Performance limits of binary annular phase masks codesigned for depth-of-field extension

Rafael Falcón, François Goudail, Caroline Kulcsár, Hervé Sauer

Submitted to *Optical Engineering* .
(Published in June 2017. http://dx.doi.org/10.1117/1.OE.56.6.065104)

Performance limits of binary annular phase masks co-designed for depth of field extension

Rafael Falcón[a,*], **François Goudail**[a], **Caroline Kulcsár**[a], **Hervé Sauer**[a]
[a]Laboratoire Charles Fabry, Institut d'Optique Graduate School, CNRS, Université Paris-Saclay, 91127 Palaiseau cedex, France

Abstract. We investigate the performance limits of hybrid imaging systems including annular binary phase masks optimized for depth-of-field (DoF) extension, as a function of the number of rings and the desired DoF range. The mask parameters are optimized taking into account deconvolution of the acquired raw image in the expression of the global performance of the imaging system. We prove that masks with a limited number of rings are sufficient to obtain near-optimal performance. Moreover, the best achievable image quality decreases as the required DoF range increases, so that for a given required image quality, the DoF extension reachable with binary phase masks is bounded. Finally, these conclusions are shown to be robust against different optical system aberrations and models of scene power spectral density. These results are important in practice to decide if annular binary phase masks are the relevant solution for a given imaging problem, and for mask manufacturing, since the number and thickness of the rings reachable at affordable cost by technology is generally limited.

Keywords: Image processing, binary phase masks, phase-only filters, optical system co-design, wavefront encoding..

***Rafael Falcón, rafael.falcon@institutoptique.fr

1 Introduction

Binary phase masks are wavefront encoding devices typically situated at the aperture stop of an optical system to engineer its point spread function (PSF). They can be used to extend the depth of field (DoF) of imaging systems without reducing the light throughput by producing a PSF that becomes more invariant to defocus.[1–3] However, the larger the DoF the more blurred the acquired raw image so that deconvolution has to be applied. The design of these masks has thus to take into account image processing in order to reach the optimal compromise between invariance of PSF to defocus and capacity to deconvolve the image. This joint design approach has been introduced by Cathey and Dowski[4] for continuous phase DoF enhancing masks, and generalized by Robinson and Stork[5] to other optical design tasks. It has been applied to the optimization of DoF enhancing phase masks

in imaging applications[1,6] and has been extended to include broad spectral ranges during the mask selection.[7,8]

DoF extended imaging systems including optimized binary annular phase masks and real-time deconvolution have been implemented successfully. In Ref.,[2] the combination of an optimized 2-ring binary phase mask with a digital deconvolution algorithm implemented on GPU made it possible to increase by a factor 3 the depth of field of a commercial LWIR camera with f-number equal to 1.3 and 25^o field of view (FoV). In Ref.,[9] the DoF of a high definition camera (1292x972 pixels), with f-number 1.2 and $16^o \times 12^o$ FoV operating in the visible spectral range was increased by a factor 2.5 using an optimized 6-ring mask and real-time FPGA-implemented deconvolution.

However, when designing an annular binary phase mask, there remain fundamental questions: what is the maximum depth of field reachable with such a mask? How many rings are necessary? Are the masks robust to aberrations or spatial frequency content of the scene? The purpose of this article is to address these issues. We show that more rings are necessary to reach higher DoF, but augmenting the number of rings above a given amount does not improve significantly the performance any more. Therefore nearly optimal performance is reached with a limited number of rings. Then, we also show that the best possible final image quality decreases as the targeted DoF range increases, showing that the reachable DoF extension for a given image quality is bounded. Finally, we show that binary annular phase mask imaging performance is robust to changes in the frequency content of the objects and to small amounts of classical aberrations (spherical, coma, astigmatism). These results, quantitatively described below, are important in practice to decide if annular binary phase masks are the relevant solution for a given imaging problem.

This paper is organized as follows: in section 2 we define a rigorous model of interaction between optical design and digital deconvolution as introduced in.[6] Then in section 3 we show that optimization of binary masks for a given DoF has many local optima corresponding to masks having different shapes but very similar performance. However, the number of required rings is limited and increases slowly with the desired DoF, so that near-optimal performance can be obtained with masks with few rings, practical examples and manufacturing considerations end this section. In section 4 we perform a more detailed study of the results presented in section 3 by analyzing the performances of each phase mask over the DoF. In section 5, we demonstrate the robustness of the optimized masks to the presence of third order spherical, coma and astigmatism aberrations and to a mismatch between the power spectral density of the captured object and that of the prior scene model used in the deconvolution algorithm.

2 Principle of the co-design

Binary annular phase masks are static spatial phase modulating optical elements consisting of a series of N concentric annular regions of phase modulation of alternatively 0 or π radians at the nominal wavelength λ (usually chosen in the middle of the working spectral range). Each annular constant phase area corresponds to a so-called ring, so that an N-ring mask of clear aperture radius R is parametrized by $N-1$ normalized radius values $\varphi = \{\rho_1, \dots \rho_{N-1}\}$ where the radius of the ith phase transition is $r_i = \rho_i R$ and satisfies the conditions $0 < \rho_1 < \cdots < \rho_{N-1} < \rho_N = 1$ (see Fig.1).

The hybrid imaging system that we consider in this paper consists of a lens and a binary phase mask placed at the aperture stop, and of a digital deconvolution algorithm that post-

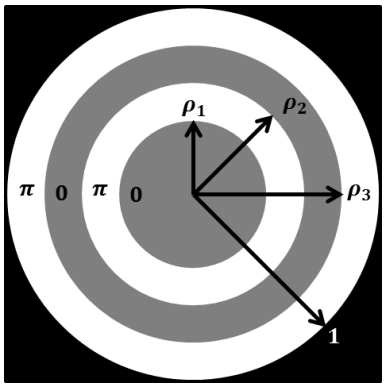

Fig 1 Example of a 4-ring annular binary phase mask with $\rho_4 = 1$. A ring is defined as an annular region with constant phase modulation. Gray areas induce a phase of $\varphi = 0$ and white areas a phase of $\varphi = \pi$ radians.

processes the raw image. For this theoretical study of binary mask DoF extension limit, a nearly monochromatic spectral range near the nominal wavelength λ is considered. We model the optical configuration as a phase mask placed in the stop or the exit pupil of a perfect optical system. This simple model has been shown to give valuable results in the context of the experimental setups described in Refs.[2] and.[9] In Section 5, to evaluate the robustness of this approach, we additionally introduce small amounts of third order aberrations in the optical system.

The final image $\hat{O}(r)$ produced by such a hybrid imaging system is modeled as:

$$\hat{O}(r) = d(r) * \left[h_\psi^\varphi(r) * O(r) + n(r)\right] \tag{1}$$

where $O(r) = I(r) - \bar{I}(r)$ is the ideal sharp image minus its mean, $*$ denotes the convolution operator, $h_\psi^\varphi(r)$ is the PSF of the optical system, $n(r)$ is the detection noise, $d(r)$ is the digital deconvolution filter, and r represent the spatial coordinates. We assume that the aperture stop of the lens is circular of radius R. The PSF $h_\psi^\varphi(r)$ is a function of the phase mask parameters φ and of the defocus factor ψ defined as an optical path difference (OPD):

$$\psi = \frac{R^2}{2}\left(\frac{1}{z_i} + \frac{1}{z_o} - \frac{1}{f}\right) \tag{2}$$

where z_o is the object distance (taken as positive here), z_i the image distance and f the effective focal length; please note that this equation can account for depth-of-field (z_i fixed and z_o variable) or depth-of-focus (z_i variable and z_o fixed) for the same value of ψ. For the sake of conciseness, all the mathematical developments are expressed using a 1D continuous formalism to represent the finite discrete 2D one, e.g., using the integral $\int \cdots dr$ to represent a finite double sum $\sum_{i=1}^{L}\sum_{j=i}^{M}\cdots$ over the pixels of the image. In order to evaluate the quality of the image provided by the hybrid imaging system for a given value of ψ and φ, we use the mean squared error (MSE) between the deconvolved image and the ideal one:[5,6]

$$MSE(\varphi,\psi) = \mathrm{E}\left[\int \left|\hat{O}(r) - O(r)\right|^2 dr\right] \tag{3}$$

where E[·] represents the mathematical expectation over the noise $n(r)$ and the sharp image $O(r)$, which are both assumed to be zero-mean, stationary random processes of power spectral density (PSD) $S_{nn}(\nu)$ and $S_{oo}(\nu)$ respectively, ν representing the spatial frequency coordinates. Let us emphasize at this point that this criterion is able to simultaneously take into account the residual blur of the final image and the noise amplification introduced by the deconvolution stage, which should both be minimized in a well-balanced way in order to get a good final image quality. Standard criteria used in conventional optical design, such as the diameter of the spot-diagrams or the diffraction MTF, are not sufficient for joint-optimization since they do not take image noise into account.

