1

Aberration balancing using an image-sharpness metric

ORESTIS KAZASIDIS^{1,*}, **SVEN VERPOORT**¹, **AND ULRICH WITTROCK**¹

¹ Photonics Laboratory, Münster University of Applied Sciences, Stegerwaldstrasse 39, 48565 Steinfurt, Germany * Corresponding author: kazasidis@fh-muenster.de

Compiled July 3, 2019

Image-sharpness metrics can be used to optimize optical systems and to control wavefront sensorless adaptive optics systems. We show that for an aberrated system the numerical value of an image-sharpness metric can be improved by adding specific aberrations. The optimum amplitudes of the additional aberrations depend on the power spectral density of the spatial frequencies of the object. © 2019 Optical Society of America

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

It is often desirable to state the performance of an optical system by a single number. This immediately allows ranking different optical systems, optimizing an optical system during its design, or finding the optimum state of an active or adaptive optics system. The wavefront variance σ^2 is such a single number, calculated from the wavefront W in the exit pupil: $\sigma^2 = \overline{W^2} - \overline{W}^2$. The wavefront variance is a useful measure for the optical performance for systems near the diffraction-limit, i.e., for wavefront aberration of less than about $\lambda/8$ RMS. For such systems, other performance metrics can be approximated from the wavefront variance. One example is the Strehl ratio *S* with the Maréchal approximation $S \approx (1 - 2\pi^2\sigma^2/\lambda^2)^2$ [1], where λ is the wavelength. Therefore, near the diffraction-limit maximizing the Strehl ratio is equivalent to minimizing the wavefront variance.

The measurement of the wavefront variance requires a wavefront sensor, and the measurement of the Strehl ratio requires a point object. Both requirements can be circumvented by using performance metrics directly calculated from the intensity distribution in the image plane $I(\mathbf{x})$, where \mathbf{x} is the position vector in the image plane.

Aberration balancing describes the deliberate addition of specific aberrations to an aberrated system in order to optimize a performance metric, commonly to minimize the wavefront variance [2]. In this paper we introduce the concept of aberration balancing for maximizing an image-sharpness metric. In Section 2 we employ Fourier optics to derive how the image-sharpness metric depends on the properties of the optical system and the object. In Sections 3 and 4 we use Zernike modes to explore the landscape of the image-sharpness metric for a point object and for an extended object. We show that for a severely aberrated system the metric can be improved by adding specific Zernike modes, although this increases the wavefront variance. Section 5 summarizes our results and discusses their implications, especially in controlling severely aberrated wavefront-sensorless adaptive and active optics systems.

2. THE IMAGE-SHARPNESS METRIC

The most common image-sharpness metric is defined as:

$$S_1 := \iint I(\mathbf{x})^2 d\mathbf{x} \,. \tag{1}$$

We found the first references to this metric in Schade [3] and Fellgett [4] for the evaluation of photographic images. We follow the notation of Muller and Buffington who first proposed to use S_1 as a feedback signal in astronomical telescopes to compensate for atmospheric seeing [5]. They used a modified Fresnel-Kirchhoff integral to prove that the global maximum of S_1 is obtained for zero aberration. Hamaker offered an alternative proof based on Fourier optics [6]. We first follow Hamaker's line of thought with a different notation. The optical transfer function *OTF* is the frequency response of an optical system:

$$\hat{I}(\mathbf{s}) = OTF(\mathbf{s})\,\hat{O}(\mathbf{s})\,,\tag{2}$$

where *I* and *O* are the intensity distributions of the image and of the object, respectively, and the symbol^denotes the Fourier transform. The vector **s** denotes the spatial frequency in the 2-D frequency domain. Applying Parseval's theorem to Eq. 1 and using Eq. 2 we get:

$$S_{1} \coloneqq \iint I(\mathbf{x})^{2} d\mathbf{x} \propto \iint |\hat{I}(\mathbf{s})|^{2} d\mathbf{s}$$
$$\propto \iint |OTF(\mathbf{s})\hat{O}(\mathbf{s})|^{2} d\mathbf{s} \propto \iint MTF(\mathbf{s})^{2} |\hat{O}(\mathbf{s})|^{2} d\mathbf{s} , \quad (3)$$

