A Physics-informed Low-rank Deep Neural Network for Blind and Universal Lens Aberration Correction

Jin Gong[1] Runzhao Yang[1] Weihang Zhang[1] Jinli Suo[1,2,3]
Qionghai Dai[1,2]
[1]Department of Automation, Tsinghua University, Beijing, China
[2]Institute of Brain and Cognitive Sciences, Tsinghua University, Beijing, China
[3]Shanghai Artificial Intelligence Laboratory, Shanghai, China

Abstract

High-end lenses, although offering high-quality images, suffer from both insufficient affordability and bulky design, which hamper their applications in low-budget scenarios or on low-payload platforms. A flexible scheme is to tackle the optical aberration of low-end lenses computationally. However, it is highly demanded but quite challenging to build a general model capable of handling non-stationary aberrations and covering diverse lenses, especially in a blind manner. To address this issue, we propose a universal solution by extensively utilizing the physical properties of camera lenses: (i) reducing the complexity of lens aberrations, i.e., lens-specific non-stationary blur, by warping annual-ring-shaped sub-images into rectangular stripes to transform non-uniform degenerations into a uniform one, (ii) building a low-dimensional non-negative orthogonal representation of lens blur kernels to cover diverse lenses; (iii) designing a decoupling network to decompose the input low-quality image into several components degenerated by above kernel bases, and applying corresponding pre-trained deconvolution networks to reverse the degeneration. Benefiting from the proper incorporation of lenses' physical properties and unique network design, the proposed method achieves superb imaging quality, wide applicability for various lenses, high running efficiency, and is totally free of kernel calibration. These advantages bring great potential for scenarios requiring lightweight high-quality photography.

1. Introduction

High-quality photography is of crucial importance for both high-fidelity visual recording (e.g., filmmaking, sports video capturing) and sophisticated computer vision tasks (e.g., surveillance, auto-piloting). High-end camera systems often employ compound lenses comprising approximately ten or more components constructed from diverse materials to compensate for geometric and photometric aberrations. Such complicated designs are proven to be effective in achieving nice image quality, but come with inherent drawbacks, including high costs, bulkiness, and fragility, making them unsuitable for scenarios with low payload capacity or limited budgets. Consequently, the demand for high-quality photography using lightweight lenses has significantly intensified.

Considering the optical aberration in image formation and the afterward processing jointly, the workload in optical design can be shifted to the later stage [23, 28], where advanced reconstruction algorithms play a crucial role. One can also utilize the physical properties in the imaging setup to facilitate reconstruction, and researchers have made some primary efforts in this direction [32, 45].

However, grand challenges lie in the variability of optical aberrations across the field of view and diverse lenses. On the one hand, the quality degeneration of a simple lens is intrinsically a convolution with non-uniform blur kernels, and the typical compensation algorithms [43] approximate the globally non-uniform deconvolution with patch-wise uniform deconvolutions, leading to a trade-off between precision and computational efficiency. This trade-off comes up with high inflexibility when adopting data-driven approaches [14, 41], which require learning a large number of models to achieve high-performance results. On the other hand, the degradation of different lenses varies significantly, so lens-specific algorithm development or parameter optimizations are required for high reconstruction performance. Furthermore, the calibration of the PSF kernels of camera lenses is quite expertise-demanding[27, 46], and blind compensation is more favorable for users. Overall, computational compensation for lens aberrations holds great promise for achieving lightweight high-quality imaging. However, there is a pressing demand for a general approach to handle spatially varying aberrations of diverse lenses in a blind manner.

In this paper, we propose a physics-informed end-to-end solution that (i) capitalizes the characteristics of lens aber-

Figure 1. An illustrative example of our lens-aberration compensation approach. (a) A camera equipped with a simple lens of a large field of view but severe optical aberration (WXSJ-H65HD) and a small-sized unmanned aerial vehicle (UAV) carrying the camera for data capture. (b) The input degenerated image (upper) and our reconstruction result (lower). (c) The zoomed-in comparison on the highlighted region (white box) in (b), where the original recording is shown in the bottom left corner and the reconstructed result in the top right corner.

ration to construct a low-dimensional representation of non-uniform blur kernels of general camera lenses, and (ii) designs a deep neural network resolving the degeneration components in the low-quality input and ensembling a set of pre-trained compensators to reverse the degeneration robustly. Specifically, we represent an arbitrary local point spread function (PSF) with a set of negative orthogonal bases, pre-train their corresponding deconvolution modules, and then retrieve their degeneration from the low-quality image captured by a low-end lens and apply the pre-trained inversion models accordingly. The proposed approach demonstrates high performance and holds high potential in lightweight photography on low-payload platforms, as shown by the impressive results captured with a small drone equipped with a compact surveillance camera in Fig. 1.

In summary, we target for general, blind, and end-to-end lens aberration correction, and make the following contributions:

1) Proposes a unified framework for lens aberration compensation with high flexibility to diverse aberrations produced by the wide range of camera lenses.
2) Builds a general low-dimensional model of lens aberrations based on Orthogonal Non-negative Matrix Factorization, utilizing the physical properties of optical lenses.
3) Designs an end-to-end network to divide and conquer the optical aberrations in the above low-dimensional space, enabling fast and blind inversion of diverse lens degeneration and ultimately lightweight high-quality imaging.
4) Demonstrates performance comparable to state-of-the-art non-blind lens-specific algorithms, validating its great potential in budget-constrained or low-capacity platforms.

2. Related Work

Lens aberration modeling. Generally, imaging lenses are physically rotational symmetric around the optical center, imparting rotation symmetry of the lens aberration [25, 37]. Ignoring the fabrication imperfections, almost all types of lens aberrations, such as spherical aberration, coma aberration, and chromatic aberration [1, 9], form a rotational symmetric pattern.

