Methods for the characterization of deformable membrane mirrors

Martin Booth, Tony Wilson, Hong-Bo Sun, Taisuke Ota, and Satoshi Kawata

We demonstrate two methods for the characterization of deformable membrane mirrors and the training of adaptive optics systems that employ these mirrors. Neither method employs a wave-front sensor. In one case, aberrations produced by a wave-front generator are corrected by the deformable mirror by use of a rapidly converging iterative algorithm based on orthogonal deformation modes of the mirror. In the other case, a simple interferometer is used with fringe analysis and phase-unwrapping algorithms. We discuss how the choice of singular values can be used to control the pseudoinversion of the control matrix.

OCIS codes: 010.1080, 010.7350, 120.2650, 220.1000.

1. Introduction

Electrostatically controlled deformable membrane mirrors were first proposed in the 1970s[1] and have more recently become available for wave-front control in low-cost adaptive optics systems for various applications.[2,3] These applications have included adaptive optics for confocal microscopy[4] and two-photon microscopy,[5,6] astronomy,[7] optical tweezers,[8–10] coupling light into optical fibers,[11] and ophthalmology.[12] Several methods have been described in the literature for the characterization and control of these mirrors.[13–15] In general, one has to obtain a control matrix that relates the voltages applied to the mirror electrodes to the deformation of the mirror surface. The control matrix would usually be based on an orthogonal set of mirror deformation modes,[14] or Zernike polynomials.[13] These methods, however, have relied on the use of a wave-front sensor for the characterization of the wave fronts generated by the mirror. We present here two methods that do not require a wave-front sensor. In the first case we use a previously calibrated aberration generator and a single photodetector to train the control system. We have already implemented adaptive optics systems trained with this method.[4,8,9] In the second case, we use a simple interferometric setup in conjunction with fringe analysis and phase-unwrapping routines. Using the latter setup, we show how appropriate representation of the acquired data provides insight into the properties of the control system and the choice of parameters.

2. Deformable Membrane Mirrors

The deformable mirror (DM) used in the study reported in this paper was a 15 mm diameter, 37 actuator model from OKO Technologies, The Netherlands. The basis of operation of these mirrors is discussed extensively elsewhere,[2,3,13] so we do not repeat that discussion here. Instead, we summarize briefly the important points. This DM consists of an aluminum-coated silicon nitride membrane suspended above a hexagonal electrode structure. The mirror surface shape is controlled by electrostatic forces that result from the application of a potential difference between each electrode and the mirror surface. To permit the movement of the mirror surface both toward and away from the electrodes, we apply a constant offset voltage to all electrodes. The curvature of the resultant mirror shape is easily compensated for by the defocusing of the wave front incident upon the DM. The voltage applied to an electrode essentially controls the curvature of the membrane in the vicinity of the electrode. It can be shown that the consequent mean deflection of the surface in the vicinity of the electrode is proportional to the square of the applied voltage. It is therefore desirable to use control signals that are proportional to the square

M. Booth (martin.booth@eng.ox.ac.uk) and T. Wilson are with the Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK. H.-B. Sun, T. Ota, and S. Kawata are with the Department of Applied Physics, Osaka University, Suita, Osaka 565-0781, Japan.

Received 23 June 2004; revised manuscript received 27 November 2004; accepted 22 December 2004.

0003-6935/05/245131-09$15.00/0

root of the electrode voltage, giving a linear relationship between control signals and displacement.

Although the DM membrane has a diameter of 15 mm, the electrode structure covers only the central portion to a diameter of 12 mm. As the circumference of the mirror is fixed and cannot be displaced, it is normal to use only the central 9–12 mm portion of the DM for aberration correction.[3] Typically, the DM will show an initial aberration, consisting mainly of astigmatism, that is due to residual stresses in the substrate that deform the supporting edge of the membrane. One should remove this initial aberration by applying appropriate control signals to the mirror before use.

The control signals can be represented by a 37-element vector **c**. Rather than driving the individual DM electrodes directly, the usual approach is to use a modal basis, such as the Zernike polynomials (see Appendix A). Each of these modes can be generated by the DM by use of an appropriate combination of control signals. In general, a phase function, $\Phi(r, \theta)$, can be described by a series of Zernike polynomials:

$$\Phi(r, \theta) = \sum_{i=1}^{\infty} a_i Z_i(r, \theta), \tag{1}$$

where a_i is the amplitude of the Zernike mode, $Z_i(r, \theta)$. If the coefficients of the Zernike modes generated by the DM are the elements of vector **a**, then we can define the following matrix-vector equation:

$$\mathbf{a} = \mathbf{B}\mathbf{c}. \tag{2}$$

As matrix **B** is generally nonsquare and possibly singular, inversion of Eq. (2) involves the use of a matrix pseudoinverse, $\mathbf{B}^{\dagger}$.[16] In this case we specify the desired Zernike modal coefficients $\mathbf{a}'$ and calculate the required control signals:

$$\mathbf{c} = \mathbf{B}^{\dagger}\mathbf{a}'. \tag{3}$$

The two techniques described in this paper essentially involve the determination of **B** or, more particularly, its inverse.

3. Wave-front Generator Method

The wave-front generator method requires a wave-front generator and a single photodetector for the training of the DM system. We implemented the experimental configuration shown in Fig. 1, which employs a wave-front generator based on a 256 × 256 pixel ferroelectric liquid-crystal spatial light modulator (FLCSLM; Displaytech, Longmont, Colo.)