because the modulation transfer function *MTF* is equal to |OTF|. We use the proportionality symbol \propto to drop any constant multiplication factor. $|\hat{O}(\mathbf{s})|^2$ is the power spectral density of the spatial frequencies of the object. Hamaker showed that the *MTF* of an aberrated system for any spatial frequency is always smaller than its diffraction-limited value [6]. This can also be

proven by using Schwarz's inequality [7]. Thus, independently of the object, S_1 is maximum for a diffraction-limited system. In this paper we further investigate Eq. 3 and discuss two aspects: 1. what this equation implicates for an aberrated system, and 2. the dependence of S_1 on the power spectral density of the object.

The *MTF* is zero for frequencies higher than the cutoff frequency s_{cut} . Also the *MTF* is an even function symmetric about the origin [1]. Thus the integration may be restricted to the first quadrant of the frequency domain and Eq. 3 can be written as:

$$S_1 \propto \iint_0^{s_{\text{cut}}} MTF(\mathbf{s})^2 \left| \hat{O}(\mathbf{s}) \right|^2 d\mathbf{s} \,. \tag{4}$$

3. BALANCING AN ABERRATED SYSTEM

The power spectral density of a point object is simply $|\hat{O}(\mathbf{s})|^2 = 1$. The substitution in Eq. 4 yields:

$$S_{1,p} \propto \iint_{0}^{s_{\text{cut}}} MTF(\mathbf{s})^2 d\mathbf{s}$$
, (5)

where the subscript p (point) distinguishes it from Eq. 4. This equation can be expressed in words as:

When imaging a point object, the image-sharpness metric S_1 is proportional to the volume of the square of the *MTF* in the first quadrant of the frequency domain.

We consider a homogeneously illuminated circular pupil. The discussion and the results can be appropriately adapted for a different pupil. We express the wavefront in terms of the Zernike modes, using the notation of Wyant and Creath [8]. The Zernike modes are orthogonal with respect to the wavefront variance σ^2 , which means that adding a Zernike mode to the wavefront or increasing the amplitude of an existing Zernike mode will always increase σ^2 . As already mentioned, the Strehl ratio *S* can be calculated from the wavefront variance if the aberration is less than about $\lambda/8$ RMS. Therefore, *S* is also orthogonal with respect to the Zernike modes near the diffraction-limit.

The situation changes if the aberration is more than $\lambda/8$ RMS. The Maréchal approximation becomes invalid and the Strehl ratio can be multiple-valued for the same wavefront variance [9]. An increase of σ^2 may lead to an increase of *S* instead of a decrease. This means that if the aberration cannot be fully compensated, the designer or the operator of the optical system may be able to improve the Strehl ratio by inducing further aberration, although this increases the wavefront variance. This technique is called aberration balancing.

To demonstrate aberration balancing with Zernike modes using the image-sharpness metric S_1 , we simulate the wavefront for different Zernike modes and calculate the point spread function (*PSF*). From the *PSF*s we then calculate the merit function (*MF*) as a normalized version of S_1 :

$$MF := \frac{\iint I(\mathbf{x})^2 d\mathbf{x}}{\left(\iint I(\mathbf{x}) d\mathbf{x}\right)^2} \iint d\mathbf{x}.$$
 (6)

This is the definition of the structural density factor of an image by Linfoot [10]. The normalization and the simulation processes are discussed in the Appendices A and B, respectively. We show characteristic results in Fig. 1 as color rendering plots of the merit function. These plots are 2-D cuts in the multi-dimensional Zernike domain, with the color indicating the magnitude of *MF*.

