COMPARING OPTICAL TO DIGITAL METRICS: WHAT IS THE OPTIMAL DEFOCUS IN A ROTATIONALLY SYMMETRIC SYSTEM?

Javier Portilla and Sergio Barbero

Instituto de Óptica
Consejo Superior de Investigaciones Científicas, Madrid.

ABSTRACT

We address the problem of finding the optimal focus of an optical system with spherical aberration, under three optimization criteria: the classical optical root-mean-square second order moment minimization, the expected mean square error on the sensor, and the expected mean square error at the output of a Wiener restoration filter. We observe that these three criteria may behave very differently, and, particularly, the classical optical criterion typically provides very poor results in comparison. This has a direct impact on the design of new hybrid optical-digital imaging systems, which account for the image quality after an optics-aware embedded digital restoration stage. We present some very encouraging results for simulations under different noise and aberration levels.

Index Terms— computational imaging; joint optical-digital design; optical metrics; digital metrics; defocus

1. INTRODUCTION

The conventional way of designing imaging optical systems is based on reducing the optical blur, as described by the Point Spread Function (PSF). However, the registered optical images can be further improved by means of computational techniques. Indeed, what is typically done in digital cameras is a sequential procedure: first to minimize a purely optical cost function, and then to deal with some remaining imperfections of the registered image by applying digital image processing. In contrast, recent confluence of optical imaging and digital computation opens a new field of applied research that aims at optimizing the image quality at the output of the digital processing stage, instead of at the image sensor plane. As a consequence, under this new hybrid joint optical-digital approach, classical quality metrics used in optical design are no longer appropriate.

Only a few authors have proposed a real hybrid joint digital-optical design strategy. A seminal work was that of Stork et al [1], where an image restoration model based on the Wiener filter was applied in conjunction with a scene power spectral density empirically derived from a collection of images. The image quality metric was the mean square error (MSE) in the restored image. A similar methodology (but using a $1/f^2$ model for the power spectral density, instead) was proposed by Vettenburg et al. [2] for depth-of-focus applications.

The optical theory establishes that the better design is wanted the more complex the optical configuration must be, a gross rule-of thumb being that one extra element is necessary to satisfactorily correct each optical aberration. Hence, the history in photographic lenses has evolved from single spherical lenses onto complicate objective lenses comprising sometimes more than ten components [3]. However, even a highly optimized optical design cannot attain perfect optical imaging, because of fundamental restrictions imposed by physics. In addition, there is an increasing demand on ultra-compact imaging systems (besides the blooming smartphone cameras industry, other embedded small cameras for drones, small robots, medical exploration devices, etc), whose robustness and/or low price limit their optical complexity. Thus, an appealing challenge posed to optical design is to obtain the optimal combination of specific aberrations, given the design constraints, to optimize a quality measurement after digital image restoration.

In this work, we restrict ourselves to on-axis analysis (small field of view) in rotationally symmetric optical systems (most ones), ignoring chromatic aberration. In these conditions, the PSF is described by two major aberrations: defocus and spherical aberration. We have empirically tested that, for the simulations implemented in Section 3, diffractive PSFs are more reliable than geometrical ones, even when confining, as we did, to the far-field Fraunhofer approximation (geometrical PSFs, though, are a valid option under certain conditions, see, e.g. [5, 4]).

Regarding optical quality metrics, one of the most used is the Root-Mean-Square (rms) deviation of the spot diagram, which converges, for increasing number of rays, to the square root of the PSF's second order moment [9]. In other occasions an Optical Transfer Function (OTF) metric is used, typically the volume under the OTF over a frequency range. Although for small values of aberration, the rms metric is highly correlated to the OTF one, Mahajan et al. [6] showed that this

This research has been supported by the Spanish Government FIS2016-75891-P grant.

correlation falls for large amounts of aberration.