This wave-front generator, described by Neil *et al.*,[17] permits the accurate generation of arbitrary phase and amplitude wave fronts. An expanded beam from a frequency-doubled Nd:YAG laser (wavelength, 532 nm) was passed through the wave-front generator. The resultant wave front passed through a beam expander and a 50–50 beam splitter onto the DM. An offset voltage was applied to all electrodes of the DM

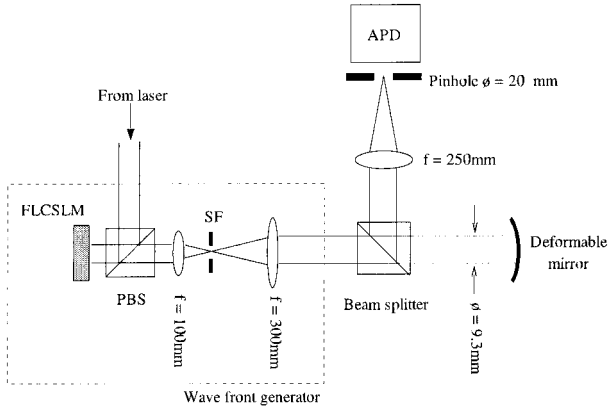

Fig. 1. Experimental setup for deformable mirror training by the wave-front generator method. PBS: polarizing beam splitter; SF, spatial filter.

such that the center of the membrane was at half its maximum deflection. To compensate for the resultant DM curvature, the lenses of the beam expander were set such that the wave fronts incident onto the DM were divergent. The wave fronts reflected from the mirror in its default position were therefore nominally flat. The beam reflected off the beam splitter was then focused onto a 20 μm diameter pinhole placed in front of an avalanche photodiode (APD). The FLCSLM wave-front generator was configured to provide an output wave front with a diameter of 9.3 mm, constant amplitude, and phase variation $\Psi(r, \theta)$, where r was the normalized radial coordinate such that the wave front had unity radius. This wave front was incident upon the DM, which effectively added the phase $\Phi(r, \theta)$ such that the wave front reflected back toward the beam splitter contained the phase variation $\Psi(r, \theta) + \Phi(r, \theta)$. The signal detected by the APD can be expressed as

$$W = \int_V \left| \iint_P \exp[i\Psi(r, \theta) + i\Phi(r, \theta)] \times \exp[i\nu r \cos(\theta - \phi)] r \mathrm{d}r \mathrm{d}\theta \right|^2 \nu \mathrm{d}\nu \mathrm{d}\phi, \tag{4}$$

where P represents the circular pupil of the final lens, V is the pinhole aperture, and (ν, ϕ) are the polar coordinates in the detector plane. We note that the pinhole diameter of 20 μm is equivalent to 56% of the diameter of the first zero in intensity in the diffraction-limited focus of the final lens.

Let the vector $\boldsymbol{\zeta}_i$ represent the DM control signals required for producing a unit amplitude of Zernike mode $Z_i(r, \theta)$. It is clear that signal W would be maximum when the aberrations introduced by the FLCSLM wave-front generator and the DM were equal and opposite, i.e., $\Psi(r, \theta) = -\Phi(r, \theta)$. We could therefore set $\Psi(r, \theta) = -Z_i(r, \theta)$ and then find $\boldsymbol{\zeta}_i$ as the set of DM control signals that maximizes W.

To do this, we used an efficient iterative procedure

based on the modal wave-front sensor (MWFS) concept.[18] In one implementation of the MWFS, a predetermined, fixed amount of a particular aberration mode (the positive bias) is added to the input wave front, and then the focal intensity is measured. An equal but opposite amount (the negative bias) is added, and the focal intensity is measured again. The difference between these two measurements is proportional to the amount of the aberration mode contained in the input wave front. Because this method effectively takes into account the mathematical properties of the signal that is being maximized,[19] it can have a more rapid convergence than many other iterative adaptation methods that use the same experimental hardware, such as model free gradient descent methods.[20] Previous implementations of MWFS have used the Zernike polynomials as the basis mode set. We introduce here a new approach based on an orthogonal set of mirror deformation modes obtained through a singular-value decomposition (SVD) process.[14] These 37 orthogonal modes describe fully the deformations that can be generated by the DM and therefore provide a more succinct modal basis than the Zernike polynomials. We denote these modes by the functions $\Xi_i(r, \theta)$ and the corresponding vectors of DM control signals by $\boldsymbol{\xi}_i$. Although the functions $\Xi_i(r, \theta)$ cannot be expressed analytically, their orthogonality means that they are a suitable basis for the MWFS.

Figure 2(a) shows measured values of W as different modes were generated by the DM. The response from each mode shows a quadratic variation about the origin, the width of which depends on the relative ability of the DM to produce that mode for a given magnitude of electrode voltage. This ability is directly related to the singular values obtained through the SVD algorithm: The larger the singular value, the greater the capability of the DM to generate that particular mode. The MWFS correction scheme proceeded as follows: For each mode in turn, a bias of amplitude $b = 80$ was applied such that the DM introduced a total aberration $\Phi_0(r, \theta) + b\Xi_i(r, \theta)$, and APD signal W_1 was measured. An equal and opposite bias was then applied, giving the total aberration $\Phi_0(r, \theta) - b\Xi_i(r, \theta)$, and APD signal W_2 was measured. If $\Phi_0(r, \theta) = a\Xi_i(r, \theta)$, output signal $\Delta W = W_2 - W_1$ is proportional to a. Figure 2(b) shows the response of such MWFS measurements to different amplitudes of the input mode. These curves were derived from the experimental results of Fig. 2(a). Each response shows clearly a linear region about $a = 0$ that gives rise to the rapid convergence of this method.

