applied optics

Modal wavefront reconstruction from its gradient

IACOPO MOCHI* AND KENNETH A. GOLDBERG

Lawrence Berkeley National Laboratory, Berkeley, California 90720, USA
**Corresponding author: iacopo.mochi@imec.be*

Received 4 February 2015; revised 24 March 2015; accepted 24 March 2015; posted 25 March 2015 (Doc. ID 233363); published 17 April 2015

Several wavefront sensing techniques provide direct or indirect measurements of the wavefront error gradient, for example the Shack–Hartmann sensor, the Foucault knife-edge test, shearing interferometry, and many others. We developed and tested a noniterative method to reconstruct the wavefront error from its gradient. The method is based on the projection of the measured gradients onto a basis derived from multiple directional derivatives that have been combined into an intermediate set of orthogonal functions. To reduce errors that arise from linear approximations, the intermediate functions can be calculated with parameters that match the known experimental conditions. This method can be implemented using any convenient set of smooth polynomials defined on a two-dimensional domain, and it is not computationally intensive. In this paper we describe the method in detail, provide an example of a possible implementation, and discuss the effect that random noise in the measured gradient has on the reconstruction.

OCIS codes: (110.7348) Wavefront encoding; (120.3180) Interferometry; (010.7350) Wave-front sensing; (220.4840) Testing; (120.5050) Phase measurement.

http://dx.doi.org/10.1364/AO.54.003780

1. INTRODUCTION

Several wavefront detection techniques are based on measurements of the wavefront gradient, relying on reconstruction algorithms to calculate the actual wavefront. This is the case for the lateral shearing interferometer [1], the Hartmann and Shack–Hartmann sensors [2], the pyramid sensor [3], and the quantitative knife-edge test [4]. All these methods provide a direct or indirect measurement of the wavefront gradient and require a reconstruction algorithm to recover the original wavefront. In some cases, like the Hartmann and Shack–Hartmann tests, measurements represent a discretely sampled version of the continuous wavefront derivative, while in shearing interferometry, measurement points may be quasi-continuous, but the underlying mechanism compares physically shifted copies of the original wavefront. Analysis in the shearing case is often performed using a linear approximation of the local derivatives, potentially missing higher-order information. Most deterministic wavefront reconstruction algorithms can be classified as modal or zonal methods. Zonal algorithms [4–6] minimize reconstruction errors locally, limiting error propagation but making them more susceptible to noise [7]. Modal algorithms fit measurement data to global shape functions, such as orthogonal polynomials, and are commonly based on Zernike or Fourier polynomials [8–12]. Fitting globally provides reduced susceptibility to noise and random errors and can deliver results that are inherently connected to physical parameters (i.e., alignment degrees of freedom) or aberration terms (defocus, astigmatism, coma, etc.) We developed a modal, noise-robust algorithm that can be used with any set of independent two-dimensional (2D) polynomials to reconstruct a wavefront from its gradient on arbitrary pupil shapes. The reconstruction is obtained by projecting the measured wavefront derivatives onto an orthonormal basis derived from the empirically calculated derivatives of the chosen polynomial set. In this method, orthogonal derivatives (e.g., ∇_x and ∇_y) are grouped together in pairs, forming a single element, reminiscent of a domino tile. In a conventional way, projection coefficients are used to calculate the original wavefront by means of a transformation matrix that maps the derivatives back to the original polynomials. One advantage of this approach is that it does not require specific assumptions or approximations of the wavefront derivatives but can be used with the same kind of difference functions (discrete or continuous) generated by the wavefront sensing method. It therefore becomes adaptable to the specific details of the measurement, including the native coordinate grid of the detector array. In the case of shearing interferometry, for example, one would use a polynomial derivative basis generated with the same discrete shear displacement used physically in the wavefront measurement. This approach avoids the potential pitfalls of point-by-point linear approximation of the derivatives, ensuring higher accuracy in the reconstruction.