The global maximum of the merit function is always obtained for zero aberration. We discussed in [11] that circular or elliptical



Fig. 1. (Color online) Color rendering plots of the merit function for pairs of Zernike modes when imaging a point object. The white contour lines near the global maximum are always circular or elliptical. The black contour lines for large aberration are circular or elliptical only in (a). (a) The pair Z_4/Z_5 . For a constant amplitude of Z_5 , *MF* has a maximum always for $Z_4 = 0$ (e.g., the X mark for $Z_5 = 0.7\lambda$), and vice versa. (b) The pair Z_3/Z_4 . When $|Z_4| \ge 0.4\lambda$, *MF* has two maxima for opposite amplitudes of Z_3 (e.g., the two X marks for $Z_4 = 0.7\lambda$). (c) The pair Z_4/Z_{11} . When $|Z_{11}| \ge 0.25\lambda$, *MF* has a maximum for a non-zero amplitude of Z_4 (e.g., the X mark for $Z_{11} = 0.7\lambda$).

contour lines in a 2-D plot of the merit function mean that the two Zernike modes influence the merit function independently. For the pair of primary astigmatisms Z_4 and Z_5 in Fig. 1a, this is true for the whole amplitude range. This is also true for every pair of Zernike modes of the same radial order and opposite azimuthal order, e.g., the two primary comas Z_6/Z_7 , and the two primary trefoils Z_9/Z_{10} (not shown).

For other Zernike pairs, the contour lines may be circular or elliptical only near the global maximum. The color rendering plots in Figs. 1b and c show that for the pairs Z_3/Z_4 and Z_4/Z_{11} of sufficiently large amplitudes, the amplitude of one Zernike mode that delivers the maximum merit function depends on the amplitude of the other Zernike mode. For the pair of defocus (Z_3) and astigmatism 0° (Z_4) in Fig. 1b, we detect an X-shaped ridge. When $|Z_4| \ge 0.4\lambda$, there exist two local maxima of the merit function which are obtained for opposite amplitudes of Z_3 . These two defocus Z_3 amplitudes correspond to the two planes of tangential and sagittal focus. Also for the pair of primary astigmatism 0° (Z_4) and secondary astigmatism 0° (Z_{11}) in Fig. 1c, the maximum MF depends on the amplitudes of both Zernike modes. When $|Z_{11}| \ge 0.25\lambda$, the maximum merit function value is obtained for a non-zero amplitude of Z_4 .

We also calculated the *MTF* from the *PSF*. The subsequent calculation of the volume of the square of the *MTF* in the first quadrant of the frequency domain (not shown) confirmed the results shown in Fig. 1 obtained with *MF*, as expected from Eq.

5. Further discussion involving the wavefront, the *PSF*, and the *MTF* can be found in [12].

4. OBJECT-DEPENDENT ABERRATION BALANCING

Equation 4 is the key to understanding aberration balancing and it can be expressed in words as:

The image-sharpness metric S_1 weightens the square of the *MTF* by the power spectral density of the object, for each spatial frequency. Maximizing S_1 , and consequently maximizing the merit function *MF* of Eq. 6, is equivalent to maximizing the weighted volume of $MTF^2 |\hat{O}|^2$ in the first quadrant of the frequency domain.

The power spectral density of an extended object is different from that of a point object. This leads to object-dependent aberration balancing. For example, if the object has low power spectral density in a particular range of spatial frequencies, the *MTF* may be low in this frequency range as well, without severely impacting the merit function. To demonstrate object-dependent aberration balancing, we create a synthetic extended object, shown in Fig. 4a of Appendix B, that comprises two bodies, which could be a planet and its moon. An extended object can be regarded as the superposition of many point objects. Therefore its image is the superposition of the images of all point objects. We simulate an aberrated image as the 2-D convolution of the planet-moon object with an aberrated *PSF*. The generation of the planetmoon object and the image simulation are further described in Appendix B.

Figure 2a shows the color rendering plot of the merit function for the Zernike modes of primary astigmatism 0° (Z_4) and secondary astigmatism 0° (Z_{11}) when imaging the planet-moon object. This plot differs from Fig. 1c, because the power spectral density of the planet-moon object differs from that of a point object. The merit function when $Z_{11} = 0.7\lambda$ is plotted in Fig. 2b (black dashed line). It corresponds to the black dashed line in Fig. 2a. The maximum *MF* is obtained for $Z_4 = 0.6\lambda$. For comparison, we also plot *MF* when imaging a point object (red dotted line in Fig. 2b, that corresponds to the black dashed line in Fig. 1c), which has a maximum for $Z_4 = 0.7\lambda$. This clearly demonstrates that, when $Z_{11} = 0.7\lambda$, the optimum amplitude of Z_4 depends on the object.