From Eqs. (3) and (1), it has been shown that, in the context of hybrid imaging,[6] the MSE has the following expression:

$$MSE(\varphi,\psi)=\int\left|\widetilde{d}(\nu)\widetilde{h}_{\psi}^{\varphi}(\nu)-1\right|^{2}S_{oo}(\nu)d\nu+\int\left|\widetilde{d}(\nu)\right|^{2}S_{nn}(\nu)d\nu \tag{4}$$

where the superscript $\sim$ denotes the Fourier transform. Our co-design goal is to find the deconvolution filter and the mask parameters that minimize $MSE(\varphi,\psi)$ over a DoF range defined by a set of K values of $\psi\in[\psi_1,\psi_2,\cdots,\psi_K]$. In the following, the maximal value of the DoF range is denoted by ψ_{max} with $|\psi_k|\leq|\psi_K|=|\psi_{max}|$. For deconvolution, we use the averaged Wiener filter that is known to minimize $\sum_k MSE(\varphi,\psi_k)$ and has the following expression:[6]

$$\widetilde{d}(\nu)=\frac{\frac{1}{K}\sum_{k=1}^{K}(\widetilde{h}_{\psi_k}^{\varphi}(\nu))^{\star}}{\frac{1}{K}\sum_{k=1}^{K}\left|\widetilde{h}_{\psi_k}^{\varphi}(\nu)\right|^{2}+\frac{S_{nn}(\nu)}{S_{oo}(\nu)}} \tag{5}$$

where $\cdot^{\star}$ stands for complex conjugate. The value of the MSE obtained for this choice of $\widetilde{d}(\nu)$ and for a given defocus ψ is denoted by $\overline{MSE}(\varphi,\psi)$, which leads to the definition of the image quality:

$$IQ(\varphi,\psi)=10\log_{10}\left(\frac{E_O}{\overline{MSE}(\varphi,\psi)}\right) \tag{6}$$

expressed in decibels (dB) and where E_O is the energy of the scene defined as $\mathrm{E}\left[\int O(r)^2dr\right]$.

In order to determine the parameters of the optimal mask, φ_{opt}, we use a "maximin" criterion. Let us define $J(\varphi)$ as:

$$J(\varphi)=\min_{k}\left[IQ(\varphi,\psi_k)\right]. \tag{7}$$

The optimal mask profile is chosen to be:

$$\varphi_{opt} = \arg\max_{\varphi} \left[J(\varphi) \right]. \tag{8}$$

Our performance criterion for the hybrid system is finally defined as:

$$IQ_{maximin} = J(\varphi_{opt}). \tag{9}$$

Here, φ_{opt} depends only on the chosen DoF range ψ_{max} and on the number of rings N. Our purpose is to investigate the evolution of $IQ_{maximin}$ as a function of ψ_{max} and N, in order to determine if there is an optimal number of rings a binary phase mask requires in order to address a given DoF range.

In sections 3 and 4, we consider that the conventional lens part is aberration-free (we only use the first order parabolic "defocus" Seidel phase term $\psi\rho^2$), whereas 3rd order aberrations are considered in Section 5. We assume throughout this paper that $n(r)$ is a white noise and that the signal-to-noise ratio (SNR) on the raw image is 34 dB with the SNR precisely defined as:

$$SNR = 10 \log_{10} \left[\frac{\int S_{oo}(\nu) d\nu}{\int S_{nn}.(\nu) d\nu} \right] \tag{10}$$

We use a generic ideal image model with a power-law PSD[10] $S_{oo}(\nu) = K\nu^{-a}$, with a being fixed to 2.5 in sections 3 and 4 and slightly varied in section 5, and the constant K being given by the SNR value assuming $\int S_{nn}(\nu) d\nu = 1$.

3 Phase mask optimization

3.1 Optimization methodology

In this section we present the optimization of the ring radii φ for a given DoF range ψ_{max} and a given number of rings N. Since binary annular masks introduce a phase modulation of 0 or π radians, with $-\pi = \pi$ modulo 2π, and as the optical system is aberration-free, the value of $IQ(\varphi, \psi)$ is symmetric around the focal point $\psi = 0$. Thus, to save computational time, our optical system is optimized only for positive values of $\psi \in \{\psi_1 = 0, \cdots, \psi_K = \psi_{max}\}$ with the underlying assumption that the total DoF range is $[-\psi_{max}, +\psi_{max}]$ around the focal point. We consider the following values of DoF range $\psi_{max} \in \{1.0,\ 1.5,\ 2.0,\ 2.5,\ 3.0\}\lambda$; for comparison, a classical diffraction limited optical system is generally considered to have a maximal wavefront error due to defocus of $\psi_{max} = \pm 0.25\lambda$, according to the so called Rayleigh's quarter wavelength rule,[11,12] The evaluation points of the DoF range are chosen such that they are evenly spaced and that $\psi_k - \psi_{k-1} = 0.5\lambda$, this step size being sufficient to obtain the optimal mask profile for a given ψ_{max}.

The optimization of φ is not simple since the function $J(\varphi)$ (see Eq.(7)) presents several local maxima and is highly nonconvex. This is clearly seen in Fig. 2.a, were we have represented the value of $J(\varphi)$ for a 2-ring mask as a function of the radius ρ_1 and for different values of ψ_{max}. It can be noticed that depending on the value of ψ_{max}, there may be several local maxima of $J(\varphi)$. Moreover, these local maxima have very similar values for a given value of ψ_{max} in the range $\psi_{max} \in \{2.0, 2.5, 3.0\}\lambda$. Similarly, we have represented in Fig. 2.b the variation of $J(\varphi)$ for a three-ring mask as a function of ρ_1 and ρ_2 for $\psi_{max} = 2.0\lambda$; the most prominent local maxima are marked with black crosses, many of them presenting

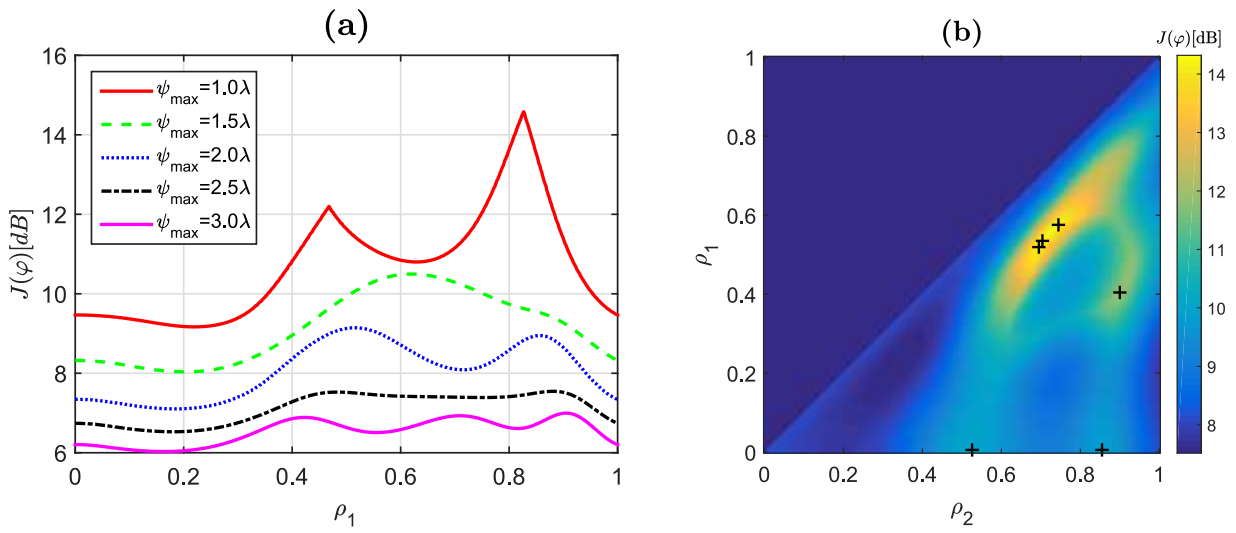

Fig 2 Evolution of $J(\varphi)$ for (a) 2-ring masks calculated for $\psi_{max} \in [1.0,\ 1.5,\ 2.0,\ 2.5,\ 3.0]\lambda$ and (b) 3-ring masks for $\psi_{max} = 2.0\lambda$; the crosses mark the most prominent local maxima.

similar values of $J(\varphi)$.