Utilizing these unique features of lens aberration, researchers have proposed some methods to simplify the degeneration by diverse lenses. For example, Rahbar et al. [29] adopt the Zernike model [24] to describe optical aberrations and can estimate the Zernike coefficients of a single channel through bicoherence and tricoherence estimation techniques, while Schuler et al. [32] represent the non-uniform aberrations with a set of orthonormal Efficient Filter Flow, which is applicable for most cases without large spherical aberration. Differently, Yue et al. [45] leverage the global rotational symmetry properties of regular lenses to transform non-uniform aberrations into uniform rings using radial splitting and warping techniques. This method capitalizes on the inherent physical properties of the imaging lens and largely simplifies the aberration model, offering inspiration for our approach to explore the unique structures of lens aberrations. However, their implementations use conventional optimization by alternatively kernel estimation and deblurring, which is time-consuming and the strong consumption of PSF uniformity within a stretched annular ring harms the performance sightly, which limits the applications demanding real-time and high-quality compensation. In contrast, we design a framework consisting of well-organized sub-networks to address all these issues decently.

Optical aberration removal. Lens aberration exists widely in optical imaging systems, and computational correction is an essential way to raise imaging quality without increasing hardware budget. One common way to model lens aberration is by convolving the image with spatially varying PSFs and compensation is naturally conducted via deconvolution [38, 44]. Existing correction methods can be broadly categorized into non-blind and blind ones.

Non-blind correction assumes known PSFs and algorithms are extensively studied [4, 22]. Researchers have proposed different algorithms to estimate kernel PSFs via, e.g., using a combination of defocused and focused images [2], a set of binary random patterns [10], aggregating degenerations in informative nature image patches [11], analyzing the spectral characteristics of the lens system [3, 19, 34]. Further, to cover PSFs of different camera lenses, Shih et al. [33] utilize interpolation methods by fitting a spatial Gaussian model. More recently, Li et al. [20] propose a data-driven deep learning approach explicitly taking the PSF as input and introducing lens-specific priors for high-quality compensation. In sum, non-blind lens aberration compensation techniques have shown promising performance, but they require expertise demanding PSF calibration or robust estimation, which are not friendly for non-expert users and require lens-specific model training.

Blind methods have gained significant attention due to their convenience for un-trained users and high flexibility to diverse lenses. The typical strategy is to estimate PSFs and conduct compensation sequentially. Among them Rahbar et al. [29] introduce Zernike moments, and Tang & Kutulakos [35] employ the Seidel model to simplify the lens aberration model; while Delbracio et al. [5] propose a robust algorithm based on empirical observations about the distribution of the gradient in clear natural images. Recently, some researchers have adopted the data-driven scheme and developed deep neural networks to compensate for the aberrations, but often focus on specific types of aberrations to ensure good convergence, such as radial lens distortion [31] and chromatic aberrations [7].

In light of these developments, the blind compensation techniques (either based on conventional optimization or deep neural networks) cannot achieve performance comparable to their non-blind counterpart. Besides, existing methods are incapable of handling diverse lenses and various types of aberrations flexibly [4, 8]. In contrast, the proposed work leverages the physical properties of camera lenses and casts the complex aberrations of diverse lenses into a unified low-dimensional space, in which we divide and conquer the degeneration via incorporating geometric priors and a small number of pre-trained modules. Benefiting from the proper use of lenses' physical properties and elegant network design, we can achieve performance comparable to non-blind techniques as well. Such a general end-to-end blind solution with superb performance holds great potential for high-quality lightweight imaging systems on portable devices or low-capacity mobile platforms.

Non-uniform deconvolution. Mathematically lens aberration compensation can be described as a non-uniform deconvolution process and shares the same formulation with other tasks such as camera shake, defocus, object motion, etc. Various techniques and approaches have been proposed to address the technical challenges posed by non-uniform blurring [39, 47]. There are three main ways to address the non-uniformity. The most intuitive and widely used way is to assume patch-wise uniform PSFs and conduct deconvolution patch by patch [36, 42]. The deconvolution can be implemented via conventional optimization previously and deep neural networks recently, and usually in a non-blind manner. There are a bunch of algorithms, and we do not list them here. The second solution is to transform the various PSFs into a low dimensional space and remove the blur along each dimension [7, 17, 32]. The third way is to adopt data-driven techniques and fed training data with varying PSFs for high generalization ability [48]. One can also make extensive use of the structure of the PSF patterns to circumvent the high complexity by introducing physical constraints, e.g., transform the spatially varying aberrations in an annular ring into uniform via warping [45], and decompose the spatially varying defocus blur into several uniform ones according to the scene depth [18, 49]. Differently, our approach focuses on lens aberration which is of different features from other degenerations and can utilize the unique properties of lenses for a better design, in a similar way to [45]. In addition, we are dedicated to an end-to-end solution working for diverse lenses without model training.

3. Physics-inspired Low Rank Lens Aberration Model

There exist various lens aberrations, such as spherical aberration and astigmatism, resulting in various non-uniform quality degenerations. In addition, the aberration models of diverse lenses differ a lot. To provide a universal solution for different lenses, it is of crucial importance to reduce the dimension of the PSFs and provide a unified representation, based on which one can design a low-rank model to address the aberrations in a divide-and-conquer manner.