First, the initial aberration of the DM, $\Phi_0(r, \theta)$ was removed as follows: The DM control signals, represented by the vector $\mathbf{c}_n$, were set to zero, and the FLCSLM wave-front generator was set such that $\Psi(r, \theta) = 0$. The lateral position of the pinhole was adjusted to ensure a maximum photodetector signal. The correction procedure described above was followed, and the updated control signals were then calculated as

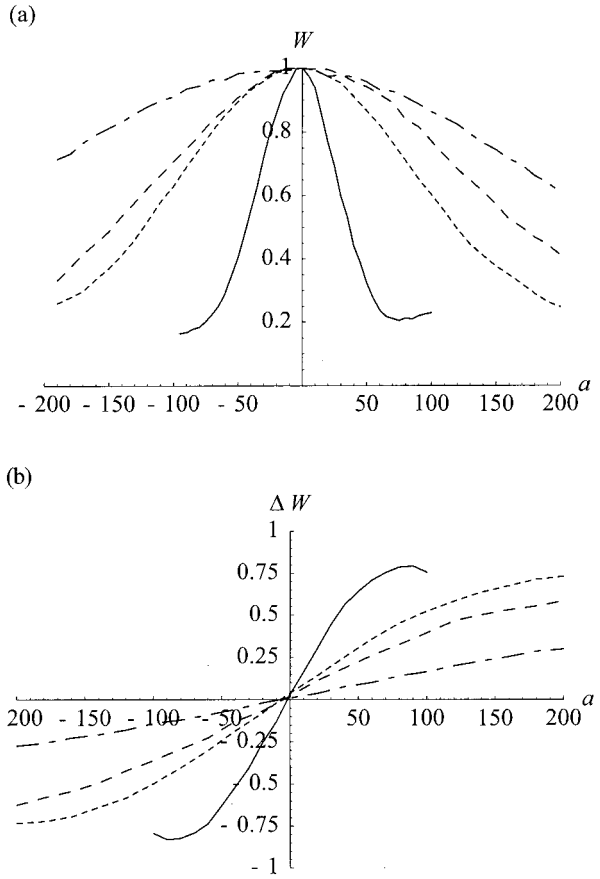

Fig. 2. (a) Photodetector responses for a selection of mirror deformation modes. (b) Output signals from modal sensor measurements based on the mirror modes. Solid curves, mode 0; dashed curves, mode 6; long-dashed curves, mode 12; dotted–dashed curves, mode 18. Mode 0 is the mode with the largest corresponding singular value. Other modes are ordered according to descending singular values.

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \gamma S_i^{-1}(W_1 - W_2)\boldsymbol{\xi}_i, \tag{5}$$

where γ was a parameter dependent on the overall intensity and the detector gain and S_i was the gradient of the sensor response, calculated from the curves of Fig. 2(b). The value of γ was chosen by the user to control the convergence rate. Sequential correction of each mode, $\Xi_i(r, \theta)$, was performed. One cycle was sufficient to remove this initial aberration.

We then obtained control signals $\boldsymbol{\zeta}_i$ by setting the output phase of the FLCSLM wave-front generator to $\Psi(r, \theta) = -aZ_i(r, \theta)$ and following the above procedure. Once the maximum APD signal had been found, we calculated vectors $\boldsymbol{\zeta}_i$ from the final control signals $\mathbf{c}$ as

$$\boldsymbol{\zeta}_i = \frac{1}{a}\,\mathbf{c}. \tag{6}$$

Vectors $\boldsymbol{\zeta}_i$ form the columns of matrix $\mathbf{B}^{\dagger}$ of Eq. (3). This method is therefore a direct determination of the matrix pseudoinverse required for obtaining control

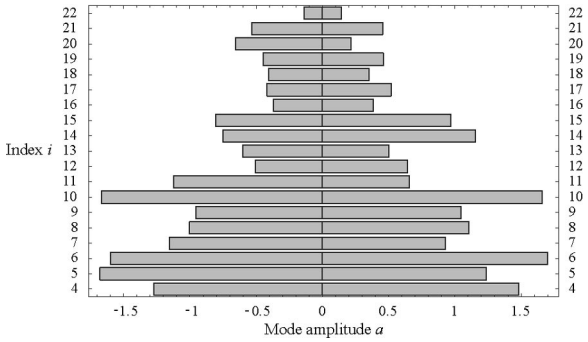

Fig. 3. Range of each Zernike mode generated by the deformable mirror for a maximum wave-front error variance of 0.05 rad^2.