To optimize binary phase masks with a larger number of rings, a graphical representation of $J(\varphi)$ is no longer possible and therefore a global optimization algorithm is needed. We have chosen the population-based Particle Swarm algorithm,[13] which consists of a collection of individual particles randomly placed on the solution landscape and moving through it with a speed based on their own position, the position of the best minimum found at that iteration and a random perturbation. After convergence of the iterative process, a gradient-based local optimization algorithm is run to reach the nearest local maximum.

3.2 Optimization results

The randomness of the algorithm added to the complexity of the optimization problem makes it not always possible to find the global maximum in a single optimization run. For this reason, a large number of optimization runs, with randomly generated starting points, were calculated in parallel on a 48 core cluster of computers, leading to up to 40 different local maxima found, depending of the complexity of the optimization landscape. After multi-hour

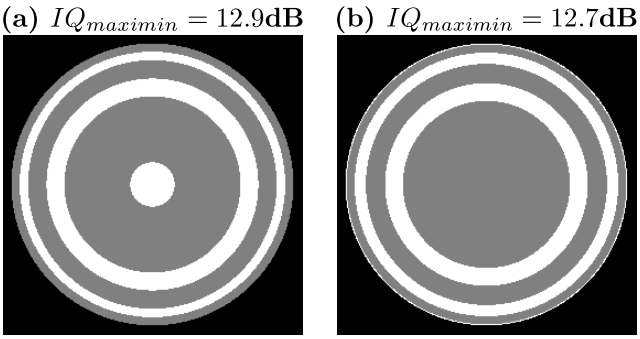

Fig 3 Comparison of two equivalent local maxima of $J(\varphi)$ obtained by optimization for 6-ring masks and for $\psi_{max} = 2.5\lambda$.

computations with computation time increasing with N, the best obtained masks for each case were then selected as optimal.

We have noticed that the number of local maxima of the optimization landscape increases along with the number of rings. It also starts to increase with the required DoF range, it reaches its maximum value for ψ_{max} around 2.0λ and then slightly decreases when ψ_{max} reaches 2.5λ as the optimization landscape becomes smoother. Similarly to the cases shown in Fig. 2, different mask shapes with very similar $J(\varphi)$ values were found. As an example, we show in Fig. 3 two of the best masks obtained for 6 rings and $\psi_{max} = 2.5\lambda$. These masks have different shapes but very similar image quality. This being said, we can observe that the last ring of Fig.3.b is very thin so that from a technological point of view the mask in Fig. 3.a would, as discussed in section 3.3, certainly be easier to manufacture.

Some global pattern can be observed on the shapes of the optimal masks as ψ_{max} and the number of rings increase. In Fig. 4 we show the profiles of the optimal masks obtained for DoF ranges $\psi_{max} = 2.0\lambda$ and 2.5λ, with 5, 6 and 7 rings each. A similar feature on all the masks is a wide annular region between two rings at $\rho_{k-1} \cong 0.60$ and $\rho_k \cong 0.75$. We can see that, for a given DoF, when the number of optimization parameters (rings) increases, the masks have similar profiles with the new rings appearing after the aforementioned annular

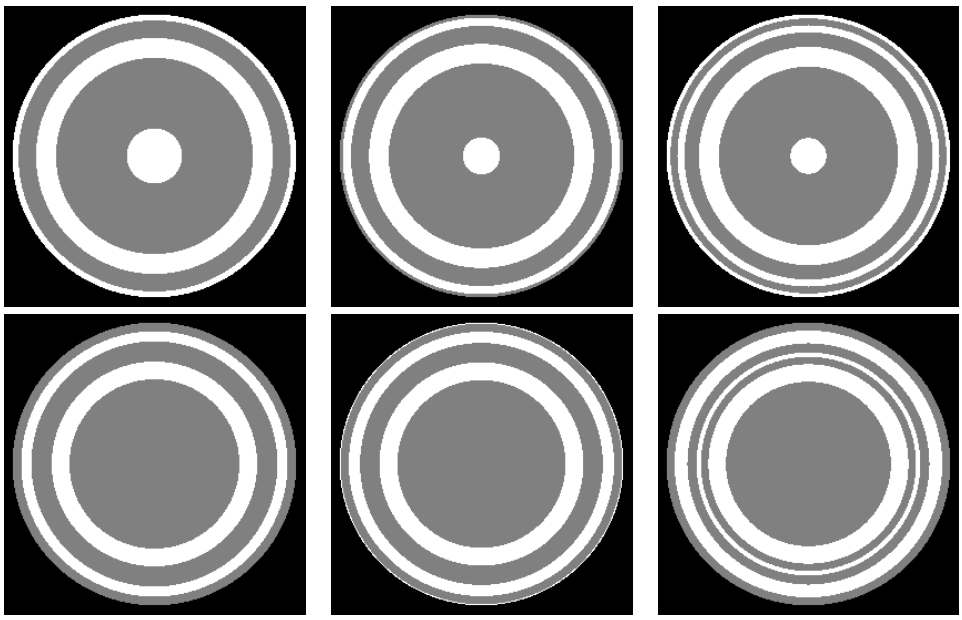

Fig 4 Masks optimized with (from left to right) 5, 6 and 7 rings and for DoF ranges $\psi_{max} = 2.0\lambda$ (top row) and $\psi_{max} = 2.5\lambda$ (bottom row).

region. This can probably be related to the parabolic shape of the "defocus" wavefront error, with the central part being relatively flat and thus not requiring to be engineered to enhance the image quality.

As an illustration of how the masks modify the impulse response of the optical system, Fig. 5.a shows the PSF of an aberration-free diffraction-limited system with a circular aperture at different focal points. Fig. 5.b shows the PSF of the same system with a binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$. It can be seen that the inclusion of the phase mask largely reduces the spread of the PSF for $\psi = 1\lambda$ and $\psi = 2\lambda$, and leads to quasi-similar PSFs over the ψ range. Even though the PSFs are not strictly invariant throughout the DoF range (as could be anticipated from the maximin optimization criterion in Eq. (9)), they induce a similar blur in the image. The deconvolution process is thus able to successfully restore the image quality in a similar way independently of the object distance, as will be seen in the following sections.

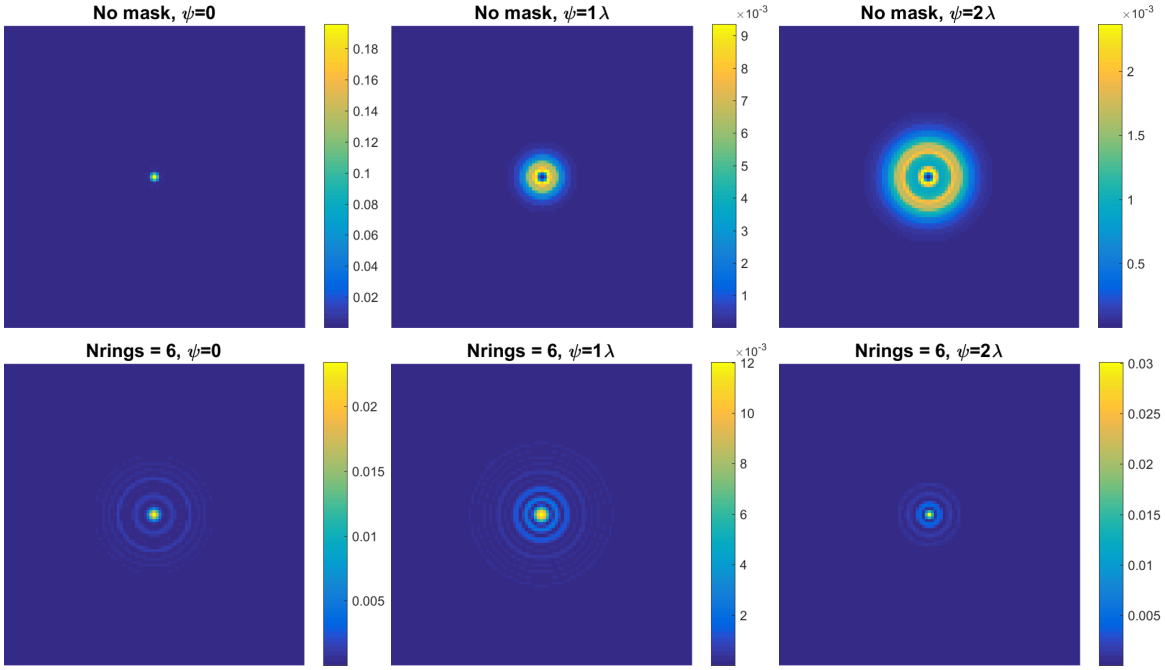

Fig 5 Point spread function at DoF (left to right) $\psi = 0$, $\psi = 1.0\lambda$ and $\psi = 2.0\lambda$ for: (top row) an aberration-free, diffraction-limited optical system and (bottom row) the same system with a 6-ring binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$.

