End-to-end optimization of optical systems with extended depth of field under wide spectrum illumination

Alice Fontbonne, Hervé Sauer, François Goudail

To cite this version:

Alice Fontbonne, Hervé Sauer, François Goudail. End-to-end optimization of optical systems with extended depth of field under wide spectrum illumination. Applied optics, 2022, 61 (18), pp.5358-5367. 10.1364/ao.457214 . hal-03699106

HAL Id: hal-03699106
https://hal.science/hal-03699106v1

End-to-end optimization of optical systems with extended depth of field under wide spectrum illumination

Alice Fontbonne,[1],* Hervé Sauer,[1] and François Goudail [1]

[1] Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127, Palaiseau, France
*alice.fontbonne@protonmail.com

Abstract

We study a way to take into account the scene illumination spectrum during end-to-end optimization of optical-digital hybrid systems that include annular binary phase masks to enhance their depth of field (DoF). We show that a phase mask specifically optimized for wide spectrum panchromatic imaging performs better under this type of illumination than phase masks optimized under monochromatic illumination assumption. Indeed, thanks to spectral averaging, the modulation transfer functions of such a phase mask are very close to each other. This guarantees a very homogeneous image quality across the DoF range, which we demonstrate theoretically and experimentally using a dedicated optical setup.

This accepted author version has been published as: Alice Fontbonne, Hervé Sauer, and François Goudail, "End-to-end optimization of optical systems with extended depth of field under wide spectrum illumination," Appl. Opt. 61, 5358-5367 (2022)
DOI : https://doi.org/10.1364/AO.457214

1 INTRODUCTION

The depth of field (DoF) of an imaging system can be enhanced by placing an appropriate phase mask in its aperture stop. This mask allows us to "engineer" the point spread function (PSF) in order to make it more invariant to defocus without reducing the light throughput. However, the larger the DoF, the more blurred the acquired raw image so that these masks must be associated with an appropriate deblurring algorithm. The DoF extension capacity can be maximized by jointly optimizing the phase mask and the processing algorithm so as to reach the optimal compromise between invariance of PSF to defocus and capacity to deblur the image. Phase masks for DOF extension were introduced by Cathey and Dowski [6]. Then, Stork and Robinson [25] introduced joint optimization of optical systems and digital restoration algorithms, and this combined approach was applied to the optimization of DOF extending masks by Diaz et al. [3, 4]. This kind of approach is now extended to many forms of phase masks [15, 22, 14, 20, 1, 30], and to the use of neural networks [24, 7, 13, 16, 26, 27].

Among the many existing types of DOF extending masks, annular binary phase masks are easy to manufacture, and can be co-optimized with a simple Wiener deconvolution filter. Their performance and robustness have been characterized theoretically and experimentally in the case of monochromatic illumination [9, 11, 12]. However, the question remains of their efficiency under wide spectrum illumination [19, 18, 8]. Indeed, at first glance, binary phase masks seem not to be adapted to this configuration since they are most often manufactured by binary etching of a dielectric plate and rely on alternating 0 and π radian dephasing, which can be reached at only one wavelength. In a recent article, we have shown with simulations and experiments that binary masks optimized for a single wavelength can be robust to their use under wide spectrum illumination [10] and suggested, by using simulations, that even better results could be obtained by optimizing the mask specifically for wide spectrum illumination.

Since then, we manufactured and experimentally studied such masks, and we thus propose in the present article a comprehensive theoretical and experimental comparison of binary phase masks using different metrics: the modulation transfer function (MTF), which represents the purely optical properties of the optical system including this mask, the effective MTF, which takes into account the image post-processing, and the image quality after deconvolution, which takes into account noise amplification by post-processing. Our approach highlights the differences brought by a change of illumination spectrum on the hybrid system and on the final image quality. Masks optimized under monochromatic illumination with different wavelengths and under wide spectrum illumination are compared on the basis of these metrics and their relative performance are experimentally validated with imaging experiments.

This article is organized as follows: first, we describe the optimization procedure of phase masks for wide illumination spectrum in Sec. 2, and apply it to the co-optimization of different phase masks for various illumination spectra. Then in Sec. 3, we analyze the theoretical behavior of DoF enhancing phase masks in panchromatic imagers in terms of Modulation Transfer Function (MTF), effective MTF and final image quality. Finally, in Sec. 4, the conclusions drawn from the theoretical study are experimentally validated with imaging experiments carried out with a setup including a commercial lens, a Vis-NIR CMOS sensor and co-optimized phase masks.

2 Phase mask optimization under wide spectrum illumination

Our goal in this article is to study theoretically and experimentally the difference between binary annular phase masks optimized under wide spectrum illumination assumption, and more classical masks optimized under monochromatic spectrum illumination assumption. In this introductory section, we first review for the interested reader the classical optimization procedure of such binary phase masks, depending on the illumination spectrum. Most of this information can also be found in our previous work [10].

2.1 Wide spectrum illumination and defocus error

Let us consider that the scene seen by the imaging system has a mean spectral radiance $R_\lambda(\lambda)$ defined over the spectral range $[\lambda_{min}, \lambda_{max}]$. We assume that the scene is globally neutral and backscatters a mean spectrum proportional to the one of the lighting source. Moreover, we consider that the sensor has a spectral sensitivity $S(\lambda)$, and we define the normalized "effective" imaging spectrum $E(\lambda)$ as the product of $R_\lambda(\lambda)$ and $S(\lambda)$, normalized so that its integral over the whole spectral range is equal to 1. The shape of the polychromatic OTF $\tilde{h}_\Lambda(\nu_x, \nu_y)$ over a spectrum Λ (interested reader can refer to [10] for details on its construction) depends on the aberrations of the optical system, and in DoF extension problems, the main one is defocus error. It is convenient to parameterize this error by the signed wavefront error at the edge of the exit pupil or, in other words, the optical path difference (OPD) ψ on the marginal rays between the actual defocused wavefront and the reference sphere. Assuming that the exit pupil is far from the image plane, and using a first non-zero order expansion in the image Numerical Aperture NA and in the image distance shift Δz_i, this parameter can be written as a function of Δz_o, the object defocus distance:

$$\psi(\Delta z_o) \approx \frac{(z_{i0} NA)^2}{2} \left(\frac{1}{z_{i0}} + \frac{1}{z_{o0} + \Delta z_o} - \frac{1}{f'} \right) \quad , \tag{1}$$

where f' is the effective focal length, z_{o0} is the distance between the nominal object position and the lens object principal plane H (taken as positive for real objects in front of the lens) and z_{i0} is the distance between the lens image principal plane H' and the nominal image plane, which is also the fixed detector position. These quantities are depicted on the scheme Fig. 1. Note that since the effective focal length and the positions of the principal planes depend on the