Considering the rotational-symmetry of optical lenses [15, 16, 40], we divide the lens's field of view into several concentric rings and warp them into rectangular stripes, each of which is of approximately uniform PSFs. This operation can largely decrease the dimension of the lens aberrations [13, 21]. Further, we crop these stripes into patches and apply ONMF to find a set of orthogonal and nonnegative bases to cover the space of the lens PSFs. If learned from a large set of lenses, the bases can represent an arbitrary PSF with high accuracy. The orthogonality can avoid ambiguity

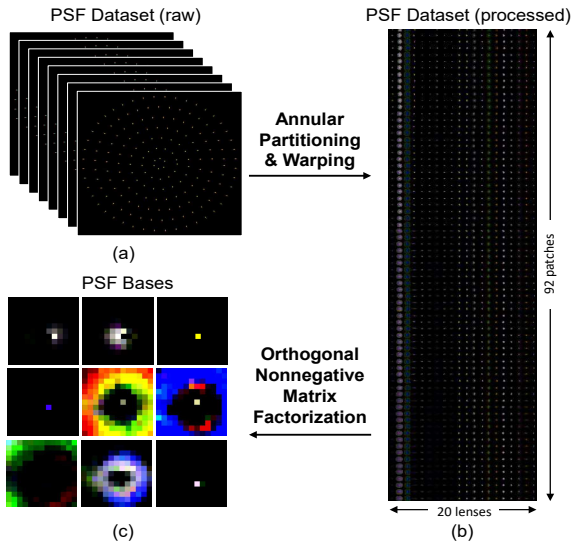

Figure 2. Illustration of the ONMF-based low-rank lens aberration model. (a) PSF dataset of 20 low-end lenses simulated by *Zemax*. (b) The PSF lattice consisting of 1840 PSF kernels, generated by dividing the images by each lens into 5 concentric rings, warping into stripes, and further cropping into 13×13-pixel patches. (c) 9 top PSF bases obtained by applying Orthogonal Nonnegative Matrix Factorization onto (b).

during decomposing the complex aberrations and the non-negativity facilitates compensating aberrations in the same way as handling conventional PSFs.

To this end, we first collect 20 representative low-end lenses from ZEBASE database and construct a high-dimensional matrix encompassing all their PSFs after annular division (5 rings), ring-to-rectangle warping and cropping. The matrix is then factorized using the ONMF model [30] and yield a set of principal bases. The workflow detailing this procedure as visualized in Fig. 2. So far, we have arrived at a low dimensional representation of the lens aberrations, which can cover the image degeneration of diverse lenses.

4. Universal Framework for Lens Aberration Correction

The above idea of building a low-rank model to compensate the lens aberration in a divide-and-conquer manner is non-trivial for several reasons: the decomposition of aberrations in a blurry input is highly imposed; the deconvolutions need to be conducted on the components which are of different intensity ranges with general natural images; the fusion should also tolerate the imperfections in both decomposition and deconvolution. To overcome these hurdles, we propose to design a jointly trained low-rank deep network that enables flexible optical aberration correction. Specifically, our network comprises three main modules trained in an end-to-end manner. The first module conducts ONMF-based decomposition, followed by several identical adaptive deconvolution modules. Finally, we incorporate a synthetic network module to further enhance the correction result. Fig. 3 summarizes the workflow of the whole network and reports the architecture of the key modules.

4.1. Decomposing Aberrations Attributed to PSF Bases

As aforementioned, after applying matrix decomposition to the PSF library, we get a basis set $\{\mathbf{B}_i\}$, which can compose an arbitrary PSF $\mathbf{k}$ with corresponding coefficients $\{\alpha_i\}$, i.e.,

$$\mathbf{k} = \sum_i \alpha_i \cdot \mathbf{B}_i. \tag{1}$$

Hence, a recorded degraded patch $\mathbf{Y}$ can be represented as

$$\mathbf{Y} = \mathbf{X} \otimes (\sum_i \alpha_i \cdot \mathbf{B}_i) + \mathbf{n} = \sum_i (\alpha_i \cdot \mathbf{X}) \otimes \mathbf{B}_i + \mathbf{n} \tag{2}$$

where $\mathbf{X}$ is the latent sharp image, $\mathbf{n}$ is the noise and $\otimes$ denotes 2D convolution.

As Fig. 3 shows, suppose we can decompose the blurry Y into components $\{\mathbf{Y}_i = (\alpha_i \cdot \mathbf{X}) \otimes \mathbf{B}_i\}$, we can pre-train deconvolution models to compensate the aberrations caused by $\{\mathbf{B}_i\}$ and estimate $\mathbf{X}$ by simply estimating the scaling factor (α_i). The decomposition is implemented with a deep neural network built on the U-net structure. Here we replace the plain convolution in original U-net with residual blocks for better convergence. Notice that the network is fully convolutional and can be applied to images with an arbitrary size. In our experiment, we divide the acquired image into equidistant annual rings with different widths based on the lens's physical properties.

In order to ensure the decomposition performance of the network, we define the following loss function

$$\mathcal{L}_{\text{decom}} = \|\mathbf{Y} - \sum_i \mathbf{Y}_i\|_2 + \sum_i \left(\mathbf{Y}_i - \tilde{\mathbf{Y}}_i\right) \tag{3}$$

with $\mathbf{Y}_i$ and $\tilde{\mathbf{Y}}_i$ denoting the decomposed blurry components and the generated version by *Zemax* software, equivalent to $\mathbf{X} \otimes (\alpha_i \cdot \mathbf{B}_i)$. Here the first term forces the summation of the decomposed components consistent with the input, and the second term ensures that each retrieved component is equal to the convolution of the sharp image $\mathbf{X}$ with the corresponding PSF basis.

