Wavefront sensing with critical sampling

Rafael Navarro,[1,*] Justo Arines,[2] and Ricardo Rivera[1]

[1]*ICMA, Universidad de Zaragoza and Consejo Superior de Investigaciones Científicas, Facultad de Ciencias, P. Cerbuna 12, 50009 Zaragoza, Spain*

[2]*Departamento de Física Aplicada, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain*

**Corresponding author: rafaelnb@unizar.es*

Received October 8, 2010; revised December 21, 2010; accepted December 21, 2010; posted January 6, 2011 (Doc. ID 136145); published January 31, 2011

Different types of nonredundant sampling patterns are shown to guarantee completeness of the basis formed by the sampled partial derivatives of Zernike polynomials, commonly used to reconstruct the wavefront from its slopes (wavefront sensing). In the ideal noise-free case, this enables one to recover double the number of modes J than sampling points I (critical sampling $J = 2I$). With real data, noise amplification makes the optimal number of modes lower $I < J < 2I$. Our computer simulations show that optimized nonredundant sampling provides a significant improvement of wavefront reconstructions, with the number of modes recovered about 2.5 higher than with standard sampling patterns.

OCIS codes: 010.7350, 080.1005, 220.1010.

Wavefront sensing is a highly successful method for obtaining the wavefront aberration of an optical system from the measurement of wavefront slopes [1,2] [i.e., transverse aberrations (x'_i, y'_i)] at some array of sampling points (ρ_i, θ_i), $i = 1, \ldots I$ at the pupil plane. The main applications are atmospheric optics and astronomy [3], the human eye [4], and optical testing [5]. The measurements are proportional to the gradient of the wavefront $(x'_i, y'_i) = f'/R_{\text{pup}} \nabla W(\rho_i, \theta_i)$; R_{pup} is the pupil (or wavefront) radius and f' is the focal length of the lens (or microlens array) of the measuring instrument [5]. To recover the wavefront W, one has to integrate the gradient using some expansion of W in terms of some derivable basis functions. For circular pupils, Zernike polynomials (ZPs) form a complete orthogonal basis (on a circle of unit radius) so that the wavefront can be represented as $W(\rho, \theta) = \sum_j^J c^j Z^j(\rho, \theta)$, where c^j are the coefficients and Z^j are the ZPs. Here, the radial order n and angular frequency m are merged into a single index $j = (n(n+2)+m)/2$ (ANSI Z80.28 standard). We can substitute the expansion of W for the discrete set of measurements:

$$m_i = R_{\text{pup}}/f' \begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \sum_j^J c^j \begin{pmatrix} Z'^j_{Xi} \\ Z'^j_{Yi} \end{pmatrix},$$

or in matrix-vector notation:

$$\mathbf{m} = \begin{pmatrix} \mathbf{Z'_X} \\ \mathbf{Z'}_Y \end{pmatrix} \mathbf{c} = \mathbf{D}\mathbf{c}, \tag{1}$$

where Z'^j_{Xi}, Z'^j_{Yi} are the partial derivatives of the jth ZP at point i. Then, the solution for the wavefront coefficients is $\mathbf{c} = \mathbf{D}^{-1}\mathbf{m}$. There are two possibilities to invert $\mathbf{D}$. The usual strategy is to apply a strong oversampling, $J \ll I$, and then compute the least-squares solution (i.e., the pseudoinverse) so that the coefficients are estimated as $\hat{\mathbf{c}} = (\mathbf{D^T D})^{-1}\mathbf{D^T m}$. The opposite strategy would be to apply critical sampling, $J = 2I$, so that $\mathbf{D}$ is square, but then the problem is that $\mathbf{D}^{-1}$ exists only if $\text{Det}(D) \neq 0$ ($\mathbf{D}$ nonsingular), which is the condition for completeness. If that happens, then Eq. (1) can be solved exactly and the computation is exact and invertible. However, matrix inversion is numerically unstable, which could have the undesirable effect of noise amplification. For this reason, orthogonality is a highly desirable additional condition, which guarantees a trivial computation of the inverse as matrix transpose. When orthogonality is not feasible, the condition number (CN) should be as close as possible to 1 to avoid excessive noise amplification. The problem is that sampled ZPs and sampled ZP derivatives are neither complete nor orthogonal for the usual sampling patterns (square, hexagonal, polar, etc. [6]). Previous studies [7] showed that some inhomogeneous sampling patterns (nodal points of Albrecht cubature) can guarantee orthogonality in the wavefront reconstruction. Recently [8], we found different types of nonredundant sampling patterns (where $\theta_i \neq \theta_k$ and $\rho_i \neq \rho_k$ $\forall i, k$) that guarantee completeness of the discrete ZP basis. In this Letter, the goal was to study whether those nonredundant patterns also guarantee completeness for the ZP derivatives so that $\text{Det}(\mathbf{D}) \neq 0$. This means that, ideally, it would be possible to construct an invertible square matrix $\mathbf{D}$ by applying critical sampling (i.e., $J = 2I$). This would permit us to pass from the wavefront coefficients to the slope measurements $\mathbf{m} = \mathbf{D}\mathbf{c}$ and vice versa $\mathbf{c} = \mathbf{D}^{-1}\mathbf{m}$. This implies important benefits: attaining the critical sampling $J = 2I$ limit means to retrieve double the number of coefficients (Zernike modes) than sampling points, exact recovery of the measurements and exact solution for the coefficients (in the ideal case of zero noise). As far as we know, critical sampling, or even the case $J > I$ was not attempted before, basically because, with standard sampling grids, $\text{Det}(\mathbf{D}) = 0$ and $\text{Rank}(\mathbf{D}) < I$.

