1], Qi Jiang[1], Shaohua Gao[1], Lei Sun[1], Kailun Yang[2,3], and Kaiwei Wang[1,†]

***Abstract*—The popularity of mobile vision creates a demand for advanced compact computational imaging systems, which call for the development of both a lightweight optical system and an effective image reconstruction model. Recently, joint design pipelines come to the research forefront, where the two significant components are simultaneously optimized via data-driven learning to realize the optimal system design. However, the effectiveness of these designs largely depends on the initial setup of the optical system, complicated by a non-convex solution space that impedes reaching a globally optimal solution. In this work, we present Global Search Optics (GSO) to automatically design compact computational imaging systems through two parts: (i) Fused Optimization Method for Automatic Optical Design (OptiFusion), which searches for diverse initial optical systems under certain design specifications; and (ii) Efficient Physic-aware Joint Optimization (EPJO), which conducts parallel joint optimization of initial optical systems and image reconstruction networks with the consideration of physical constraints, culminating in the selection of the optimal solution. Extensive experimental results on the design of three-piece (3P) sphere computational imaging systems illustrate that the GSO serves as a transformative end-to-end lens design paradigm for superior global optimal structure searching ability, which provides compact computational imaging systems with higher imaging quality compared to traditional methods. The source code will be made publicly available at https://github.com/wumengshenyou/GSO.**

***Index Terms*—Computational imaging, end-to-end lens design, image reconstruction, global optimization**

I. Introduction

We are heading to a new era of mobile vision, characterized by increasingly compact imaging optical systems. These advancements shift more correction tasks from traditional optical design to image reconstruction algorithms, a process central to computational imaging [1]. Traditionally, the optical system and the image reconstruction model in computational imaging have been designed sequentially and separately, as depicted in Fig. 1(a), which may show curtailed performance induced by incompatibility between the two components. Recent years have seen the rise of joint design pipelines that effectively bridge the gap between optical design and algorithmic development [2]–[4]. These paradigms leverage differentiable imaging simulation models within an automatic differentiation (AD) framework, enabling the joint optimization of optical systems and image reconstruction models. Focusing on final image quality metrics, the joint design mode moves away from traditional optical design measures like the Point Spread Function (PSF) or the Modulation Transfer Function (MTF).

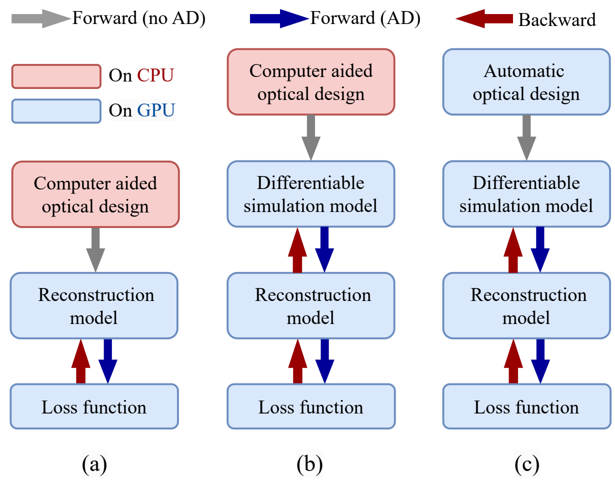


Fig. 1. Comparison of the design modes for computational imaging systems. (a) shows the separate, sequential design mode. (b) shows the joint design mode that requires manual determination of the initial structures. (c) shows the proposed GSO paradigm (joint design mode in which the algorithm automatically provides the initial structures).

This paradigm has been widely applied successfully to the design of single-element optical systems, *e.g.*, Diffractive Optical Element (DOE) or metasurface [2], [5], [6]. Meanwhile, considerable efforts have been made to expand the paradigm to compound optical systems composed of multiple refractive optical elements [3], [4], [7] and further expand the optimization variables to the full set of lens parameters [8], [9]. However, the design of compound lenses presents a significant challenge due to their highly non-convex nature, making it almost impossible to commence with a random set of parameters. Typically, a preliminary design exhibiting basic functional performance is developed first. This initial design is then refined through a process of joint optimization. As illustrated in Fig. 1(b), even

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 12174341 and in part by Hangzhou SurImage Technology Company Ltd.

[1]Y. Gao, Q. Jiang, S. Gao, L. Sun, and K. Wang are with the State Key Laboratory of Extreme Photonics and Instrumentation, Zhejiang University, Hangzhou 310027, China (e-mail: gaoyao@zju.edu.cn; qijiang@zju.edu.cn; gaoshaohua@zju.edu.cn; leo_sun@zju.edu.cn; wangkaiwei@zju.edu.cn).

[2]K. Yang is with the School of Robotics, Hunan University, Changsha 410012, China (E-mail: kailun.yang@hnu.edu.cn).

[3]K. Yang is also with the National Engineering Research Center of Robot Visual Perception and Control Technology, Hunan University, Changsha 410082, China.

[†]Corresponding author: Kaiwei Wang.

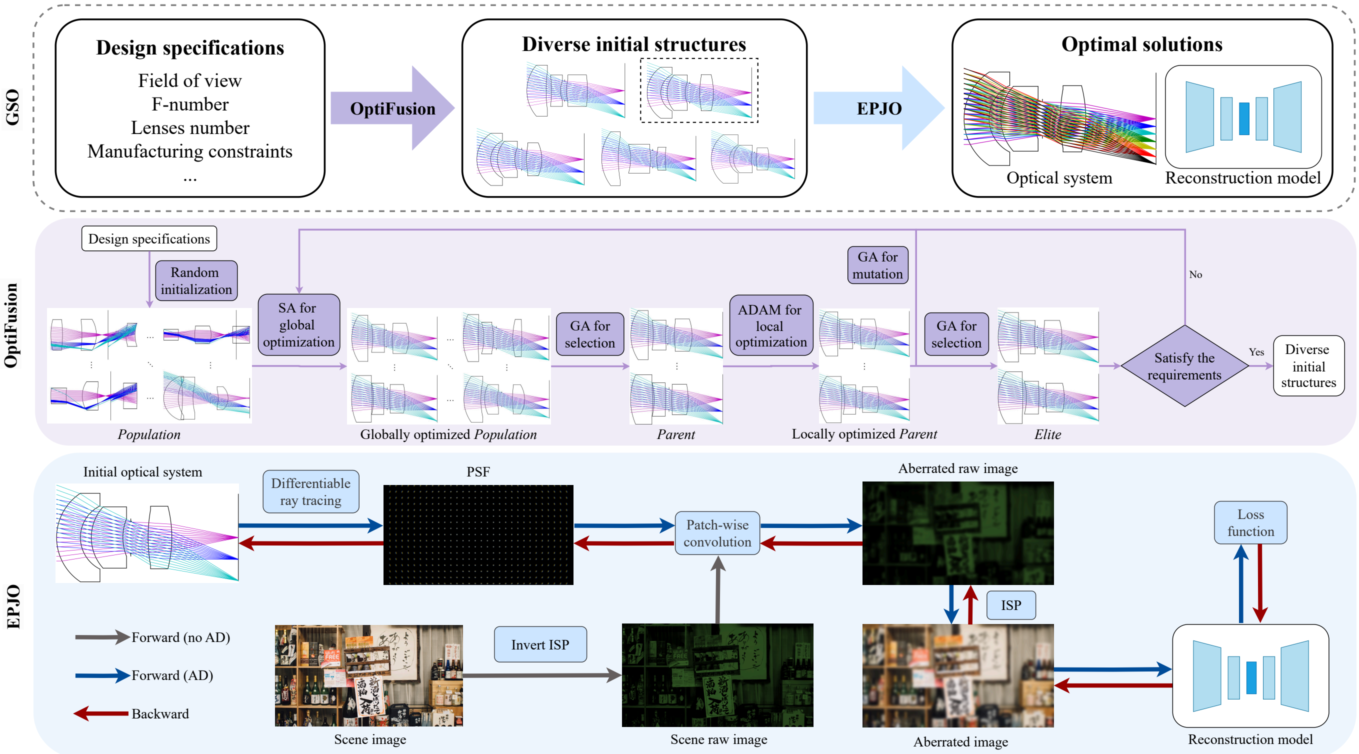


Fig. 2. Overview of our compact computational imaging systems design method. GSO (Global Search Optics) includes OptiFusion (Fused Optimization Method for Automatic Optical Design) and EPJO (Efficient Physic-aware Joint Optimization). OptiFusion fuses Simulated Annealing (SA), Genetic Algorithm (GA), and ADAM to automatically search for initial structures with sufficient diversity based on traditional optical design metrics. EPJO includes an enhanced differentiable simulation model that incorporates differentiable ray tracing, patch-wise convolution, as well as an Image Signal Processing (ISP) pipeline. Additionally, EPJO incorporates customized memory-efficient techniques that enable parallel joint optimization of the initial structures discovered by OptiFusion and image reconstruction models, within reasonable computational resources. This approach allows us to select the jointly optimal solutions based on the final reconstructed image quality metrics.

with the involvement of the image reconstruction network, the traditional method of manually restricting the overall design space based on optical design principles, does not obviate the necessity for skilled personnel. To this intent, this paper highlights a key issue: the traditional methods often fails to identify an initial lens design that closely approximates the global optimum. This shortcoming primarily arises from the potential disconnect between traditional optical design metrics and the final reconstructed image quality metrics.

To address this issue, recent studies on joint optimization propose to start from randomly initialized configurations, leveraging curriculum learning to reduce dependence on an initial design [10], [11]. Nevertheless, these approaches often overlook the intricate and stringent manufacturing constraints associated with optical systems, potentially leading to the limitations of the optimized results in practical applications. Additionally, the direct joint optimization of both the optical system and the image reconstruction model consumes considerably more computational resources than traditional optical design. Joint optimization starting from random configurations could, therefore, prolong the design process and restrict the breadth of the solution space available for exploration.

In this work, we introduce Global Search Optics (GSO), a comprehensive end-to-end lens design framework as shown in Fig. 1(c), which bypasses the requirement for manual initial setting determination and features robust global search capabilities. For the sake of design efficiency and performance, we believe that establishing sound initial structures based on traditional optical design metrics remains essential for our joint design approach. Although some studies have proposed a Deep Neural Network (DNN) framework to automatically infer lens design starting points tailored to the desired specifications, the diversity of output lens designs for the same specifications is limited [12], [13]. Uniquely, GSO includes the Fused Optimization Method for Automatic Optical Design (OptiFusion), which combines Simulated Annealing (SA), Genetic Algorithm (GA), and ADAM to autonomously find initial structures with adequate diversity rooted in traditional optical design metrics. GSO also includes Efficient Physic-aware Joint Optimization (EPJO), featuring an advanced differentiable simulation model and customized memory-efficient techniques. This allows for parallel joint optimization of initial structures identified by OptiFusion and image reconstruction networks, efficiently using computational resources to select the optimal solution based on the final image quality metrics. Furthermore, EPJO considers the intricate physical constraints of optical systems and the categorical nature of glass materials to ensure practical manufacturability. The overview of GSO is shown in Fig. 2.

To manage the complexity of the solution space and achieve the compact design of the computational imaging systems, we highlight GSO's enhanced global search capability by designing Three-Piece (3P) spherical computational imaging systems, illustrating a marked improvement over traditional manual methods for initial structure determination. Additionally, our experimental findings suggest a lack of direct correlation between traditional optical design metrics and the final reconstructed image quality metrics, reinforcing the rationale for increasing the diversity of initial optical designs to avoid overlooking optimal solutions. We have also conducted a rigorously controlled comparative study, which shows that the joint design approach consistently outperforms the traditional separate and sequential design methods in achieving superior image quality. To summarize, our key contributions are:

- Introduction of Global Search Optics (GSO), a comprehensive end-to-end lens design framework that thoroughly and autonomously explores the solution space for compact computational imaging systems.
- Validation of GSO's superior global search capability through comparison with traditional method of initial structure determination, and demonstration of the weak correlation between traditional optical design metrics and final reconstructed image quality metrics.
- Execution of definitive experiments that establish the joint design approach's substantial enhancement in the performance of computational imaging systems over the separate design mode.