4.2. Adaptive Feature-Domain Wiener Deconvolution

The ONMF-based decomposition module can extract the blurry components caused by the corresponding PSF bases, so we design matching compensation modules and embed them into our joint framework. However, both the range and pattern of the intensities in the decomposed components differ from natural images, and deconvolving these components in the spatial domain is prone to ringing artifacts and over-smoothness. A recent study [6] demonstrates that high-quality image reconstruction can be achieved by performing

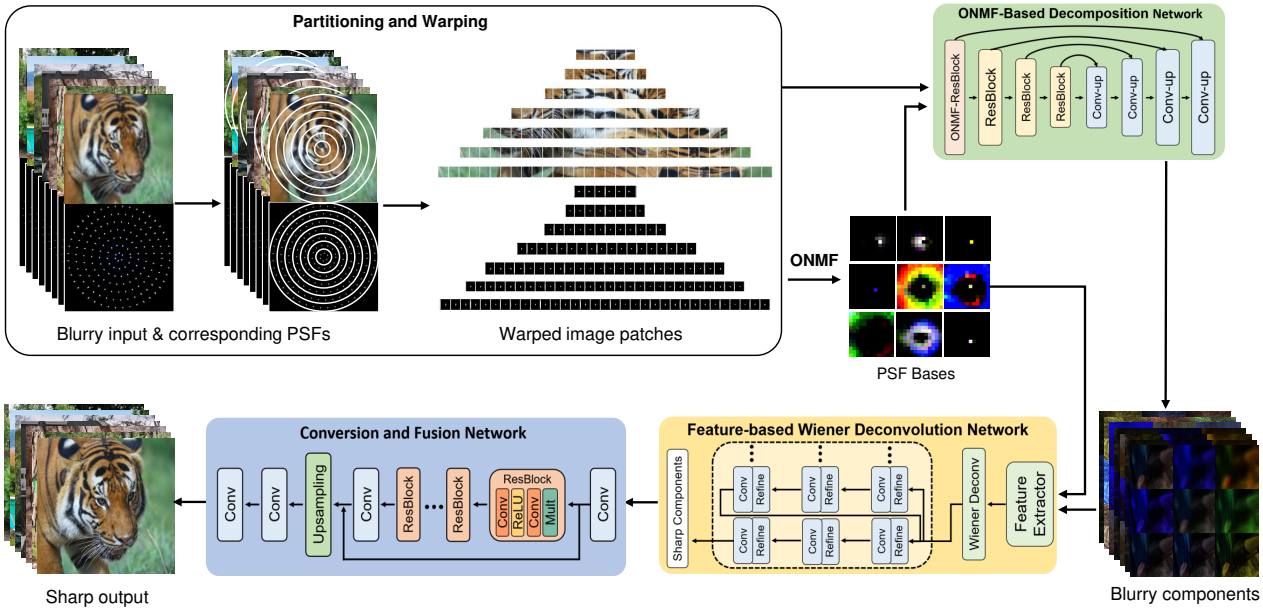

Figure 3. The framework of the proposed approach. The whole pipeline comprises preprocessing to build the low-rank aberration model and a divide-and-conquer compensation. The preprocessing involves annular partitioning of images and corresponding lens PSFs, ring-to-rectangle warping, and learning a low-dimensional representation of the PSFs. The compensation consists of three primary modules: (i) Decomposition, a neural network based on Orthogonal Non-Negative Matrix Factorization (ONMF), decomposing the blurry image into components corresponding to the representation bases of lens aberration; (ii) Deconvolution, implemented as a cascaded encoder-decoder network to map the decomposed components into the feature domain and conducts pre-trained Wiener deconvolution sequentially; (iii) Fusion, which aggregates the multiple scaled versions of the latent sharp image from the previous deconvolution modules to get the final output. The whole network is trained in an end-to-end manner.

deconvolution in the feature domain. Thus, we propose an adaptive feature-domain Wiener deconvolution module to recover the lost high frequencies better.

Specifically, for a specific patch $\mathbf{X}_i = \alpha_i \mathbf{X}$, we reconstruct it from the i-th blurry decomposed component $\mathbf{Y}_i$

$$\mathbf{X}_i^* = \arg\min \|\mathbf{Y}_i - \mathbf{X}_i \otimes \mathbf{B}_i\|, \tag{4}$$

where $\mathbf{X}_i$ and $\mathbf{Y}_i$ respectively represent the sharp and blurry image components matching the i-th PSF basis. In implementation, we build a feature-based Wiener adaptive deconvolution network. We denote f_i as a set of learnable linear filters and convolve $\mathbf{Y}_i$ with them to extract useful features and obtain the relationship among the blurry input, PSF, and the high-quality output in the feature domain. According to the properties of convolution, Eq. 2 turns into

$$\mathbf{F}_i \mathbf{Y}_i = \mathbf{F}_i (\mathbf{X}_i \otimes \mathbf{B}_i) + \mathbf{F}_i \mathbf{n}, \tag{5}$$

where multiplying with $\mathbf{F}_i$ is equivalent to convolving with f_i. Correspondingly, the above optimization in Eq. 4 is equivalent to finding a set of feature-based Wiener deconvolution operators $\mathbf{G}_i$ (which can be obtained based on the conclusion in [6]) to reverse the aberration by $\mathbf{B}_i$

$$\mathbf{X}_i^* = \arg\min \|\mathbf{G}_i \mathbf{F}_i \mathbf{Y}_i - \mathbf{F}_i \mathbf{X}_i\|. \tag{6}$$

The compensation is implemented as a deep neural network with the network structure and learned in a data-driven manner. During training the loss function is designed in an intuitive manner:

$$\mathcal{L}_{\text{deconv}} = \|\mathbf{F}_i \mathbf{X}_i - \mathbf{G}_i \mathbf{F}_i \mathbf{Y}_i\|_2. \tag{7}$$

The network is designed to share the parameters across all scales except for the first encoder block at the first cascade, which helps achieve fast and high-quality deconvolution.