We compared five different sampling patterns with a fixed number of $I = 91$ points ($n = 12$): standard hexagonal (H), perturbed hexagonal (PH), random (R), and two types of spirals (S1 and S2). The details of how to build these nonredundant patterns are given in [8]. The P is obtained by adding a small random shift to the coordinates of the sampling points of the order of $10^{-3}R_{\text{pup}}$. Random sampling was proposed previously [9]. In both spirals, the sampling of the angular coordinate is regular and homogeneous $\theta_i = \theta_{\max}/(I-1)i$; for S1,

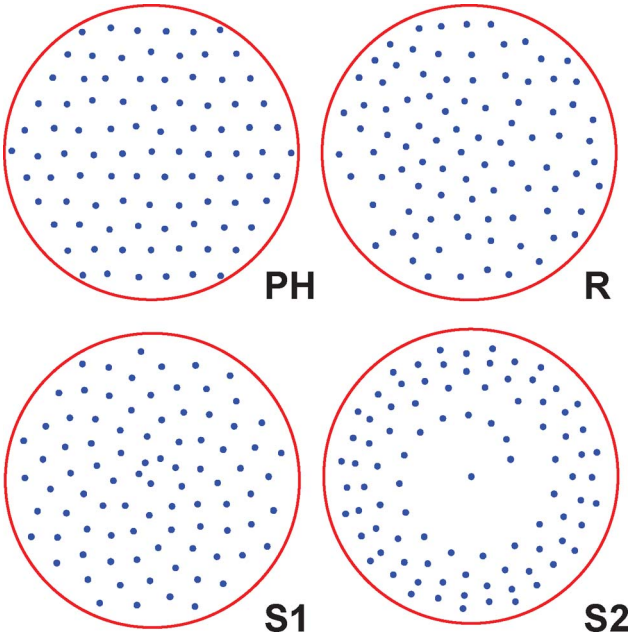

Fig. 1. (Color online) Sampling patterns PH, R, S1, and S2 with $I = 91$.

$\rho_i = \rho_{\max}\sqrt{\theta_i/\theta_{\max}}$, which provides approximately homogenous density of samples across the pupil, whereas, for S2, $\rho_i = \rho_{\max}\sqrt[4]{\theta_i/\theta_{\max}}$ and, hence, the density increases toward the periphery. These sampling patterns are compared in Fig. 1. The reason for using the S2 sampling pattern is that it provides increasing density toward the periphery, in a similar fashion as other sampling patterns proposed to guarantee orthogonality of discrete ZPs [7]. Nevertheless, we compared different spirals where ρ was proportional to θ, $\theta^{1/2}$, $\theta^{1/3}$, $\theta^{1/4}$, etc., finding that $1/4$ provided the best CN to invert $\mathbf{D}$.

The first crucial result is that the same sampling patterns, which guarantee completeness of the ZPs [8], namely, R, PH, S1, and S2 also guarantee completeness: the two sampled partial derivatives of ZPs form a complete basis for the set of measurements m. For the particular case of $I = 91$ and critical sampling, then $J = 182$ and $\mathbf{D}$ is a 182×182 square matrix. The rank was maximum 182 for all nonredundant samplings, but it was much lower $89 < I$ for the hexagonal (redundant) one. Therefore, it is possible to perform wavefront sensing with critical sampling to recover $2J$ Zernike modes of the wavefront.