2. DESCRIPTION OF THE METHOD

The first step is to choose a suitable set of 2D orthogonal polynomials that can effectively represent the wavefront error that we want to reconstruct. Let us consider a 2D measurement domain R and a basis set of orthogonal polynomials $\{Z_n\}$. Let W be the wavefront error to reconstruct and a_n the coefficients of the expansion of W on the polynomial basis $\{Z_n\}$, so that

$$W = \sum_{n=0}^{N} a_n Z_n. \tag{1}$$

The measured gradient of W is expressed as the two orthogonal components (S_x, S_y):

$$(S_x, S_y) = \vec{\nabla} W. \tag{2}$$

The two orthogonal directions x and y are taken to be parallel to the derivatives. Here, the derivative itself may be either continuous or discrete, calculated to match the physical details of the measurement. For example, in the case of shearing interferometry, the interfering wavefronts may have an inherent shear distance that could be equal to some number of pixels in the discrete measurement domain. In that case, the discrete derivatives should be calculated from the difference between displaced versions of W, with a displacement matching the shear distance. From the original basis, we create a new basis of paired difference functions to decompose the measured derivatives:

$$(D_x, D_y)_j = \vec{\nabla} Z_j, \tag{3}$$

where $j > 0$. In this definition, each element of this basis is comprised of two matrices matching the size of S_x and S_y (see Fig. 1). Note that in many basis definitions the first term, Z_0, is constant and both of its directional derivatives are zero, so the series description can skip the 0th element without loss of generality. To make the projection onto the new basis, we first orthogonalize and normalize our basis using a modified Gram–Schmidt (MGS) procedure [13], obtaining a new set of elements $(\hat{D}_x, \hat{D}_y)$. Note that, in practice, if the wavefront is defined over a domain R, the derivatives are often calculated over smaller regions R_x and R_y, and the wavefront reconstruction can be performed only on R_2, the intersection of R_x and R_y. Now we can project the paired measured wavefront-error gradient onto the new basis, summing over the combined domain, R_2:

$$m_j = \sum_{R_2} (S_x \hat{D}_x + S_y \hat{D}_y). \tag{4}$$

The coefficients $\{m_j\}$ define a linear combination of the polynomial gradient basis terms that approximates the measured wavefront-error gradient. The approximation is due to the fact that in any practical implementation of the method, the gradient basis has a finite number N of elements:

$$(S_x, S_y) \simeq \sum_{j=1}^{N} m_j (\hat{D}_x, \hat{D}_y)_j. \tag{5}$$

The last step is to remap the coefficients $\{m_j\}$ into the corresponding polynomial coefficients from the basis used in Eq. (1). The elements of the transformation matrix arise from the Gram–Schmidt orthonormalization calculation. Alternately, we can find them by projecting the derivatives of the Z polynomials onto the orthonormal basis, obtaining a matrix of coefficients $P_{n,j}$ that satisfies

$$(D_x, D_y)_n = \sum_{j=1}^{N} P_{n,j} (\hat{D}_x, \hat{D}_y)_j, \tag{6}$$

for each element in the set of the Z polynomials gradients. Inverting the matrix P enables us to express the elements of the orthonormal basis in terms of the original derivatives $(D_x, D_y)_j$:

$$(\hat{D}_x, \hat{D}_y)_j = \sum_{n=1}^{N} P^{-1}_{j,n} (D_x, D_y)_n. \tag{7}$$

Now, combining Eqs. (1–3), we have a series representation of the gradient on the basis of the derivatives of Z:

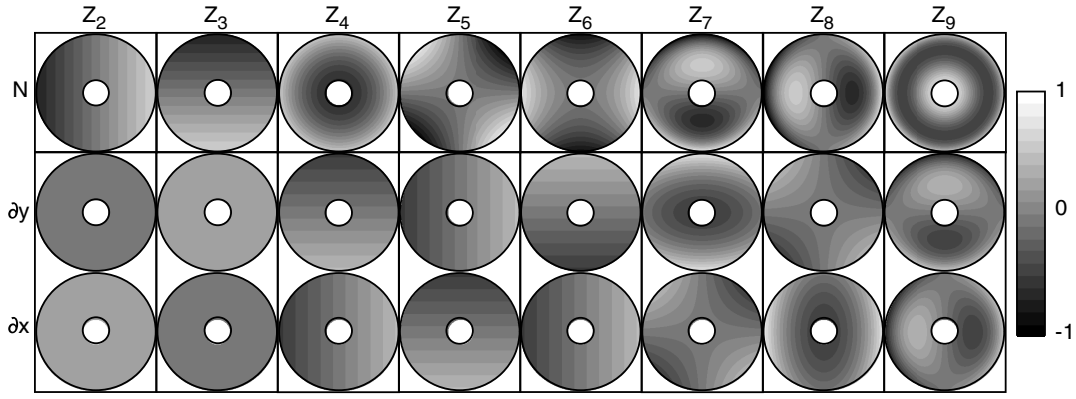

Fig. 1. Top row shows eight Zernike annular polynomials (excluding the piston term). With the outer radius normalized to 1, the inner radius is 0.2. (Note that the choice of this particular basis set is arbitrary.) Below each term, its derivatives D_y and D_x are shown. The polynomial set Z_j and the derivative set $(D_x, D_y)_j$ have been both orthogonalized and scaled for display between −1 and 1.

$$\vec{\nabla} W = (S_x, S_y) = \sum_{n=1}^{N} a_n (D_x, D_y)_n. \quad (8)$$

Substituting Eqs. (5) and (7) into Eq. (8), we find

$$\sum_{j=1}^{N} m_j \sum_{n=1}^{N} P^{-1}_{j,n} (D_x, D_y)_n \simeq \sum_{n=1}^{N} a_n (D_x, D_y)_n, \quad (9)$$

and the expression for the fitting coefficients on Z that produce the reconstructed wavefront, W:

$$a_n \simeq \sum_{j=1}^{N} m_j P^{-1}_{j,n}. \quad (10)$$

3. PRACTICAL IMPLEMENTATION

To summarize, the necessary steps for a correct implementation of this reconstruction algorithm are as follows:

1. Select a finite set of 2D polynomials that are expected to accurately describe the measured wavefront error on an appropriate pixel grid corresponding to the measurement domain.
2. Numerically calculate the derivatives of the polynomials with an algorithm that models the physical measurement procedure.
3. Trim the edges of the domain to remove edge artifacts, if necessary. In some cases, like in shearing interferometry, the derivatives are calculated on a subdomain of the original pupil, and the reconstructed wavefront is defined on the intersection of the subdomains of the two derivatives (see Fig. 2).
4. Group the orthogonal derivatives from Steps 2 and 3 into pairs (D_x, D_y), creating new basis functions $(\hat{D}_x, \hat{D}_y)$ that represent both derivatives in one array.
5. From those paired polynomial derivatives, create a new basis of orthonormal "derivatives" using an MGS procedure to facilitate accurate fitting.
6. Calculate the transformation matrix, P, between the paired derivatives and the corresponding orthogonalized elements. The elements of P may arise from the MGS procedure, or they can be calculated from the projection of each derivative pair onto the new, orthonormalized basis.
7. Project the experimentally measured derivatives onto the orthonormalized basis to find the coefficients $\{m_j\}$.
8. Calculate the coefficients of the reconstructed wavefront $\{m_j\}$ using Eq. (10).
9. The reconstruction accuracy can be evaluated by comparing the calculated gradients of the reconstructed wavefront to the measured derivatives (S_x, S_y) used as input.

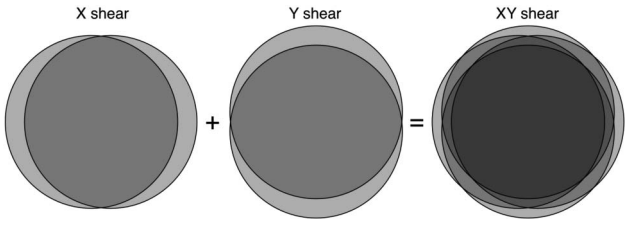

Fig. 2. In shearing interferometry the wavefront derivatives are calculated on a domain R_2 defined by the intersection of the sheared pupil. Without additional information, the wavefront can be reconstructed only over the intersection domain of the two derivatives.