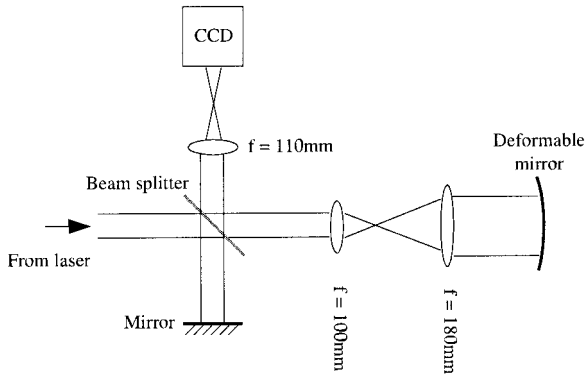

Fig. 4. Experimental setup for the interferometer method.

signals **c** from Zernike coefficients **a**. It was found that for $a = 1$ cycling through each mode $\Xi_i(r, \theta)$, twice was sufficient to complete the iteration. As explained by Paterson *et al.*,[14] the higher-order mirror modes can be generated only in small amplitudes by the DM and it is often advantageous to discard these modes to improve controllability. We found that including more than 21 mirror modes had little effect on the results when we attempted to fit Zernike modes 4–22. Indeed, for most Zernike modes, as few as 10 mirror modes were sufficient for accurate representation.

Using the DM to correct aberrations from the wave-front generator, we can measure the accuracy of the Zernike modes produced by the DM. If the output from the wave-front generator is set such that $\Psi(r, \theta) = -aZ_i(r, \theta)$ and control signals $a\zeta_i$ are applied to the DM, then, assuming that the DM produces perfect Zernike modes, there should be no aberration in the final wave front. In practice, however, there will be some error in the modes generated by the DM, so, as a is increased from zero, the photodetector signal should drop. If the error in the final wave front, $\Delta\Phi(r, \theta)$, is small, then the photodetector signal is given by[21]

$$W = W_0[1 - \mathrm{var}(\Delta\Phi)], \tag{7}$$

where W_0 is the signal when no aberration is present and var($\Delta\Phi$) is the variance of the wave-front error over the pupil. Hence, if we choose 0.05 rad^2 as the maximum acceptable variance of the wave-front error, for each Zernike mode we can define the useful range of the DM as the values of a for which $W \geq 0.95\ W_0$. This range is shown by the bars in Fig. 3. We note that, using this criterion, if the DM surface were maintained perfectly flat while the output of the wave front were varied in the same manner, the range would extend from $a = -0.22$ to $a = 0.22$ for any chosen Zernike mode. We see a similar trend to that described by other authors, whereby the DM produces lower-order Zernike modes more efficiently than higher-order modes. We note that for mode 22 the range is worse than if the mirror had been kept flat, indicating that the training method failed. This result is compatible with the results of others that have indicated that this DM is not capable of generating this mode.

4. Interferometer Method

We describe a method of DM characterization that is based on an interferometer and software so that we may analyze the DM's shape by using fringe pattern analysis and phase unwrapping. If the DM is placed in one arm of an interferometer, the interference pattern produced will encode the phase introduced by the DM. Fringe pattern analysis permits the extraction of wave-front phase information from a single such interferogram, albeit modulo 2π. The full unwrapped phase information, which corresponds directly to the DM shape, can be calculated by phase-unwrapping techniques. The following results show that relatively straightforward fringe pattern analysis and phase unwrapping can be used to characterize and control the DM.

The DM was incorporated into a Twyman–Green interferometer, as shown in Fig. 4. The expanded beam from a 633 nm wavelength He–Ne laser was incident upon a 50–50 beam splitter that separated the beam into an object and a reference path. The reference beam was reflected off a plane mirror and returned to the beam splitter. The object beam was expanded by two positive lenses such that the beam filled the mirror aperture. The axial position of the second lens was adjusted to compensate for the curvature of the DM when the defocus bias was applied. The surface shape of the DM introduced a phase modulation into the reflected wave fronts that passed back through the beam expander. The resultant interferogram was acquired by use of a CCD camera placed in an image plane conjugate to the DM. Adjustment of the orientation of the mirror in the reference path permitted control of the number of fringes in the interferogram.

A. Fringe Pattern Analysis

Fringe pattern analysis of interferograms was introduced in the early 1980s[22,23] and has been followed by many advanced variations. Within certain practical limitations, the phase distortion introduced by the DM can be extracted from a single interferogram. When sufficient tilt is introduced into the reference

beam, many fringes appear in the interferogram. The Fourier transform (FT) of this fringe pattern then contains three orders that are spatially separated, the first order of which is isolated with a mask and shifted by an integer number of pixels such that it lies near the zero-frequency position. The inverse FT of the resultant image yields

$$I(x,\ y) = \mathcal{F}\{M(u,\ v)\} \otimes \{A_r A_0 \exp[i(\phi + \Delta\tau)]\}, \quad (8)$$

where $A_0(x, y)$ and $\phi(x, y)$ are, respectively, the amplitude and phase distributions in the object wave front, $A_r(x, y)$ is the amplitude distribution in the reference wave front, $\mathcal{F}\{\ \}$ is the FT operator, and $\otimes$ is a two-dimensional convolution operator. The Fourier plane mask is represented by $M(u, v)$, where u and v are the coordinates that describe the Fourier plane. The small phase tilt $\Delta\tau(x, y)$ arises when an integer pixel shift, rather than a more-accurate noninteger shift, is used to move the masked Fourier plane order to the zero-frequency position. The convolution of Eq. (8) can be thought of as a local averaging of the complex pixel values for which the extent of the averaging depends on the form of the mask, $M(u, v)$. As long as the mask is appropriately chosen, we can derive the approximate equality