Here we analyze the optimal balance between defocus and first order spherical aberration when considering an MSE metric in the digitally restored image, and its differences with respect to optical and sensor's image quality metrics. As we show, results are specially relevant for a joint digital-optical design approach, because the optimal aberration balance significantly differs depending on which metric is used, and because substantial final image quality improvement derives from using digital metrics. Finally, a crucial aspect when using digital metrics (either at the sensor or restored) is sensor noise variance, as it has a strong influence on both the cost function and the spectral profile of the restoration filters.

2. METHODS

2.1. Signal, blur and noise: the observation model

In order to make a full evaluation of the quality corresponding to each optical configuration, we make some assumptions:

- An observation model describing the image at the output of the sensor:

$$\mathbf{y} = \mathbf{h} * \mathbf{x} + \mathbf{w},$$

  where the symbol "*" denotes convolution, and $\mathbf{h}$ is the PSF, assumed constant in all the field of view. $\mathbf{x}$ is the ideal image, and $\mathbf{w}$ is the noise. This model does not consider either optical distortion.

- A power spectral density (PSD) model for typical images. We used $P_x(f) = k/f^2$, for $f \neq 0$ and 0 in the origin (see, e.g., [7]), where $f = ||\mathbf{f}||$ is the radial frequency. k is computed to satisfy that the variance (PSD's integral, excluding zero) is $\sigma_x^2 = 1$.

- A spectral model for the sensor's noise. For simplicity we assume white noise: $P_w = \sigma_w^2$. We consider three sensor noise scenarios: low, medium, and high, corresponding to σ_w $\{1/128, 1/32, 1/8\}$. If there was no optical blur, that would correspond to a SNR in the sensor image of 42, 30 and 18 dB, respectively.

2.2. PSFs computation

Optical aberrations (within third-order aberration theory) of a rotationally symmetric optical system are described by the wave aberration function: $W(x,y) = B_d \left(\frac{\rho}{R}\right)^2 + A_s \left(\frac{\rho}{R}\right)^4$, where B_d and A_s are defocus and spherical aberration respectively, R is the pupil radius, (x,y) are pupil plane coordinates and $\rho = \sqrt{x^2+y^2}$. Within Fraunhofer diffraction theory [8], the PSFs $h(x',y';B_d,A_s)$ can be approximated from $W(x,y)$ using the Fourier Transform ($\mathcal{F}\{\}$):

$$h(x',y') = \left|\mathcal{F}\left\{A(\hat{x},\hat{y})\exp(-i2\pi W(\hat{x},\hat{y})/\lambda)\right\}\right|^2, \quad (1)$$

where $A()$ is a uniform radiance pupil transmittance, λ is the wavelength, and $\hat{x},\hat{y} = (x,y)/(\lambda F)$, being F the F-number.

2.3. Three quality metrics

2.3.1. Second-order moment of PSF

The rms value of the spot diagram converges, for large number of rays, to the square root of the second order moment of the PSF (geometrical and diffractive). The latter, for a circularly symmetric system, is:

$$E_O^2 = 2\pi \int h(r) r^2 r dr = 2\pi \int h(r) r^3 dr, \quad (2)$$

where $r = \sqrt{x'^2 + y'^2}$. An advantage of using this metric is that it is a quadratic function of the coefficients of the wave aberration expansion [9]. The iso-E_O^2 curves in the $B_d - A_s$ plane are rotated ellipses centered at $(0,0)$. As a consequence, the B_d values for a given A_s minimizing E_O^2 ($\hat{B}_d^{(O)}(A_s)$) are on a straight line passing by the origin. In particular, it is straightforward to derive that $\hat{B}_d^{(O)}(A_s) = -4/3A_s$ [9]. Thus, in presence of spherical aberration, introducing some (negative) defocus results in improving the rms optical quality. It is interesting to note that, under the more elaborated OTF volume metric (which correlates better to the MSE on the sensor, excluding the noise effect), the optimal ratio $\frac{B_d}{A_s}$ differs from the predicted for the spot rms, decreasing in amplitude as A_s increases [6].