4.3. Attention-based Fusion

According to Eq. 2, each deconvolution model can provide an estimation of the latent high-quality image, multiplied by a scaling factor. However, the potential inaccuracy in decomposition and artifacts in deconvolution would harm the quality of the final output. To overcome this challenge, we propose an effective strategy to streamline the reconstruction process. Specifically, we decompose more components from the blurry input and secondly apply the corresponding deconvolution to obtain multiple aberration-compensated versions, then fuse them together via a weight-trainable fusion network to raise robustness.

The above strategy introduces a substantial increase in running time. To accelerate the training of the deconvolution modules, we adopt a coarse-to-fine strategy, i.e., training a base model and subsequently fine-tuning it, which is largely faster than training all these basis-specific networks from

scratch. Moreover, our investigation reveals that the errors in the decomposition module can harm the successive deconvolution. Consequently, we introduce the decomposition confidence to serve as a valuable indicator of the decomposition accuracy/reliability of the decomposition process, and use it to guide the fusion.

5. Experiments

In this section, after describing the details of model training in Sec. 5.1, we first analyze the advantages of specific designs and key parameter settings in our approach (Subsection 5.2). Then, we demonstrate our superior performance against the state-of-the-art (SOTA) algorithms on synthetic (Subsection 5.3) and real data (Subsection 5.4).

5.1. Implementation Details

For model training, we gather 300 images from the Flickr2K dataset for model training, ensuring wide applicability for diverse natural scenes. Specifically, we select 20 common commercial lenses and simulate their spatially varying PSF using *Zemax* software. Then we simulate the lens aberration via convolving 100 high-definition (2K) images from Flickr2K with the generated PSF and apply the successive operations to train the model, including annular decomposition, ring-to-rectangle warping, patch cropping, etc. In total, we actually obtained the sharp-blurry pair and PSF of 9200 (100×92) patches.

During model training, we adopt the Adam optimizer with default parameters. The learning rate is initialized as 10^{-4}, which is halved every 100 epochs. PyTorch code and trained models are available on our project page.

5.2. Influences of the Key Parameters/Settings

Partitioning strategies. By capitalizing the rotational symmetry of camera lenses, we employ annular partitioning to divide the image into a sequence of concentric rings, where patches in each ring share a highly similar PSF after warping into a stripe. This strategy substantially reduces the spatial variance among PSFs across the field of view, consequently reducing the required number of bases for accurate PSF representation. The illustration and performance of the proposed partitioning method are depicted in Fig. 4(a), whereas Fig. 4(b) shows the counterpart of conventional grid partitioning. From the middle row, one can notice the highly consistent PSFs in the same annular ring in (a), in contrast to the large PSF difference among the patches in (b).

We also compare the final performance of two partitioning strategies on the synthetic dataset. Remarkably, when utilizing an equal number of bases (9 bases), the annular splitting reaches an impressive PSF representation accuracy of 0.93 and performs 30.16dB in the final reconstruction. In contrast, the conventional grid splitting yields a largely lower accuracy of 0.69 and the reconstruction achieves 27.77dB.

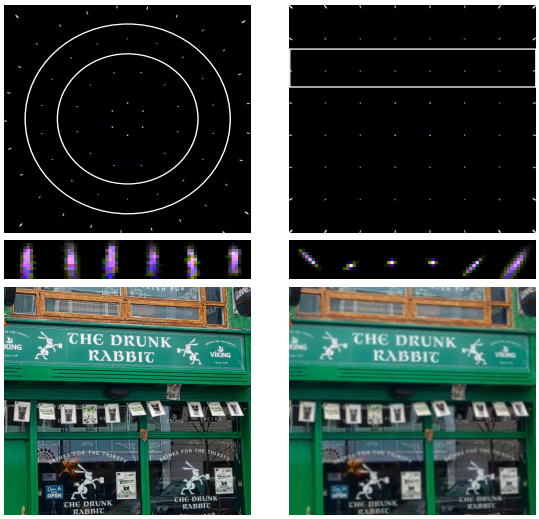

Figure 4. Comparison between the annular partitioning in our model (a) and conventional grid partitioning (b). Top row: the illustration of partitioning. Middle row: the PSFs within the highlighted regions in the top row, with the left one stretched to a rectangle. Bottom row: the aberration compensation results with the same number of bases (9 in our experiment).

Visually, we show an example in the 3rd row of Fig. 4, which shows that annular partitioning exhibits noticeable advantageous enhancement.

The number of annular rings. From the previous experiment, we came to the conclusion that annular partitioning is a better option. Further, we study the proper setting of the number of concentric rings, which will directly affect the amount of calculation and the precision of PSF representation. Although increasing the number of split rings can handle the radical difference of lens aberration better, it also brings higher computational complexity. Hence, we traverse several levels of radical division, i.e., the number of annular rings to find a good balance between precision and efficiency. The results are shown in Table 1, which shows that for usual commercial lenses, the performance improvement becomes marginal when the number grows beyond five. Therefore, we compromised and chose to split into 5 rings in our experiments.

Table 1. The performance at varying numbers of rings.

# of Rings	Patch Size (pixels)	Data Volume	PSNR (dB)	SSIM
3	214×214	75	29.12	0.914
5	**128×128**	**92**	**30.96**	**0.952**
7	92×92	132	30.99	0.953
9	72×72	160	30.91	0.950

The number of bases. We applied ONMF to obtain a low dimensional representation of the lens aberration, i.e., PSF. Intuitively, more bases provide a higher PSF representation accuracy that helps the reconstruction but tends to increase the ill-posedness of the decomposition network and the num-

ber of required deconvolution modules, harming the final performance as well as the running efficiency on the contrary. Therefore, pursuing the optimal number of bases is of crucial importance. As Tab. 2 shows, we test the performance with different numbers of bases and obtain a good balance at 9 bases, with high quality and low computational complexity.

Table 2. The performances using different numbers of bases.