$$\psi(x,\ y) \approx \arg[I(x,\ y)] - \Delta\tau(x,\ y), \quad (9)$$

where $\psi(x, y)$ is the wrapped phase, i.e., the continuous phase $\phi(x, y)$ modulo 2π. For this approximation to be reasonable, we require first that variations in the amplitude functions $A_r(x, y)$ and $A_0(x, y)$ be small and second that phase function $\phi(x, y)$ be such that the necessary spatial frequencies for its retrieval are within the passband of the mask. The first condition is usually satisfied, except at the edge of the image of the mirror aperture where $A_0(x, y)$ drops to zero. The masking process leads to inaccuracies in the retrieved phase over this region. Methods to overcome this problem, for example, use of analytic continuation,[24] have been developed. However, as the usable diameter of the DM is not usually more than 80% of the full aperture, phase errors in the region of the aperture edge are immaterial. Furthermore, the phase features that can be introduced by the DM are limited by the size of the electrodes, which are approximately 1.7 mm in diameter. This equivalently places an upper limit on the spatial frequencies that one can expect to be present and is therefore a useful criterion for the choice of masking function, $M(u, v)$.

A potential limitation of fringe analysis is insensitivity to tip and tilt modes that is due to the arbitrary amount of tip and tilt that is added to produce the interference fringes. In this experiment, however, the edge of the DM is fixed such that a single calibration step provides a reference for further measurements. Any residual tip or tilt relative to this reference must be introduced by the deformation of the DM.

B. Phase Unwrapping

The fringe pattern analysis provides the wrapped phase function $\psi(x, y)$. For further analysis and Zernike mode fitting we required the unwrapped phase $\phi(x, y) = \psi(x, y) - 2\pi k(x, y)$, where $k(x, y)$ takes the appropriate integer values. Although many advanced methods have been developed for this purpose, we found that in the present case, in which the phase variations were smooth and contained no singular points, a single step fast FT method was sufficient. This method, adapted from Ref. 25, is outlined in Appendix B.

C. Measurement of the Control Matrix

In this section we describe the steps taken to determine control matrix **B**. First we measured in turn each influence function $\psi_k(x, y)$, the characteristic shape produced when a unit control signal is applied to the kth DM electrode. The mirror deformation $z(x, y)$ for arbitrary control signals is a linear superposition of these functions scaled by the corresponding control signal. We applied a control signal c to each of the $K = 37$ electrodes in turn and recorded the resultant phase function $\psi_k(x, y)$, where the subscript k corresponds to the kth electrode. These functions were measured over the 12 mm diameter functional aperture of the deformable mirror rather than the full aperture. In the following discussion, the radius of this functional aperture is normalized to 1. The mirror deformation $z(x, y)$ is therefore given by

$$z(x,\ y) = \sum_{k=1}^{K} c\psi_k(x,\ y). \quad (10)$$

Each of the influence functions can in turn be represented by a series of P Zernike polynomials denoted $Z_p(x, y)$:

$$\psi_k(x,\ y) = \frac{1}{c}\sum_{p=1}^{P} \beta_{kp} Z_p(x,\ y). \quad (11)$$

The coefficients β_{kp} were calculated from the influence functions by use of the orthogonality properties of the Zernike polynomials:

$$\beta_{kp} = \frac{1}{\pi}\iint_{x^2+y^2\le 1} \psi_k(x,\ y) Z_p(x,\ y)\mathrm{d}x\mathrm{d}y. \quad (12)$$

This process was repeated for a range of values of control signal c, and the corresponding β_{kp} were calculated. A least-squares linear fit gave $\beta_{kp} = \hat{b}_{kp} + b_{kp}c$, where the offset $\hat{b}_{kp}$ corresponded to the initial aberration figure of the DM. The coefficients b_{kp} formed the elements of matrix **B** in Eq. (2).

D. Control Matrix Inversion

Several methods have been suggested for the pseudoinversion of **B** to obtain $\mathbf{B}^{\dagger}$. Claflin and Baraket[26] used an analytic model of a DM to calculate the

pseudoinverse, whereas Dayton *et al.*[27] followed the equivalent process by using calculated influence functions. Fernández and Artal[12] employed SVD to invert the matrix[16] based on measurements with a Hartmann–Shack wave-front sensor. Zhu *et al.*[13] also used such a sensor with an iterative scheme to obtain the control signals for a required combination of Zernike coefficients. By combining interferometric measurements with fringe analysis and SVD matrix inversion, we were able to implement an equivalent system by using standard optics equipment without a wave-front sensor. This method involves decomposition of matrix **B** into an orthogonal set of deformation modes. Nominally, these modes are equivalent to the theoretically derived modes used earlier that were based on an idealized model of the DM. However, the modes that arise from the SVD of our experimentally derived **B** showed some differences, probably owing to the orientation and positioning of the DM relative to the beam. The advantage of matrix pseudoinversion by SVD is the ability to remove the modes that can lead to problems. For example, higher-order modes can be generated only inefficiently by the DM and require high control signals. As the electrode voltages, and hence the control signals, are limited by the design of the DM, including these modes in the inversion can easily lead to nonlinear operation because of saturation of the electrode voltages. Fernández and Artal removed such modes, which have singular values smaller than an arbitrary value, before calculating $\mathbf{B}^{\dagger}$.[12] This process does not, however, take into account the relative importance of the different modes that the system might need to correct. We extend the concept to take this into account and to gain further insight into the performance of the DM.