2.3.2. Expected MSE in the sensor's image space

The *rms* metric is a simple way to quantify the optical blur. However, we are rather interested in assessing the impact of the optical blur on the image quality. This is strongly dependent on PSF's *shape*, not just on its second order moment. Although OTF-based metrics provide more information about the sensor's image quality, they do not consider the two other factors affecting the quadratic error: the PSDs of (uncorrupted) signal and noise. However, under the model assumptions in section 2.1, computing the expected MSE on the sensor's output image is straightforward:

$$E_S^2 = \int_{f\neq 0} P_e^{(S)}(\mathbf{f}) d\mathbf{f}, \quad (3)$$

where $P_e^{(S)}(\mathbf{f}) = \mathbb{E}\{|Y(\mathbf{f}) - X(\mathbf{f})|^2\}$. Substituting $Y(\mathbf{f}) = H(\mathbf{f})X(\mathbf{f}) + W(\mathbf{f})$ and operating, we obtain:

$$\begin{aligned} P_e^{(S)}(\mathbf{f}) &= P_w(\mathbf{f}) + P_x(\mathbf{f})|1 - H(\mathbf{f})|^2 \\ &= \sigma^2 + |1 - H(\mathbf{f})|^2 / f^2. \end{aligned} \quad (4)$$

2.3.3. Expected MSE in the restored image space

Under a hybrid optical-digital design perspective, the image at the sensor's output is just an intermediate result. How can

we improve this image applying image restoration? Although the answer depends on the particular restoration algorithm, one can use as a useful guide the output of the optimal linear filter (Wiener) [1, 2]. Under this simplifying choice:

$$E_R^2 = \int_{f\neq 0} P_e^{(R)}(\mathbf{f}) d\mathbf{f},$$

where $P_e^{(R)}(\mathbf{f}) = \mathbb{E}\{|H_w(\mathbf{f})Y(\mathbf{f}) - X(\mathbf{f})|^2\}$, and $H_w(\mathbf{f}) = P_x(\mathbf{f})H^*(\mathbf{f})/(|H(\mathbf{f})|^2 P_x(\mathbf{f}) + P_w(\mathbf{f}))$ is the Wiener filter. Operating, we obtain:

$$\begin{aligned} P_e^{(R)}(\mathbf{f}) &= \frac{P_x(\mathbf{f})P_w(\mathbf{f})}{P_x(\mathbf{f})|H(\mathbf{f})|^2 + P_w(\mathbf{f})} \\ &= \left(|H(\mathbf{f})|^2/\sigma^2 + f^2\right)^{-1}. \end{aligned} \quad (5)$$

3. SIMULATIONS, RESULTS AND DISCUSSION

We have studied the joint space of (B_d, A_s) , with $B_d \in [-4, 0]$, $A_s \in [0, 3]$, in 0.1 intervals of each aberration (all in μm). We have assumed $\lambda = 0.5\mu$m, focal length $f = 50$ mm, and F-number of 10, which determine the pupil radius and the normalization factor in $\hat{x}, \hat{y}$ (Eq. 1). We have considered a high-resolution sampling, corresponding to a 512×512 virtual sensor grid with 2.5 μm virtual detectors. This sampling guarantees an aliasing-free representation [9, 4].

3.1. Numerical optimization

We first computed $E_O^2(A_s, B_d)$ (independent of the noise level), $E_{S_0}^2(A_s, B_d)$ for no noise (then we just add σ_w^2 for each noise level, according to Eqs. (3) and (4)), and $E_R^2(A_s, B_d, \sigma_w^2)$. Then we applied a bicubic spline interpolation, doubling the sampling density along each aberration. We have found, for each given pair (A_s, σ_w^2), the marginal optimal defocus values $\hat{B}_D^{(O)}(A_s)$, $\hat{B}_D^{(S)}(A_s, \sigma_w^2)$, $\hat{B}_D^{(R)}(A_s, \sigma_w^2)$. Then, we obtained sub-sample accuracy in the minimization by means of a 3-point parabolic fitting.