II. Related Work

A. Computational Imaging for Compact Optical Systems

The aberration-induced image blur is inevitable for Compact Optical Systems (COS) due to insufficient lens groups for aberration correction [14], [15]. To this intent, computational imaging methods [16], [17] appear as a preferred solution, where optical designs with necessary optical components are equipped with an image reconstruction model. Early efforts have been made to solve the inverse image reconstruction problem through model-based methods [18], [19]. Recently, learning-based methods [20]–[29] have been widely explored for delivering more impressive results of computational imaging for COS, which benefits from the blooming development of image restoration [30], [31], image super-resolution [30], [32], [33] and image deblur [31], [34], [35] methods. Further research has developed deep learning frameworks for the joint optimization of COS and reconstruction models, aiming to perfectly align them and thus enhance overall imaging performance [2], [3], [11], [36], [37]. Traditionally, joint design has relied on manually crafted lenses as initial points [2], [3], [36] or employed strategies like curriculum learning [11] for optimizing random initial lenses and reconstruction models, somewhat restricting the breadth of global search capability. Considering these limitations, this work introduces Global Search Optics (GSO), a novel framework for the design of compact computational imaging systems, to automatically generate a variety of starting points for joint optimization, and efficiently achieve joint optimization of all starting points and reconstruction models.

B. Automatic Optical Design

In the field of joint optimization of optical systems and post-processing models, generating a variety of initial optical system structures is essential. This need highlights the importance of automatic optical design, which seeks to develop algorithms capable of minimizing or even eliminating manual intervention in the design process. The Damped Least Squares (DLS) method, introduced by Kenneth Levenberg [38], has been favored in engineering for its rapid convergence. However, DLS often becomes trapped in local minima, and it requires considerable expertise to establish a robust initial structure, limiting the potential for full automation. Efforts have been made to automate the inference of lens design starting points using Deep Neural Networks (DNN) tailored to specific requirements [13], [39]. Yet, the lack of a comprehensive optical system database restricts the diversity of the outputs, and the model is confined to basic specifications like effective focal length, F-number, and field of view, without accommodating more complex physical constraints. As algorithms and computational power have evolved, various heuristic global search algorithms, *e.g.*, Simulated Annealing (SA), Genetic Algorithm (GA), Ant Colony Algorithm (ACA), Particle Swarm Optimization (PSO), and Tabu Search (TS), have become prevalent in automatic optical design [40]–[45]. Nevertheless, the purpose of the above works is still to automatically design the optimal optical system under traditional design metrics, and the diversity of design results cannot be guaranteed. Consequently, we propose the Fused Optimization Method for Automatic Optical Design (OptiFusion), which combines Simulated Annealing (SA), Genetic Algorithm (GA), and ADAM to automatically search for diverse initial structures.

C. Joint Optimization of Optical Systems and Image Processing Algorithms

The joint optimization of optical systems and image processing algorithms represents a groundbreaking paradigm that has gained traction in recent years [2]–[5], [46], [47]. This paradigm has been applied successfully to the design of single-element optical systems composed of a single Diffractive Optical Element (DOE) or metasurface [2], [5], [48]–[53] and has also been applied to the design of hybrid systems composed of an idealized thin lens combined with a DOE as an encoding element [6], [47], [54]–[58]. Most recently, there has been an effort to expand the paradigm to compound optical systems composed of multiple refractive optical elements [3], [4], [7]–[9], [11], [13], [59], [60]. However, many of these studies have neglected the intricate physical constraints inherent in real-world applications of optical systems [3], [4], [7], [59], [60]. Some efforts have merely imposed basic constraints, like ray angle [8], [11], which do not adequately address manufacturability concerns. Furthermore, the substantial computational memory required for joint optimization continues to be a significant challenge, with some researchers questioning the feasibility of fully optimizing with the available computational resources [9], [23], [28]. This work proposes Efficient Physic-aware Joint Optimization (EPJO) to address these challenges. EPJO not only takes into

Algorithm 1 Implementation steps of OptiFusion

Input: Design specifications, number of generations in GA (N), number of $Individuals$ (m)
Output: The last generation of $Elite$ (Z_N)

```
X_1 ← Initialization()                ▷ Random Initialization
for g = 1, 2, ..., N do
    X'_g ← SA(X_g)                    ▷ Global Optimization
    Y_g ← SelectParent(X'_g)          ▷ Select Parent
    Y'_g ← ADAM(Y_g)                  ▷ Local Optimization
    Z_g ← SelectElite(Y'_g)           ▷ Select Elite
    M_g ← Mutate(Y'_g)                ▷ Mutate Parent
    X_{g+1} ← Merge(M_g, Z_g)         ▷ Next Generation
end for
return Z_N
```

account the complex physical constraints of optical systems and the categorical nature of glass materials to ensure their manufacturability but also achieves efficient joint optimization through customized memory-efficient techniques.

III. OptiFusion: Proposed Method for Automatic Optical Design

OptiFusion is an evolutionary algorithm designed to automatically generate diverse initial optical systems for subsequent joint optimization. This method combines Genetic Algorithms (GA), Simulated Annealing (SA), and ADAM to optimize spot size and meet physical constraints (Sec. III-A). The foundational concept of OptiFusion is based on evolutionary theory, where all optical systems constitute a $Population$, and each system within is considered an $Individual$. OptiFusion begins with random initialization of the $Population$ (Sec. III-B). Throughout the evolutionary process, each generation applies SA for preliminary global optimization (Sec. III-C), followed by selection of a subset of the globally optimized $Population$ as the $Parent$, using GA's selection mechanism (Sec. III-D). ADAM then performs local optimization on the $Parent$ (Sec. III-E). A select portion of this locally optimized $Parent$ group is designated as $Elite$. Should the evolutionary process continue, the $Parent$ undergoes mutation and is merged with the $Elite$ for further optimization in the subsequent generation (Sec. III-F). If the evolutionary process concludes, the $Elite$ is finalized as the output. The specific procedures are outlined in Algorithm 1.

A. Loss Function of OptiFusion

OptiFusion models a compact lens as a stack of several spherical glass elements, characterized by their curvatures (c), glass and air spacings (s), and the refractive indexes (n) and Abbe numbers (v) at the "d" Fraunhofer line ($587.6nm$). Following [3], we employ the approximate dispersion model $n(\lambda) \approx A + B/\lambda^2$ to retrieve the refractive index at any wavelength λ, where A and B follow the definition of the "d"-line refractive index and Abbe number. Once the field of view and aperture size are set, ensuring no vignetting occurs, we express the normalized lens parameters — including curvatures, spacings, refractive indexes, and Abbe numbers — as an n-dimensional vector

$$\boldsymbol{x} = (x^{(1)}, x^{(2)}, ..., x^{(n)})^T \in \mathbb{R}^n. \tag{1}$$

The primary objective is to optimize $\boldsymbol{x}$ to minimize a specific loss function $\mathcal{L}(\boldsymbol{x})$.

Imaging Quality Loss. Traditional lens designs focus on straightforward metrics such as the spot RMS radius, where memory demands for computing are relatively moderate. In OptiFusion, to expedite the search for viable initial structures, we integrate a spot loss ($\mathcal{L}_s$) and a longitudinal chromatic aberration loss ($\mathcal{L}_{lc}$) to assess an optical system:

$$\mathcal{L}_{iq} = \mathcal{L}_s + \lambda_{lc}\mathcal{L}_{lc}. \tag{2}$$

Here, $\mathcal{L}_s$ quantifies the average spot RMS radius across all sampled fields of view and wavelengths. And $\mathcal{L}_{lc}$ accounts for the average longitudinal chromatic aberration, not addressed by $\mathcal{L}_s$. We typically set λ_{lc} at 0.25 to maintain an optimal balance.

Physical Constraint Loss. For basic parameters $\boldsymbol{x}$, we straightforwardly constrain their normalized values within the range $[0, 1]$. However, for key physical properties, *e.g.*, effective focal length and total system length, which are derived from $\boldsymbol{x}$, it's imperative to incorporate a soft physical constraint loss ($\mathcal{L}_{pc}$) to align with design specifications using the Lagrangian approach. Suppose there are n_i physical quantities to be constrained, with each quantity q_i subject to a lower threshold $q^{(i)}_{min}$ and an upper threshold $q^{(i)}_{max}$, along with a specified weight α_i. The $\mathcal{L}_{pc}$ is then expressed as:

$$\mathcal{L}_{pc} = \frac{1}{n_i}\sum_i \alpha_i[\max(q^{(i)}_{min} - q_i, 0) + \max(q_i - q^{(i)}_{max}, 0)]. \tag{3}$$

This formulation implies a linear penalty for any deviation of q_i from the interval $[q^{(i)}_{min}, q^{(i)}_{max}]$, ensuring a smooth optimization process under these constraints.

OptiFusion Loss. The overall loss function for OptiFusion, denoted as $\mathcal{L}_{OF}$, is formulated as:

$$\mathcal{L}_{OF} = \mathcal{L}_{pc} + \lambda_{iq}\mathcal{L}^2_{iq}. \tag{4}$$

Here, λ_{iq} is set to 20, balancing the emphasis on imaging quality with other design considerations, such as effective focal length and total system length. And the unit of $\mathcal{L}_{OF}$ is millimeters. When multiple working object distances are specified in the design, the average value of $\mathcal{L}_{OF}$ across all distances serves as the aggregate loss for optimization purposes.

B. Initialization

To reduce reliance on manual input from optical designers and enable fully automated design, OptiFusion begins with the random initialization of the $Population$, which comprises m $Individuals$ expressed as:

$$X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_m\}. \tag{5}$$

Each $\boldsymbol{x}_i$ is a randomly initialized individual, structured as per Eq. (1). In terms of the parameters, normalized curvatures and spacings within $\boldsymbol{x}_i$ are randomized within the range $[0, 1]$. Differently, the normalized refractive indexes and Abbe numbers are set to either 0 or 1, based on established optical design insights suggesting that extreme values of refractive indexes and Abbe numbers often enhance the imaging performance of simple optical systems.

C. SA for Preliminary Global Optimization of Population

Simulated Annealing (SA) is a heuristic algorithm that mimics the thermodynamic process of cooling to achieve global optimization by potentially accepting suboptimal solutions to escape local minima. Unlike gradient-based methods such as ADAM or DLS, SA does not require derivative information, thereby reducing computational demands. Thus, SA is particularly suited for a preliminary global search when facing a large number of highly inferior $Individuals$ to be optimized.

SA iteratively optimizes $Population$. During each iteration, assuming that $\forall \boldsymbol{x}_i \in X$, we calculate the loss $\mathcal{L}_i$ based on Eq. (4) and adjust the annealing temperature to improve adaptability:

$$T_{SA} = \lambda_{SA}\mathcal{L}_i, \tag{6}$$

where λ_{SA} is predefined as 0.1. A random perturbation $\Delta\boldsymbol{x} \in (-0.1, 0.1)$ is applied to $\boldsymbol{x}_i$, yielding a new $Individual$ $\boldsymbol{x}'_i$ and its loss $\mathcal{L}'_i$. The change in loss, $\Delta\mathcal{L} = \mathcal{L}'_i - \mathcal{L}_i$, determines the acceptance probability of $\boldsymbol{x}'_i$:

$$P_{SA} = \min(e^{-\Delta\mathcal{L}/T_{SA}}, 1). \tag{7}$$

A random number $\epsilon \in (0, 1)$ is drawn; if $\epsilon < P_{SA}$, $\boldsymbol{x}_i$ is updated to $\boldsymbol{x}'_i$; otherwise, it remains unchanged. Furthermore, SA tracks the best historical solution and its loss $\mathcal{L}_i^{best}$ for each $Individual$, utilizing this information to gauge the progress towards convergence. In general, we define the mean loss value of $Population$ as

$$\mathcal{L}_{mean} = \frac{1}{m}\sum_{i=1}^{m}\mathcal{L}_i^{best}. \tag{8}$$

When the rate of decrease of $\mathcal{L}_{mean}$ is less than the threshold, which is typically set to 0.025, it is considered that the global optimization has reached convergence, and we output the set of historical optimal $Individuals$ for further selection:

$$X' = \{\boldsymbol{x}_1^{best}, \boldsymbol{x}_2^{best}, ..., \boldsymbol{x}_m^{best}\}. \tag{9}$$

D. Selection of Parent

The $Parent$ is selected as a subset of X', denoted as Y:

$$Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{m'}\} \subset X', \tag{10}$$

where $m' = r(0.06m)$ and $r(\cdot)$ represents rounding to the nearest integer. To curate a collection of high-quality and diverse Y from X', we refine the Genetic Algorithm's (GA) selection process to better suit optical design. We begin by defining:

$$\mathcal{L}_{all} = \{\mathcal{L}_1^{best}, \mathcal{L}_2^{best}, ..., \mathcal{L}_m^{best}\}. \tag{11}$$

We then sort X' based on $\mathcal{L}_{all}$ and select Y from X' prioritizing from highest to lowest quality. To prevent the selection of overly similar optical systems and maintain diversity within the $Parent$ group, we measure the Euclidean distance $d = \|\boldsymbol{x}'_i - \boldsymbol{x}'_j\|$ between $\forall \boldsymbol{x}'_i, \boldsymbol{x}'_j \in X'$. If $d \leq 0.2$, only the superior individual is chosen for inclusion in Y.