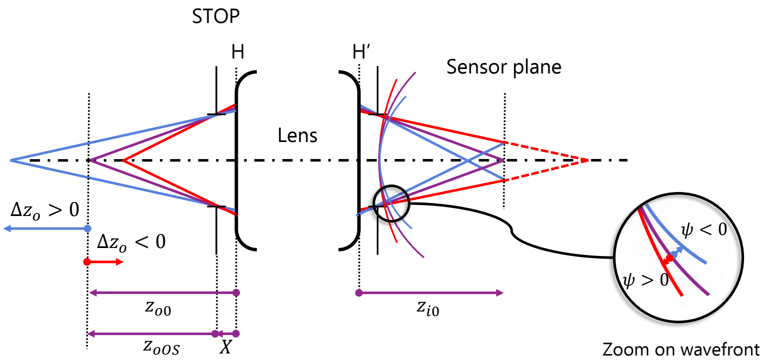

Figure 1: Diagram linking the object defocus distance Δz_o to the defocus parameter ψ. Notation is defined in the text on Sec. 2.1 and Sec. 4.1.

wavelength, the value of the defocus parameter may also be wavelength-dependent. We have chosen to neglect this dependency in this article as we have checked that for the commercial lens used in the experimental setup described in Sec. 4, this dependency is indeed negligible. Under this hypothesis, the parameter ψ is independent of the wavelength. Nevertheless, the underlying optical wave dephasing at wavelength λ is equal to $2\pi\psi/\lambda$ and the imaging properties depend of the wavelength through this parameter. This is why, in the present article, we always express the value of the OPD ψ as a real-valued factor times a wavelength (somewhat arbitrarily) chosen to be noted λ_G. For example, according to the Rayleigh criterion, at the given wavelength of illumination λ_G, it can be said that the defocus parameter range in which the image quality of a conventional optical system can be considered nearly as good as if it were focused is $[\psi_{min}, \psi_{max}] = [-\frac{\lambda_G}{4}, \frac{\lambda_G}{4}]$. Our goal is to extend its DoF range up to $[-2\lambda_G, 2\lambda_G]$ for a panchromatic imaging system, that is, multiply the DoF range by eight. We will also compare the DoF extension performance obtained under wide spectrum illumination with the performance obtained under green and red monochromatic illumination.

2.2 Co-designed phase masks for DoF-extended panchromatic imaging

We extend the DoF of an imaging system by placing a binary annular phase mask in its aperture stop. These masks are glass plates inducing a phase modulation of alternatively 0 or π radians at a chosen nominal wavelength, thanks to the engraving of a series of N concentric annular regions. Each annular constant-phase area is called a "ring", so that a N-ring mask of clear aperture radius R is parameterized by $N-1$ normalized radius values $\phi = \{\rho_1, ..., \rho_{N-1}\}$, where the radius of the i'th phase transition is $r_i = \rho_i R$ and ϕ satisfies the conditions $0 < \rho_1 < ... < \rho_{N-1} < \rho_N = 1$.

In practice, the binary phase masks we use are manufactured as binary depth modulation on fused silica substrates using UV photolithography associated with Ion Beam Etching (IBE) followed by Inductive Coupled Plasma (RIE - ICP) etching. This process allows to easily etch rings with a lateral precision of a few micrometers, and a precision on the etching depth of ±20 nm (the quality of the manufactured mask has been shown to be sufficient in [11]). The etching depth e has a decisive influence on the mask properties. Under monochromatic spectrum illumination at wavelength λ_0, choosing a step height corresponding to a phase shift of π radians is natural since it is the only one that leads to a symmetrical behavior of the hybrid system imaging on either side of the focal point (i.e., the OTF is identical for defocus parameter ψ and $-\psi$), making it possible to double the reachable DoF. For that purpose, the etching depth must have the following value:

$$e = \frac{\lambda_0}{2(n_{\lambda_0} - 1)} \quad , \tag{2}$$

where n_{λ_0} is the refractive index of the material (in our case, fused silica) at λ_0. For example, for green monochromatic illumination at $\lambda_0 = \lambda_G = 545$ nm, one has $e = 592$ nm, and for red illumination at $\lambda_0 = \lambda_R = 781$ nm, the required etching depth is $e = 864$ nm.

In contrast, under wide spectrum illumination, the choice of the etching depth is not so obvious and it is necessary to consider the etching depth as a free optimization parameter, in the same way as the normalized radius values ϕ.

Co-designing a panchromatic imaging system thus consists in jointly determining the deconvolution filter, the mask radii ϕ and the etching depth e that minimize a co-designed criterion. Although there is much discussion about the best way to evaluate the quality of a restored image [33, 29, 32], we choose the mean square error (MSE) because it leads to analytical expressions of the optimal deconvolution filter and of the optimization criterion [5]. It can be shown that the MSE may be expressed as follows:

$$MSE(\phi, e, \psi) = \iint \Big(|\widetilde{h}^{\phi,e}_{\Lambda,\psi}(\nu_x,\nu_y)\tilde{w}^{\phi,e}_{\Lambda,\Psi}(\nu_x,\nu_y) - 1|^2 S_{oo}(\nu_x,\nu_y) + \\ |\tilde{w}^{\phi,e}_{\Lambda,\Psi}(\nu_x,\nu_y)|^2 S_{nn}(\nu_x,\nu_y)\Big) d\nu_x d\nu_y \quad (3)$$

where $S_{oo}(\nu_x,\nu_y)$ and $S_{nn}(\nu_x,\nu_y)$ are the power spectral density (PSD) of the signal and the noise respectively, and $\tilde{w}^{\phi,e}_{\Lambda,\Psi}(\nu_x,\nu_y)$ is a unique deconvolution filter. We model the object PSD $S_{oo}(\nu_x,\nu_y)$ as a function proportional to $\nu^{-2.5}$ with $\nu = \sqrt{\nu_x^2 + \nu_y^2}$, which is a realistic statistical model of natural scenes[21, 28, 9, 11]. In addition, we will consider in the remainder of the paper that, unless otherwise stated, the PSD of the noise $S_{nn}(\nu_x,\nu_y)$ is white. The goal is to make this MSE as low as possible over the targeted defocus range, defined by a discrete set of K defocus parameter values so that $\psi \in \Psi = \{\psi_1, \psi_2, \cdots, \psi_K\}$. For a given set Ψ and configuration (ϕ, e), the linear deconvolution filter that minimizes the average MSE is the average Wiener filter, which has the following expression in the Fourier space [5]:

$$\tilde{w}^{\phi,e}_{\Lambda,\Psi}(\nu_x,\nu_y) = \frac{\frac{1}{K}\sum_{k=1}^{K}[\widetilde{h}^{\phi,e}_{\Lambda,\psi_k}(\nu_x,\nu_y)]^\star}{\frac{1}{K}\sum_{k=1}^{K}\left|\widetilde{h}^{\phi,e}_{\Lambda,\psi_k}(\nu_x,\nu_y)\right|^2 + \frac{S_{nn}(\nu_x,\nu_y)}{S_{oo}(\nu_x,\nu_y)}} \quad , \quad (4)$$

where the superscript $\star$ denotes the complex conjugate. The optimal parameters $\{\phi_{opt}, e_{opt}\}$ of the mask are obtained by optimizing the following "minimax" criterion:

$$\{\phi_{opt}, e_{opt}\} = \arg\min_{\phi,e}\left\{\max_k\left[\mathrm{MSE}(\phi, e, \psi_k)\right]\right\} \quad , \quad (5)$$

with an intensive particle swarm global numerical optimization algorithm [23].