In Fig. 3 we simulate two aberrated images of the planetmoon object when $Z_{11} = 0.7\lambda$. In Fig. 3a, $Z_4 = 0$, which corresponds to the best (smallest) wavefront variance. In Fig. 3b, $Z_4 = 0.6\lambda$, which corresponds to worse (higher) wavefront variance but to the best (maximum) merit function for uncompensated $Z_{11} = 0.7\lambda$ (the X mark of the black dashed line in Fig. 2b). Adding primary astigmatism 0° (Z_4) in the presence of secondary astigmatism 0° (Z_{11}) leads to a sharper image. In Fig. 3a the planet and the moon are indistinguishable. In Fig. 3b the moon is distinguished from the planet as a small blob. The moon is better identified as a secondary peak in Fig. 3c which shows the pixel profiles along the red lines of Figs. 3a and b. In addition, structural characteristics of the planet become recognizable in Fig. 3b. These characteristics are clearer in the diffraction-limited image (Fig. 4b).

Maximizing the MTF for $0.1 s_{cut}$, which corresponds to the separation between planet and moon, would lead to a worse (smaller) merit function but an even sharper moon peak. However, this would not be appropriate because imaging is about finding something that is not known a priori. By maximizing



Fig. 2. (Color online) (a) Color rendering plot of the merit function for the pair Z_4/Z_{11} , when imaging the planet-moon object of Fig. 4a. The white contour lines near the global maximum are elliptical. For large aberration the black contour lines are no longer elliptical. (b) The black dashed line plotted against the left y-axis is a cut through the black dashed line in (a) when $Z_{11} = 0.7\lambda$. The red dotted line plotted against the right y-axis corresponds to *MF* when imaging a point object and $Z_{11} = 0.7\lambda$. It is a cut through the plot in Fig. 1c. The maxima of the two plots (the X marks) are obtained for different amplitudes of Z_4 .



Fig. 3. Simulated images with 0.7λ of secondary astigmatism 0° (Z_{11}). (a) The amplitudes of all the other Zernike modes are zero. The moon is hidden in the halo of the planet. (b) Adding 0.6λ of primary astigmatism 0° (Z_4) the moon is distinguished from the planet as a small blob. (c) Pixel profiles along the lines of (a) and (b). The moon appears as a secondary intensity peak for the dotted line that corresponds to (b). The image becomes sharper, although the wavefront variance increases by adding Z_4 .

MF, the moon detection is possible without the pre-knowledge of its existence.

5. CONCLUSIONS

We have demonstrated aberration balancing with Zernike modes using an image-sharpness metric defined as merit function in Eq. 6. To date, the term "aberration balancing" has been used to describe the minimization of the wavefront variance by deliberately adding one or more Seidel modes to an aberrated system. Unlike Zernike modes, Seidel modes are not balanced with respect to the wavefront variance. The Zernike modes are orthogonal with respect to our merit function only near the diffraction limit, i.e., for wavefront aberration of less than about $\lambda/8$ RMS. We have shown that for large aberration the merit function may be improved by adding specific Zernike modes. We also call this effect "aberration balancing", but instead of the wavefront variance we link it to the maximization of the image-sharpness metric. An experimental demonstration was presented in [11] for balancing spherical aberration with defocus.

To the best of our knowledge, this is the first study exploring the landscape of an image-sharpness metric away from the global optimum, i.e., in a severely aberrated system. When imaging a point object, the merit function is proportional to the volume of the square of the *MTF* in the first quadrant of the frequency domain, because the power spectral density of a point object is a constant function. When imaging an extended object, the square of the *MTF* is weighted by the power spectral density of the object, for each spatial frequency. Therefore, the optimum amplitudes of the additional Zernike modes depend on the power spectral density of the object, which leads to object-dependent aberration balancing.

Object-dependent aberration balancing is of interest for severely aberrated adaptive and active optics systems, where low-order Zernike modes can partially balance uncorrectable high-order Zernike modes, with amplitudes depending on the object. Using this technique, previously unrecognizable characteristics of the object can become detectable. In the simplified example of Fig. 3, inspired by astronomy, a moon that was hidden in the halo of its planet comes into sight.