3.3 Mask performance evaluation

Let us now evaluate the changes in the performance obtained by using an increasing amount of rings at a given DoF range. Fig. 6 represents the value of $IQ_{maximin}$ as a function of the number of rings for different values of ψ_{max}. We observe that the image quality naturally increases as more rings are added, but rapidly levels off in all cases. When $\psi_{max} = 1.0\lambda$, the increase on IQ from 2 to 3 rings is of $[\Delta IQ]_{2\to3} = 0.92$dB, a considerable gain in this IQ range, but adding extra rings does not increase significantly the image quality. When $\psi_{max} = 1.5\lambda$, the last significant gain of quality happens when passing from 3 to 4 rings, but is only of $[\Delta IQ]_{3\to4} = 0.55$dB. The increase in performance between 3 and 4 rings becomes larger when the DoF increases, $[\Delta IQ]_{3\to4} = 1.4$dB for $\psi_{max} = 2.0\lambda$, $[\Delta IQ]_{3\to4} = 1.8$dB for $\psi_{max} = 2.5\lambda$ and $[\Delta IQ]_{3\to4} = 3.0$dB for $\psi_{max} = 3.0\lambda$. This trend clearly shows a need for extra rings as ψ_{max} increases.

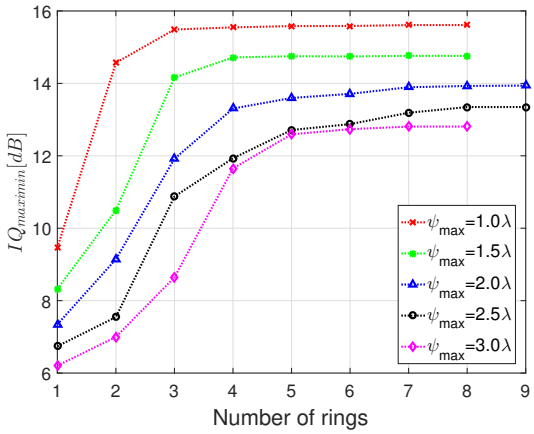

Fig 6 Performance comparison of optimal annular phase masks with different amount of rings for $\psi_{max} \in [1.0,\ 1.5,\ 2.0,\ 2.5,\ 3.0]\lambda$. The case where the number of rings is equal to 1 is equivalent to having no mask.

However, for all values of ψ_{max}, a saturation of the performance with the number of rings is observed. When $\psi_{max} = 1.0\lambda$, it occurs after 3 rings and for $\psi_{max} = 1.5\lambda$, after 4 rings. When $\psi_{max} = 2.0\lambda$, the behavior is slightly different, since the growth of $IQ_{maximin}$ slows down gradually, until it stops at 7 rings. The case of $\psi_{max} = 2.5\lambda$ is somehow more intricate, but, again, shows no practical IQ improvement after 7 rings.

This behavior seems to be due to the optimization landscape being the most nonconvex and prone to local maxima at values of ψ_{max} around 2.0λ. On the other hand, when $\psi_{max} = 3.0\ \lambda$, the IQ stops increasing as soon as with 5 rings, with a rather low IQ maximal value. This shows that we have reached a limit on how much the DoF of a system can be extended with an annular binary phase mask and a mean Wiener filter, regardless of how complex that mask is.

3.4 Typical examples and manufacturing considerations

It is interesting to note that the optimal masks obtained in the previous paragraphs depend only on the ψ_{max} and N (number of rings) parameters. They can thus be used to extend

the DoF of many different optical systems of different dimensions. For example, let us consider an optical system with effective focal length $f = 20$mm and F-number $F\# = 2$. If the spectral operating range is in the visible ($\lambda = 587$nm), a mask optimized for $\psi_{max} = 2.5\lambda$, as represented in Fig. 3, provides a depth-of-focus (DoFs) for an object at infinity, of $\Delta z_i = \pm 2(\psi_{max}/\lambda)\lambda f^2/R^2 = \pm 8(\psi_{max}/\lambda)\lambda(F\#)^2$, that is ≈ 47µm. This value should be compared to the DoFs of a diffraction limited standard lens of the same F-number ($\Delta z_{standard} = \pm 2\lambda(F\#)^2$) of ≈ 4.7µm, showing, as expected, a tenfold increase of the DoFs, since the DoFs of a diffraction limited standard lens corresponds to $\psi_{max} = 0.25\lambda$ according to the Rayleigh's quarter wavelength rule. The same generic masks of Fig. 3 could also be used to extend the DoF of a system with longer effective focal length $f = 200$mm and $F\# = 4$. In this case, the DoFs would be ≈ 190µm while the DoFs of a standard lens of the same F-number is ≈ 19µm.

The question is now how to manufacture these masks. Manufacturing of real masks will strongly depend on the user's specific application, e.g., the spectral range and the usable materials, the pupil size of the targeted system and the cost of the manufacturing (for one prototype or for a commercial series). Many manufacturing methods can be envisaged, such as direct diamond turning, molding, photolithography associated with chemical, plasma or reactive ion etching, effective index sub-wavelength patterning, among others. Each method has its own limitations in terms of materials that can be etched, on the minimal lateral size of the pattern that can be imprinted, on the maximum depth of the etch, and on the overall area that can be worked at affordable cost or within a realistic duration.

Let us take an example. Consider again the $f = 20$mm, $F\# = 2$ system, and a DoF extension of $\psi_{max} = 2.5\lambda$. The radius of the mask should thus be $R = 5$mm. A mask

with such a radius can be manufactured with many different technologies, such as diamond turning or photolithography. Let us now assume that the mask in Fig.3.a is used: the smallest normalized ring width is $\Delta\rho = \rho_k - \rho_{k-1} = 0.05$, which corresponds to a real width of $\Delta r = R \cdot \Delta\rho = 0.25\text{mm} = 250\mu\text{m}$. This value is a priori affordable by all the above-mentioned manufacturing methods. On the other hand, the mask of Fig.3.b, which provides the same DoF extension as the one in Fig.3.a, has a minimal normalized ring width of $\Delta\rho = 0.01$, which leads to an actual width of $\Delta r = 50\mu\text{m}$, which is too thin for usual diamond turning but clearly affordable for photolithography or subwavelength patterning. The generic mask of Fig.3.a could also be used to extend the DoF of the long focal system ($f = 200\text{mm}$, $F\# = 4$). In this case, the real diameter of the mask would be 50mm, and some technologies, such as microlithography or subwavelength patterning, would be difficult to use at an affordable cost. On the other hand, the minimum ring width would be $\Delta r = 1.5\text{mm}$, well within the capabilities of diamond turning.

More generally, we have found that for $\psi_{max} \leq 3\lambda$ and $N \leq 7$, there is always a mask with near-optimal performance having a smallest normalized ring width greater than 0.05. They thus could all be easily manufactured for the two optical systems we have taken as examples.

We have thus shown that the mask optimized in the previous section without any constraint on the minimal ring width can be manufactured and used in optical systems with realistic focal lengths and F-numbers. However, specific constraints on ring width could be introduced at the optimization stage if required by the manufacturing limitations for a given application.

4 Performance of the optimal masks through the DoF range

To understand better the behavior of the hybrid imaging system throughout the DoF range, Fig. 7 shows the variation of $IQ(\varphi_{opt}, \psi)$ as a function of ψ for the optimal masks obtained for the DoF ranges of $\psi_{max} = 1.5$, 2.0, 2.5 and 3.0λ, and for different numbers of rings.