If we combine Eqs. (2) and (3) we obtain an expression that relates the actual Zernike coefficients generated by the DM, **a**, to the desired coefficients, **a**′:

$$\mathbf{a} = \mathbf{B}\mathbf{B}^{\dagger}\mathbf{a}'. \tag{13}$$

The similarity matrix $\mathbf{M} = \mathbf{B}\mathbf{B}^{\dagger}$ therefore tells us not only the generation efficiency of each Zernike mode (the diagonal elements of **M**) but, in addition, what other, undesired modes are also present (within the range of aberration modes represented by **a**). Because the form of **M** depends on the SVD pseudoinversion, it is a useful tool with which to compare the operation of the DM when different numbers of modes are included in the inversion process.

The inclusion of modes that require large control signals in the SVD pseudoinversion leads to a greater likelihood of electrode voltage saturation and nonlinear operations. Although the onset of saturation would depend on the exact values of **a**′, we can monitor the general behavior by considering the vector norms of the columns of $\mathbf{B}^{\dagger}$. A larger norm indicates a larger influence of the corresponding Zernike mode's amplitude on the control signals. Because the infinity norm is equivalent to the largest element in the vector, it is most appropriate for estimating the onset of control signal saturation. We define saturation likelihood Q as the largest of the infinity norms of the columns of $\mathbf{B}^{\dagger}$. This is, of course, equivalent to the maximum element of $\mathbf{B}^{\dagger}$:

$$Q = \max(\mathbf{B}^{\dagger}). \tag{14}$$

Figure 5 shows representations of the similarity matrices obtained when different numbers of modes, N, are excluded from the SVD matrix inversion. Each column of **M** shows the Zernike modes that would be produced by the DM when the corresponding element

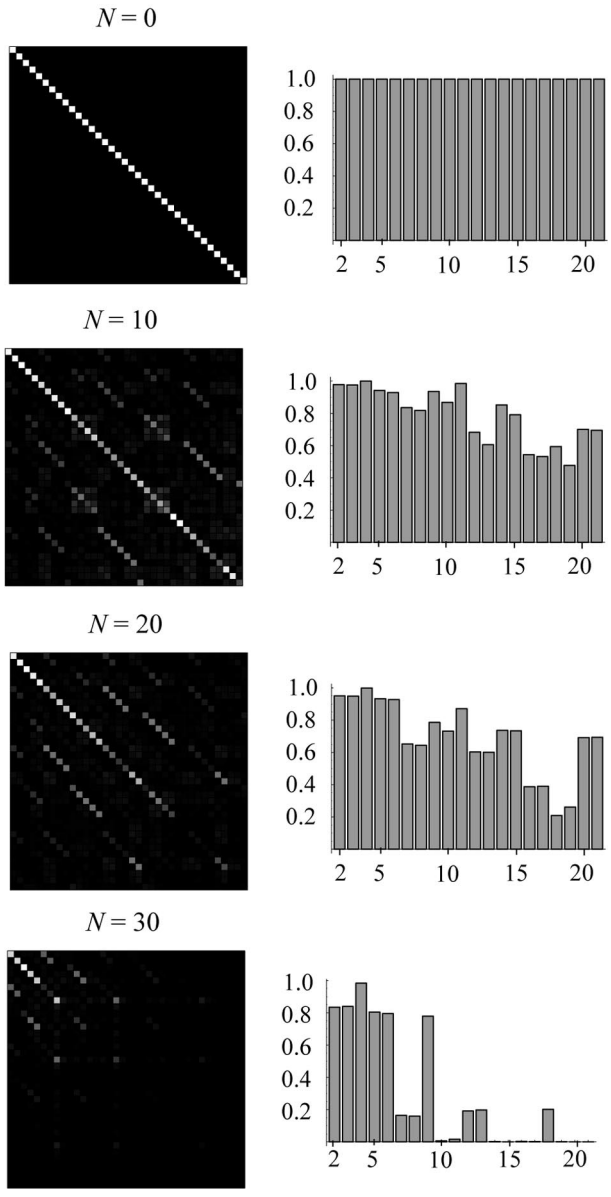

Fig. 5. Graphic representations of similarity matrices **M** obtained when N singular values are removed during the SVD inversion. The gray-scale shading represents the absolute value of the matrix element. The Zernike modes represented range from mode 2 (top left) to 37 (bottom right). The histograms at the right show the values of the first 20 diagonal elements, corresponding to Zernike modes 2–21.

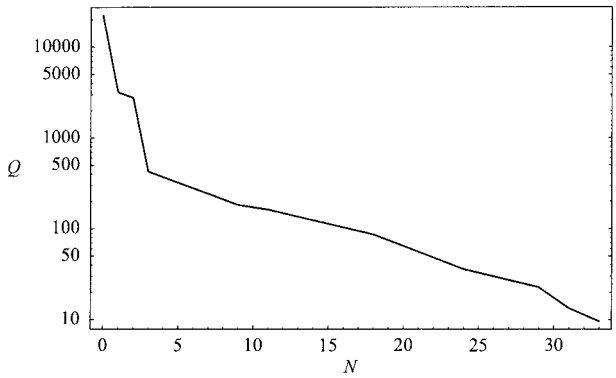

Fig. 6. Saturation likelihood Q after N singular values had been removed during the SVD matrix inversion.

of $\mathbf{a}'$ is set to 1, the other elements being 0. Figure 6 shows the equivalent values of Q when different numbers of modes are excluded.