We have used this optimization method to design three binary phase masks. We considered a SNR equal to 34 dB since it corresponds to the case of a well contrasted image of a natural scene acquired with a CMOS camera, where the dominant source of noise is photon shot noise. The first one, denoted "$2\lambda_G$-mono mask", is the optimal mask [9] for a targeted DoF range of $[-2\lambda_G, 2\lambda_G]$, which is 8 times the natural DoF range of the lens used with a monochromatic illumination at $\lambda_G = 545$ nm. The second one, denoted "$2\lambda_R$-mono mask", is the optimal mask [9] for a larger targeted DoF range of $[-2\lambda_R, 2\lambda_R]$, which is 8 times the natural DoF of the lens used with a monochromatic illumination at $\lambda_R = 781$ nm. It has exactly the same parameters ϕ as the $2\lambda_G$-mono mask, except the etching depth which is adapted to the wavelength (Eq. 2). The third phase mask is optimized to be used under wide illumination spectrum: we used the sun spectrum approximated as a simple model of a black body at 5778 K for $R_\lambda(\lambda)$ and the sensitivity of a silicon sensor for $S(\lambda)$. The targeted defocus parameter range is $[-2\lambda_G, 2\lambda_G]$, so this phase mask is denoted "$2\lambda_G$-wide mask". The optimal parameters of these three masks (etching depth and normalized outer radii) are given in Fig. 2 and the spectra used for their optimization are displayed in Fig. 3. As in [9], the chosen number of rings corresponds to a mask giving the best performance with the lowest number of rings.

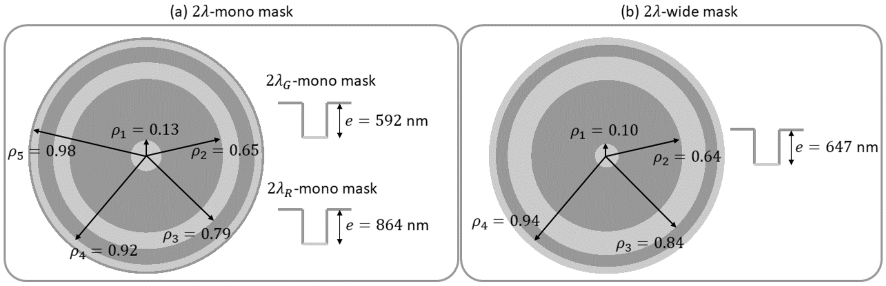

Figure 2: Step height and normalized radii of the optimal phase masks obtained under monochromatic spectrum assumption ((a) 2λ_G-mono mask and 2λ_R-mono mask) and under wide spectrum assumption ((b) 2λ_G-wide mask). The mask substrate is made of fused-silica.

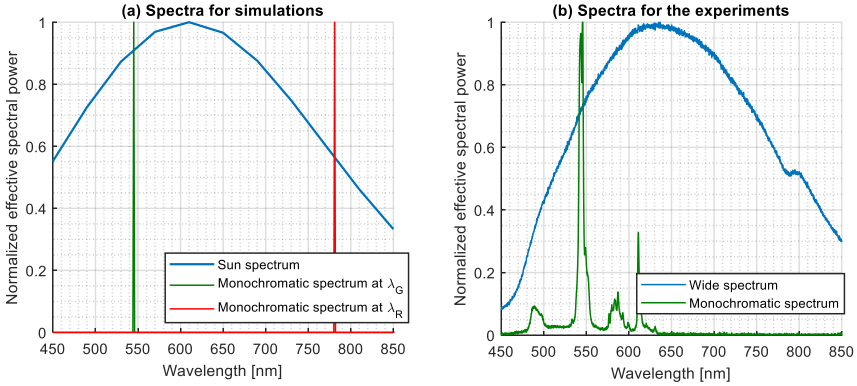

Figure 3: (a) Effective theoretical spectrum $E(\lambda) = R_\lambda(\lambda) \times S(\lambda)$ used during the optimization process. (b) Effective spectrum of a fluorescent lamp with green filter ("Monochromatic spectrum") and an halogen lamp ("Wide spectrum") based on their measured spectrums used during experimentation. For the sake of clarity on the plots here, all the spectrums are represented normalized to one at their maxima.

3 PREDICTED PERFORMANCE OF PANCHROMATIC AND MONOCHROMATIC PHASE MASKS

The robustness of the 2λ_G-mono mask to a wide spectrum has already been verified in [10]. In this present paper, we theoretically investigate whether this robustness depends on the etching step height and check in detail whether a phase mask specifically optimized with a wide spectrum can improve the performance. We carry out such a study by analysing the MTF (Sec. 3.1), the effective MTF (Sec. 3.2) and by quantitatively evaluating the image quality (Sec. 3.3).

3.1 Analysis of MTFs

The first column of Fig. 4 displays the MTFs for several values of ψ within the targeted defocus range and when the illumination spectrum is the one for which they were optimized. The MTFs have been computed by Fourier optics, assuming that the lens is diffraction limited (without phase mask) and the only defect is defocus. The x-scale of the graphs is the reduced spatial frequency relative to the fixed Nyquist frequency of the sensor that will be used in the experimental validation of Sec. 4. First, we observe in Fig. 4(a) that the MTFs of the 2λ_G-mono mask are exactly the same for a defocus value ψ_k and its opposite $-\psi_k$. This was expected since the phase step between the rings of the mask is exactly equal to π. The second main observation is that for all values of ψ in $[-2\lambda_G, 2\lambda_G]$, the MTFs are all very close to each other. This means that the images obtained with an imaging system using this mask have the same quality over a large range of defocus parameters ψ. Moreover, as the MTFs do not exhibit zeros at low and medium spatial frequencies, image restoration can be expected to be effective, leading to a significant DoF enhancement.