# of bases	5	7	**9**	11	13
SSIM	0.901	0.927	**0.952**	0.955	0.955
PSNR (dB)	28.14	29.12	**30.96**	31.01	31.00
Training time	~9hrs	~11hrs	**~12hrs**	~16hrs	~18hrs

The applicability to various lenses. After determining the optimal number of bases, we test on 5 low-end commercial lenses (not included in the lens dataset) to verify the generalization ability of our model. By calculating the representation accuracy of the new lenses and the final lens correction result in Fig. 5, we arrive at two conclusions: the reconstruction quality is directly proportional to the accuracy of PSF representation; our approach is widely applicable for diverse lenses.

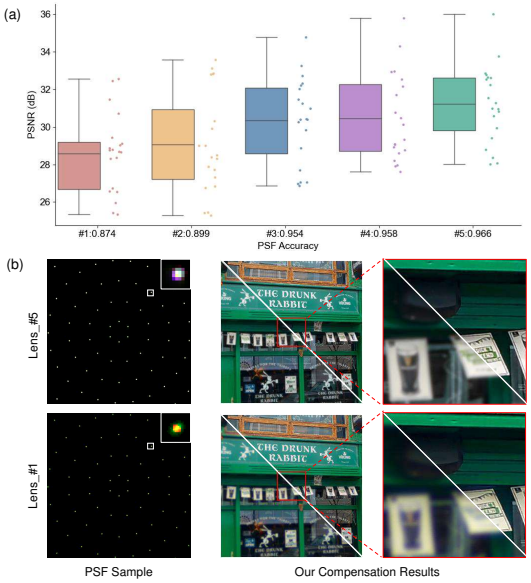

Figure 5. Our performance on diverse lenses. (a) The representation precision (horizontal axis) and their performance (vertical axis) on 5 test lenses, in terms of PSNR. Here the scattered box plot is drawn from the compensation results of 20 test images synthesized from the Flickr2K dataset. (b) PSF (left column) and our compensation result (right column, with the bottom-left and top-right insets being input and output respectively) of two lenses with the best and worst performance in (a).

Balancing both effectiveness and efficiency, we empirically divide the original low-quality images (1280×1280 pixels) into 5 concentric rings and select 9 bases (preserving an impressive 96.6% of the PSF variation in our lens database) in the final implementations of our model, based on the above quantitative experiment results.

5.3. Performance Comparison

We compare our approach with SOTA blind (DeblurGANv2 [17], MPRNet [26]) and non-blind (DPIR [12], DWDN [6]) deblurring methods, including both optimization-based and CNN-based algorithms. The results are shown in Tab. 3 and Fig. 6, from which one can observe the following trends: (i) Blind deblurring algorithms generally exhibit good performance in addressing motion blur problems, but in terms of lens aberration correction, our approach performs better than SOTAs and yields notably clearer images with finer details and fewer artifacts. (ii) By employing our low-rank PSF learning model, the PSF can be efficiently characterized, facilitating blind deconvolution to attain performance on par with or even superior non-blind algorithms. Overall, we achieve blind lens aberration correction, surpassing the SOTA blind deblurring methods, and perform on par with non-blind approaches.

Table 3. Quantitative performance comparison with SOTAs

	DeblurGANv2	MPRNet	Eboli's	DPIR	DWDN	Ours
Blind ?	✓	✓	✓	×	×	✓
PSNR (dB)	23.04	28.67	29.42	31.86	31.78	30.96
SSIM	0.726	0.927	0.934	0.962	0.960	0.952

5.4. Real experiments

To test the performance on real data, we use several low-end compact commercial cameras composed of a simple lens to capture low-quality images and computationally raise their visual quality with our model. Besides the results in Fig. 1, we show two more examples in Fig. 7, with the photo of the cameras in the left column, blurry input in the 2nd column, and the final reconstructed high-quality outputs in the rightmost column. We also show the side-by-side zoomed-in comparison of highlighted regions in the 3rd column for better visualization. One can see that the details in both the center and the periphery are recovered decently. Also, the consistently improved quality validates the wide applicability of the proposed approach to various lenses.

6. Conclusion

We have reported a versatile scheme capable of compensating lens aberrations of various lenses in an end-to-end manner and without model retraining or refinement. The universality of our approach stems from two key designs: Firstly, we incorporate the key physical properties inherent in camera lenses, such as rotational symmetry and low-dimensional structure after ring-to-rectangle warping; (ii) we integrate a deep neural network to reverse the aberration in a divide-and-conquer manner, i.e., decompose the low-quality input into basic components corresponding to the low-dimensional compositions of the aberration model, and

Figure 6. Performance comparison with SOTA methods. (a) The input blurry image. (b) The result after aberration correction by our model. (c-i) The comparison among the results produced by different SOTA algorithms (d-h), in contrast to the blurry input (c) and ground-truth sharp version (i), with PSNR and SSIM scores presented in Tab. 3. Here we compare three ROIs, which are cropped from different locations to demonstrate the performance on non-uniform lens aberrations. Note that (d-h) are blind compensation results and (f)(g) are non-blind.

Figure 7. Results on real data captured by two compact cameras with a low-end lens. (a) The photo of two lenses, with the smaller one being only around 1cm, which is highly portable but exhibits significant aberrations. (b)(d) Raw images captured using the cameras in (a) and the results after compensation. (c) The zoomed-in view of the highlighted regions in (b) and (d), distinctly showcases the performance at both regions closer to the center (red boxes) and toward the corners (yellow boxes) of the sensor's field of view.

then adopt pre-trained compensation modules to reconstruct the high-quality image with high robustness. The proposed approach offers high generalization ability to diverse lenses and requires no expertise-demanding calibration. Moreover, we achieve performance comparable to existing methods with careful calibration and lens-specific model training.