The first observation is that, as ψ_{max} increases from one graph to the next, all curves get lower values for a given number of rings. This indicates that, as ψ_{max} increases, the problem becomes more difficult and the reachable optimal performance decreases. When $\psi_{max} = 1.5\lambda$, it is seen that all the curves for more than 4 rings yield a quasi-constant performance throughout the DoF range and are almost identical. This is consistent with the fact, observed in Fig. 6, that for this DoF range, the increase of performance levels off after 4 rings. When $\psi_{max} = 2.0\lambda$, the variation of IQ as a function of ψ is wider; we observe a drastic change of the curve shape between 3-ring and 4-ring masks, then the curves are smoother and become flatter, to reach a stable behavior from 7 rings onwards. When $\psi_{max} = 2.5\lambda$, the performance of the 5-ring and 6-ring masks is very similar, but adding a 7th ring makes it possible to suppress the valley around $\psi = 0.72\lambda$ while following the path of the 6-ring mask afterward; this leads to the small increase of $IQ_{maximin}$ observed for this value of ψ_{max} when going from 6 to 7 rings. When $\psi_{max} = 3.0\lambda$, we see that the mask with 3 rings is clearly insufficient to compensate for such a large value of the defocus, and all the optimized masks have a similar performance after 5 rings. Finally, it is interesting to notice that in all cases, increasing the number of rings tends to reduce the peak to valley variation of $IQ(\varphi, \psi)$ even if it does not increase considerably the minimum of $IQ(\varphi, \psi)$ compared to the situation with less rings. This fact can be of importance for some applications, and may request to

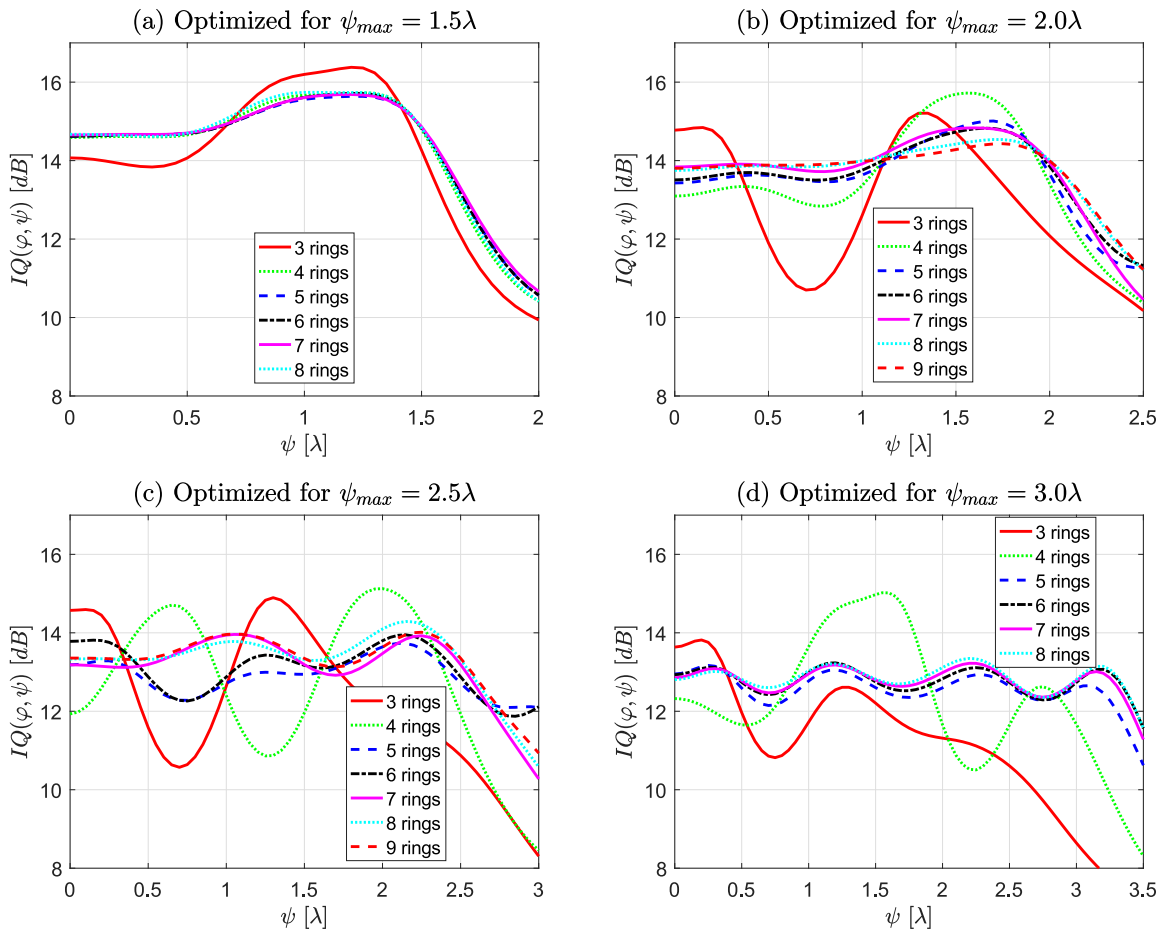

Fig 7 Performance comparison of masks optimized for different ψ_{max} and number of rings as a function of ψ.

be taken into account in the whole hybrid system design.

5 Robustness of the performance

Until now, the masks and deconvolution filters were optimized by assuming a generic image model with power-law PSD $S_{oo}(\nu) \propto \nu^{-2.5}$ and an aberration-free system. To check the robustness of our results and conclusions to these assumptions, we now evaluate the performance of the masks optimized for this model on systems that possess some amount of 3rd order aberrations, or that act on images whose PSDs follow a power law with different coefficients, or also on real-world images.

5.1 Sensitivity to aberrations

The binary phase masks obtained in the previous section were optimized considering an aberration-free system. If these masks were to be used in a real system, they could have to deal with a small amount of residual aberrations additionally to the defocus. To study the sensitivity of the mask performance to these residual aberrations, we first include in our simulated optical system a known amount of aberrations with a wavefront error expressed by the following Seidel terms: $W_s = \alpha\rho^4$ for spherical aberration, $W_c = \beta\rho^3\cos(\theta)$ for coma and $W_a = \gamma\rho^2\cos(2\theta)$ for astigmatism. As a reference, a conventional optical system was simulated to show the effects on $IQ(\psi, \varphi)$ of each aberration in addition to defocus. As mentioned before, according to the Raleigh's quarter wavelength rule, a peak-to-valley wavefront error due to aberration larger than $\lambda/4$ in a conventional optical system is considered to be visually noticeable and designers may aim for lower values.

In Fig.8.a, Fig.8.c, Fig.8.e and Fig.8.g, the image quality is shown as a function of the DoF obtained with the 5-ring binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$ after spherical, coma, astigmatism, or a combination of the three were added to the pupil function of the optical system. Each aberration was neither taken into account when optimizing the mask, nor when calculating the averaged Wiener filter. On the other hand Fig.8.b, Fig.8.d, Fig.8.f and Fig.8.h show the values of the image quality as a function of the DoF on a traditional system when affected by the aforementioned aberrations. Notice that we now use the full DoF range from $-\psi_{max}$ to $+\psi_{max}$ in the graphs, including the negative values, since some aberrations break the symmetry of the PSF around $\psi = 0$.

We can observe in Fig.8.a that for spherical aberration, the values of $IQ(\varphi, \psi)$ close to

$\psi = 0$ stay very close to the aberration-free case, but as $|\psi|$ increases, the values of $IQ(\varphi, \psi)$ are reduced. We can see that for $\alpha > 0.2\lambda$, the minimum values of $IQ(\varphi, \psi)$ drop by almost 2dB, and drop by 3.3dB for $\alpha = 0.5\lambda$. Comparatively, for the classical system shown in Fig.8.b, $IQ(\varphi, \psi)$ drops by 2.9dB from $\alpha = 0$ to $\alpha = 0.5\lambda$ at $\psi = 0$. We can also notice the shift of the "point of best focus" towards $-\psi$ as α grows and we can see that, without a mask, the values of $IQ(\psi, \varphi)$ drop rapidly outside of this point. This shows that the systems optimized solely for DoF correction are somewhat sensitive to small spherical aberrations, and the drop in image quality is comparable to conventional systems at the focal point $\psi = 0$.

For coma, we can see in Fig.8.c that when passing from $\beta = 0$ to $\beta = 0.5\lambda$, $IQ(\varphi, \psi)$ has a drop of less than 1dB between $\psi = -1\lambda$ and $\psi = 1\lambda$ but drops by 2.9dB outside of this range. In comparison, in Fig.8.d, the drop is 4.2dB for the conventional system at $\psi = 0$, showing that the system with the binary phase mask is more robust to coma aberration, even if the mask optimization did not take it into account.

For astigmatism, shown in Fig.8.e and Fig.8.f, we can see that the drop in $IQ(\varphi, \psi)$ is 3.6dB for the system with a binary phase mask and 4.0dB for the conventional system. The behavior is, similarly to coma, more stable between $\psi = -1\lambda$ and $\psi = 1\lambda$ and varies more dramatically beyond this range.

Finally, for the combined effects of the aberrations in Fig.8.g and Fig.8.h, we can see a maximum drop of $IQ(\psi, \varphi)$ of 2.7dB for the system with a binary mask and deconvolution and of 2.7dB for the conventional system at $\psi = 0$. The spherical aberration seems to dominate $IQ(\psi, \varphi)$ for the system with a phase mask, as we can see from the strong dissymetry in the range $1 < |\psi| < 2$, that occurs for spherical aberration (Fig.8.a). On the other hand, we see in the conventional system a more pronounced drop of $IQ(\psi, \varphi)$ at the "point of

best focus" than in the case with only spherical aberration, meaning that it is affected more strongly by the other two aberrations. From this analysis, we can conclude that systems with binary annular phase masks are robust to residual aberrations. The corresponding performance loss is then comparable to conventional systems affected by similar residual aberrations.