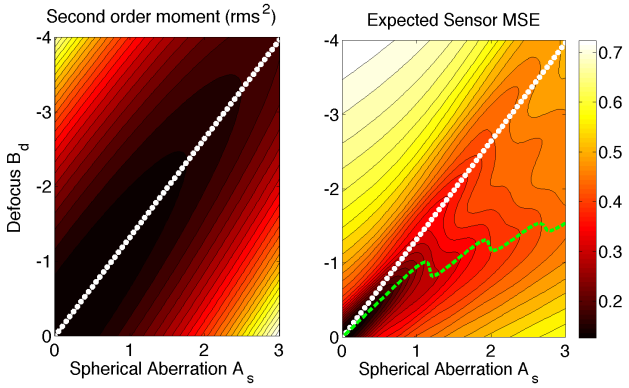

Fig. 1. Optical (E_O^2, left) and sensor (E_S^2, right) cost functions, drawing the optimal B_d for each A_s value (in μm).

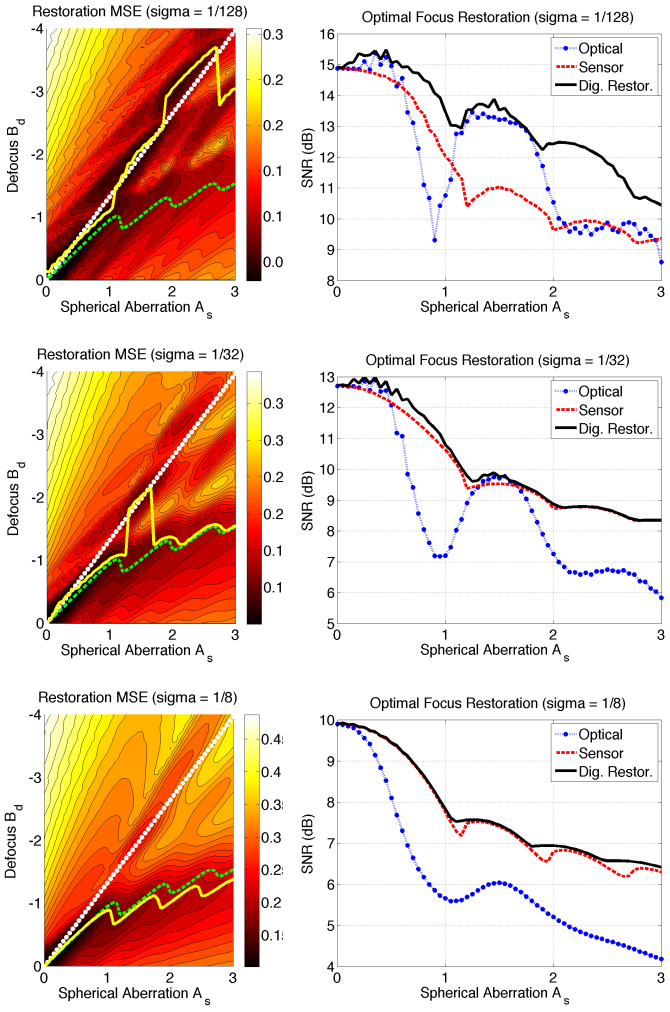

Fig. 2. Left: E_R^2 (expected Mean Square Error in the Wiener restoration) as a function of B_d and A_s in μm, including the optimal $\hat{B}_d^{(R)}(A_s, \sigma_w)$ (in continuous lines - other lines correspond to $\hat{B}_d^{(O)}$ and $\hat{B}_d^{(S)}$, included here for comparison), for the 3 noise levels. Right: Corresponding expected SNR.