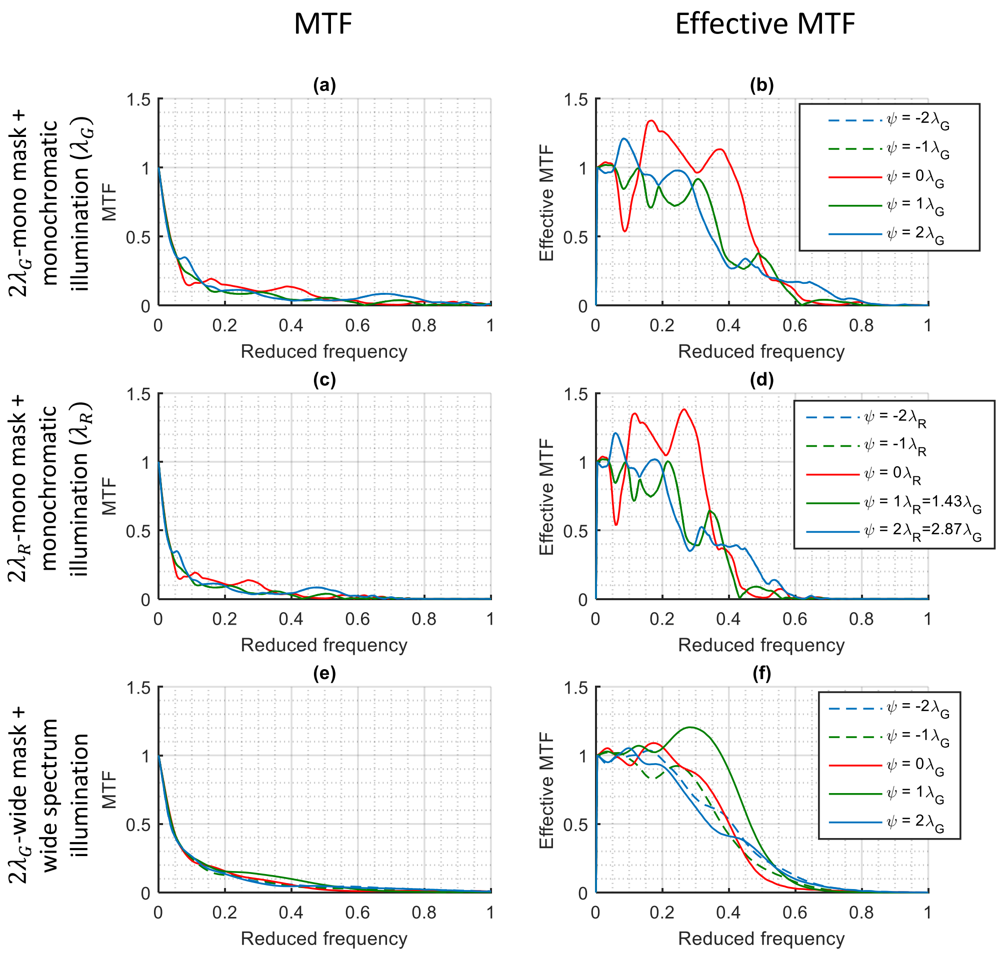

Figure 4: MTFs (first column) and effective MTFs (second column) for several defocus parameter values and different configurations. (a-b) For the $2\lambda_G$-mono mask, when the illumination is monochromatic at λ_G. (c-d) For the $2\lambda_R$-mono mask, when the illumination is monochromatic at λ_R. (e-f) For the $2\lambda_G$-wide mask with a wide spectrum illumination. The legend for (a-c-e) is the same as for (b-d-f). The reduced frequency corresponds to the spatial frequency, normalized by the Nyquist frequency of the sensor used for the experimentation (see Sec. 4).

Figure 4(c) displays the MTFs of the $2\lambda_R$-mono mask used under λ_R monochromatic illumination, for different values of ψ spanning the $[-2\lambda_R, 2\lambda_R]$ DoF range. They are very similar to those of Fig. 4(a), except for a λ_G/λ_R scale factor on the spatial frequency axis. This shape similarity was expected since the ring diameters ϕ_{opt} of both masks are the same. The observed scale factor is the standard scaling of the OTF with wavelength in Fourier optics, which is the same whatever the defocus. Since, for a given spatial frequency, the $2\lambda_R$-mono mask leads to lower MTF than the $2\lambda_G$-mono mask, noise issues may appear for the former mask because the signal is drowned by noise (and thus cannot be recovered) at lower spatial frequencies.

In the Fig. 4(e), we can see that the $2\lambda_G$-wide mask exhibits MTFs that are also similar to each other and somewhat similar to those of the $2\lambda_G$-mono mask. There are however some differences. Firstly, the MTFs are no longer exactly equal on either side of the DoF range. Furthermore, thanks to spectral averaging, they are less "bumpy", and thus even more similar to each other. Therefore, it can be anticipated that this phase mask will perform better than the $2\lambda_G$-mono mask when a single average Wiener filter is used for deconvolution [10].

3.2 Analysis of effective MTFs

The MTFs are not sufficient to fully describe the behavior of an entire optical/digital hybrid system as they do not take into account the digital post-processing. A better metric in that respect is the *effective* MTF, which includes the (deterministic) effect of the deconvolution filter:

$$|\tilde{h}_{\mathit{eff}}^{\Lambda,\phi,e,\psi_k}(\nu_x,\nu_y)| = |\tilde{h}_{\Lambda,\psi_k}^{\phi,e}(\nu_x,\nu_y) \times \tilde{w}_{\Lambda,\Psi}^{\phi,e}(\nu_x,\nu_y)| \quad . \tag{6}$$

The effective MTF can be seen as the "transfer function" between each spatial frequency component of the ideal image and those of the post-processed image. Ideally, it should be uniformly equal to 1, which means that the restored image is perfectly equal to the clean one in the absence of noise. However, this is not the case when using the average Wiener filter: values greater than 1 are observed at some spatial frequencies, which means that the signal at these frequencies is over-amplified, leading to stronger contrasts than in the clean image and possibly to ringing effect (artefact) at sharp edges. Note that this over amplification is "required" for the average Wiener filter to minimize the MSE criterion [5]. This is due to the fact that the average Wiener filter is the result of an averaging over the all the MTFs within the targeted defocus range, and for this reason, it is not perfectly adapted to the MTF at a given defocus. The effective MTFs of the $2\lambda_G$-mono mask used under a λ_G-monochromatic illumination are represented on Fig. 4(b). They are all close to one for low spatial frequencies, which means that images can correctly be restored for all defocus values within the targeted DoF range. Fig. 4(d) shows that the effective MTFs of the $2\lambda_R$-mono mask used under a monochromatic illumination at λ_R are also close to one for low spatial frequencies, but they collapse more quickly. Indeed, since the optical cutoff frequency at λ_R is lower than at λ_G, the MTFs are lower relatively to the noise and the average Wiener filter can restore less spatial frequencies. Finally, it is observed in Fig. 4(f) that the best effective MTFs are obtained with the $2\lambda_G$-wide mask. Indeed, they are all very close to one for the lower spatial frequencies, and much smoother than with the other masks. This is due to the smoothness of the MTFs, that we have noticed in Fig. 4(e), and which is due to spectral averaging.

In practice, the illumination spectrum of a scene is rarely monochromatic. That is why it is interesting to check the robustness of the $2\lambda_G$-mono mask and the $2\lambda_R$-mono mask to their use in the presence of wide spectrum illumination. This is done on Fig. 5, where we display the MTFs and effective MTFs obtained in this situation. The $2\lambda_G$-mono mask is quite robust to wide spectrum illumination, because its MTFs are similar to each other over the whole DoF range (Fig. 5(a)). Nevertheless, we can see in Fig. 5(b) that the effective MTF for the defocus parameter $\psi = 0\lambda_G$ is higher than the others, and significantly larger than 1, which could create deconvolution artifacts [10, 2]. This is even more visible with the $2\lambda_R$-mono mask, which means that this phase mask is even less robust when used under wide visible illumination spectrum. Moreover, in this case, the effective MTFs for other values of ψ collapse rapidly. This behavior can be explained by the fact that the used spectrum is more centered on the green wavelengths than on the red wavelengths (See Fig. 3.a).