5.2 Robustness to image model

To evaluate the robustness of the optimized masks and filters to different scenes, let us consider the minimal value of image quality obtained with the masks optimized for $S_{oo}(\nu) \propto \nu^{-2.5}$ on an image having another PSD. This image quality will be defined as:

$$IQ_{min} = \min_{k} \left[IQ(\varphi_{opt}^{2.5}, \psi_k) \right] . \tag{11}$$

where $\varphi_{opt}^{2.5}$ defines the optimal mask for $S_{oo}(\nu) \propto \nu^{-2.5}$. It has to be noted that IQ_{min} is computed using the averaged Wiener filter based on the nominal PSD $S_{oo}(\nu) \propto \nu^{-2.5}$, which is different of the PSD of the observed image.

We have plotted in Fig. 9.a the value of IQ_{min} obtained with a generic image with PSD $S_{oo}(\nu) \propto \nu^{-2}$, for different values of ψ_{max}, as a function of the number of rings. Since this PSD has a more high-frequency content, the images are more difficult to deconvolve, and we indeed observe that the values of IQ_{min} are globally smaller than those of $IQ_{maximin}$ obtained in Fig. 6. However, the global behavior of IQ_{min} is similar: it grows with the number of rings until, at a certain amount of rings, it levels off and has a much slower growth. Even so, it levels off for a larger number of rings than in Fig. 6.

On the other hand, we have represented in Fig. 9.b the value of IQ_{min} obtained with a generic image with PSD $S_{oo}(\nu) \propto \nu^{-3}$. Since the PSD now has less high-frequency content, the images are easier to deconvolve and we indeed observe that the values of IQ_{min} are globally larger than in Fig 6. The value of IQ_{min} still grows with the number of rings until it levels off, but it saturates for smaller number of rings than in Fig. 6 and Fig. 9.a. It is to be noted that the slight decrease of image quality with the number of rings observed in some curves (for example, for $\psi_{max} = 1\lambda$) is due to the fact that the deconvolution filter, based on the nominal PSD, is not perfectly adapted to the actual PSD of the observed image. We can conclude from these simulations that as the high-frequency content of the image increases, the overall hybrid imaging performance decreases and more rings are needed to saturate the performance.

Let us now consider the performance of the optimized masks on real-world images. Fig. 10.a (referred to in the following as "Lena") has a small high-frequency content, and we have checked that its PSD $S_{oo}(\nu)$ falls between $\nu^{-2.5}$ and ν^{-3}. On the other hand, Fig. 10.b (referred to in the following as "Sea"), contains more high spatial frequency details and we have checked that its PSD falls between $\nu^{-2.5}$ and ν^{-2}.

In Figs. 11.a and 11.c, we have represented the values of IQ_{min} obtained on Lena and on Sea respectively. The conclusions are similar to those drawn from Fig. 9, showing that the scene PSD is indeed the main factor influencing the final imaging performance of the hybrid system. On Fig. 11.b and Fig. 11.d, we have represented the values of $IQ_{maximin}$ obtained on Lena (b) and Sea (d) with masks optimized for the PSD model best adapted to each image, that is, $S_{oo}(\nu) \propto |\widetilde{O}(\nu)|^2$, the square modulus of the Fourier transform of the sharp image. By comparing Fig. 11.b with Fig. 11.a, we can see that the behavior of the curves and the

obtained values of $IQ_{maximin}$, obtained with the optimal masks, are very close to the values of IQ_{min} obtained with mask optimized with $S_{oo}(\nu) \propto \nu^{-2.5}$. The same conclusion can be drawn from the Sea image (Fig. 11.d and Fig. 11.c). This shows that the optimization of the masks is robust to the chosen PSD model of the scene.

5.3 Visual performance comparison

Let us now visually assess the performance of the hybrid imaging systems based on the optimized masks on images perturbed with a simulated defocus blur. For simplicity, the deconvolutions are performed in the Fourier domain directly using Eq.5 with no truncation of the deconvolution filter. We have displayed on Fig. 12.a the simulated image of Lena that would be obtained by a well-focused diffraction limited optical system followed by deconvolution with a Wiener filter adapted to the observed image. This "best possible image" will be our reference. On Fig. 12.b, we have represented a simulated image of the same scene that would be obtained with an imaging system defocused of $\psi = \psi_{max} = 2.0\lambda$ followed by deconvolution with an averaged Wiener filter as defined in Eq. 5. We can see that for this level of defocus the image is strongly blurred.

In Fig. 13 are displayed the simulated images obtained with a hybrid imaging system with masks optimized at $\psi_{max} = 2.0\lambda$ and composed of 3 to 8 rings and zoom-ins of the brim of the hat, its feathers and the eye of Lena so we can better appreciate the differences. These masks have been optimized with the generic power-law PSD model $S_{oo}(\nu) \propto \nu^{-2.5}$ and not the PSD of the scene, $S_{oo}(\nu) \propto |\widetilde{O}(\nu)|^2$. The averaged Wiener filter used for deconvolution is also based on the generic power-law PSD. For all the deconvolved images in Fig. 13 we reach the goal of having a better sharpness and image quality than for the maskless system

in Fig. 12.b. Of course, the sharpness is inferior to that obtained in Fig. 12.a, as the phase mask alters the optical transfer function of the system to extend the DoF. We note that the quality of the deconvolved images increases with the number of rings; with the 3-ring mask, all the features of the scene can be identified, but there is an important loss of sharpness, as well as ringing artifacts around high spatial frequency details, like the brim of the hat or the feathers. With the 4-ring mask, the ringing is greatly reduced, but there is a slight increase in blur as can be seen on feathers. With the 5-ring mask the sharpness is increased and ringing is reduced even more; we can now identify finer feathers on the hat. With the 6-ring mask, there is a very slight increase in visual quality, mostly in contrast: this is consistent with the shape of the curve corresponding to $\psi_{max} = 2$ in Fig. 11.a. The image obtained with a 7-ring mask corresponds to a more significant sharpness improvement, which is again consistent with the increase of IQ_{min} observed in Fig. 11.a. Finally, passing from 7 to 8 rings decreases IQ_{min} by 0.1dB. This somehow counter-intuitive result, also observed in Fig. 11.a, is due to the fact that the masks were optimized for a generic PSD model and not for the PSD of the sharp image. Indeed, this drop disappears in Fig. 11.b.

In this section we have shown that the visual quality of the images as a function of the number of rings is consistent with the value of the image quality criterion used to optimize the masks.

6 Conclusion

We have studied the performance of DoF-extending annular binary phase masks as a function of the number of rings and the desired DoF extension. This study was done using a co-design approach with a hybrid imaging system that takes into account deconvolution in the global

image quality. We have shown that the mask optimization landscape is highly nonconvex and that it is plagued with many local maxima. However, some clear general tendencies have been identified.

In particular, we have shown that more rings are necessary to reach higher DoF, but the best possible image quality decreases as the required DoF range increases, so that the DoF extension reachable with binary phase masks is bounded. Moreover, for a given value of the required DoF, there is a point where augmenting the amount of rings does not improve the performance anymore, indicating that a limited number of rings is enough to obtain the best possible performance. This result is important for mask manufacturing, since the number and thickness of the rings reachable by technology may be limited given a numerical aperture and a target DoF.

We have also shown that binary annular phase masks, optimized for a nominal aberrationless system, are relatively robust to small amounts of spherical, coma and astigmatism aberrations and as these aberrations grow, the drop on image quality is comparable to the drop on a conventional system. Finally, we have empirically validated that using a generic power-law PSD model is sufficient to obtain masks with a predictable performance regardless of the observed scene.

This work has many perspectives, such as applying the same co-design approach to other types of DoF enhancing masks with different profiles and complexities, or taking into account the ratio between the pixel size and the diffraction-limited spot size.[14] In these studies, it will be also interesting to investigate how co-design can help correcting for aberrations as well as extending the DoF, as in Ref.[15] for example, but with the added originality of taking into account deconvolution in the final image quality. Moreover, since it is a common practice

in embedded imaging systems to truncate the deconvolution filter to save processing time and power consumption, it will be important to study the effects of this truncation on the final image and to investigate the compromise between computational power and restored image quality for real-time video processing on hybrid system usage. Finally, it would be interesting to study the relationship between IQ and more conventional optical system design quality metrics, like the MTF, even-though the later cannot quantify the raw image noise amplification but only the deterministic scene frequency contains restoration.