For accurate generation of individual Zernike modes, $\mathbf{M}$ should be diagonally dominated and in the ideal case would be an identity matrix. Although such is indeed the case when $N = 0$, Q is prohibitively large, meaning that control signal saturation is highly likely. As N increases, Q drops dramatically; this drop is accompanied by a drop in magnitude of some of the diagonal elements of $\mathbf{M}$ and an increase in magnitude of certain off-diagonal terms, indicating modal cross talk. As N increases further, the contributions of some modes disappear entirely from $\mathbf{M}$. The choice of N is therefore a compromise between keeping the ability to generate different modes and reducing the likelihood of saturation.

Figure 7 shows an experimentally determined $\mathbf{M}$ for $N = 20$, which we obtained by setting each element of $\mathbf{a}'$ in turn to 1 and then applying the control signals according to Eq. (3). The columns of $\mathbf{M}$ were found as the Zernike mode coefficients of the generated wave front. It is clear that the experimentally determined matrix is almost identical to the calculated matrix. The root-mean-square error between the two matrices was calculated as 0.019.

We can illustrate the usefulness of matrix $\mathbf{M}$ by considering an example in which the Zernike aberration modes 2–8 in addition to mode 11 are required. Such a combination of modes is common, for example, in aberrations encountered in biological microscopy.[28] We could specify that, to produce each of these modes with sufficient accuracy, the diagonal elements of $\mathbf{M}$ that correspond to those modes should be greater than 0.8, whereas Q should be as small as possible. We find that these conditions are satisfied when 14 singular values are removed. The relevant diagonal elements of $\mathbf{M}$ then range from 0.8099 for mode 8 to 0.9998 for mode 4.

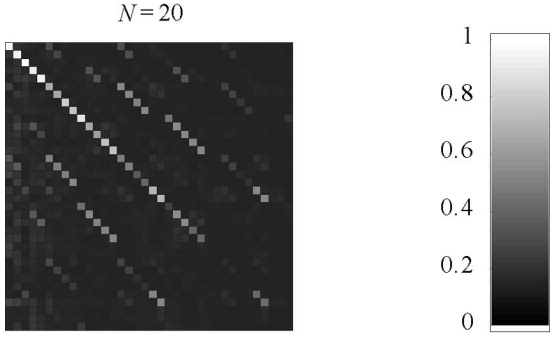

Fig. 7. Experimentally determined similarity matrix after 20 singular values had been removed during the matrix pseudoinversion. The scale represents the absolute value of the matrix element.

5. Conclusions

We have demonstrated two independent methods that can be used to characterize deformable membrane mirrors for use in adaptive optics systems. These differ from previously published methods in that they do not require the use of a wave-front sensor in the experimental setup. The first method permitted the training of a control system by use of a previously calibrated wave-front generator and a single photodetector and used modal wave-front measurements based on orthogonal mirror deformation modes. The second method used an interferometer with fringe analysis and phase unwrapping to measure the phase of the wave fronts generated by the deformable mirror. With each system we were able to measure the accuracy of Zernike modes produced by the mirror. The interferometer procedure employed pseudoinversion of the control matrix by singular-value decomposition, permitting the removal of modes likely to cause saturation of electrode voltages. We described a technique that provides insight into this inversion process as a compromise between inclusion of Zernike modes and the magnitude of elements in the matrix pseudoinverse. Experimental measurements showed excellent correspondence with the calculations.

The choice of characterization scheme would normally be dictated by availability of hardware. As the interferometric method described in this paper is based on standard optics laboratory components, it is an attractive method, especially if a wave-front sensor is not available. In comparison with the wave-front generator method, the interferometric method also permits more-detailed analysis of the performance of the DMs.

Appendix A. Zernike Polynomials

Various definitions and indexing schemes for Zernike polynomials have been used by a number of authors. In this paper we use the normalization of Noll[29] and a single index i that is explained by Neil *et al.*[18] This single index is interchangeable with the dual indices m and n that represent the azimuthal and radial orders, respectively. The Zernike polynomials are defined as