In Figure 1 we show $E_O^2(A_s, B_d)$ and $E_{S_0}^2(A_s, B_d)$, including their marginal optimal defocus values. Differences between these two criteria are already very important, especially for high aberration levels. In Figure 2 we extend the analysis to the final goal, i.e., to minimize the error in the digitally restored image. Here differences between the two previous criteria and the final one are very strong, especially for low noise levels. On the right panels we show the expected SNR obtained doing a sequential optimization: first, to obtain the optimal defocus value, according to each of the three metrics, and then to restore the resulting image, using Wiener for the involved PSF and noise level. We can see that for moderate and high noise levels, MSE sensor and restoration metrics become very similar, whereas the optical metric is still very

far below in performance.

3.2. Visual results (Wiener and beyond)

In this subsection we address three questions: (1) how the numerical differences shown above translate into visual results; (2) how accurate is the model used to predict the quality of the Wiener restoration; and (3) is the Wiener filter a reliable *oracle* to guide the optical parameters when using more powerful (non-linear) restoration methods instead? These questions deserve a more detailed analysis; here we just present an example that may be representative of the general behavior. Figure 3 presents simulations of a degraded image, and its restoration, using the three criteria and two different restoration methods: Wiener and a non-linear, state-of-the-art method (ConDy10 L2-relaxed L0 [10]).

We used in this example the *Pirate* image with $A_s = 0.8\mu m$ and $\sigma_w = 1/128$ (high SNR, e.g. for an average half full-well capacity level of 12500 [11]). We see in Fig. 2 (top) how the optical criterion introduces the strongest defocus ($-1.06\mu m$), and the opposite for the MSE sensor criterion ($-0.78\mu m$), the MSE restoration lying in-between ($-0.94\mu m$). In Fig 3 we show (first column) the corresponding sensor images (left), and after having applied a Wiener filter (center). The optical criterion is clearly the worst in image quality (use $\times 4$ zoom). Slightly less noisy image results from using Restored-MSE, as compared to Sensor-MSE.

In order to explore our third question, we also examine the results obtained with a more powerful non-linear restoration method [10] (right column in figure). The improvement is very noticeable for all three cases. Still, the relative order in image quality is preserved: the optically optimized image is clearly the worst. Wiener-optimized and sensor-optimized provide both excellent results. Finally, Table 4 shows a quantitative comparison of these results in terms of the SNR, in dB, and also in terms of the model prediction. We see that, although not accurate in absolute numbers, the model correctly predicted the relative position in actually achieved image quality. Furthermore, the correct relative order is maintained even when using non-linear restoration. We have also applied the Wiener restoration directly to the $B_d = 0$ case (no defocus), and obtained 14.74 dB, that is, 1.1 dB better than applying the optical defocus correction.

Discussing the causes of previous results, it is well-known that the error after linear restoration greatly increases when the OTF goes close or below the noise level, especially in frequencies having high signal energy. This is a very different criterion to minimizing the second order moment of the PSF, which may favor a smaller effective spatial support of the PSF at the prize of having more abrupt boundaries. Abrupt boundaries (as produced with strong defocus), in turn, give rise to very damaging OTF oscillations, which often get below noise level, and may even cross zero-level (contrast inversion).

4. CONCLUSIONS

The results presented here indicate that, far from well-corrected optical systems, it may be inappropriate (even counterproductive) to use classical optical design criteria such as the rms spot, when optical blur is digitally compensated in the camera acquisition loop. Instead, it may suffice to minimize a simple sensor spectral error model for drastically improving results, for moderate to high noise regimes. However, it is preferable to directly minimizing the expected Wiener error under the same spectral model, so producing more robust results, both in high and low aberration, and high and low noise regimes, for a similar computational complexity. Although a more extensive study is necessary, this latter criterion seems to hold as a useful approximation when, instead of Wiener, more powerful non-linear restoration techniques are applied.