E. ADAM for Local Optimization of Parent

Despite the quick convergence offered by the Damped Least Squares (DLS) method, its computational speed can significantly decrease as the number of variables increases, due to the necessity for matrix inversion. Alternatively, ADAM [61], known for its efficient local optimization and adaptive learning rate adjustments, is more apt for automatic optical design. ADAM does require gradient information for parameter optimization; however, in cases of the relatively simple $\mathcal{L}_{OF}$, effective optimization can be achieved using the first-order difference quotient as a gradient approximation. This approach avoids the need for differentiable simulation models and substantially reduces memory usage.

Thus, we employ ADAM to optimize the $Parent$ group Y, selected as Sec. III-D, towards local optima. We also implement a cosine annealing learning rate schedule to enhance the robustness of ADAM's optimization process. The optimization steps and convergence criteria align with those described in Sec. III-C, leading to the optimization of the $Parent$, now denoted as Y'.

F. Selection of Elite and Mutation of Parent

The $Elite$ group is selected from the subset of Y' and is denoted as Z:

$$Z = \{\boldsymbol{z}_1, \boldsymbol{z}_2, ..., \boldsymbol{z}_{m''}\} \subset Y', \tag{12}$$

where $m'' = r(0.02m)$. This selection process follows the mechanisms outlined in Sec. III-D. If the process has surpassed the predetermined number of generations, the $Elite$ becomes the final output; otherwise, it is carried over to the next generation to ensure that high-quality optical systems are not discarded through the evolutionary process. Additionally, to expand the exploration of potential solutions for subsequent generations, mutation operations are applied to Y'. $\forall \boldsymbol{y}'_i \in Y'$, a number n_{mut} of variables are randomly altered within the range $[0, 1]$, where n_{mut} is set to $r(0.3n)$. Moreover, the total length of the optical system is kept constant pre- and post-mutation to ensure the rationality of the mutation results. The mutated results, represented as M, along with the $Elite$ Z, are then merged to form the $Population$ X for the next generation.

IV. EPJO: Proposed Pipeline for Joint Optimization

This section outlines the differential imaging simulation model presented in Sec. IV-A, which facilitates simultaneous optimization of the optical system and the image reconstruction network. Sec. IV-B defines the loss function of EPJO. Then we introduce a customized adjoint back-propagation strategy for memory-efficient in Sec. IV-C. Finally, we described the detailed steps of EPJO for joint optimization in Sec. IV-D.

A. Differentiable Imaging Simulation Model

We establish an accurate differentiable simulation model suitable for compact optical systems with large aberrations, which achieves gradient back-propagation from image reconstruction network parameters to optical system parameters.

Differentiable PSF Formation Model. In our differentiable imaging simulation pipeline, the aberration-induced degradation is represented through the energy dispersion of the Point Spread Function (PSF). We employ a ray-tracing-based model for PSF formation that enables accurate and differentiable results. Differentiable ray tracing is achieved by alternating between updating the coordinates of the rays from one interface to the next using the Newton iteration method and updating the direction cosines following Snell's Law as in [3] and [4]. Rays are initially positioned at the entrance pupil, and a ray-aiming correction step [8] is applied to ensure precise simulation of optical systems, particularly those affected by pupil aberrations. Then, rays can be traced to the image plane to obtain the PSF. In compact optical systems with significant geometric aberrations, where diffraction effects can be neglected, the PSF is calculated by Gaussianizing the intersection of the ray with the image plane [36]. Specifically, the intensity distribution of a ray at the image plane intersection is modeled by a Gaussian function:

$$E(m,n) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{r(m,n)^2}{2\sigma^2}). \tag{13}$$

Here, $r(m,n)$ represents the distance between the pixel (m,n) and the ray's center at the image plane, and $\sigma{=}\sqrt{\Delta x^2{+}\Delta y^2}/3$. The superposition of each Gaussian spot results in the final differentiable PSF.

After obtaining PSFs for all sampled fields and wavelengths using the aforementioned methods, we synthesize them into three-channel RGB PSFs. This synthesis utilizes the spectral sensitivity characteristics of the simulated CMOS sensor, as follows:

$$PSF_c(\theta) = \sum_{\lambda_c} W_c(\lambda_c) \cdot PSF(\theta, \lambda_c). \tag{14}$$

Here, θ represents the sampled fields of view, and c represents one of R, G, and B channels. λ_c represents sampling wavelengths of the corresponding channel and $W_c(\lambda_c)$ represents the corresponding normalized wavelength response coefficient. Moreover, it is essential to account for the influence of longitudinal chromatic aberration on the central positioning of each channel within the three-channel RGB PSFs. Therefore, we designate the center of the G-channel PSF as the reference point for the RGB PSFs, adjusting the PSFs of the R and B channels based on their actual central positions. Consequently, we generate the integrated three-channel RGB PSFs across all sampled fields of view.

Patch-wise Convolution and ISP Pipeline. To facilitate the construction of more realistic aberrated images, an Image Signal Processing (ISP) pipeline is employed [62]. Initially, the scene image I_S undergoes sequential applications of inverse Gamma Correction (GC), inverse Color Correction Matrix (CCM), and inverse White Balance (WB) to transform it into the scene raw image I'_S. The inverse ISP pipeline is expressed as:

$$I'_S = P_{WB}^{-1} \circ P_{CCM}^{-1} \circ P_{GC}^{-1}(I_S), \tag{15}$$

where $\circ$ denotes the composition operator, and P_{WB}, P_{CCM}, and P_{GC} represent the procedures for WB, CCM, and GC, respectively.

Subsequently, patch-wise convolution is applied to I'_S. I'_S is divided into $n_h{\times}n_w$ patches, each measuring $s{\times}s$ pixels. It is assumed that PSFs within these patches are spatially uniform. Convolution is then performed between the image patches and their corresponding PSFs, which are then recompiled into the degraded raw image I'_A. Each patch of I'_S is designated as $I'_S(c,h,w)$, where c indicates one of the R, G, and B channels, and h and w denote the patch's position on the image plane. The associated PSF, $PSF(c,h,w)$, is derived by interpolating PSFs across all sampled fields of view and adjusting them by rotating to the correct angle:

$$PSF(c,h,w) = P_{rot}(\sum_{\theta} W(\theta) \cdot PSF(c,\theta)), \tag{16}$$

where P_{rot} represents the rotation process, $PSF(c,\theta)$ is the PSF from a specific field of view and $W(\theta)$ is the normalized interpolation weight determined by the inverse square formula. The degraded raw image patch $I'_A(c,h,w)$ is approximated as:

$$I'_A(c,h,w) \approx PSF(c,h,w) * I'_S(c,h,w). \tag{17}$$

After that, we mosaic the degraded raw image I'_A before adding shot and read noise to each channel. Moreover, we sequentially apply the demosaic algorithm, WB, CCM, and GC to the R-G-G-B noisy raw image, where the aberration-degraded image I_A in the sRGB domain is obtained. The ISP pipeline can be defined as:

$$I_A = P_{GC} \circ P_{CCM} \circ P_{WB} \circ P_{demosaic} \circ (P_{mosaic}(I'_A) + N), \tag{18}$$

where N represents the Gaussian shot and read noise. P_{mosaic} and $P_{demosaic}$ represent the procedures of mosaicking and demosaicking, respectively.

B. Loss Function of EPJO

We define the loss function of EPJO balancing the emphasis on final reconstructed image quality with consideration of intricate physical constraints.

Imaging Quality Loss. We reconstruct aberration-degraded images I_A through an image reconstruction network to produce reconstructed images I_R. To extend the depth of field in compact computational imaging systems, we segment the continuous object distance range into three training depths. The imaging quality loss function is formulated as:

$$\begin{aligned} \mathcal{L}'_{iq} = \frac{1}{3}\sum_j [\mathcal{L}_{mse}(I_{Rj}, I_S) + \lambda_1 \mathcal{L}_{perc}(I_{Rj}, I_S)] \\ + \sum_{j \neq 2} \lambda_2 \mathcal{L}_{mse}(I_{Rj}, I_{R2}), \end{aligned} \tag{19}$$

where $\mathcal{L}_{mse}$ denotes the MSE loss, and $\mathcal{L}_{perc}$ indicates the perceptual loss function based on the pre-trained VGG16 network [63], enhancing alignment with human perception. And $\mathcal{L}_{mse}(I_{Rj}, I_{R2})$ means that we take I_{R2} as a reference to keep reconstructed images depth-invariant. We set λ_1=0.01, λ_2=0.1.

Physical Constraint Loss. Our joint optimization process, EPJO, also imposes strict constraints on relevant physical quantities and aligns glass variables with catalog glasses to ensure manufacturability. The physical constraint loss function is given by:

$$\begin{aligned}\mathcal{L}'_{pc} = \frac{1}{n_i}\sum_i \alpha_i[(\max(q_{min}^{(i)} - q_i, 0) \\ + \max(q_i - q_{max}^{(i)}, 0))]^2 + \mathcal{L}_{gv},\end{aligned} \tag{20}$$

where $\mathcal{L}_{gv}$ minimizes the squared distance between each set of continuous glass variables and the nearest catalog glass:

$$\mathcal{L}_{gv} = \frac{1}{p}\sum_{i=1}^{p} \min(\lambda_n \|n_i - \boldsymbol{n_{cat}}\|_2^2 + \lambda_v \|v_i - \boldsymbol{v_{cat}}\|_2^2), \tag{21}$$

where p is number of lenses and empirically we set λ_n=100, λ_v=0.0004. Unlike Eq. (3), Eq. (20) implies a more severe quadratic penalty instead of a linear penalty for any deviation of q_i from the interval $[q_{min}^{(i)}, q_{max}^{(i)}]$, which is more suitable for optical systems that have already roughly met the specifications.

EPJO Loss. To balance imaging quality and physical constraints effectively, we define the EPJO loss as:

$$\mathcal{L}_{EPJO} = \mathcal{L}'_{pc} + \lambda'_{iq}\mathcal{L}'_{iq}, \tag{22}$$

in which λ'_{iq} is empirically set to 100.

C. Adjoint Back-propagation for Memory Savings

When the loss function is in the image space (*e.g.* Eq. (19)) which involves calculating a large number of PSFs, simulating high-resolution images, and going through image reconstruction networks, straightforward back-propagation requires unaffordable device memory. The work of [4] has proposed an adjoint back-propagation approach that splits forward computations into multiple passes to alleviate the back-propagation memory issue. Unfortunately, our differentiable imaging simulation model is based on the convolution of PSFs and images rather than relying on image rendering in which many millions of Monte Carlo rays are sampled [4], which makes existing adjoint methods not directly applicable. Therefore, we propose a customized adjoint back-propagation method for our differentiable imaging simulation model.

Fundamentally, the device memory of our differentiable simulation model is mainly consumed in storing intermediate variables for calculating a large number of PSFs. Therefore, we propose a novel adjoint approach to manually separate the calculation of PSFs from subsequent steps. Given the loss function $\mathcal{L}_{EPJO}$, our goal is to evolve variable parameters θ iteratively towards an optimal θ' using gradient-based optimization, and this requires computing $\partial\mathcal{L}_{EPJO}/\partial\theta$, the partial derivatives that indicate how design parameters affect the error metric locally. Assuming $F(\theta)$ is a continuous function of θ for calculating PSFs, $\partial\mathcal{L}_{EPJO}/\partial\theta$ can be represented by the chain rule as:

$$\frac{\partial\mathcal{L}_{EPJO}}{\partial\theta} = \frac{\partial\mathcal{L}_{EPJO}}{\partial F(\theta)}\frac{\partial F(\theta)}{\partial\theta}. \tag{23}$$

According to Eq. (23), after calculating PSFs, we perform the first back-propagation to obtain $\partial F(\theta)/\partial\theta$, the partial derivatives of PSFs with respect to the optical system parameters. Then, we store $F(\theta)$ and $\partial F(\theta)/\partial\theta$ while clearing the computation graph and intermediate variables because the memory consumption for storing $F(\theta)$ and $\partial F(\theta)/\partial\theta$ is much smaller. Subsequently, we take PSFs as a differentiable input to calculate $\mathcal{L}_{EPJO}$. Finally, we conduct a second back-propagation to obtain $\partial\mathcal{L}_{EPJO}/\partial F(\theta)$, and thus we can obtain $\partial\mathcal{L}_{EPJO}/\partial\theta$ according to Eq. (23). Since the computation time in joint optimization is mainly spent on the image reconstruction network rather than calculating PSFs, the additional time required to perform the first back-propagation to calculate $\partial F(\theta)/\partial\theta$ can be ignored. Therefore, such an adjoint back-propagation approach significantly reduces memory consumption to an affordable level without affecting the optimization time.