In the ideal noise-free case, the reconstruction will be exact, but the lack of orthogonality could affect the matrix inversion, which could amplify the noise of real measurements. We studied the problem of potential noise amplification in two ways. First, we obtained the singular value decomposition of matrix $\mathbf{D}$ as a metric to predict the noise amplification. In particular, the CN that is the ratio between the maximum and minimum singular values is a good metric for the numerical instability associated with the matrix inversion and also gives a rough estimation of the expected noise amplification. The optimum value is 1, but, for large nondiagonal matrices, this usually requires orthogonality. The CNs obtained for the square 182×182 matrices improve progressively: ∞ for H, 4.3×10^7 for PH, 1.6×10^7 for S1, 4×10^6 for R, and 1.710^5 for S2. To have a more realistic estimation of the performance, including the effects of noise amplification, we conducted a series of computer simulations. The coefficients (up to 182) of a wavefront were generated randomly according to the statistics (mean and variance) found in a population of human eyes [10]. Coefficients for higher orders were assumed to be zero, which is likely to happen in normal eyes. In this way, we consider only the effect of noise, while we avoid possible cross coupling and aliasing [11] owing to potential spectral overlapping. This synthetic wavefront has an rms value of 0.54 μm ($\sim 1\lambda$). For each condition, 30 different measurements were simulated using the expression $\mathbf{m}_k = \mathbf{D}\mathbf{c} + \mathbf{n}_k$ for the kth realization, where $\mathbf{n}_k$ is (Gaussian zero mean) random noise. Then, we computed the mean and standard deviation (error bars in figures). The noise variance was adjusted to simulate different levels of signal-to-noise ratio (SNR) from 1 to ∞ (zero noise). The results for the different sampling patterns are plotted in Fig. 2 for the case of SNR $= 30$, computed as the ratio between the rms values of measurements and noise, respectively. The SNR is within the range of typical values in ocular aberrometers [12].

The vertical axis corresponds to the rms difference between the original (ideal) wavefront and that reconstructed from the noisy measurements. The horizontal axis represents the number of modes J considered in the matrix $\mathbf{D}$, which means that, for the reconstructions, the expression $\mathbf{c} = \mathbf{D}^{-1}\mathbf{m}$ applies only to the last point $J = 182$. For the other points ($J < 182$), the matrix is not square, $182 \times J$, and then the coefficients were computed using the pseudoinverse $\hat{\mathbf{c}} = (\mathbf{D}^{\mathbf{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathbf{T}}\mathbf{m}$. In fact, we applied the QR factorization (where Q is orthogonal and R is triangular) to the pseudoinverse to improve its CN (typically by a factor of 2.) As we can see, all the sampling patterns show a similar performance for $J \leq 62$, but, for $J > 62$, the noise amplification increases rapidly for the redundant H pattern. For the nonredundant sampling patterns (R, PH, S1, and S2), the effect of noise amplification becomes patent for higher values of J; as J increases, S2 shows the best behavior. For this sampling pattern, and SNR $= 30$, the optimal performance is obtained for $J \approx 122$, significantly greater than $I = 91$.

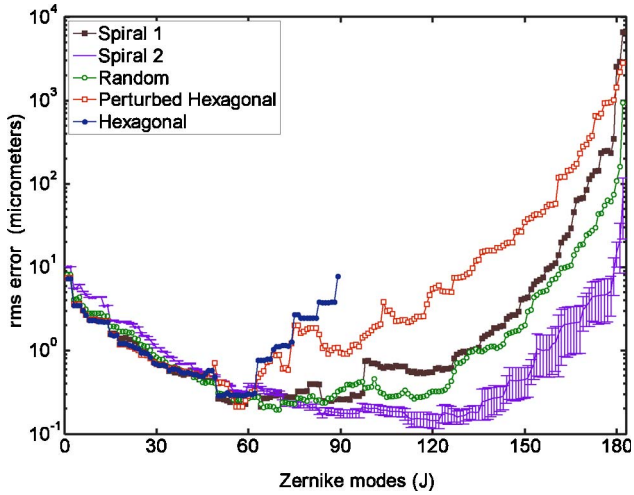

Fig. 2. (Color online) Root-mean-square error of the reconstructed wavefront for different sampling patterns.

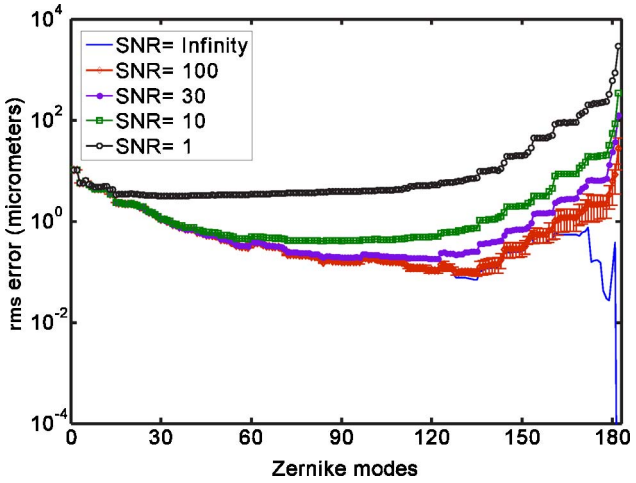

Fig. 3. (Color online) Root-mean-square error of the reconstructed wavefront for spiral S2 sampling and for different SNRs of the input measurements.