The image-sharpness metric is calculated from the intensity distribution in the image plane and therefore no additional sensor is required. This also means that this control method is free from non-common path errors which can affect systems that rely on wavefront sensors.

Our results can be used to immediately improve the performance of wavefront-sensorless adaptive and active optics systems that use the Zernike modes to control the wavefront corrector and have to correct large aberration, i.e., more than $\lambda/8$ RMS. In the simplest case, these systems sequentially maximize the merit function in separate Zernike modes, e.g., in [13]. Figure 1 indicates that in the case of large aberration a single iteration of the Zernike modes is not sufficient to reach the global maximum of this image-sharpness metric. Therefore, the algorithm should be adapted to account for the merit function landscape.

Using this merit function is equivalent to maximizing the power spectral density of the image. Aberration balancing can be also performed using Fourier-based metrics, which allow noise filtering and contrast maximization for a particular range of spatial frequencies.

Spurious resolution, appearing as contrast reversal, leads to decreased *MTF* for a wide range of spatial frequencies, because the *MTF* is a continuous function. Our merit function cannot

distinguish between positive and negative *MTF*. Therefore, contrast reversal may lead to a better merit function value. However, this can happen only for objects with unnaturally sharp and narrow peaks in their power spectral density.

APPENDIX A: METRIC NORMALIZATION

We define our merit function in Eq. 6 by dividing S_1 of Eq. 1 by the square of the total energy in the image plane, and by multiplying by the area of the image plane. Conservation of energy ensures that the total energy in the image plane remains constant independently of the aberration. This is valid for an infinitely large detector with infinite dynamic range. Due to the finite detector area, the total energy is not constant. It is therefore appropriate to use its square as normalization factor. This normalization can also account for intensity variations and sensitivity fluctuations.

The multiplication by the area of the image plane renders the merit function a dimensionless quantity. After the normalization, $MF \ge 1$. MF is minimum (MF = 1) if the intensity distribution is constant, which means that there is no contrast.

APPENDIX B: PSF AND IMAGE SIMULATION

We simulate the wavefront in the exit pupil of an isoplanatic imaging system with a homogeneously illuminated circular pupil in MATLAB over a 300 pixel \times 300 pixel grid. Phase aberration is expressed as Zernike modes, with 0.05λ step size for their amplitudes. The PSF is calculated by applying the 2-D fast Fourier transform to the exit pupil function. We choose the width of the diffraction-limited PSF to be 40 pixels, significantly larger than the required width of 5 pixels according to the Nyquist sampling theorem. Thus, the simulations are free from sampling artifacts. We limit the total grid of the PSF to 1201 pixel \times 1201 pixel, 30 times larger than the diffractionlimited *PSF*. The error caused by this truncation is negligible. The aberrated *PSFs* are normalized to the maximum of the diffraction-limited PSF, to allow comparison among PSFs with different aberration. Each pixel is sampled with 16-bit depth, which matches the analog-to-digital converter used in the Near Infrared Camera of the upcoming James Webb Space Telescope [14].

The planet-moon object is shown in Fig. 4a. It is a synthesis of two bodies: of an image of the Earth taken by the Earth Polychromatic Imaging Camera on board NASA's Deep Space Climate Observatory satellite [15] and of an image of the Moon taken by the cameras aboard NASA's Lunar Reconnaissance Orbiter spacecraft [16]. The images were converted to grayscale to simulate monochromatic imaging. The separation between the centers of the bodies is 80 pixels, which corresponds to 0.1 of the cutoff frequency. Their diameter ratio is about 3.7, equal to the Earth-Moon diameter ratio. The ratio of the mean brightnesses of the bodies is 3.6, equal to the ratio of the geometric albedos of the Earth and of the Moon [17]. The size of the planetmoon object is 300 pixel \times 300 pixel. For the image simulation we calculate the convolution of the zero-padded planet-moon object with a PSF. The aberrated images are normalized to the maximum of the diffraction-limited image, to allow comparison among images with different aberration. Figure 4b shows the diffraction-limited image, where some structural characteristics of the planet are apparent. Figure 4c shows the pixel profiles along the red lines of Figs. 4a and b.