Algorithm 2 Implementation steps of EPJO

Input: Lenses number (p), initial optical system (O) and randomly initialized image reconstruction model (R)

Output: Jointly optimized optical system (O'_p) and image reconstruction model (R'_p)

```
{O'_0, R'_0} ← JointOptimize({O, R})      ▷ Continuous Glass
for j = 1, 2, ..., p do
    O_j ← SelectGlass(O'_{j-1}, j)        ▷ Select Catalog Glass
    R_j ← R'_{j-1}
    {O'_j, R'_j} ← JointOptimize({O_j, R_j})
end for
return {O'_p, R'_p}
```

D. Implementation Steps of Joint Optimization

Unlike training individual image reconstruction networks, joint optimization requires a focused approach that takes into account the distinct characteristics of both optical systems and networks. Therefore, we have tailored exclusive steps specifically for joint optimization, as outlined in Algorithm 2.

Conditions for Convergence. In each epoch, the optimization process involves n_O iterations for adjusting the optical system parameters. Within each iteration dedicated to the optical system, n_R iterations are performed to fine-tune the network parameters, ensuring their adaptability to change in the optical system. After each iteration of optimizing the optical system, we evaluate its performance on the validation dataset. The combination of optical system and network parameters that yields the best performance among the n_O iterations of each epoch is selected as the optimal configuration for that epoch. If the best performance of the current epoch fails to surpass the best performance of the previous epoch, it is considered that the joint optimization has converged to an optimal state. We empirically set n_O=5, n_R=1000 to ensure the smooth progress of the entire optimization process.

Replacing Continuous Glass Variables With Real Glass. Given the discrete nature of glass materials, our initial approach involves optimizing the refractive index and Abbe number of the material as continuous variables. Using Eq. (21) as a guiding principle, we gradually move towards the realization of actual materials within the solution space. Subsequently, to translate these continuous variables into the desired catalog glass material, we employ a step-wise substitution method. This involves systematically selecting the glass materials that require replacement in a prescribed order. Once the computational imaging system is optimized to satisfy convergence conditions, we proceed to replace the chosen continuous variables with the closest matching material from our glass library. This replacement is based on Eq. (21), and the corresponding variables are subsequently fixed. The process continues with retraining until convergence is achieved, followed by the replacement of the next glass component, until all glass materials have been replaced.

V. Experiments And Results

Under two distinct specifications for three-piece (3P) spherical lens designs, we compare the designs generated by OptiFusion with those produced by the optical design software CODE V and evaluate the proficiency of OptiFusion in automatically crafting diverse initial structures for EPJO in Sec. V-A. In Sec. V-B, we furthermore evaluate the global search capability of GSO by comparing it with the method of manually determining the initial structures, aided by CODE V. Additionally, we investigate the interrelationship between traditional optical design metrics and the final reconstructed image quality metrics. Finally, in Sec. V-C, we test the enhancement in reconstructed image quality achieved by the joint design mode, in contrast to the separate design mode.

A. Experiments on Automatic Optical Design

We establish two representative specifications for three-piece (3P) spherical lens designs, as outlined in Table I. Specifically, 3P-I necessitates a narrow field of view coupled with a substantial aperture, whereas 3P-II requires a broad field of view and a smaller aperture. To ensure a comprehensive depth of field, we establish three distinct working distances for each design specification. $\mathcal{L}_{OF}$ (Eq. (4)) serves as the unifying evaluation metric for optical systems. The physical attributes that necessitate soft constraints include effective focal length, distortion, air edge spacing, glass edge thickness, back working distance, total length, and image height. The respective weights assigned to these attributes, based on their constraint ranges, are $\{0.1, 1, 0.1, 0.1, 0.05, 0.01, 1\}$. Subsequently, we utilize OptiFusion and CODE V to independently design optical systems that adhere to the specified design requirements. When utilizing OptiFusion, we set the number of generations N to 14 and the number of $Individuals$ m to 2000. The final $Elite$ set comprises 40 distinct optical systems, from which we select the average values of the top-4 evaluation metrics. Given the incorporation of heuristic random search algorithms, we employ OptiFusion to design each specification three times to mitigate the effects of randomness. For comparative purposes, we design 4 distinct lenses for each specification, leveraging the assistance of CODE V. The average values of their evaluation metrics are then considered for analysis.

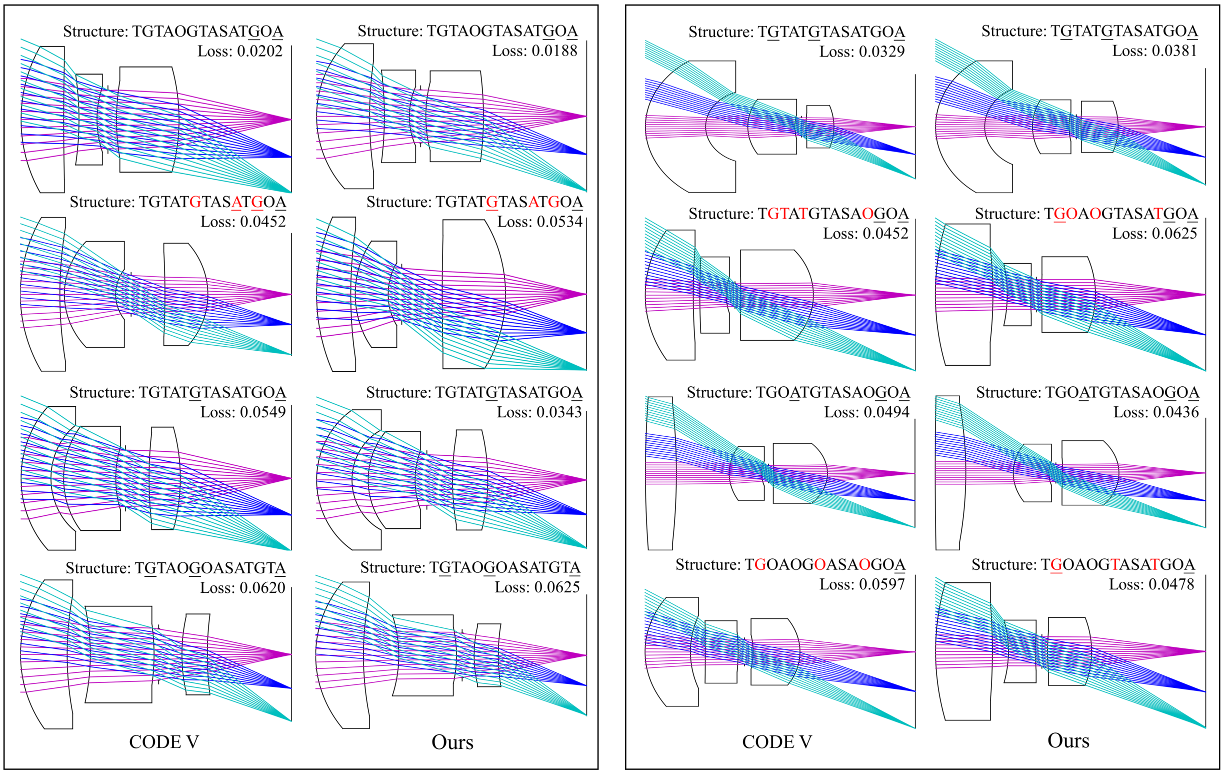
TABLE I
DESIGN SPECIFICATIONS FOR TWO TYPES OF THREE-PIECE (3P) SPHERICAL LENSES.

Parameters	3P-I	3P-II
Field of view	$40°$	$64°$
F-number	2.5	4.0
Effective focal length	$38mm \sim 42mm$	$21mm \sim 25mm$
Working distance	$100m, 10m, 5m$	$5m, 1m, 0.5m$
Distortion	$-2\% \sim 2\%$	$-8\% \sim 8\%$
Curvature	$-0.1 \sim 0.1$	
Semi-diameter	$0mm \sim 20mm$	
Air center spacing	$1mm \sim 15mm$	
Air edge spacing	$1mm \sim 15mm$	
Glass center thickness	$4mm \sim 15mm$	
Glass edge thickness	$5mm \sim 15mm$	
Refractive index	$1.51 \sim 1.76$	
Abbe number	$27.5 \sim 71.3$	
Wavelength	$486nm, 588nm, 656nm$	
Back working distance	$18mm \sim 30mm$	
Total length	$40mm \sim 60mm$	
Image height	$14.16mm \sim 14.44mm$	
Aperture stop position	Between the second and third lenses	

Experimental Results. As outlined in Table II, the design results generated by OptiFusion, across all three iterations, consistently outperform those achieved with the assistance of CODE V software. This superiority is noteworthy, given that OptiFusion demonstrates overall stability despite incorporating a random search mechanism. Fig. 3 offers a more vivid comparison of the design outcomes. It is evident that OptiFusion and CODE V can design similar initial structures when adhering to identical design specifications. Furthermore, the optimal designs arrived at by both methods exhibit a high degree of congruence. Given the intricate relationship between software-assisted design time and the proficiency of designers, we offer only an approximate reference for the average design time. Typically, CODE V assisted designs require approximately 3 hours, whereas OptiFusion significantly reduces this timeframe to just about 1 hour. Additionally, CODE V relies heavily on human design expertise, often commencing with an existing structure as a starting point. In contrast, OptiFusion is a fully automated design approach that can initiate the design process from scratch, solely relying on the provided design specifications. Therefore, as an innovative method for automatic optical design, OptiFusion not only offers a diverse array of initial structures for joint design mode but also holds the potential to emerge as a superior alternative to traditional optical design software, particularly in the realm of simple and compact optical designs.

B. Experiments on Searching for Optimal Solutions

We employ optical systems with evaluation metrics $\mathcal{L}_{OF}$ less than 0.065 searched by OptiFusion in Sec. V-A as the



3P-I 3P-II

Fig. 3. Visual comparison between OptiFusion and CODE V under two design specifications: 3P-I and 3P-II. For each specification, the left-hand side showcases the design results achieved with the assistance of CODE V, whereas the right-hand side displays the most closely matching structure discovered by OptiFusion. The structure of each optical system is named as a sequence of G (glass thickness), A (air spacing), T (convex surface), O (concave surface), and S (aperture stop). The underlined G and A represent corresponding thickness or spacing greater than $10mm$. The different parts are highlighted in red.

TABLE II
COMPARISON BETWEEN OPTIFUSION AND CODE V.

	CODEV	OptiFusion
3P-I	0.0456	**0.0435, 0.0433, 0.0433**
3P-II	0.0468	**0.0384, 0.0379, 0.0385**

initial structures of EPJO for further joint optimization. The number of initial structures we obtain for 3P-I is $\{5, 7, 6\}$ and the number of initial structures we obtain for 3P-II is $\{15, 10, 17\}$. As a comparison, we also perform the same joint optimization on the 2×4 initial structures designed with the assistance of CODEV in Sec. V-A.

Differentiable Imaging Simulation. To match the image height in Table I, we employ a virtual sensor with diagonal d=$28.6mm$. The sensor resolution is set to 1920×1280 pixels, which means the pixel size should be calculated as $12.394\mu m$. We sample 5 wavelengths for each channel based on a quantum efficiency curve that follows the Sony IMX172 sensor similar to [8]. To reasonably control the speed and memory consumption of differentiable imaging simulation, we uniformly sample the PSF of 7 fields of view and get the PSFs of the non-sampling field point by interpolation. We assume that the PSFs in the range of 64×64 pixels are spatially uniform, so every image is split into 30×20 patches that are 64×64 in size.

Data Preparation. We adopt DIV2K [64] which contains 900 images of $2K$ resolution as ground truths and divide these images into the training set and validation set at 8:1. Then, images of different sizes are center-cropped and rotated to 1920×1280 pixels to match the sensor resolution, and images with length or width less than that of the sensor resolution will be discarded. Finally, we have obtained a training set containing 697 images and a validation set containing 92 images.

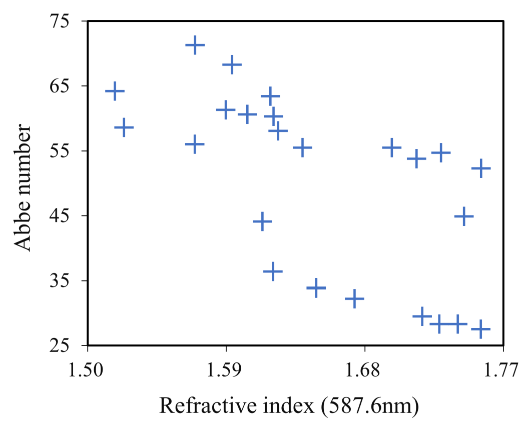


Fig. 4. Catalog glasses that meet the design specifications and are available in stock all year round from the Chengdu Guangming Optoelectronic Corporation in China.