Even though the effective MTF is a valuable tool to characterize the properties of deconvolved images in terms of object spatial frequency filtering, it is not representative of noise amplification and artefacts due to deconvolution. Since these deleterious phenomena alter the final image perceptive quality, they are best assessed by evaluating quantitatively the quality of deconvolved images. This is done in the next paragraph.

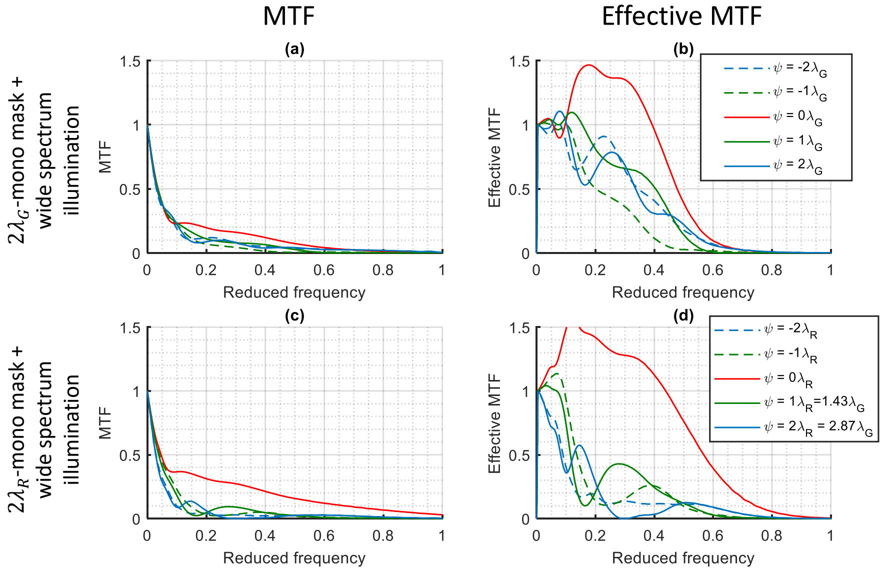

Figure 5: MTFs (first column) and effective MTFs (second column) for several defocus parameter values and different configurations. (a-b) For the $2\lambda_G$-mono mask, with a wide spectrum illumination. (c-d) For the $2\lambda_R$-mono mask, with a wide spectrum illumination. The legend for (a-c) is the same as for (b-d). The reduced frequency corresponds to the spatial frequency, normalized by the Nyquist frequency of the sensor used for the experimentation (see Sec. 4).

3.3 Image quality

The previously introduced MSE (Eq. 3) allows one to quantify the global difference between the final restored image and the ideal image. In the following, we will define the image quality (IQ) as:

$$IQ(\psi) = 10 \log_{10} \frac{1}{\mathrm{MSE}(\phi, e, \psi)} \quad , \tag{7}$$

It depends on the ratio between the total energy of the image (which has been normalized to 1) and the MSE. It thus increases as the MSE decreases, i.e., as the quality of the restored image increases. We omit the ϕ, e (mask parameters) and Λ (spectrum wavelengths)arrer qualifiers on the IQ symbol to simplify notation.

The IQ for the three phase masks, calculated under the assumption that the used spectra are the same as those used during their respective optimizations, is plotted in Fig. 6(a) as a function of ψ. We first notice that, over the $[-2\lambda_G, 2\lambda_G]$ range, the $2\lambda_G$-wide mask (used with a wide spectrum illumination) has a better IQ than the $2\lambda_G$-mono mask (used with a λ_G-monochromatic illumination). This means that, even if binary phase masks seem to be specifically suited for use with a monochromatic spectrum (since in this case, the phase shift between two rings is exactly equal to π rad), they can reach better performance under wide spectrum, thanks to spectral averaging, and provided that appropriate optimization (taking into account a wide spectrum) is performed. On average over the $[-2\lambda_G, 2\lambda_G]$ defocus parameter range, the IQ of the $2\lambda_R$-mono mask is smaller than the IQ of the $2\lambda_G$-mono mask and the $2\lambda_G$-wide mask, which was expected from the analysis of the effective MTF in Fig. 4. Outside this range, the IQ collapses for the two masks optimized for this DoF range (the $2\lambda_G$-mono mask and the $2\lambda_G$-wide mask). It is then the $2\lambda_R$-mono mask which performs better, because it has been optimized for the wider range $[-2\lambda_R, 2\lambda_R] = [-2.87\lambda_G, 2.87\lambda_G]$.

Figure 6(b) shows the IQ of the phase masks in the case where the illumination spectrum is wide. Note that the blue curves on Fig. 6(a) and Fig. 6(b) are therefore the same. However, the other curves are modified by this change of illumination spectrum. In particular, they are no

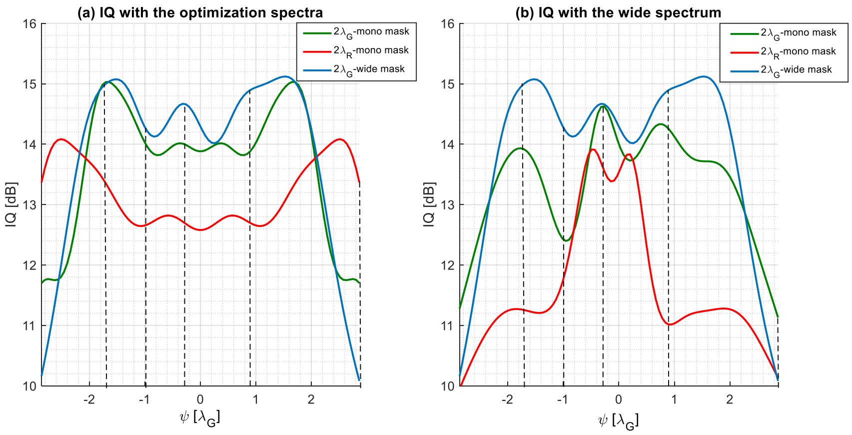

Figure 6: (a) IQ as a function of the defocus parameter for the three studied phase masks when they are used with the spectrum for which they have been optimized. (b) IQ as a function of the defocus parameter for the three studied phase masks when they are used with a wide spectrum illumination (blue curve is the same as in (a)). Vertical dotted lines indicates specific defocus parameter values of interest (see Sec. 4) .

longer symmetrical with respect to the focal plane ($\psi = 0$). The $2\lambda_G$-mono mask still performs quite well on the $[-2\lambda_G, 2\lambda_G]$ DoF range, which confirms that this phase mask is robust to this change of illumination spectrum. The $2\lambda_R$-mono mask, on the other hand, is not robust at all: its IQ collapses over the targeted DoF range because the illumination spectrum is far from the monochromatic red wavelength used for its optimization. For $\psi = 2.87\lambda_G = 2\lambda_R$, the $2\lambda_R$-mono mask and the $2\lambda_G$-wide mask have the same performance.

Therefore, several important conclusions can be drawn from these simulations. The first one is that the robustness of a monochromatically optimized phase mask strongly depends on the difference between the nominal wavelength and the characteristics of the wide spectrum (its width and barycenter wavelength). The second conclusion is that it is important to take into account the spectrum of use when optimizing a phase mask. The third one is that the phase mask optimized considering a wide spectrum illumination (the $2\lambda_G$-wide mask) and used with this wide spectrum performs significantly better than a mask optimized considering a monochromatic spectrum and used with this monochromatic illumination. In other words, even when using a binary phase mask, DoF enhancement is easier under wide spectrum than under monochromatic one thanks to spectral averaging.