So far, our experiments assume sufficient exposure and high pixel count, but we acknowledge the potential for future enhancements, such as accounting for more realistic noise models and addressing other degradations like downsampling. Looking ahead, we envision further advancements of our model, focusing on the development of a lightweight network and on-chip implementation. As computational aberration compensation continues to progress, our method serves as a promising step towards enabling practical and cost-effective optical aberration correction for a wide range of applications.

References

[1] Samuel Arba-Mosquera, Shwetabh Verma, and Shady T Awwad. Theoretical effect of coma and spherical aberrations translation on refractive error and higher order aberrations. In *Photonics*, page 116. MDPI, 2020. 2

[2] Johannes Brauers, Claude Seiler, and Til Aach. Direct psf estimation using a random noise target. In *Digital Photography VI*, pages 96–105. SPIE, 2010. 3

[3] Ayan Chakrabarti, Todd Zickler, and William T Freeman. Analyzing spatially-varying blur. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2512–2519. IEEE, 2010. 3

[4] Jinlin Cui and Wei Huang. Optical aberration correction for simple lenses via sparse representation. *Optics Communications*, pages 201–213, 2018. 3

[5] Mauricio Delbracio, Ignacio Garcia-Dorado, Sungjoon Choi, Damien Kelly, and Peyman Milanfar. Polyblur: Removing mild blur by polynomial reblurring. *IEEE Transactions on Computational Imaging*, pages 837–848, 2021. 3

[6] Jiangxin Dong, Stefan Roth, and Bernt Schiele. DWDN: deep wiener deconvolution network for non-blind image deblurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 9960–9976, 2021. 4, 5, 7

[7] Thomas Eboli, Jean-Michel Morel, and Gabriele Facciolo. Fast two-step blind optical aberration correction. In *European Conference on Computer Vision*, pages 693–708. Springer, 2022. 3

[8] T Furieri, A Bassi, and S Bonora. Large field of view aberrations correction with deformable lenses and multi conjugate adaptive optics. *Journal of Biophotonics*, page e202300104, 2023. 3

[9] O García-Liévanos and S Vázquez-Montiel. Free system of spherical and coma aberrations by use aspherical and diffractive surfaces. In *AIP Conference Proceedings*, pages 659–664. American Institute of Physics, 2008. 2

[10] Felix Heide, Mushfiqur Rouf, Matthias B Hullin, Bjorn Labitzke, Wolfgang Heidrich, and Andreas Kolb. High-quality computational imaging through simple lenses. *ACM Transactions on Graphics*, pages 1–14, 2013. 3

[11] Michael Hirsch and Bernhard Scholkopf. Self-calibration of optical lenses. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 612–620, 2015. 3

[12] Zhanli Hu, Hengzhi Xue, Qiyang Zhang, Juan Gao, Na Zhang, Sijuan Zou, Yueyang Teng, Xin Liu, Yongfeng Yang, Dong Liang, et al. DPIR-Net: Direct pet image reconstruction based on the wasserstein generative adversarial network. *IEEE Transactions on Radiation and Plasma Medical Sciences*, pages 35–43, 2020. 7

[13] Qi Jiang, Hao Shi, Lei Sun, Shaohua Gao, Kailun Yang, and Kaiwei Wang. Annular computational imaging: Capture clear panoramic images through simple lens. *IEEE Transactions on Computational Imaging*, 8:1250–1264, 2022. 3

[14] Jaihyun Koh, Jangho Lee, and Sungroh Yoon. Single-image deblurring with neural networks: A comparative survey. *Computer Vision and Image Understanding*, page 103134, 2021. 1

[15] Amit Kohli, Anastasios Angelopoulos, Sixian You, and Laura Waller. Shift-variant deblurring for rotationally symmetric systems. In *Computational Optical Sensing and Imaging*, pages CTh5A–4. Optica Publishing Group, 2021. 3

[16] Amit Kohli, Anastasios Angelopoulos, Sixian You, Kyrollos Yanny, and Laura Waller. Linear revolution-invariance: Modeling and deblurring spatially-varying imaging systems. *arXiv preprint arXiv:2206.08928*, 2022. 3

[17] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8878–8887, 2019. 3, 7

[18] Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics*, pages 70–81, 2007. 3

[19] Weili Li, Xiaoqing Yin, Yu Liu, and Maojun Zhang. Computational imaging through chromatic aberration corrected simple lenses. *Journal of Modern Optics*, pages 2211–2220, 2017. 3

[20] Xiu Li, Jinli Suo, Weihang Zhang, Xin Yuan, and Qionghai Dai. Universal and flexible optical aberration correction using deep-prior based deconvolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2613–2621, 2021. 3

[21] Esther YH Lin, Zhecheng Wang, Rebecca Lin, Daniel Miau, Florian Kainz, Jiawen Chen, Xuaner Cecilia Zhang, David B Lindell, and Kiriakos N Kutulakos. Learning lens blur fields. *arXiv preprint arXiv:2310.11535*, 2023. 3

[22] Ting Lin, ShiQi Chen, Huajun Feng, Zhihai Xu, Qi Li, and Yueting Chen. Non-blind optical degradation correction via frequency self-adaptive and finetune tactics. *Optics Express*, pages 23485–23498, 2022. 3

[23] Alice Lucas, Michael Iliadis, Rafael Molina, and Aggelos K Katsaggelos. Using deep neural networks for inverse problems in imaging: Beyond analytical methods. *IEEE Signal Processing Magazine*, pages 20–36, 2018. 1

[24] Virendra N Mahajan. Zernike annular polynomials for imaging systems with annular pupils. *Journal of the Optical Society of America*, pages 75–85, 1981. 2