Catalog Glasses. To convert continuous glass variables into catalog glasses, we use glasses that meet the design specifications and are available in stock all year round from the

TABLE III
QUANTITATIVE COMPARISON BETWEEN CODE V ASSISTED JOINT DESIGN METHOD AND GSO.

	CODE V			GSO1			GSO2			GSO3		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
3P-I	29.99	0.8533	0.0772	30.38	0.8636	0.0719	30.24	0.8614	0.0746	30.31	0.8605	0.0727
	29.93	0.8513	0.0809	29.91	0.8556	0.0820	30.12	0.8606	0.0755	29.92	0.8508	0.0806
	29.64	0.8466	0.0833	29.84	0.8527	0.0806	29.80	0.8500	0.0818	29.59	0.8444	0.0838
	29.28	0.8339	0.0929	29.42	0.8406	0.0885	29.67	0.8465	0.0848	29.54	0.8423	0.0878
3P-II	29.43	0.8389	0.0881	29.68	0.8438	0.0839	29.83	0.8484	0.0814	29.70	0.8506	0.0814
	29.32	0.8304	0.0933	29.57	0.8399	0.0879	29.69	0.8398	0.0848	29.68	0.8418	0.0867
	29.10	0.8279	0.0964	29.46	0.8391	0.0881	29.58	0.8396	0.0855	29.60	0.8419	0.0882
	29.12	0.8266	0.0968	29.45	0.8396	0.0881	29.50	0.8430	0.0879	29.57	0.8393	0.0882

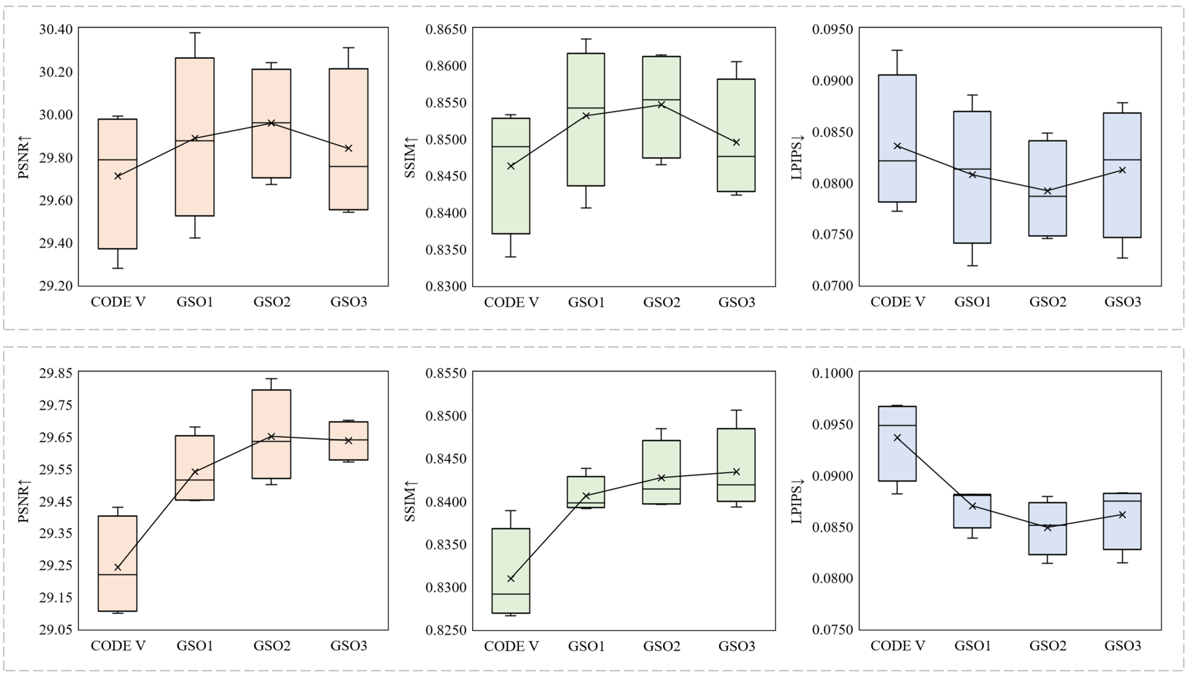


Fig. 5. Visualization of comparison between CODE V assisted joint design method and GSO. The design results of 3P-I and 3P-II are presented from top to bottom and we use PSNR, SSIM, and LPIPS as evaluation metrics which are displayed from left to right.

Chengdu Guangming Optoelectronic Corporation in China, as shown in Fig. 4.

Training Details. We use SwinIR [30] as the image reconstruction network without modifying the architecture. The Residual Swin Transformer Blocks (RSTB) number, Swin Transformer Layer (STL) number, window size, channel number, and attention head number are generally set to 5, 6, 8, 96, and 6, respectively. During training, the patch size and batch size are set to 256×256 and 12 respectively, in which each of the 3 working distances occupies 4 batch size. The ADAM optimizer with different learning rates is utilized, considering the respective characteristics of optical system variables and network variables. Specifically, the learning rates for curvature, spacing, refractive index, and Abbe number are set to 0.0002, 0.02, 0.001, and 0.2 respectively, and the learning rate of the network is 0.0001. To deal with a total of 68 initial structures, we implement EPJO in PyTorch [65] on 30 NVIDIA GeForce RTX 3090 GPUs for 64 hours, in which the average optimization time required for each initial structure is about 32 hours.

Experimental Results. After EPJO completes the joint optimization, we evaluate the results using PSNR, SSIM [66], and LPIPS [63]. The final ranking basis is determined by averaging the rankings obtained from these metrics. Subsequently, we select the top-4 results and compare them to the CODE V assisted joint design outcomes. Table III reveals that, compared to the CODE V assisted joint design method, GSO achieves significant improvements. For the optimal solution of 3P-I, GSO improves PSNR by $0.25dB \sim 0.38dB$, SSIM by $0.0072 \sim 0.0103$, and LPIPS by $0.0026 \sim 0.0053$. Similarly, for the optimal solution of 3P-II, GSO enhances PSNR by $0.25dB \sim 0.40dB$, SSIM by $0.0049 \sim 0.0115$, and LPIPS by $0.0042 \sim 0.0067$. Furthermore, Fig. 5 visually demonstrates that the design results obtained through GSO are overall superior.

The key distinction between the two methods is that GSO incorporates OptiFusion, enabling the automatic generation of a vast array of diverse optical systems that adhere to the design

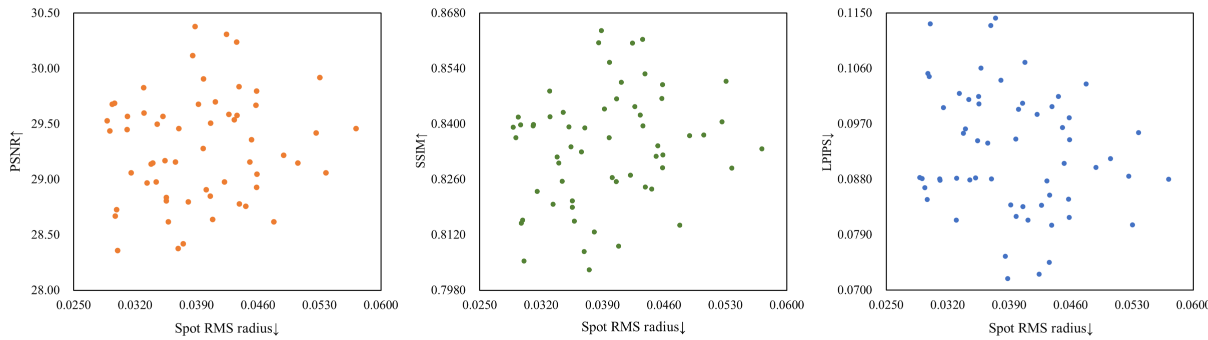


Fig. 6. Correlation between spot RMS radius and reconstructed image quality after joint optimization. The horizontal axis represents the average spot RMS radius of all working distances, sampling fields of view, and sampling wavelengths after joint optimization, and the vertical axis represents the corresponding PSNR, SSIM, and LPIPS.

requirements. Conversely, with assistance of CODE V, only a limited number of high-quality optical systems can be crafted within a constrained timeframe. Consequently, in contrast to manually scouring for a small number of initial structures, GSO significantly enlarges the solution space by embracing a greater diversity of initial structures. This broadening is necessitated by the weak correlation between traditional optical design metrics and the ultimate reconstructed image quality metrics. For instance, while it is imperative for the reconstructed image quality that the spot size remains small, this does not automatically translate to an initial structure with a smaller spot size outperforming others after joint optimization with the reconstruction network. This suggests that the image reconstruction network may have unique and often unforeseen preferences. Therefore, exploring a broader range of diverse initial structures is crucial to prevent overlooking the optimal initial structure and significantly enhancing the likelihood of discovering the global optimal solution.

To further prove our point, Fig. 6 demonstrates the weak correlation between spot RMS radius and reconstructed image quality, in which the horizontal axis represents the average spot RMS radius of all working distances, sampling fields of view and sampling wavelengths after joint optimization, while the vertical axis represents the corresponding PSNR, SSIM, and LPIPS of the reconstructed images. The scattered and irregular data distribution in Fig. 6 indicates that the relationship between spot RMS radius and final reconstructed image quality is unclear, and even if the spot of the optical system is smaller, the final reconstruction quality of the images may still be better. Please refer to the Appendix for detailed results.

C. *Comparison of Joint Design and Separate Design Methods*

Although in theory, the joint design method can explore the solution space more comprehensively compared to the separate design method due to its ability to synchronously optimize the optical system and image reconstruction model, the presence of complex physical constraints does not necessarily guarantee better performance of the joint design method, and the two methods have always lacked fair experiments for quantitative comparison. In this section, we design experiments to investigate the benefit of joint design methods in improving the upper limit of computational imaging system performance. To ensure the fairness of the experiment, the initial structures are consistent with the initial structures of EPJO in Sec. V-B and the loss function is also set to Eq. (22). The difference lies in that the separate design method replaces the reconstructed image in Eq. (22) with the degraded image for optimization and then fixes the designed optical system before training the reconstruction network. In other words, the optical system is independently designed without the reconstruction network, and then the reconstruction network is independently optimized. In addition, all training strategies are consistent with Sec. V-B, ensuring that the only factor affecting the final result is whether the optical system is co-designed with the reconstruction network. As the results of the joint design method have been obtained from Sec. V-B, we only implement the separate design method on 30 NVIDIA GeForce RTX 3090 GPUs for about 72 hours.

Experimental Results. As shown in Fig. 7, the PSNR, SSIM, and LPIPS of degraded and reconstructed images are presented from top to bottom. The horizontal axis of each subgraph represents the order of joint optimization metrics and the vertical axis compares the quality of degraded or reconstructed images between the joint design method and the separate design method. Overall the degraded image quality of the 43 out of 60 separately designed structures is better than that of the jointly designed structures, bringing improvements of $0.04dB$ in average PSNR, 0.0027 in average SSIM, and 0.0076 in average LPIPS. This is because the separate design method aims to design optical systems with degraded image quality as the goal. In contrast, the reconstructed image quality of the 53 out of 60 jointly designed structures is better than that of the separately designed structures, bringing improvements of $0.30dB$ in average PSNR, 0.0081 in average SSIM, and 0.0082 in average LPIPS. Please refer to the Appendix for detailed results. Despite the presence of other interference factors, such as physical constraints and the substitution of continuous glass variables with actual glass during the training process, the joint design method can still significantly enhance the final reconstructed image quality in most instances, even if the initial degraded image quality does not reach the optimal level. The joint design method is particularly well-suited for visual tasks like image reconstruction because it employs

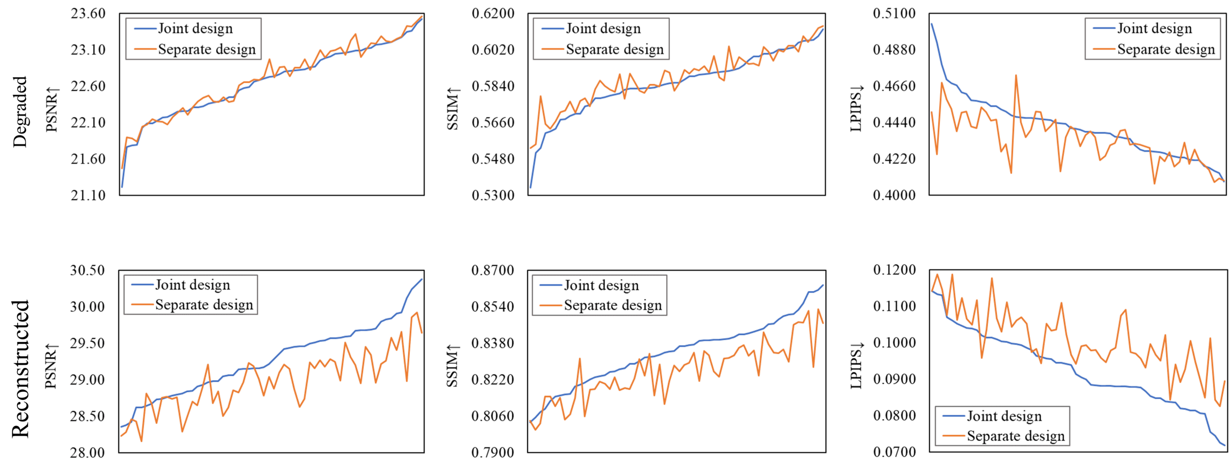


Fig. 7. Quantitative comparison of joint design mode and separate design mode. The imaging quality of the degraded and the reconstructed images are presented from top to bottom. We use PSNR, SSIM, and LPIPS as imaging quality metrics, which are displayed from left to right. The horizontal axis represents the order of joint optimization metrics and the vertical axis compares the quality of degraded or reconstructed images between joint design and separate design methods.

the final reconstructed image quality derived post-algorithm, rather than traditional optical design metrics, as the training objective. This approach allows for the structural parameters of the optical systems to be finely tuned for visual tasks. Moreover, the algorithm parameters for these tasks can undergo profound adjustments that are often unpredictable. This element of unpredictability underscores the critical importance and unique advantage of joint optimization.