4 EXPERIMENTAL VALIDATION OF PANCHROMATIC DOF ENHANCEMENT

The conclusions of Sec. 3 have been obtained with numerical simulations. They now need to be experimentally validated. In this section, we experimentally verify, with a real imaging setup, the superiority of a phase mask optimized considering wide spectrum illumination over a phase mask optimized considering monochromatic spectrum illumination.

4.1 Methodology

In order to experimentally investigate the behavior of the systems under monochromatic and wide spectrum illumination, we select two light sources whose spectra have roughly similar properties to those used for optimization and simulation (Sec. 2 and Sec. 3). Figure 3(b) gives the two effective illumination spectra used for the experimental validation. The so-called "monochromatic spectrum" is obtained with a fluorescent lamp and a green filter. The so-called "wide spectrum" is obtained with a halogen lamp. These two spectra are used to illuminate a scene which is observed with a hybrid optical-digital setup consisting of a commercial 50 mm lens (Edmund Optic #59-873), a co-designed phase mask put at few millimeters in front of the lens (which changes the STOP position, but with, in practice here at F/11.1, a negligible impact on the lens imaging performance), a Jai GO camera equipped with a 2464×2056 pixel, 3.45 μm pitch CMOS sensor and a digital post-processing. The characteristics of the setup are given in Tab. 1.

Focal length f'	50 mm
Phase mask position	New STOP of the optical system (Ø = 4.5 mm) placed in front of the lens (see Fig. 8(a))
F-number / Image numerical aperture	$F/11.1$ / 0.045 (with the new STOP)
Sensor size	8.5 mm ×7.09 mm
Pixel pitch / Nyquist spacial frequency	3.45 μm / 145 lp.mm^{-1}

Table 1: Characteristics of the hybrid system used for the experimentation.

To experimentally validate the DoF enhancement under wide spectrum illumination and the superiority of the $2\lambda_G$-wide mask, we place objects at different distances of the lens, corresponding to different defocus parameter values. The lens is focused at 58 cm (with this distance taken up to the front STOP position). Let us denote z_{oOS} the distance between the well-focused object and the STOP, which is easy to measure experimentally, and X the distance between the STOP and the image principal plane (Fig. 1). Using Eq. 1 with $z_{o0} = z_{\mathrm{oOS}} + X$ and $NA = \left(\frac{Ø/2}{z_{\mathrm{oOS}}} \times \frac{z_{o0}}{z_{i0}}\right)$ (Ø being the diameter of the STOP), one obtains the relation between the defocus parameter ψ and the object defocus distance Δz_o as a function of these parameters:

$$\psi(\Delta z_o) \approx \frac{Ø^2}{8}\left(1 + \frac{X}{z_{oOS}}\right)^2 \left(\frac{1}{z_{i0}} + \frac{1}{z_{\mathrm{oOS}} + X + \Delta z_o} - \frac{1}{f'}\right) \quad , \tag{8}$$

with $\frac{1}{z_{i0}} + \frac{1}{z_{\mathrm{oOS}}+X} - \frac{1}{f'} = 0$. We estimate the distance X between the entrance pupil (also the STOP here) and the object principal plane thanks to a "Zemax Black Box" of the Edmund Optics® lens. We obtain $X \approx 45$ mm, which allows us to plot the defocus parameter ψ as a function of the object defocus distance, Fig. 7.

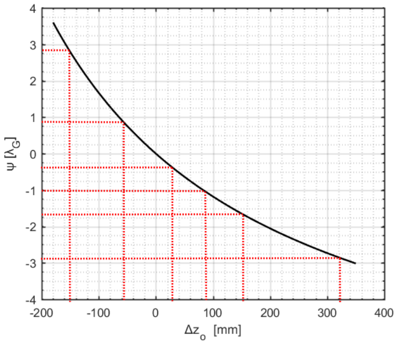

Figure 7: Defocus parameter ψ as a function of the object defocus distance Δz_o when the used lens is nominally focused at 58 cm.

In the experimental setup, we have chosen to place objects at 42.7 cm, 52.2 cm, 60.4 cm, 66.4 cm, 73.7 cm and 90.4 cm of the lens, which correspond (according to Fig. 7) to defocus parameter values ψ equal to $2.87\lambda_G = 2\lambda_R$, $0.90\lambda_G$, $-0.32\lambda_G$, $-1\lambda_G$, $-1.7\lambda_G$ and $-2.87\lambda_G$ respectively. These defocus parameter values are represented with dotted lines on Fig. 6 and Fig. 7. The considered objects are small pieces of paper with a spoke target and "Lorem Ipsum" sentences consisting of random Latin words, requiring a real reading of the constituting characters to determine whether the sentence is readable or not. A picture of the setup is given on Fig. 8(a).

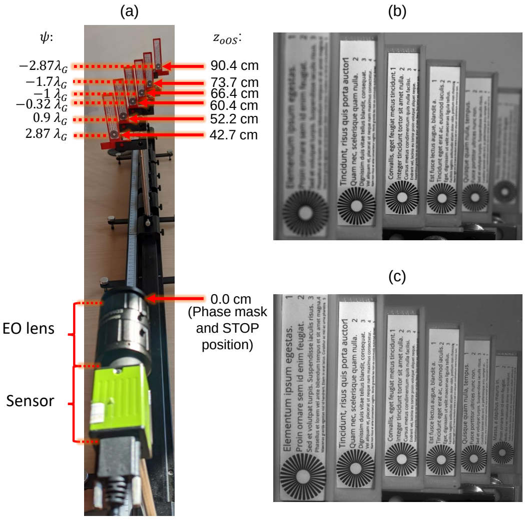

Figure 8: (a) Picture of the experimental setup and distances of interest. (b) Image of the scene without phase mask. (c) Deconvolved image of the scene with the $2\lambda_G$-wide mask.

Figure 8(b) is the image of the scene without phase mask nor deconvolution. The DoF is clearly limited, and only the third plane, placed at 60.4 cm of the lens, and corresponding to $\psi = -0.32\lambda_G$ can be considered sharp (because it is close to the focus plane, at 58 cm). The first plane and the last plane appear very blurry. Figure 8(c) is the image of the scene with the $2\lambda_G$-wide mask under wide illumination spectrum, with deconvolution. Since this phase mask has been optimized for a targeted DoF range of $[-2\lambda_G, 2\lambda_G]$ (Fig. 2), the second to fifth planes, which belong to this range, appear sharp and well contrasted. The first and last planes, which do not belong to the targeted DoF range, also have their image quality significantly improved. This image allows us to validate the DoF extension provided by the $2\lambda_G$-wide mask under wide spectrum illumination. To verify that it is the best mask among those proposed in this article, we now look at some details of the image.