4. DEMONSTRATION

To demonstrate the capabilities of the algorithm, we reconstructed a wavefront generated using a random combination of the first 32 Zernike annular polynomials on a domain of 256×256 pixels. In Fig. 1 we show a nine-element subset of the polynomial basis we used for the reconstruction and the corresponding subset of the derivative basis. In this case the derivatives have been calculated as the finite-difference function that would arise from a shearing interferogram with a shear magnitude of 2 pixels. Note that the domain of the wavefront derivatives is not strictly annular, as shown in Fig. 2. The reconstruction algorithm will generate a wavefront that can be extended to the full size of the chosen polynomials, but the validity of the reconstruction is limited to the domain of the derivatives. In the absence of noise, we compared the reconstructed wavefront and the original one over the derivatives' domain, and we found that the residual difference is a floating point round-off error with a magnitude of 10^{-13} waves peak-to-valley, as shown in Fig. 3. In any practical implementation of the method, when we do not have the original wavefront available *a priori* for comparison, the best way to test the quality of the reconstruction is to differentiate the reconstructed wavefront and compare the calculated derivatives with the measured ones, as shown in Fig. 4. Once again, it is essential to calculate the derivatives using the same method employed in the computation of the derivatives' basis.

5. RECONSTRUCTION ERRORS

In principle, the reconstruction algorithm we presented reproduces the original wavefront over the R_2 domain, with the exception of the piston term that is not present in the measured difference data. In practice, the accuracy of the reconstruction is limited by the finite number of polynomials used and by their characteristics. Ideally, best results are achieved when using a polynomial basis whose elements accurately describe the aberrations that are expected to be present in the optical system under investigation [14]. This of course requires some insight about the wavefront error that is being reconstructed, including

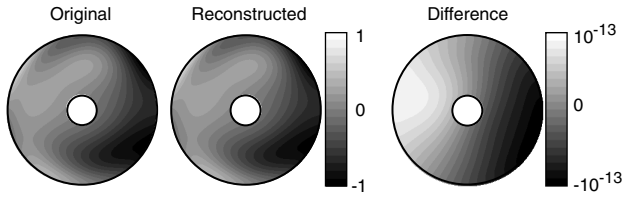

Fig. 3. Example of a wavefront obtained as a random combination of the first 32 Zernike annular polynomials and its reconstruction in the absence of noise. The peak-to-valley value is two waves. The difference between the original and the reconstructed wavefront is limited by floating point round-off error to an order of magnitude of 10^{-13} waves, peak-to-valley.

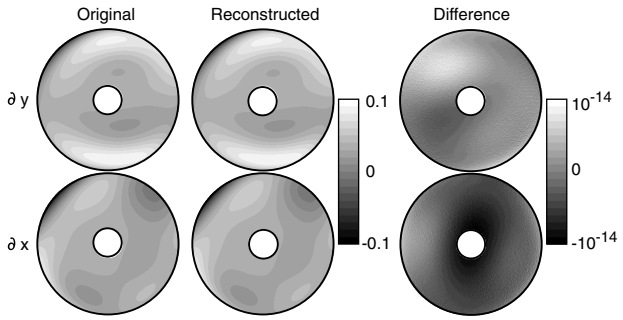

Fig. 4. Comparison between the derivatives of the wavefront shown in Fig. 3 and the derivatives calculated from the reconstruction. The peak-to-valley of the value across the domain is 0.2 waves/pixel. The difference between the original and the reconstructed derivatives is limited by floating point round-off error, with an order of magnitude of 10^{-14} waves/pixel, peak-to-valley.

the presence of high-spatial-frequency features or imperfections in the elements under test. To show how the choice of the basis can affect the reconstruction, we generated a random wavefront error using the first 24 Chebyshev polynomials restricted to a circular domain, we calculated its discrete gradient to simulate measurement, and then we reconstructed it using different sets of Chebyshev and Zernike polynomials. When the reconstruction is performed with the same polynomials used to generate the wavefront, the reconstruction converges quickly as the number of polynomials reaches 24 and produces the correct, original wavefront. Not surprisingly, when a different polynomial basis is used, the reconstruction requires a much larger number of terms to obtain a good reconstruction, as shown in Fig. 5. In this specific case, this behavior depends on the fact that each Chebyshev polynomial expresses spatial frequencies with x or y periodicity while Zernike polynomials express radial and angular frequencies.