Figure 3 compares the results for SNR = 1, 10, 30, 100, and ∞ (noise-free case) for the best sampling pattern S2. As expected, for SNR = ∞, the wavefront reconstruction is perfect for $J = 2I$; otherwise, noise amplification is present and very clear for $J > 140$. Beyond that point, all curves (except ∞) are roughly parallel. The main difference is the minimum value, which is clearly dependent on the SNR of the measurements. The position of the absolute minimum moves to the right as the SNR increases, but curves (especially those corresponding to realistic values of SNR, 10 or 30) show a rather flat valley, indicating that the optimal value of the number of modes J is not critical.

In summary, standard (redundant) patterns, such as H, exhibit a more limited performance when compared to the nonredundant ones. The absolute limit for the number of modes reconstructed in the ideal case (SNR = ∞) is equal to Rank(**D**), which is $86 < I$ for H, whereas all proposed nonredundant sampling patterns give perfect reconstruction for critical sampling ($J = 2I$). For the case of SNR = 30, the best reconstruction provided by the H pattern yields a 0.29 μm rms error for a number of modes $J = 50$ (55% of sampling points). For this SNR, the best reconstruction is obtained for the spiral S2, with a 0.145 μm rms error for $J = 122$ (134% of sampling points). That optimal value of J corresponds to a trade off between spectral subsampling (increase of rms error when J decreases) and noise amplification (when J increases). These computer simulations are limited to a fixed number of sampling points ($I = 91$), but the conclusions are more general. For a lower number of sampling points, the CN (associated to the inversion of large matrices) improves [8], so the expected noise amplification is lower. For very large values of I (sampling points), other basis functions (such as Fourier series) could be more appropriate than ZPs in practice.

To our knowledge, this is the first empirical demonstration of wavefront reconstruction with critical sampling. This permits the invertible computation of the measurements **m** from the wavefront W and vice versa, which could be important in different applications, such as optimization in optical design, inverse problems in optics, iterative computations, etc. In the presence of noise (wavefront sensing), nonredundant patterns enable one to recover more modes than sampling points ($J > I$) and hence an improved wavefront reconstruction. Two conditions are necessary for that: the sampling pattern must be nonredundant to guarantee completeness of the basis of sampled ZPs, and, to optimize the CN of **D**, the density of sampling must increase toward the periphery, which is consistent with sampling grids proposed previously [7] to guarantee orthogonality. Regarding practical applications, these inhomogeneous spiral sampling patterns are difficult to implement in conventional monolithic microlens arrays used in Hartmann–Shack sensors. However, highly flexible and reconfigurable (almost in real time) sampling grids can be easily implemented using laser ray tracing [5] or liquid-crystal spatial light modulators [13].

This work was supported by the Commisión Interministerial de Ciencia y Tecnología (CICyT), Spain, grant FIS2008-00697. R. Rivera was supported by the European Union Program of High Level Scholarships for Latin America, Alban, E07D402088CL. J. Arines was supported by the program Isidro Parga Pondal (Xunta de Galicia 2009).

References

1. R. J. Noll, J. Opt. Soc. Am. **68**, 139 (1978).
2. R. Cubalchini, J. Opt. Soc. Am. **69**, 972 (1979).
3. E. E. Silbaugh and B. M. Welsh, J. Opt. Soc. Am. A **13**, 2453 (1996).
4. J. Liang, B. Grimm, S. Goelz, and J. F. Bille, J. Opt. Soc. Am. A **11**, 1949 (1994).
5. R. Navarro and E. Moreno-Barriuso, Opt. Lett. **24**, 951 (1999).
6. V. Voitsekhovich, L. Sanchez, V. Orlov, and S. Cuevas, Appl. Opt. **40**, 1299 (2001).
7. S. Rios, E. Acosta, and S. Bara, Opt. Commun. **133**, 443 (1997).
8. R. Navarro, J. Arines, and R. Rivera, Opt. Express **17**, 24269 (2009).
9. O. Soloviev and G. Vdovin, Opt. Express **13**, 9570 (2005).
10. L. N. Thibos, X. Hong, A. Bradley, and X. Cheng, J. Opt. Soc. Am. A **19**, 2329 (2002).
11. J. Herrmann, J. Opt. Soc. Am. **71**, 989 (1981).
12. P. Rodriguez, R. Navarro, J. Arines, and S. Bará, J. Refract. Surg. **22**, 275 (2006).
13. J. Arines, V. Durán, Z. Jaroszewicz, J. Ares, E. Tajahuerce, P. Prado, J. Lancis, S. Bará, and V. Climent, Opt. Express **15**, 15287 (2007).