4.2 Details of the enhanced performance

Figure 9 represents a word from each plane of the scene for various defocus parameter values and various imaging configurations. In the configuration without phase mask (Figure 9, column (1)), only the Lorem Ipsum word "Endrerit" is clearly readable. This is because, without phase mask and with a λ_G-monochromatic illumination spectrum, the DoF range is reduced to the Rayleigh distance on either side of the focal plane, that is, $[-\lambda_G/4, \lambda_G/4]$. If we consider a larger depth of field, the words are more and more blurred, so that we cannot even distinguish the word "ornare" (Fig. 9(F)) obtained for $\psi = -2.87\lambda_G$.

The columns (2), (3) and (4) correspond to three different "mask / illumination" configurations for which the targeted DoF parameter range was $[-2\lambda_G, 2\lambda_G]$. It is observed that the $2\lambda_G$-wide mask with wide spectrum illumination (Fig. 9, column (3)) gives an image quality quite homogeneous over the DoF range (row (B) to (E)), which well agrees with theoretical prediction. The quality is better (greater IQ) than the $2\lambda_G$-mono mask / monochromatic spectrum

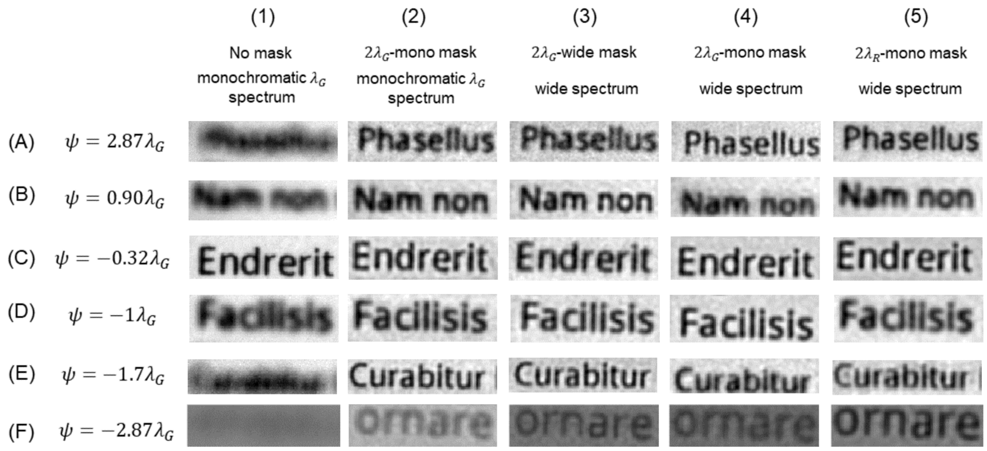

Figure 9: Details of the scene for different defocus parameter values and different configurations.

configuration (Fig.9, column (2)), even though the picture seems slightly blurrier to the human vision (compared to an, in fact, over-contrasted picture). In particular, row (B), the text "Nam non" is sharper with the $2\lambda_G$-wide mask (Fig. 9, column (3)) than with the $2\lambda_G$-mono mask (Fig. 9, column (2)) under their respective nominal illuminations. This is coherent with the better IQ obtained on Fig. 6(a) for $\psi = 0.90\lambda_G$. A similar difference in IQ (of about 0.7 dB) is observable on Fig. 6(a) at $\psi = -0.32\lambda_G$, but its origin is different: when the illumination is monochromatic, the $2\lambda_G$-mono mask presents over-contrast at the focal plane (or next to it, Fig. 9(2C)). This over-contrast explains the lower IQ for the $2\lambda_G$-mono mask. We therefore can see from these details that the $2\lambda_G$-wide mask performs better (in terms of IQ, i.e. in other words, of pixel fidelity to the ideal image of the scene) than the $2\lambda_G$-mono mask under their respective nominal illumination conditions, which translates into more homogeneous visual image quality over the targeted DoF range.

We can also study the behavior of phase masks optimized under monochromatic conditions and used with a wide spectrum illumination. Figure 9, column (4), the $2\lambda_G$-mono mask is robust to its use with a wide spectrum illumination, since we can still read the words. There is less over-contrast on Fig. 9(4C) than on Fig. 9(2C) thanks to spectral averaging, but on the other hand, the image is more blurred for some defocus parameters (Figure 9(4B) and (4E)). It is the $2\lambda_R$-mono mask (Figure 9, column (5)) which has the worst performance, as predicted by the simulations (Fig. 6(b)). In particular, although the word "Phasellus" is readable under wide spectrum illumination (Figure 9(5A)), it is much less sharp than it would be under monochromatic red spectrum illumination (according to the theoretical prediction of Fig. 6). Here, the sharpness is similar to that obtained with the $2\lambda_G$-wide mask, although this mask was optimized on a smaller targeted DoF range not including $2\lambda_R$. This could be have been expected from the IQ study (the performance of these two masks is equal on Fig. 6(b), at $\psi = 2.87\lambda_G$). The $2\lambda_R$-mono mask is therefore not robust to its use with the studied wide spectrum illumination.

To conclude, this experimental study validates the conclusions obtained by simulations in Sec. 3. It shows that the robustness of a phase mask optimized under monochromatic spectrum illumination assumption depends on the nominal monochromatic wavelength and its difference with the considered wide spectrum illumination. Above all, it shows the benefit of optimizing binary phase masks by taking into account the spectrum of use, to obtain a better imaging quality.

5 CONCLUSION

We have studied optimization and performance of DoF enhancing phase masks in the presence of monochromatic or wide spectrum illuminations. Through a theoretical study, we have shown that the performance of a phase mask optimized under wide spectrum assumption and used under this type of illumination is higher in terms of IQ than that of a phase mask optimized and used under monochromatic spectrum illumination. This interesting effect is partly due to the fact that deconvolution artifacts are reduced by wide spectrum illumination thanks to spectral averaging. Experiments have validated this theoretical result. Moreover, we have shown that the robustness of a phase mask optimized under monochromatic assumption to wide spectrum illumination highly depends on the nominal wavelength of optimization. These two results are important since they show that it is possible and useful to take into account the spectral width of illumination together with the scene characteristics and the targeted DoF range to efficiently co-design DoF enhancing imaging systems.

This work has many perspectives. The main one is to generalize the present approach to DoF extension of RGB color imaging systems. The general performance of binary phase masks with RGB sensors has already been proven experimentally [31, 17], but the limitations of such an approach remain to be studied in detail. Another important perspective is to consider other optimization criteria than the MSE, such as SSIM (Structural SIMilarity), in order to get even better results in terms of visual quality of the final image.

Acknowledgments The work reported in this study is supported in part by the Agence de l'Innovation de Défense (AID) that provides half of a PhD fellowship to Alice Fontbonne. We thank Eloi Lemaire for the setting up of the optical bench. **Disclosures** The authors declare no conflicts of interest.

Data Availability Statement Data is available upon request.

References

[1] Po-Sheng Chiu, Kurt Vonmetz, Federico Canini, and H. Paul Urbach. Investigation of extended depth-of-field f/8 camera with optimized cubic phase mask and digital restoration. *EPJ Web of Conferences*, 238:03008, 2020.

[2] Mads Demenikov and Andrew R. Harvey. Image artifacts in hybrid imaging systems with a cubic phase mask. *Opt. Express*, 18(8):8207–8212, Apr 2010.