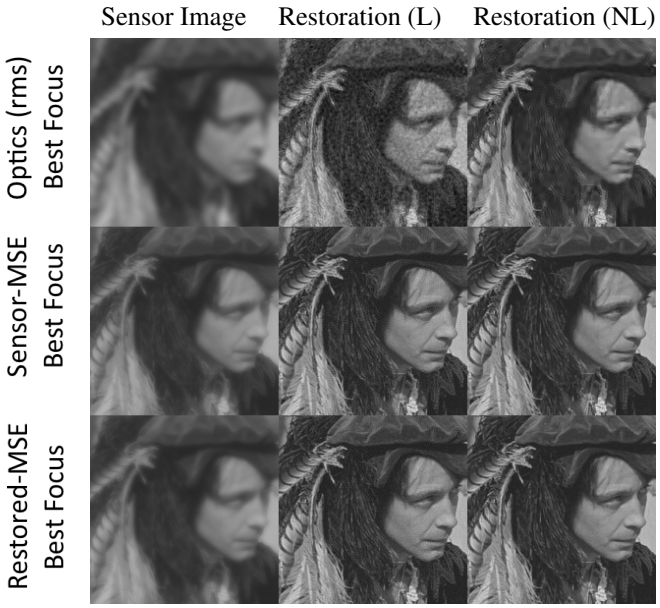

Fig. 3. Simulated observation crops (left), results with Wiener (center) and L2-relaxed L0 [10] (right), for $\sigma_w = 1/128$ and $A_s = 0.8\mu m$.

Optimized metric $\rightarrow$	E_O^2	E_S^2	E_R^2
Model Prediction	11.10	13.40	14.41
Actual (Wiener)	13.64	19.23	19.95
Actual (Non-Linear)	16.78	22.22	22.61

Table 1. SNR values, in dB, of the restored images from Fig. 3, predicted by the model, restored with a linear method (Wiener) and a non-linear one (L2-relaxed L0 [10]).

5. REFERENCES

[1] D. G. Stork and M. D. Robinson, "Theoretical foundations for joint digital-optical analysis of electro-optical imaging systems," *Applied Optics*, vol. 47, no. 10, pp. B64–B75, 2008.

[2] T. Vettenburg, D. N. Bustin, and Harvey A.R., "Fidelity optimization for aberration-tolerant hybrid imaging systems," *Optics Express*, vol. 18, no. 9, pp. 9220–9228, 2010.

[3] R. Kingslake, *A history of the photographic lens*, Academic Press, Boston, 1989.

[4] A. E. Savakis and Trussell H. J., "On the accuracy of psf representation in image restoration," in *IEEE Trans Image Process*, 1993, vol. 2, pp. 252–259.

[5] J. Portilla and S. Barbero, "Accuracy of geometric point spread function estimation using the ray-counting method," in *Proc. SPIE*, 2012, pp. 855003–855003.

[6] V. N. Mahajan and J. A. Diaz, "Comparison of geometrical and diffraction imaging in the space and frequency domains," *Applied Optics*, vol. 55, no. 12, pp. 3241–3250, 2016.

[7] D L Ruderman and W Bialek, "Statistics of natural images: Scaling in the woods," *Phys. Rev. Letters*, vol. 73, no. 6, pp. 814–817, 1994.

[8] J. W. Goodman, *Introduction to Fourier optics*, McGraw-Hill, San Francisco, 1968.

[9] V. N. Mahajan, *Optical imaging and aberrations. Part II: Wave diffraction Optics*, SPIE Optical Engineering Press, Bellingham, Washington, USA., 2012.

[10] J. Portilla, A. Tristan-Vega, and I.W. Selesnick, "Efficient and robust image restoration using multiple-feature L2-relaxed sparse analysis priors," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5046–5059, Dec 2015, ®Matlab code publicly available at DOI: 10.13140/RG.2.1.3531.5280.

[11] Roger N. Clark, "Digital camera reviews and sensor performance summary," http://www.clarkvision.com/articles/digital.sensor.performance.summary/, Last Update: 2016-10-14.