Fig. 4. (a) The synthetic extended object, comprising a planet and its moon. (b) The simulated diffraction-limited image. The contrast of both images in (a) and (b) has been adjusted with $\gamma = 0.5$ for illustration purposes only. (c) Pixel profiles along the lines of (a) and (b). The high-frequency spatial information of the planet-moon object (solid line) is not transmitted by the system. Nevertheless, the planet and the moon are clearly distinguishable in the diffraction-limited image (dotted line).

The discrete version of the merit function defined in Eq. 6 is:

$$MF_{\text{discrete}} \coloneqq \frac{\sum \sum I^2}{\left(\sum \sum I\right)^2} N_x N_y$$

where $N_x N_y$ is the image area in pixel². The summation occurs over 1201 pixel × 1201 pixel for the point object and over 300 pixel × 300 pixel for the planet-moon object.

FUNDING.

This work is financially supported by the programme "Qualifizierungsstelle" of Münster University of Applied Sciences.

ACKNOWLEDGMENT.

We thank Michel Verhaegen, Gleb Vdovin, and Oleg Soloviev for discussion and support.

REFERENCES

- 1. C. S. Williams and O. A. Becklund, *Introduction to the Optical Transfer Function* (SPIE Press, Bellingham, Wash., 2002).
- V. N. Mahajan, Aberration Theory Made Simple (SPIE Press, Bellingham, Washington USA, 2011), 2nd ed.
- O. H. Schade, "Image Gradation, Graininess and Sharpness in Television and Motion Picture Systems: Part II: The Grain Structure of Motion Picture Images—An Analysis of Deviations and Fluctuations of the Sample Number," J. Soc. Motion Pict. Telev. Eng. 58, 181–222 (1952).
- 4. P. Fellgett, "Concerning Photographic Grain, Signal-to-Noise Ratio, and Information," J. Opt. Soc. Am. **43**, 271–282 (1953).
- R. A. Muller and A. Buffington, "Real-time correction of atmospherically degraded telescope images through image sharpening," J. Opt. Soc. Am. 64, 1200–1210 (1974).
- J. P. Hamaker, J. D. O'Sullivan, and J. E. Noordam, "Image sharpness, Fourier optics, and redundant-spacing interferometry," J. Opt. Soc. Am. 67, 1122–1123 (1977).

5

- 7. J. W. Goodman, Introduction to Fourier Optics (McGraw-Hill, 1996).
- J. C. Wyant and K. Creath, "Basic wavefront aberration theory for optical metrology," in *Applied Optics and Optical Engineering*, vol. 11 R. R. Shannon and J. C. Wyant, eds. (Academic Press, Inc., New York, 1992), chap. 1.
- W. B. King, "Dependence of the strehl ratio on the magnitude of the variance of the wave aberration," J. Opt. Soc. Am. 58, 655–661 (1968).
- 10. E. H. Linfoot, *Fourier methods in optical image evaluation* (The Focal Press, London and New York, 1964).
- O. Kazasidis, S. Verpoort, and U. Wittrock, "Algorithm design for imagebased wavefront control without wavefront sensing," Proc. SPIE 10695, 1069502 (2018).
- O. Kazasidis, S. Verpoort, and U. Wittrock, "Image-based wavefront correction for space telescopes," presented at International Conference on Space Optics 2018, proceedings paper to be published in Proc. SPIE (2019).
- S. Bonora and R. J. Zawadzki, "Wavefront sensorless modal deformable mirror correction in adaptive optics: optical coherence tomography," Opt. Lett. 38, 4801–4804 (2013).
- James Webb Space Telescope User Documentation, "NIRCam Detectors," https://jwst-docs.stsci.edu/display/JTI/NIRCam+Detectors.
- NASA, "NASA Captures "EPIC" Earth Image," https://www.nasa.gov/ image-feature/nasa-captures-epic-earth-image.
- NASA, "The Near Side of the Moon," https://moon.nasa.gov/resources/ 77/the-near-side-of-the-moon/.
- NASA, "Moon Fact Sheet," https://nssdc.gsfc.nasa.gov/planetary/ factsheet/moonfact.html.