[25] James P McGuire and Russell A Chipman. Polarization aberrations. 1. rotationally symmetric optical systems. *Applied optics*, pages 5080–5100, 1994. 2

[26] Armin Mehri, Parichehr B Ardakani, and Angel D Sappa. MPRNet: Multi-path residual network for lightweight image super resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2704–2713, 2021. 7

[27] Ali Mosleh, Paul Green, Emmanuel Onzon, Isabelle Begin, and JM Pierre Langlois. Camera intrinsic blur kernel estimation: A reliable framework. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4961–4968, 2015. 1

[28] Gregory Ongie, Ajil Jalal, Christopher A Metzler, Richard G Baraniuk, Alexandros G Dimakis, and Rebecca Willett. Deep learning techniques for inverse problems in imaging. *IEEE*

Journal on Selected Areas in Information Theory, pages 39–56, 2020. 1

[29] Kambiz Rahbar and Karim Faez. Blind correction of lens aberration using zernike moments. In *IEEE International Conference on Image Processing*, pages 861–864. IEEE, 2011. 2, 3

[30] Abderrahmane Rahiche and Mohamed Cheriet. Forgery detection in hyperspectral document images using graph orthogonal nonnegative matrix factorization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 662–663, 2020. 4

[31] Jiangpeng Rong, Shiyao Huang, Zeyu Shang, and Xianghua Ying. Radial lens distortion correction using convolutional neural networks trained with synthesized images. In *Asian Conference on Computer Vision*, pages 35–49. Springer, 2017. 3

[32] Christian J Schuler, Michael Hirsch, Stefan Harmeling, and Bernhard Schölkopf. Blind correction of optical aberrations. In *European Conference on Computer Vision*, pages 187–200. Springer, 2012. 1, 2, 3

[33] Yichang Shih, Brian Guenter, and Neel Joshi. Image enhancement using calibrated lens simulations. In *European Conference on Computer Vision*, pages 42–56. Springer, 2012. 3

[34] Tiancheng Sun, Yifan Peng, and Wolfgang Heidrich. Revisiting cross-channel information transfer for chromatic aberration correction. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3248–3256, 2017. 3

[35] Huixuan Tang and Kiriakos N Kutulakos. What does an aberrated photo tell us about the lens and the scene? In *IEEE International Conference on Computational Photography*, pages 1–10. IEEE, 2013. 3

[36] Kaiyi Tang, Shuangyang Zhang, Yang Wang, Xiaoming Zhang, Zhenyang Liu, Zhichao Liang, Huafeng Wang, Lingjian Chen, Wufan Chen, and Li Qi. Learning spatially variant degradation for unsupervised blind photoacoustic tomography image restoration. *Photoacoustics*, page 100536, 2023. 3

[37] Berge Tatian. Aberration balancing in rotationally symmetric lenses. *Journal of the Optical Society of America*, pages 1083–1091, 1974. 2

[38] Chao Wang, Juan Chen, Hongguang Jia, Baosong Shi, Ruifei Zhu, Qun Wei, Linyao Yu, and Mingda Ge. Parameterized modeling of spatially varying psf for lens aberration and defocus. *Journal of the Optical Society of Korea*, pages 136–143, 2015. 3

[39] Pei Wang, Wei Sun, Qingsen Yan, Axi Niu, Rui Li, Yu Zhu, Jinqiu Sun, and Yanning Zhang. Non-uniform motion deblurring with blurry component divided guidance. *Pattern Recognition*, page 108082, 2021. 3

[40] Esther Whang, David McAllister, Ashwin Reddy, Amit Kohli, and Laura Waller. Seidelnet: an aberration-informed deep learning model for spatially varying deblurring. In *AI and Optical Data Sciences IV*, pages 276–281. SPIE, 2023. 3

[41] Chudan Wu, Yan Wo, Guoqing Han, Zhangyong Wu, and Jiyun Liang. Non-uniform image blind deblurring by two-stage fully convolution network. *IET Image Processing*, pages 2588–2596, 2020. 1

[42] Zhenhua Xu, Huasong Chen, and Zhenhua Li. Fast blind deconvolution using a deeper sparse patch-wise maximum gradient prior. *Signal Processing: Image Communication*, page 116050, 2021. 3

[43] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma. Image super-resolution as sparse representation of raw image patches. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008. 1

[44] Kyrollos Yanny, Kristina Monakhova, Richard W Shuai, and Laura Waller. Deep learning for fast spatially varying deconvolution. *Optica*, pages 96–99, 2022. 3

[45] Tao Yue, Jinli Suo, Jue Wang, Xun Cao, and Qionghai Dai. Blind optical aberration correction by exploring geometric and visual priors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1684–1692, 2015. 1, 2, 3

[46] Dazhi Zhan, Weili Li, Xiaoqing Yin, Caiyun Niu, and Jin Liu. Psf estimation method of simple-lens camera using normal sinh-arcsinh model based on noise image pairs. *IEEE Access*, pages 49338–49353, 2021. 1

[47] Kaihao Zhang, Wenqi Ren, Wenhan Luo, Wei-Sheng Lai, Björn Stenger, Ming-Hsuan Yang, and Hongdong Li. Deep image deblurring: A survey. *International Journal of Computer Vision*, pages 2103–2130, 2022. 3

[48] Zhihong Zhang, Yuxiao Cheng, Jinli Suo, Liheng Bian, and Qionghai Dai. INFWIDE: Image and feature space wiener deconvolution network for non-blind image deblurring in low-light conditions. *IEEE Transactions on Image Processing*, pages 1390–1402, 2023. 3

[49] Changyin Zhou, Stephen Lin, and Shree K Nayar. Coded aperture pairs for depth from defocus and defocus deblurring. *International Journal of Computer Vision*, pages 53–72, 2011. 3