VI. Conclusion

We have introduced the GSO, an end-to-end design framework tailored for compact computational imaging systems, capable of autonomously exploring optimal solutions. We have established that GSO's improvement in global search capability stems from the increased diversity of initial structures, highlighting the lack of significant correlation between traditional optical design metrics and the jointly optimized reconstructed image quality. Our conclusive experiments illustrate that the joint design approach markedly boosts the performance of computational imaging systems over the separate design method.

Looking ahead, there is potential to expand OptiFusion's capabilities to autonomously design more intricate optical systems. The extensive number of initial structures required for more complex optical systems can impede the design efficiency of GSO when all options are explored. Furthermore, the development of a comprehensive lens library through OptiFusion could facilitate the analysis of visual task model preferences, substantially narrowing the range of initial structures to be screened. This enhancement would allow the extension of GSO to more intricate computational imaging design tasks. Additionally, such a lens library could also serve to train network models for lens generation, accelerating the inference of suitable initial structures.

References

[1] J. Suo, W. Zhang, J. Gong, X. Yuan, D. J. Brady, and Q. Dai, "Computational imaging and artificial intelligence: The next revolution of mobile vision," *Proceedings of the IEEE*, vol. 111, no. 12, pp. 1607–1639, 2023. 1

[2] V. Sitzmann *et al.*, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," *ACM Transactions on Graphics*, vol. 37, no. 4, pp. 1–13, 2018. 1, 3

[3] Q. Sun, C. Wang, Q. Fu, X. Dun, and W. Heidrich, "End-to-end complex lens design with differentiate ray tracing," *ACM Transactions on Graphics*, vol. 40, no. 4, pp. 1–13, 2021. 1, 3, 4, 6

[4] C. Wang, N. Chen, and W. Heidrich, "dO: A differentiable engine for deep lens design of computational imaging systems," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 905–916, 2022. 1, 3, 6, 7

[5] E. Tseng *et al.*, "Neural nano-optics for high-quality thin lens imaging," *Nature Communications*, vol. 12, no. 1, p. 6493, 2021. 1, 3

[6] J. Chang and G. Wetzstein, "Deep optics for monocular depth estimation and 3D object detection," in *Proc. ICCV*, 2019, pp. 10 192–10 201. 1, 3

[7] E. Tseng *et al.*, "Differentiable compound optics and processing pipeline optimization for end-to-end camera design," *ACM Transactions on Graphics*, vol. 40, no. 2, pp. 1–19, 2021. 1, 3

[8] G. Côté, F. Mannan, S. Thibault, J.-F. Lalonde, and F. Heide, "The differentiable lens: Compound lens search over glass surfaces and materials for object detection," in *Proc. CVPR*, 2023, pp. 20 803–20 812. 1, 3, 6, 9

[9] Y. Zhang *et al.*, "Large depth-of-field ultra-compact microscope by progressive optimization and deep learning," *Nature Communications*, vol. 14, no. 1, p. 4118, 2023. 1, 3

[10] X. Yang, Q. Fu, and W. Heidrich, "Automatic lens design based on differentiable ray-tracing," in *Computational Optical Sensing and Imaging*, 2022, pp. CTh4C–2. 2

[11] ——, "Curriculum learning for ab initio deep learned refractive optics," *arXiv preprint arXiv:2302.01089*, 2023. 2, 3

[12] G. Côté, J.-F. Lalonde, and S. Thibault, "Extrapolating from lens design databases using deep learning," *Optics Express*, vol. 27, no. 20, pp. 28 279–28 292, 2019. 2

[13] ——, "Deep learning-enabled framework for automatic lens design starting point generation," *Optics Express*, vol. 29, no. 3, pp. 3841–3854, 2021. 2, 3

[14] E. Kee, S. Paris, S. Chen, and J. Wang, "Modeling and removing spatially-varying optical blur," in *Proc. ICCP*, 2011, pp. 1–8. 3

[15] V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," *Applied Optics*, vol. 33, no. 34, pp. 8121–8124, 1994. 3

[16] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, "Non-stationary correction of optical aberrations," in *Proc. ICCV*, 2011, pp. 659–666. 3

[17] F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb, "High-quality computational imaging through simple lenses," *ACM Transactions on Graphics*, vol. 32, no. 5, pp. 1–14, 2013. 3

[18] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, "Blind correction of optical aberrations," in *Proc. ECCV*, vol. 7574, 2012, pp. 187–200. 3

[19] T. Yue, J. Suo, J. Wang, X. Cao, and Q. Dai, "Blind optical aberration correction by exploring geometric and visual priors," in *Proc. CVPR*, 2015, pp. 1684–1692. 3

[20] Y. Peng, Q. Sun, X. Dun, G. Wetzstein, W. Heidrich, and F. Heide, "Learned large field-of-view imaging with thin-plate optics," *ACM Transactions on Graphics*, vol. 38, no. 6, pp. 1–14, 2019. 3

[21] S. Chen, H. Feng, K. Gao, Z. Xu, and Y. Chen, "Extreme-quality computational imaging via degradation framework," in *Proc. ICCV*, 2021, pp. 2612–2621. 3

[22] S. Chen, H. Feng, D. Pan, Z. Xu, Q. Li, and Y. Chen, "Optical aberrations correction in postprocessing using imaging simulation," *ACM Transactions on Graphics*, vol. 40, no. 5, pp. 1–15, 2021. 3

[23] S. Chen, J. Zhou, M. Li, Y. Chen, and T. Jiang, "Mobile image restoration via prior quantization," *Pattern Recognition Letters*, vol. 174, pp. 64–70, 2023. 3

[24] Q. Jiang, H. Shi, L. Sun, S. Gao, K. Yang, and K. Wang, "Annular computational imaging: Capture clear panoramic images through simple lens," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 1250–1264, 2022. 3

[25] Q. Jiang *et al.*, "Minimalist and high-quality panoramic imaging with PSF-aware transformers," *arXiv preprint arXiv:2306.12992*, 2023. 3

[26] ——, "Real-world computational aberration correction via quantized domain-mixing representation," *arXiv preprint arXiv:2403.10012*, 2024. 3

[27] ——, "Computational imaging for machine perception: Transferring semantic segmentation beyond aberrations," *IEEE Transactions on Computational Imaging*, 2024. 3

[28] J. Zhou, S. Chen, Z. Ren, W. Zhang, J. Yan, H. Feng, Q. Li, and Y. Chen, "Revealing the preference for correcting separated aberrations in joint optic-image design," *Optics and Lasers in Engineering*, vol. 178, p. 108220, 2024. 3

[29] J. Luo, Y. Nie, W. Ren, X. Cao, and M.-H. Yang, "Correcting optical aberration via depth-aware point spread functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 3

[30] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image restoration using swin transformer," in *Proc. ICCVW*, 2021, pp. 1833–1844. 3, 10

[31] L. Chen, X. Chu, X. Zhang, and J. Sun, "Simple baselines for image restoration," in *Proc. ECCV*, vol. 13667, 2022, pp. 17–33. 3

[32] X. Wang *et al.*, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. ECCVW*, vol. 11133, 2018, pp. 63–79. 3

[33] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. ECCV*, vol. 11211, 2018, pp. 294–310. 3

[34] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proc. CVPR*, 2022, pp. 5718–5729. 3

[35] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general U-shaped transformer for image restoration," in *Proc. CVPR*, 2022, pp. 17 662–17 672. 3

[36] Z. Li, Q. Hou, Z. Wang, F. Tan, J. Liu, and W. Zhang, "End-to-end learned single lens design using fast differentiable ray tracing," *Optics Letters*, vol. 46, no. 21, pp. 5453–5456, 2021. 3, 6

[37] T. Yang, H. Xu, D. Cheng, and Y. Wang, "Design of compact off-axis freeform imaging systems based on optical-digital joint optimization," *Optics Express*, vol. 31, no. 12, pp. 19 491–19 509, 2023. 3

[38] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944. 3

[39] G. Côté, J.-F. Lalonde, and S. Thibault, "Extrapolating from lens design databases using deep learning," *Optics Express*, vol. 27, no. 20, pp. 28 279–28 292, 2019. 3

[40] D. Guo, L. Yin, and G. Yuan, "New automatic optical design method based on combination of particle swarm optimization and least squares," *Optics Express*, vol. 27, no. 12, pp. 17 027–17 040, 2019. 3

[41] J. Sun and X. Li, "Automatic design of machine vision lens based on improved genetic algorithm and damped least squares," in *Proc. SPIE*, vol. 11895, 2021, pp. 188–202. 3

[42] J. Zhang, Z. Cen, and X. Li, "Automated design of machine vision lens based on the combination of particle swarm optimization and damped least squares," in *Proc. SPIE*, vol. 11548, 2020, pp. 261–272. 3

[43] W. Yue, G. Jin, and X. Yang, "Adaptive particle swarm optimization for automatic design of common aperture optical system," *Photonics*, vol. 9, no. 11, p. 807, 2022. 3

[44] Z. Tang, M. Sonntag, and H. Gross, "Ant colony optimization in lens design," *Applied Optics*, vol. 58, no. 23, pp. 6357–6364, 2019. 3

[45] C. Reichert, T. Gruhonjic, and A. Herkommer, "Development of an open source algorithm for optical system design, combining genetic and local optimization," *Optical Engineering*, vol. 59, no. 5, p. 055111, 2020. 3

[46] G. Wetzstein *et al.*, "Inference in artificial intelligence with deep optics and photonics," *Nature*, vol. 588, no. 7836, pp. 39–47, 2020. 3

[47] Q. Sun, E. Tseng, Q. Fu, W. Heidrich, and F. Heide, "Learning rank-1 diffractive optics for single-shot high dynamic range imaging," in *Proc. CVPR*, 2020, pp. 1383–1393. 3

[48] D. S. Jeon *et al.*, "Compact snapshot hyperspectral imaging with diffracted rotation," *ACM Transactions on Graphics*, vol. 38, no. 4, pp. 1–13, 2019. 3

[49] X. Dun, H. Ikoma, G. Wetzstein, Z. Wang, X. Cheng, and Y. Peng, "Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging," *Optica*, vol. 7, no. 8, pp. 913–922, 2020. 3

[50] S.-H. Baek *et al.*, "Single-shot hyperspectral-depth imaging with learned diffractive optics," in *Proc. ICCV*, 2021, pp. 2631–2640. 3

[51] I. Chugunov, S.-H. Baek, Q. Fu, W. Heidrich, and F. Heide, "Mask-ToF: Learning microlens masks for flying pixel correction in time-of-flight imaging," in *Proc. CVPR*, 2021, pp. 9116–9126. 3

[52] L. Li, L. Wang, W. Song, L. Zhang, Z. Xiong, and H. Huang, "Quantization-aware deep optics for diffractive snapshot hyperspectral imaging," in *Proc. CVPR*, 2022, pp. 19 748–19 757. 3

[53] J. Bacca, T. Gelvez-Barrera, and H. Arguello, "Deep coded aperture design: An end-to-end approach for computational imaging tasks," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 1148–1160, 2021. 3