The wavefront reconstruction accuracy can be affected by random and systematic errors in the measured derivatives. Such errors propagate to the reconstructed wavefront and limit the accuracy of the final result. For example, the way photon shot noise affects the wavefront reconstruction depends on how the measured data is processed to obtain the wavefront derivatives. We will show here how Gaussian noise present in the (measured) wavefront derivatives influences the quality of the reconstruction. We simulated a random, circular wavefront on a 64 × 64 pixel domain, using the first 32 Zernike polynomials. We calculated its discrete derivatives to simulate measurement and added different levels of Gaussian noise to them to test the reconstruction. The RMS wavefront reconstruction error, E, was calculated by subtracting the computed wavefront from the known, input wavefront and normalizing the resultant RMS to that of the input wavefront:

$$E = \frac{\mathrm{RMS}(W - W_{\mathrm{rec}})}{\mathrm{RMS}(W)} = \frac{\mathrm{RMS}(\Delta W)}{\mathrm{RMS}(W)}. \quad (11)$$

To get a statistically meaningful estimate of E, we repeated this procedure for 1000 random wavefronts and averaged the results. Figure 6 shows that with a signal-to-noise ratio higher than 5 we can expect a normalized RMS reconstruction error E below 0.04. Here the signal-to-noise ratio is defined as the ratio of the signal mean and the standard deviation of the Gaussian noise. Systematic errors are another common occurrence in the wavefront reconstruction. In many cases, such errors are specific to the method used to obtain the derivatives. In shearing interferometry, for instance, one source of error comes from imperfect knowledge of the shear fraction, defined as the ratio between the shear size and the wavefront diameter. To quantify the role of this kind of error, we generated a random wavefront and calculated its discrete derivatives with a fixed shear fraction. Then we performed the reconstruction assuming a range of shear fraction values above and below the known value. Figure 7 shows the calculated error E as a function of the assumed shear fraction values. At the correct shear fraction value, the error is zero, as expected. The results of this test demonstrate that uncertainty in the shear fraction can lead to significant errors in the reconstructed wavefront. However, in the

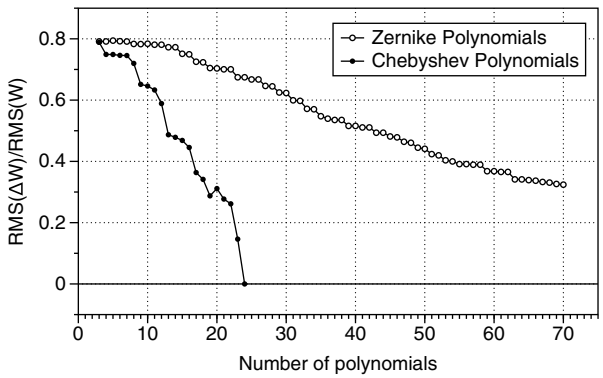

Fig. 5. We generated 200 random wavefronts using the first 24 Chebyshev polynomials, calculated their gradient, and reconstructed them using Zernike polynomial bases and Chebyshev polynomial bases with increasing size. The plot shows the average reconstruction error for the two polynomial types. Using Zernike polynomials, it takes a much larger basis to obtain an accurate reconstruction.

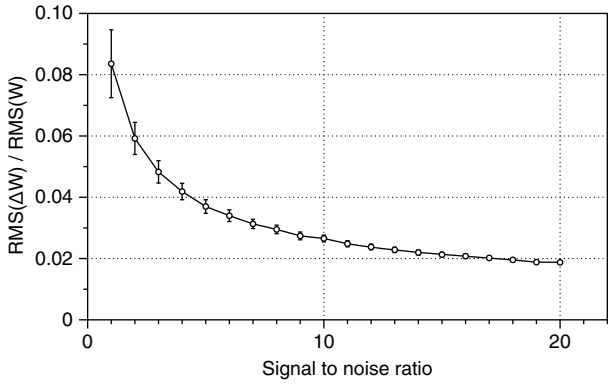

Fig. 6. Plot shows the value of E averaged over 1000 randomly generated wavefronts as a function of the signal to noise ratio in the measured derivatives. The value of E was calculated using wavefronts obtained by the random combination of the first 32 Zernike polynomials. The error bars in the plot represent a 2σ confidence interval.