Disclosures

Authors have no conflict of interest related to the work leading to the results presented in this manuscript.

Acknowledgments

We thank Marie-Anne Burcklen for fruitful discussions. The work reported in this paper is supported in part by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA Grant Agreement no. PITN-GA-2013-608082 (ADOPSYS) that provides a Ph.D fellowship to Rafael Falcón.

References

1 E. Ben-Eliezer, N. Konforti, B. Milgrom, *et al.*, "An optimal binary amplitude-phase mask for hybrid imaging systems that exhibit high resolution and extended depth of field," *Optics express* **16**(25), 20540–20561 (2008).

2 F. Diaz, M.-S. L. Lee, X. Rejeaunier, *et al.*, "Real-time increase in depth of field of an

uncooled thermal camera using several phase-mask technologies," *Optics letters* **36**(3), 418–420 (2011).

3 T. Zhao, A. Liu, Q. Liu, *et al.*, "Axial intensity distribution analysis for a depth-of-field-extended optical system using a low-frequency binary phase mask," *Applied optics* **53**(17), 3782–3786 (2014).

4 W. T. Cathey and E. R. Dowski, "New paradigm for imaging systems," *Applied optics* **41**(29), 6080–6092 (2002).

5 M. D. Robinson and D. G. Stork, "Joint design of lens systems and digital image processing," in *Contract Proceedings 2006*, 63421G–63421G, International Society for Optics and Photonics (2007).

6 F. Diaz, F. Goudail, B. Loiseaux, *et al.*, "Increase in depth of field taking into account deconvolution by optimization of pupil mask," *Optics letters* **34**(19), 2970–2972 (2009).

7 B. Milgrom, N. Konforti, M. A. Golub, *et al.*, "Pupil coding masks for imaging polychromatic scenes with high resolution and extended depth of field," *Optics express* **18**(15), 15569–15584 (2010).

8 S. Elmalem, N. Konforti, and E. 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 M.-A. Burcklen, F. Diaz, F. Leprêtre, *et al.*, "Experimental demonstration of extended depth-of-field f/1.2 visible high definition camera with jointly optimized phase mask and real-time digital processing," *Journal of the European Optical Society-Rapid publications* **10** (2015).

10 D. L. Ruderman, "Origins of scaling in natural images," *Vision research* **37**(23), 3385–3398 (1997).

11 M. Born, E. Wolf, and A. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Cambridge University Press (1999).

12 A. Maréchal, "Étude des effets combinés de la diffraction et des aberrations géométriques sur l'image d'un point lumineux," *Rev. Opt* **26**, 257–277 (1947).

13 Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, 69–73, IEEE (1998).

14 T. Vettenburg, *Optimal design of hybrid optical digital imaging systems*. PhD thesis, Heriot-Watt University (2010).

15 J. Ojeda-Castañeda, P. Andres, and A. Diaz, "Annular apodizers for low sensitivity to defocus and to spherical aberration," *Optics Letters* **11**(8), 487–489 (1986).

Rafael Falcón is a PhD. fellow at Institut d'Optique Graduate School (Palaiseau) under the ADOPSYS European program. He received his BS degree in physics from Universidad Nacional Autónoma de México (Mexico City) and his MS degrees in Optics from Polytechnica Warszawska (Warsaw) and Institut d'Optique Graduate School under an Erasmus Mundus Grant.

François Goudail graduated from the École Supérieure d'Optique, (Orsay), in 1992 and obtained his PhD in 1997 from the University of Aix-Marseille III. He has been associate professor at Fresnel Institute, (Marseille), until 2005. He is now professor at the Institut

d'Optique Graduate School (Palaiseau). His research topics include information extraction in images from different types of passive and active sensors (hyperspectral, SAR, polarimetric), wavefront engineering and joint design of optical systems and image processing algorithms.

Hervé Sauer is a former student of the École Normale Supérieure de Fontenay/St-Cloud. He got the Agrégation degree in physics in 1985, and a MSc and PhD in physics from Université Paris Sud-XI (Orsay) in 1985 and 1990 respectively. He is an associate professor at Institut d'Optique Graduate School (Palaiseau) since 1990. His main research topics deal with simulations of optical systems and optical system design.

Biographies of the other authors are not available.

List of Figures

5 Point spread function at DoF (left to right) $\psi = 0$, $\psi = 1.0\lambda$ and $\psi = 2.0\lambda$ for: (top row) an aberration-free, diffraction-limited optical system and (bottom row) the same system with a 6-ring binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$.

6 Performance comparison of optimal annular phase masks with different amount of rings for $\psi_{max} \in [1.0,\ 1.5,\ 2.0,\ 2.5,\ 3.0]\lambda$. The case where the number of rings is equal to 1 is equivalent to having no mask.

7 Performance comparison of masks optimized for different ψ_{max} and number of rings as a function of ψ.

8 Performance comparison of a system with a 5-ring binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$ and deconvolution and conventional systems with the addition of: a)-b) spherical aberration corresponding to $\alpha = \{0, 0.1, ..., 0.5\}\lambda$; c)-d) coma corresponding to $\beta = \{0, 0.1..., 0.5\}\lambda$; e)-f) astigmatism corresponding to $\gamma = \{0, 0.1, ..., 0.5\}\lambda$; e)-f) aberrations such that $\alpha = \beta = \gamma = \{0, 0.05, ..., 0.25\}\lambda$

9 Performance comparison of optimal annular phase masks obtained for the model $S_{oo}(\nu) \propto \nu^{-2.5}$ and applied to scenes following the model (a) $S_{oo}(\nu) \propto \nu^{-2}$ and (b) $S_{oo}(\nu) \propto \nu^{-3}$.

10 Natural images used as scenes to test the performance of the hybrid imaging system

11 Variation of IQ_{min} and $IQ_{maximin}$ for different values of ψ_{max} and different numbers of rings, using two different scenes: Lena (a and b) and Sea (c and d). The masks are optimized for the model with $S_{oo}(\nu) \propto \nu^{-2.5}$ (a and c respectively) and for the scenes themselves where $S_{oo}(\nu) \propto |\widetilde{O}(\nu)|^2$ (b and d).

12 Images produced by a diffraction limited system with $SNR = 34dB$ after deconvolution at (a) $\psi = 0\lambda$ and (b) $\psi = 2.0\lambda$.

13 Image and zoom-ins of Lena obtained for defocus $\psi = \psi_{max}$ with the simulated hybrid imaging system with masks optimized for $\psi_{max} = 2.0\lambda$ and different numbers of rings.