[3] Frédéric Diaz, François Goudail, Brigitte Loiseaux, and Jean-Pierre Huignard. Increase in depth of field taking into account deconvolution by optimization of pupil mask. *Opt. Lett.*, 34(19):2970–2972, Oct 2009.

[4] Frédéric Diaz, François Goudail, Brigitte Loiseaux, and Jean-Pierre Huignard. Comparison of depth-of-focus-enhancing pupil masks based on a signal-to-noise-ratio criterion after deconvolution. *J. Opt. Soc. Am. A*, 27(10):2123–2131, Oct 2010.

[5] Frédéric Diaz, François Goudail, Brigitte Loiseaux, and Jean-Pierre Huignard. Increase in depth of field taking into account deconvolution by optimization of pupil mask. *Optics letters*, 34(19):2970–2972, 2009.

[6] Edward R. Dowski and W. Thomas Cathey. Extended depth of field through wave-front coding. *Appl. Opt.*, 34(11):1859–1866, Apr 1995.

[7] Shay Elmalem, Raja Giryes, and Emanuel Marom. Learned phase coded aperture for the benefit of depth of field extension. *Opt. Express*, 26(12):15316–15331, Jun 2018.

[8] Shay Elmalem, Naim Konforti, and Emanuel Marom. Polychromatic imaging with extended depth of field using phase masks exhibiting constant phase over broad wavelength band. *Applied optics*, 52(36):8634–8643, 2013.

[9] Rafael Falcón, François Goudail, Caroline Kulcsár, and Hervé Sauer. Performance limits of binary annular phase masks codesigned for depth-of-field extension. *Optical Engineering*, 56(6):065104, 2017.

[10] Alice Fontbonne, Hervé Sauer, and François Goudail. Theoretical and experimental analysis of co-designed binary phase masks for enhancing the depth of field of panchromatic cameras. *Optical Engineering*, 60(03), mar 2021.

[11] Alice Fontbonne, Hervé Sauer, Caroline Kulcsár, Anne-Lise Coutrot, and François Goudail. Experimental validation of hybrid optical–digital imaging system for extended depth-of-field based on co-optimized binary phase masks. *Optical Engineering*, 58(11):1, nov 2019.

[12] Alice Fontbonne, Hervé Sauer, Caroline Kulcsár, Anne-Lise Coutrot, and François Goudail. What is the depth of field reachable in practice with generic binary phase masks and digital deconvolution? In Corinne Fournier, Marc P. Georges, and Gabriel Popescu, editors, *Unconventional Optical Imaging II*. SPIE, mar 2020.

[13] A. Halé, P. Trouvé-Peloux, and J.-B. Volatier. End-to-end sensor and neural network design using differential ray tracing. *Optics Express*, 29(21):34748, 2021.

[14] Vladimir Katkovnik, Mykola Ponomarenko, and Karen Egiazarian. Lensless broadband diffractive imaging with improved depth of focus: wavefront modulation by multilevel phase masks. *Journal of Modern Optics*, 66(3):335–352, 2019.

[15] L. Ledesma-Carrillo, C. M. Gómez-Sarabia, M. Torres-Cisneros, R. Guzmán-Cabrera, C. Guzmán-Cano, and J. Ojeda-Casta neda. Hadamard circular masks: high focal depth with high throughput. *Opt. Express*, 25(15):17004–17020, Jul 2017.

[16] Yuankun Liu, Chongyang Zhang, Tingdong Kou, Yueyang Li, and Junfei Shen. End-to-end computational optics with a singlet lens for large depth-of-field imaging. *Optics Express*, 29(18):28530, 2021.

[17] Benjamin Milgrom, Roy Avrahamy, Tal David, Avi Caspi, Yosef Golovachev, and Shlomo Engelberg. Extended depth-of-field imaging employing integrated binary phase pupil mask and principal component analysis image fusion. *Opt. Express*, 28(16):23862–23873, Aug 2020.

[18] Benjamin Milgrom, Naim Konforti, Michael A. Golub, and Emanuel Marom. Novel approach for extending the depth of field of barcode decoders by using rgb channels of information. *Opt. Express*, 18(16):17027–17039, Aug 2010.

[19] Benjamin Milgrom, Naim Konforti, Michael A Golub, and Emanuel Marom. Pupil coding masks for imaging polychromatic scenes with high resolution and extended depth of field. *Optics express*, 18(15):15569–15584, 2010.

[20] Minhnghia Pham, Vannhu Le, Huykien Le, and Synam Truong. Invariant-ability of the PSF of wide-field microscopy to the DOF by using cubic phase mask. In *2020 4th International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)*. IEEE, 2020.

[21] Daniel L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37(23):3385 – 3398, 1997.

[22] Suho Ryu and Chulmin Joo. Design of binary phase filters for depth-of-focus extension via binarization of axisymmetric aberrations. *Opt. Express*, 25(24):30312–30326, Nov 2017.

[23] Y. Shi and R. Eberhart. A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*. IEEE, 1998.

[24] Vincent Sitzmann, Steven Diamond, Yifan Peng, Xiong Dun, Stephen Boyd, Wolfgang Heidrich, Felix Heide, and Gordon Wetzstein. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Transactions on Graphics*, 37(4):1–13, 2018.

[25] David G. Stork and M. Dirk Robinson. Theoretical foundations for joint digital-optical analysis of electro-optical imaging systems. *Appl. Opt.*, 47(10):B64–B75, Apr 2008.

[26] Qilin Sun, Congli Wang, Qiang Fu, Xiong Dun, and Wolfgang Heidrich. End-to-end complex lens design with differentiable ray tracing. *ACM Trans. Graph.*, 40(4), 2021.

[27] Ethan Tseng, Ali Mosleh, Fahim Mannan, Karl St-Arnaud, Avinash Sharma, Yifan Peng, Alexander Braun, Derek Nowrouzezahrai, Jean-François Lalonde, and Felix Heide. Differentiable compound optics and processing pipeline optimization for end-to-end camera design. *ACM Transactions on Graphics*, 40(2):1–19, 2021.

[28] A. van der Schaaf and J.H. van Hateren. Modelling the power spectra of natural images: Statistics and information. *Vision Research*, 36(17):2759 – 2770, 1996.

[29] Zhou Wang and Alan C. Bovik. Mean squared error: Love it or leave it? a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, 2009.

[30] Jiaqian Yu, Shouqian Chen, Fanyang Dang, Xueshen Li, Xiaotian Shi, Hui Wang, and Zhigang Fan. The suppression of aero-optical aberration of conformal dome by wavefront coding. *Optics Communications*, 490:126876, 2021.

[31] Zeev Zalevsky, Amir Shemer, Alexander Zlotnik, Eyal Ben Eliezer, and Emanuel Marom. All-optical axial super resolving imaging using a low-frequency binary-phase mask. *Opt. Express*, 14(7):2631–2643, Apr 2006.

[32] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. A comprehensive evaluation of full reference image quality assessment algorithms. In *2012 19th IEEE International Conference on Image Processing*, pages 1477–1480, 2012.

[33] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57, mar 2017.