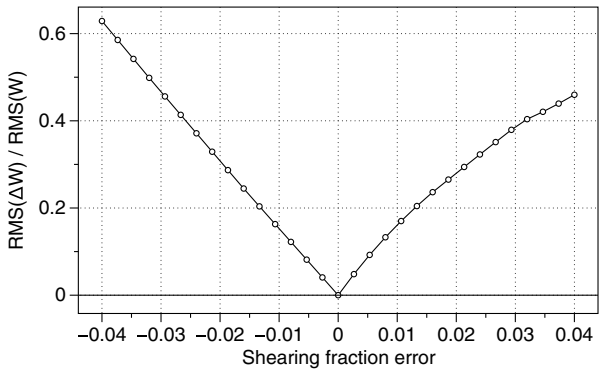

Fig. 7. Plot shows the value of E averaged over 100 wavefronts obtained as a random combination of the first 32 Zernike polynomials. We calculated the derivatives, shearing the wavefront with a shear fraction of 0.0625, and reconstructed the wavefront assuming a shearing fraction from 0.0225 to 0.1025.

majority of cases, the shear fraction s can be measured directly from the CCD camera image as

$$s = \frac{q}{2r}, \tag{12}$$

where r is the radius of the interfering beams and q is their displacement. In these cases the uncertainty on the shearing fraction is determined by the image sampling [15] and can be calculated as

$$\frac{\Delta s}{s} = \left[\left(\frac{\Delta r}{r}\right)^2 + \left(\frac{\Delta q}{q}\right)^2\right]^{1/2}. \tag{13}$$

6. COMPUTATIONAL PERFORMANCE

The method does not rely on any iterative procedure and is therefore both fast and deterministic. In particular, in the case of repeated measurements over the same domain, Steps 1 to 5 of the reconstruction procedure (building the basis and orthonormalizing it) need only to be performed only once, and each wavefront can be reconstructed with a simple projection and a matrix operation (Steps 6 and 7). The method is also capable of handling big wavefront maps without making use of large amounts of memory [6]. We measured the reconstruction time on a 2.6 GHz processor in Matlab. We performed the reconstruction of random wavefronts with bases of different sizes and element numbers. Figure 8 shows how the reconstruction time scales quadratically with the number of elements in the basis and with the wavefront sampling. The performance of the algorithm is in line with other similar reconstruction methods [16], and it can be improved by using a more powerful processor or a parallel computing architecture.

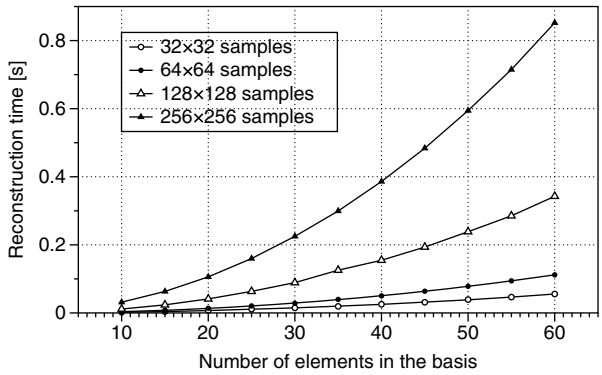

Fig. 8. Wavefront reconstruction computation time for wavefront maps with different sampling as a function of the number of elements in the polynomial basis. For a wavefront with 32 × 32 samples and a basis of 10 polynomials, we measured a reconstruction time of 2 ms. This time grew to 32 ms as the number of samples increased to 256 × 256 with the same number of basis terms.

7. CONCLUSION

We developed a method for the reconstruction of a wavefront error from its gradient. The method is deterministic and is based on the projection of the measured gradient onto an orthonormal basis of polynomials derived from the gradient of a specific polynomial set. To reconstruct the wavefront, we map this decomposition back onto the chosen polynomial basis. This method can be applied with any set of orthogonal polynomials. This reconstruction technique is based on the direct projection of the data on an orthonormal basis; therefore it works correctly regardless of the number of terms used to generate the test wavefront error and the number of terms considered in the reconstruction. In other words it is possible to calculate separately the contribution to the wavefront error of any element of the chosen basis. The number of polynomials that should be used in the reconstruction depends on the accuracy needed in the reconstructed wavefront. We tested the speed of the reconstruction algorithms for different wavefront sampling and polynomial basis sizes, and we found that it is in line with other modal reconstruction methods [16]. The reconstruction time scales quadratically with the size of the basis and the wavefront sampling. We evaluated the performance of this method in the presence of Gaussian noise and with shear-magnitude uncertainty. Simulations show that the effect of uncorrelated Gaussian noise in the two components of the measured gradient has limited effect on the reconstruction accuracy. In particular, we found that as long as the signal-to-noise ratio is higher than 5 we can expect less than 4% error in the reconstructed wavefront. The systematic error induced by a shear-magnitude error can be significant: a shearing fraction error of 0.01 can lead to a wavefront RMS reconstruction error E of 0.1. We believe that the main advantage of this method is that it can be applied quite generally, without relying on approximations to the wavefront differences obtained with any technique or relying on the properties of particular polynomial sets. In this the proposed approach is different from other modal reconstruction techniques like the difference Zernike polynomials fitting method [10,17,18]. This makes this method both flexible and reliable. Furthermore, the quality of the result can be directly evaluated by comparing the measured derivatives to those calculated from the reconstructed wavefront.