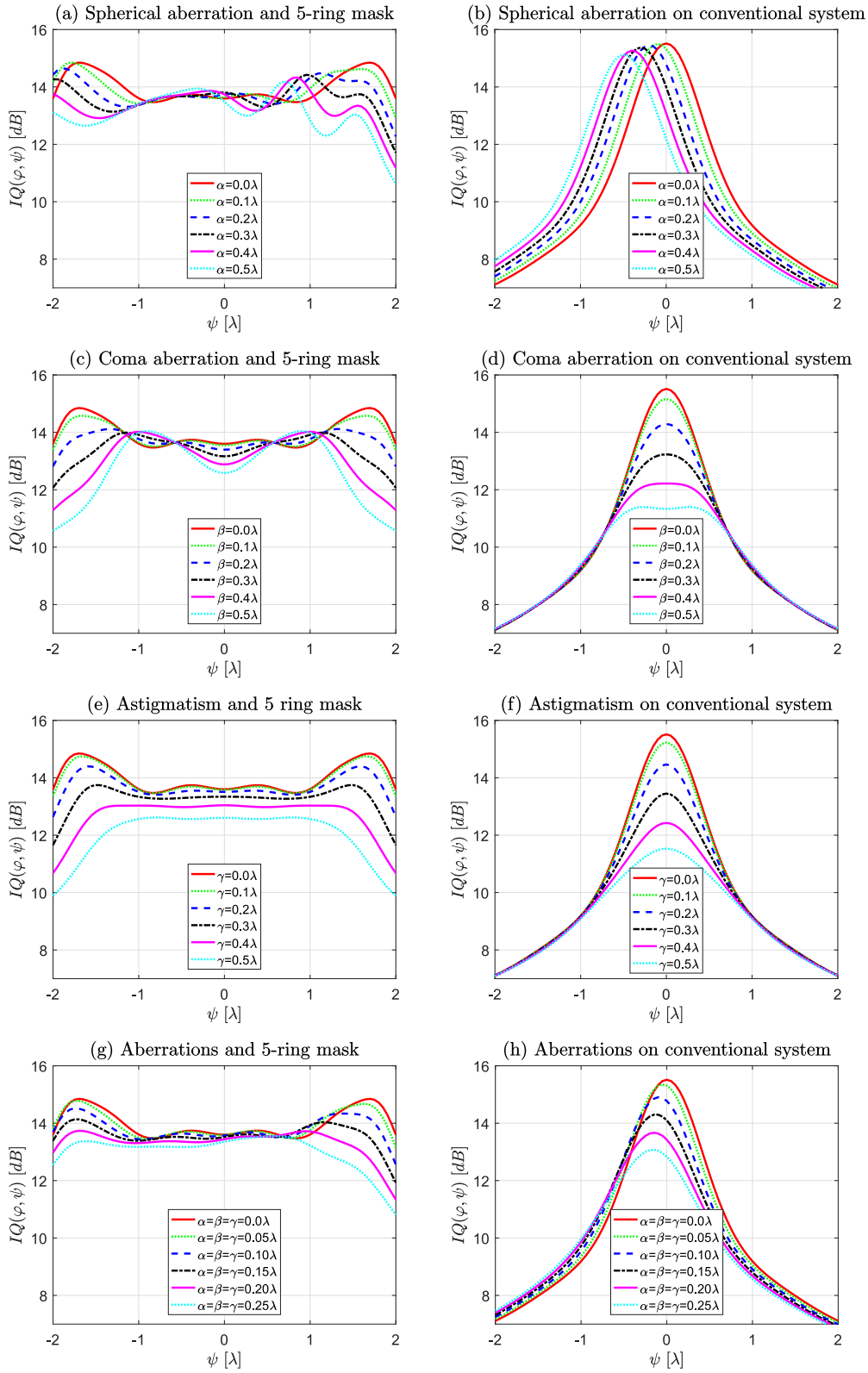

Fig 8 Performance comparison of a system with a 5-ring binary annular phase mask optimized for $\psi_{max} = 2.0\lambda$ and deconvolution and conventional systems with the addition of: a)-b) spherical aberration corresponding to $\alpha = \{0, 0.1, ..., 0.5\}\lambda$; c)-d) coma corresponding to $\beta = \{0, 0.1..., 0.5\}\lambda$; e)-f) astigmatism corresponding to $\gamma = \{0, 0.1, ..., 0.5\}\lambda$; e)-f) aberrations such that $\alpha = \beta = \gamma = \{0, 0.05, ..., 0.25\}\lambda$

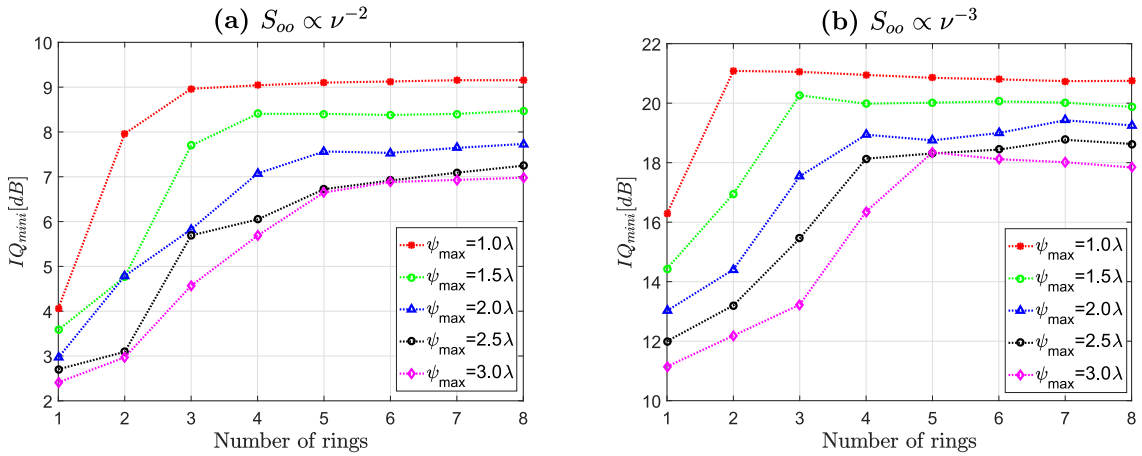

Fig 9 Performance comparison of optimal annular phase masks obtained for the model $S_{oo}(\nu) \propto \nu^{-2.5}$ and applied to scenes following the model (a) $S_{oo}(\nu) \propto \nu^{-2}$ and (b) $S_{oo}(\nu) \propto \nu^{-3}$.

Fig 10 Natural images used as scenes to test the performance of the hybrid imaging system

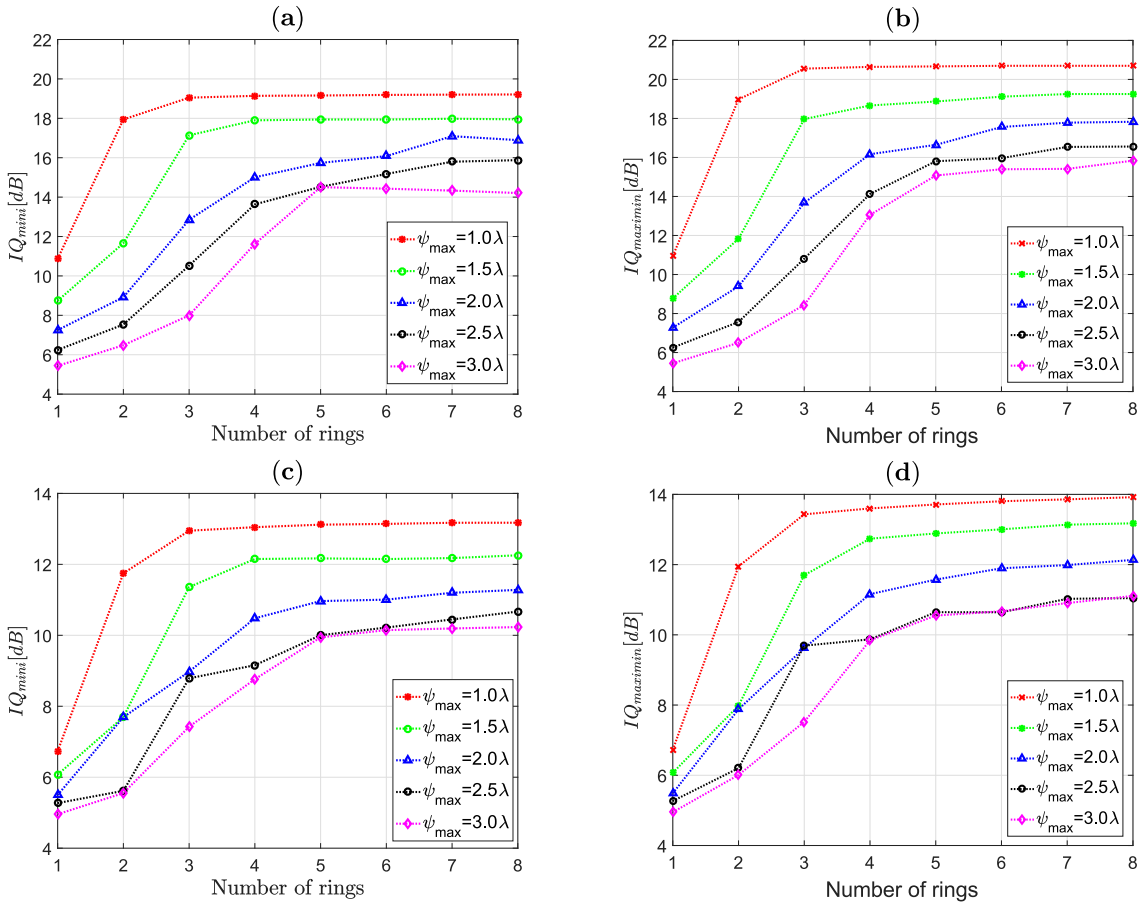

Fig 11 Variation of IQ_{min} and $IQ_{maximin}$ for different values of ψ_{max} and different numbers of rings, using two different scenes: Lena (a and b) and Sea (c and d). The masks are optimized for the model with $S_{oo}(\nu) \propto \nu^{-2.5}$ (a and c respectively) and for the scenes themselves where $S_{oo}(\nu) \propto |\widetilde{O}(\nu)|^2$ (b and d).

Fig 12 Images produced by a diffraction limited system with $SNR = 34dB$ after deconvolution at (a) $\psi = 0\lambda$ and (b) $\psi = 2.0\lambda$.

Fig 13 Image and zoom-ins of Lena obtained for defocus $\psi = \psi_{max}$ with the simulated hybrid imaging system with masks optimized for $\psi_{max} = 2.0\lambda$ and different numbers of rings.