[54] Q. Sun, J. Zhang, X. Dun, B. Ghanem, Y. Peng, and W. Heidrich, "End-to-end learned, optically coded super-resolution SPAD camera," *ACM Transactions on Graphics*, vol. 39, no. 2, pp. 1–14, 2020. 3

[55] C. A. Metzler, H. Ikoma, Y. Peng, and G. Wetzstein, "Deep optics for single-shot high-dynamic-range imaging," in *Proc. CVPR*, 2020, pp. 1372–1382. 3

[56] H. Ikoma, C. M. Nguyen, C. A. Metzler, Y. Peng, and G. Wetzstein, "Depth from defocus with learned optics for imaging and occlusion-aware depth estimation," in *Proc. ICCP*, 2021, pp. 1–12. 3

[57] Z. Shi *et al.*, "Seeing through obstructions with diffractive cloaking," *ACM Transactions on Graphics*, vol. 41, no. 4, pp. 1–15, 2022. 3

[58] S. Pinilla, S. R. M. Rostami, I. Shevkunov, V. Katkovnik, and K. Egiazarian, "Hybrid diffractive optics design via hardware-in-the-loop methodology for achromatic extended-depth-of-field imaging," *Optics Express*, vol. 30, no. 18, pp. 32 633–32 649, 2022. 3

[59] A. Fontbonne, H. Sauer, and F. Goudail, "Comparison of methods for end-to-end co-optimization of optical systems and image processing with commercial lens design software," *Optics Express*, vol. 30, no. 8, pp. 13 556–13 571, 2022. 3

[60] X. Yang, Q. Fu, Y. Nie, and W. Heidrich, "Image quality is not all you want: Task-driven lens design for image classification," *arXiv preprint arXiv:2305.17185*, 2023. 3

[61] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015. 5

[62] T. Brooks, B. Mildenhall, T. Xue, J. Chen, D. Sharlet, and J. T. Barron, "Unprocessing images for learned raw denoising," in *Proc. CVPR*, 2019, pp. 11 036–11 045. 6

[63] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. CVPR*, 2018, pp. 586–595. 7, 10

[64] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. ICCV*, vol. 2, 2001, pp. 416–425. 9

[65] A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. NeurIPS*, vol. 32, 2019, pp. 8024–8035. 10

[66] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. 10

APPENDIX A
DETAILED RESULTS

Initially, we obtained $\{5, 7, 6\}$ structures for 3P-I and $\{15, 10, 17\}$ structures for 3P-II, totaling 60 structures. In Table IV, we present a comprehensive evaluation of the jointly optimized final structures, including their spot loss and corresponding performance metrics such as PSNR, SSIM, and LPIPS. The spot loss is calculated as the average spot RMS radius across all 3 working distances, 7 sampling fields of view, and 15 sampling wavelengths after the joint optimization process. For clarity and ease of analysis, we have sorted all structures in descending order of their spot loss. Obviously, a smaller spot loss does not necessarily mean better imaging quality after the final reconstruction. In addition, we present in Table V the PSNR, SSIM, and LPIPS of all jointly designed and separately designed compact computational imaging systems.

TABLE IV
SPOT RMS RADIUS AND CORRESPONDING RECONSTRUCTED IMAGE QUALITY.

	Spot loss↓	PSNR↑	SSIM↑	LPIPS↓
S1	0.0288	29.53	0.8392	0.0883
S2	0.0291	29.44	0.8366	0.0882
S3	0.0294	29.68	0.8418	0.0867
S4	0.0297	29.69	0.8398	0.0848
S5	0.0297	28.67	0.8150	0.1052
S6	0.0299	28.73	0.8158	0.1047
S7	0.0300	28.36	0.8054	0.1133
S8	0.0311	29.45	0.8396	0.0881
S9	0.0311	29.57	0.8399	0.0879
S10	0.0315	29.06	0.8230	0.0997
S11	0.0330	29.83	0.8484	0.0814
S12	0.0330	29.60	0.8419	0.0882
S13	0.0333	28.97	0.8198	0.1020
S14	0.0338	29.14	0.8317	0.0955
S15	0.0340	29.15	0.8302	0.0962
S16	0.0344	28.98	0.8256	0.1010
S17	0.0345	29.50	0.8430	0.0879
S18	0.0352	29.57	0.8393	0.0882
S19	0.0354	29.17	0.8343	0.0943
S20	0.0355	28.81	0.8206	0.1015
S21	0.0355	28.84	0.8190	0.1003
S22	0.0358	28.62	0.8155	0.1061
S23	0.0366	29.16	0.8330	0.0939
S24	0.0369	28.38	0.8078	0.1130
S25	0.0370	29.46	0.8391	0.0881
S26	0.0375	28.42	0.8032	0.1142
S27	0.0380	28.80	0.8128	0.1041
S28	0.0386	30.12	0.8606	0.0755
S29	0.0388	30.38	0.8636	0.0719
S30	0.0392	29.68	0.8438	0.0839
S31	0.0397	29.28	0.8366	0.0946
S32	0.0398	29.91	0.8556	0.0820
S33	0.0401	28.91	0.8265	0.0994
S34	0.0405	28.85	0.8255	0.1004
S35	0.0406	29.51	0.8464	0.0836
S36	0.0408	28.64	0.8092	0.1070
S37	0.0411	29.70	0.8506	0.0814
S38	0.0422	28.98	0.8271	0.0986
S39	0.0424	30.31	0.8605	0.0727
S40	0.0427	29.59	0.8444	0.0838
S41	0.0433	29.54	0.8423	0.0878
S42	0.0436	30.24	0.8614	0.0746
S43	0.0436	29.58	0.8396	0.0855
S44	0.0438	29.84	0.8527	0.0806
S45	0.0439	28.78	0.8241	0.0998
S46	0.0446	28.76	0.8236	0.1015
S47	0.0451	29.16	0.8319	0.0965
S48	0.0453	29.36	0.8345	0.0906
S49	0.0457	29.67	0.8465	0.0848
S50	0.0458	29.80	0.8500	0.0818
S51	0.0458	28.93	0.8290	0.0980
S52	0.0459	29.05	0.8322	0.0945
S53	0.0478	28.62	0.8145	0.1035
S54	0.0489	29.22	0.8371	0.0900
S55	0.0505	29.15	0.8373	0.0914
S56	0.0526	29.42	0.8406	0.0885
S57	0.0530	29.92	0.8508	0.0806
S58	0.0537	29.06	0.8289	0.0956
S59	0.0572	29.46	0.8338	0.0880
S60	0.0717	28.74	0.8221	0.1040

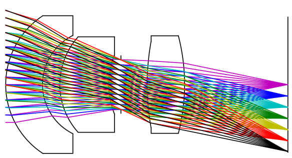
APPENDIX B
VISUALIZATION COMPARISON

In this section, we undertake a comprehensive comparison of the Global Search Optics (GSO) method, the CODE V assisted joint design approach, and the separate design method. For the design specifications 3P-I and 3P-II, we present the optimal solutions discovered by both GSO and the CODE V assisted joint design method. Notably, these two methods rely on distinct initial structures. Additionally, we include the results of the separate design method, which utilizes the same initial structure as GSO. The diagram, detailed lens specifications, and final reconstructed image quality metrics are all included in Fig. B.1, providing a comprehensive overview of the designed lens characteristics. The optimal solutions found by GSO significantly outperform those discovered by the other two methods, demonstrating that GSO is indeed capable of conducting a more comprehensive search of the solution space for computational optical systems and enhancing the imaging capabilities of these systems to their upper limits.

Furthermore, examples of image reconstruction corresponding to 3P-I and 3P-II are presented in Fig. B.2 and Fig. B.3, respectively, enabling a direct comparison of the performance achieved by each method. It is noteworthy that even when the final results of GSO sometimes result in more blurred degraded images, the quality of the reconstructed images remains superior to that achieved by the other two methods. This is attributed to the fact that the combination of the optical system and image reconstruction network discovered by GSO is closer to the global optimal solution, rather than the optimal solution obtained when the two are designed independently or the diversity of the initial structures is insufficient.

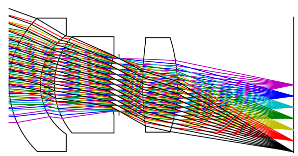
GSO | CODE V assisted joint design method | Separate design method

3P-I



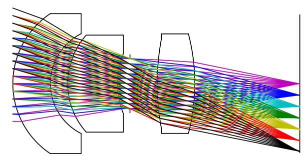
PSNR: 30.38
SSIM: 0.8636
LPIPS: 0.0719

Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{100000, 10000, 5000}	AIR	0
S	17.430	7.812	H-K51	14.59
S	11.621	3.674	AIR	10.41
S	15.095	10.387	H-LaK53B	10.10
S	16.629	2.595	AIR	6.22
S	Infinity	5.569	AIR	5.28
S	51.370	7.938	H-LaK53B	9.15
S	-40.990	21.910	AIR	10.34
I	Infinity	59.887	AIR	14.30



PSNR: 29.99
SSIM: 0.8533
LPIPS: 0.0772

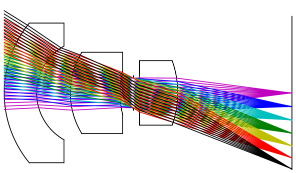
Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{100000, 10000, 5000}	AIR	0
S	19.637	6.604	H-ZF3	14.07
S	12.705	2.948	AIR	10.43
S	15.761	11.485	H-LaK53B	10.18
S	19.209	2.134	AIR	6.25
S	Infinity	4.985	AIR	5.56
S	69.690	7.296	H-LaK53B	8.79
S	-33.276	24.477	AIR	9.97
I	Infinity	59.929	AIR	14.30



PSNR: 29.64
SSIM: 0.8468
LPIPS: 0.0896

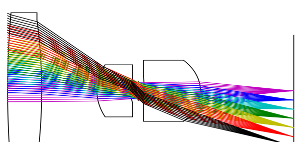
Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{100000, 10000, 5000}	AIR	0
S	17. 467	7.846	H-K51	14.57
S	11.625	3.609	AIR	10.39
S	15.100	10.401	H-LaK53B	10.09
S	16.637	2.616	AIR	6.20
S	Infinity	5.674	AIR	5.24
S	50.589	7.838	H-LaK53B	9.20
S	-40.673	21.902	AIR	10.36
I	Infinity	59.886	AIR	14.30

3P-II



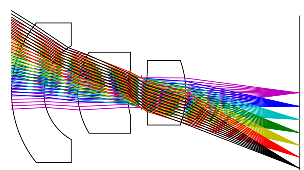
PSNR: 29.83
SSIM: 0.8484
LPIPS: 0.0814

Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{5000, 1000, 500}	AIR	0
S	20.527	6.057	H-K51	13.05
S	10.047	6.296	AIR	8.74
S	14.521	9.285	H-ZF5	7.58
S	16.930	2.514	AIR	4.16
S	Infinity	0.995	AIR	2.81
S	62.388	7.278	H-LaK53B	3.84
S	-18.019	21.323	AIR	6.02
I	Infinity	53.748	AIR	14.30



PSNR: 29.43
SSIM: 0.8389
LPIPS: 0.0881

Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{5000, 1000, 500}	AIR	0
S	306.883	6.622	H-LaK53B	20.100
S	-182.684	10.734	AIR	18.370
S	14.862	6.474	H-ZF5	6.703
S	10.050	1.570	AIR	2.969
S	Infinity	1.139	AIR	2.178
S	-54.660	10.940	H-LaK53B	3.579
S	-11.089	18.130	AIR	7.875
I	Infinity	55.609	AIR	14.30



PSNR: 29.26
SSIM: 0.8337
LPIPS: 0.0948

Type	Radius	Distance	Material	Semi-Diameter
O	Infinity	{5000, 1000, 500}	AIR	0
S	20.656	6.037	H-K51	13.01
S	10.029	6.303	AIR	8.71
S	14.545	9.276	ZF6	7.56
S	16.959	2.534	AIR	4.15
S	Infinity	1.026	AIR	2.78
S	62.360	7.317	H-LaK53B	3.83
S	-17.867	21.223	AIR	6.02
I	Infinity	53.716	AIR	14.30

Fig. B.1. The optimal solutions found by GSO, CODE V assisted joint design method and separate design method. The design results of 3P-I and 3P-II are presented from top to bottom. The optical system diagram, final reconstructed image quality metrics, and lens data are presented.

TABLE V
DETAILED COMPARISON RESULTS BETWEEN JOINT DESIGN AND SEPARATE DESIGN METHODS.