LDRD Program of Lawrence Berkeley National Laboratory under US Department of Energy (DE-AC02-05CH11231).

REFERENCES

1. M. Strojnik, G. Paez, and M. Mantravadi, "Lateral shear interferometers," in *Optical Shop Testing* (Wiley, 2007), Chap. 4.
2. D. Malacara-Doblado and I. Ghozeil, *Hartmann, Hartmann–Shack, and Other Screen Tests* (Wiley, 2007), Chap. 10.
3. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," J. Mod. Opt. **43**, 289–293 (1996).
4. D. E. Vandenberg, W. D. Humbel, and A. Wertheimer, "Quantitative evaluation of optical surfaces by means of an improved Foucault test approach," Opt. Eng. **32**, 1951–1954 (1993).
5. M. Rosensteiner, "Wavefront reconstruction for extremely large telescopes via CuRe with domain decomposition," J. Opt. Soc. Am. A **29**, 2328–2336 (2012).
6. M. P. Rimmer, "Method for evaluating lateral shearing interferograms," Appl. Opt. **13**, 623–629 (1974).
7. W. H. Southwell, "Wave-front estimation from wave-front slope measurements," J. Opt. Soc. Am. **70**, 998–1006 (1980).
8. A. Talmi and E. N. Ribak, "Wavefront reconstruction from its gradients," J. Opt. Soc. Am. A **23**, 288–297 (2006).
9. E. Clemens and W. Ingolf, "Solution to the shearing problem," Appl. Opt. **38**, 5024–5031 (1999).
10. G. Harbers, P. J. Kunst, and G. W. R. Leibbrandt, "Analysis of lateral shearing interferograms by use of Zernike polynomials," Appl. Opt. **35**, 6162–6172 (1996).
11. B. Bravo-Medina, G. Garcia-Torales, R. Legarda-Saenz, and J. L. Flores, "Wavefront recovery fourier-based algorithm used in a vectorial shearing interferometer," Proc. SPIE **8867**, 88670Z (2013).
12. R. Legarda-Saenz, "Robust wavefront estimation using multiple directional derivatives in moiré deflectometry," Opt. Lasers Eng. **45**, 915–921 (2007).
13. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University, 1996).
14. J. C. Wyant and K. Creath, "Basic wavefront aberration theory for optical metrology," in *Applied Optics and Optical Engineering*, R. R. Shannon and J. C. Wyant, eds. (Academic, 1992), Vol. **11**, p. 2.
15. G. Leibbrandt, G. Harbers, and P. J. Kunst, "Wave-front analysis with high accuracy by use of a double-grating lateral shearing interferometer," Appl. Opt. **35**, 6151–6161 (1996).
16. F. Dai, F. Tang, X. Wang, P. Feng, and O. Sasaki, "Use of numerical orthogonal transformation for the Zernike analysis of lateral shearing interferograms," Opt. Express **20**, 1530–1544 (2012).
17. F. Dai, F. Tang, X. Wang, O. Sasaki, and P. Feng, "Modal wavefront reconstruction based on Zernike polynomials for lateral shearing interferometry: comparisons of existing algorithms," Appl. Opt. **51**, 5028–5037 (2012).
18. G.-M. Dai, "Modal wave-front reconstruction with Zernike polynomials and Karhunen–Loève functions," J. Opt. Soc. Am. A **13**, 1218–1225 (1996).