	Joint Design						**Separate Design**					
	Degraded			**Reconstructed**			**Degraded**			**Reconstructed**		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
S1	23.46	0.6088	0.4148	29.53	0.8392	0.0883	23.50	0.6129	0.4078	29.27	0.8346	0.0980
S2	23.35	0.6065	0.4167	29.44	0.8366	0.0882	23.43	0.6088	0.4154	29.15	0.8285	0.0979
S3	23.53	0.6123	0.4082	29.68	0.8418	0.0867	23.57	0.6138	0.4090	29.45	0.8361	0.0932
S4	23.27	0.6022	0.4211	29.69	0.8398	0.0848	23.28	0.6021	0.4209	29.34	0.8313	0.0951
S5	23.08	0.5909	0.4376	28.67	0.8150	0.1052	23.23	0.5990	0.4238	28.67	0.8142	0.1063
S6	22.83	0.5832	0.4402	28.73	0.8158	0.1047	22.86	0.5847	0.4393	28.41	0.8069	0.1123
S7	22.71	0.5781	0.4469	28.36	0.8054	0.1133	22.74	0.5825	0.4351	28.23	0.8000	0.1188
S8	23.12	0.6005	0.4276	29.45	0.8396	0.0881	23.10	0.5967	0.4305	28.87	0.8256	0.0987
S9	23.36	0.6069	0.4135	29.57	0.8399	0.0879	23.42	0.6094	0.4103	29.28	0.8315	0.0971
S10	23.01	0.5901	0.4375	29.06	0.8230	0.0997	23.08	0.5940	0.4303	28.86	0.8179	0.1061
S11	22.81	0.5910	0.4267	29.83	0.8484	0.0814	22.74	0.5886	0.4297	29.26	0.8337	0.0948
S12	23.24	0.6067	0.4175	29.60	0.8419	0.0882	23.26	0.6059	0.4183	29.31	0.8371	0.0947
S13	23.09	0.5915	0.4378	28.97	0.8198	0.1020	23.32	0.6037	0.4212	29.21	0.8314	0.0958
S14	23.06	0.5987	0.4266	29.14	0.8317	0.0955	23.03	0.5953	0.4288	28.83	0.8237	0.1034
S15	23.18	0.6000	0.4255	29.15	0.8302	0.0962	23.29	0.6034	0.4206	29.23	0.8310	0.0944
S16	23.20	0.6028	0.4261	28.98	0.8256	0.1010	23.23	0.6041	0.4233	28.84	0.8230	0.1068
S17	23.13	0.6032	0.4229	29.50	0.8430	0.0879	23.19	0.6039	0.4202	29.24	0.8348	0.0974
S18	23.06	0.6000	0.4216	29.57	0.8393	0.0882	23.13	0.6011	0.4189	29.25	0.8328	0.0959
S19	23.09	0.5989	0.4229	29.17	0.8343	0.0943	23.00	0.5942	0.4318	28.78	0.8213	0.1023
S20	22.26	0.5678	0.4538	28.81	0.8206	0.1015	22.30	0.5721	0.4458	28.29	0.8058	0.1178
S21	23.05	0.5930	0.4342	28.84	0.8190	0.1003	23.10	0.5985	0.4305	28.49	0.8138	0.1112
S22	22.04	0.5617	0.4565	28.62	0.8155	0.1061	22.04	0.5631	0.4531	28.16	0.8045	0.1188
S23	22.09	0.5631	0.4679	29.16	0.8330	0.0939	22.15	0.5668	0.4522	29.01	0.8288	0.0968
S24	21.80	0.5511	0.4699	28.38	0.8078	0.1130	21.84	0.5553	0.4580	28.28	0.8029	0.1147
S25	23.17	0.6023	0.4240	29.46	0.8391	0.0881	23.19	0.6011	0.4261	28.74	0.8215	0.1074
S26	22.77	0.5786	0.4539	28.42	0.8032	0.1142	22.86	0.5841	0.4450	28.46	0.8039	0.1142
S27	22.96	0.5890	0.4458	28.80	0.8128	0.1041	23.10	0.5927	0.4385	28.76	0.8147	0.1066
S28	22.18	0.5693	0.4485	30.12	0.8606	0.0755	22.08	0.5766	0.4135	28.98	0.8276	0.1013
S29	22.45	0.5849	0.4435	30.38	0.8636	0.0719	22.38	0.5918	0.4144	29.64	0.8468	0.0896
S30	22.13	0.5674	0.4623	29.68	0.8438	0.0839	22.12	0.5712	0.4499	28.95	0.8237	0.1021
S31	22.83	0.5944	0.4355	29.28	0.8366	0.0946	22.97	0.5966	0.4316	28.88	0.8256	0.1037
S32	22.37	0.5782	0.4520	29.91	0.8556	0.0820	22.39	0.5868	0.4263	29.41	0.8471	0.0892
S33	22.74	0.5858	0.4407	28.91	0.8265	0.0994	22.72	0.5855	0.4416	28.65	0.8174	0.1071
S34	22.67	0.5832	0.4437	28.85	0.8255	0.1004	22.70	0.5808	0.4459	28.70	0.8179	0.1031
S35	21.79	0.5608	0.4612	29.51	0.8464	0.0836	21.88	0.5653	0.4507	29.16	0.8377	0.0942
S36	22.73	0.5804	0.4570	28.64	0.8092	0.1070	22.97	0.5900	0.4405	28.81	0.8147	0.1076
S37	22.21	0.5747	0.4467	29.70	0.8506	0.0814	22.18	0.5742	0.4398	28.96	0.8300	0.1005
S38	22.86	0.5904	0.4382	28.98	0.8271	0.0986	22.83	0.5907	0.4385	28.68	0.8184	0.1052
S39	22.41	0.5828	0.4507	30.31	0.8605	0.0727	22.45	0.5902	0.4308	29.92	0.8523	0.0826
S40	22.54	0.5856	0.4573	29.59	0.8444	0.0838	22.61	0.5907	0.4415	29.51	0.8428	0.0844
S41	22.31	0.5742	0.4461	29.54	0.8423	0.0878	22.39	0.5782	0.4504	29.18	0.8325	0.0956
S42	22.39	0.5898	0.4233	30.24	0.8614	0.0746	22.38	0.5918	0.4173	29.86	0.8529	0.0844
S43	23.21	0.6056	0.4212	29.58	0.8396	0.0855	23.21	0.6012	0.4275	28.99	0.8232	0.0996
S44	22.25	0.5875	0.4262	29.84	0.8527	0.0806	22.24	0.5921	0.4070	29.58	0.8474	0.0850
S45	22.58	0.5801	0.4469	28.78	0.8241	0.0998	22.66	0.5813	0.4443	28.74	0.8202	0.1044
S46	22.59	0.5829	0.4455	28.76	0.8236	0.1015	22.66	0.5817	0.4420	28.77	0.8212	0.1037
S47	22.32	0.5829	0.4344	29.16	0.8319	0.0965	22.44	0.5861	0.4397	29.19	0.8334	0.0984
S48	22.45	0.5841	0.4561	29.36	0.8345	0.0906	22.40	0.5839	0.4503	29.10	0.8282	0.0990
S49	22.17	0.5704	0.4666	29.67	0.8465	0.0848	22.11	0.5768	0.4387	29.20	0.8339	0.0967
S50	22.31	0.5833	0.4432	29.80	0.8500	0.0818	22.32	0.5847	0.4351	29.22	0.8354	0.0928
S51	22.81	0.5913	0.4351	28.93	0.8290	0.0980	22.87	0.5868	0.4390	28.90	0.8227	0.0975
S52	22.82	0.5889	0.4378	29.05	0.8322	0.0945	22.86	0.5903	0.4351	28.50	0.8150	0.1110
S53	22.26	0.5704	0.4793	28.62	0.8145	0.1035	22.21	0.5711	0.4682	28.43	0.8104	0.1118
S54	22.99	0.5970	0.4303	29.22	0.8371	0.0900	22.99	0.5951	0.4313	29.04	0.8295	0.0993
S55	21.21	0.5338	0.5036	29.15	0.8373	0.0914	21.47	0.5534	0.4502	28.97	0.8338	0.0942
S56	21.77	0.5536	0.4923	29.42	0.8406	0.0885	21.90	0.5791	0.4248	29.21	0.8328	0.0979
S57	22.09	0.5795	0.4395	29.92	0.8508	0.0806	22.08	0.5823	0.4297	29.66	0.8461	0.0899
S58	22.87	0.5922	0.4384	29.06	0.8289	0.0956	22.95	0.5923	0.4363	28.62	0.8179	0.1053
S59	22.36	0.5823	0.4475	29.46	0.8338	0.0880	22.48	0.5778	0.4728	28.63	0.8114	0.1091
S60	22.69	0.5857	0.4463	28.74	0.8221	0.1040	22.69	0.5818	0.4506	28.75	0.8176	0.1050

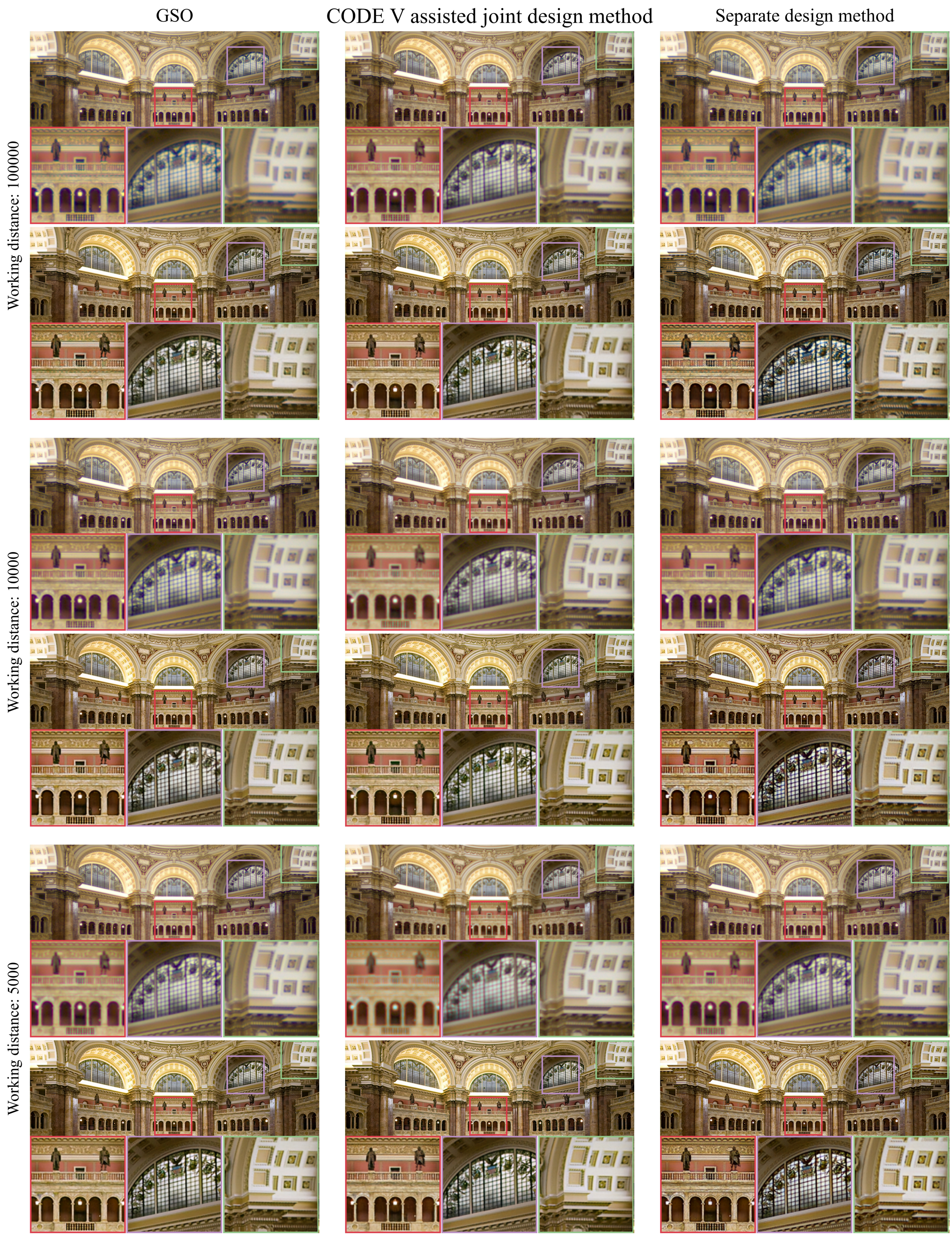


Fig. B.2. Image reconstruction examples for GSO, CODE V assisted joint design method, and the separate design method under 3P-I.

Fig. B.3. Image reconstruction examples for GSO, CODE V assisted joint design method, and the separate design method under 3P-II.