Table 1. Zernike Polynomials

Index				
i	n	m	$Z_i(r, \theta)$	Aberration Term
1	0	0	1	Piston
2	1	1	$2r\cos(\theta)$	Tip
3	1	−1	$2r\sin(\theta)$	Tilt
4	2	0	$\sqrt{3}(2r^2 - 1)$	Defocus
5	2	2	$\sqrt{6}r^2\cos(2\theta)$	Astigmatism
6	2	−2	$\sqrt{6}r^2\sin(2\theta)$	Astigmatism
7	3	1	$2\sqrt{2}(3r^3 - 2r)\cos(\theta)$	Coma
8	3	−1	$2\sqrt{2}(3r^3 - 2r)\sin(\theta)$	Coma
9	3	3	$2\sqrt{2}r^3\cos(3\theta)$	
10	3	−3	$2\sqrt{2}r^3\sin(3\theta)$	
11	4	0	$\sqrt{5}(6r^4 - 6r^2 + 1)$	Spherical (1st)
12	4	2	$\sqrt{10}(4r^4 - 3r^2)\cos(2\theta)$	
13	4	−2	$\sqrt{10}(4r^4 - 3r^2)\sin(2\theta)$	
14	4	4	$\sqrt{10}r^4\cos(4\theta)$	
15	4	−4	$\sqrt{10}r^4\sin(4\theta)$	
16	5	1	$2\sqrt{3}(10r^5 - 12r^3 + 3r)\cos(\theta)$	
17	5	−1	$2\sqrt{3}(10r^5 - 12r^3 + 3r)\sin(\theta)$	
18	5	3	$2\sqrt{3}(5r^5 - 4r^3)\cos(3\theta)$	
19	5	−3	$2\sqrt{3}(5r^5 - 4r^3)\sin(3\theta)$	
20	5	5	$2\sqrt{3}r^5\cos(5\theta)$	
21	5	−5	$2\sqrt{3}r^5\sin(5\theta)$	
22	6	0	$\sqrt{7}(20r^6 - 30r^4 + 12r^2 - 1)$	Spherical (2nd)

$$Z_n^{\,m}(r,\ \theta) = \begin{cases} m < 0, & \sqrt{2}R_n^{\,-m}(r)\sin(-m\theta) \\ m = 0, & 0 \\ m > 0, & \sqrt{2}R_n^{\,m}(r)\cos(m\theta) \end{cases}; \tag{A1}$$

$$R_n^{\,m}(r) = \sqrt{n+1}\sum_{s=0}^{(n-m)/2} \times \frac{(-1)^s(n-s)!}{s!\,[(n+m)/2 - s]!\,[(n-m)/2 - s]!}\, r^{n-2s},$$

where the indices m and n are restricted to the conditions $n \geq |m|$ and where $n - |m|$ is even. The first 22 polynomials are listed in Table 1.

Appendix B. Phase Unwrapping

From the fringe pattern analysis we obtain a wrapped phase function ψ that is equivalent to the desired phase function ϕ modulo 2π. In Ref. 25 it was shown that ϕ can be obtained from ψ in a least-squares-error sense by solution of the equation

$$\nabla^2\phi = \nabla^2\psi. \tag{B1}$$

We outline here the steps that we employed to solve this equation based on an algorithm described in Ref. 25.

Let us represent the wrapped and unwrapped phase functions in terms of values on a $(N + 1) \times (N + 1)$ square grid as $\psi_{i,j}$ and $\phi_{i,j}$, where the indices i and j represent row and column numbers, respectively. We create the periodic function $\tilde{\psi}_{i,j}$ by reflecting $\psi_{i,j}$ about the lines $i = N$ and $j = N$ such that $\tilde{\psi}_{i,j}$ is defined on a $2N \times 2N$ grid. We then calculate the discrete derivatives of $\tilde{\psi}_{i,j}$ in the x and y directions as

$$\Delta_{i,j}^{\,x} = \mathcal{W}(\tilde{\psi}_{i+1,j} - \tilde{\psi}_{i,j}), \tag{B2}$$

$$\Delta_{i,j}^{\,y} = \mathcal{W}(\tilde{\psi}_{i,j+1} - \tilde{\psi}_{i,j}).$$

The wrapping operator $\mathcal{W}(f)$ wraps all values of the argument f into the range $-\pi < f \leq \pi$ by adding or subtracting integer multiples of 2π. We then define the function

$$\tilde{\rho}_{i,j} = (\Delta_{i,j}^{\,x} - \Delta_{i-1,j}^{\,x}) + (\Delta_{i,j}^{\,y} - \Delta_{i,j-1}^{\,y}). \tag{B3}$$

This function is equivalent to the right-hand side of Eq. B1. We can therefore write a discrete approximation to Equation (B1) as

$$(\tilde{\phi}_{i+1,j} - 2\tilde{\phi}_{i,j} + \tilde{\phi}_{i-1,j}) + (\tilde{\phi}_{i,j+1} - 2\tilde{\phi}_{i,j} + \tilde{\phi}_{i,j-1}) = \tilde{\rho}_{i,j}, \tag{B4}$$

where we define $\tilde{\phi}_{i,j}$ by reflecting the unknown function $\phi_{i,j}$ in the same manner used to obtain $\tilde{\psi}_{i,j}$. Taking the two-dimensional FT of both sides of Eq. (B4) leads to the equation

$$\Phi_{m,n} = \frac{P_{m,n}}{2\cos(\pi m/N) + 2\cos(\pi n/N) - 4}, \tag{B5}$$

where $\Phi_{m,n}$ and $P_{m,n}$ are the two-dimensional FTs of $\tilde{\phi}_{i,j}$ and $\tilde{\rho}_{i,j}$, respectively. The inverse FT of $\Phi_{m,n}$ yields $\tilde{\phi}_{i,j}$, and the desired function $\phi_{i,j}$ is easily extracted as the subregion of $\tilde{\phi}_{i,j}$ for which $0 \leq i \leq N$ and $0 \leq j \leq N$.

Some additional steps were required for reliable performance of the phase-unwrapping algorithm. We note that in our case the wrapped phase $\psi_{i,j}$ was meaningfully defined only over the area of the array that corresponds to the circular DM aperture. The values of $\psi_{i,j}$ outside this region carried no useful information but contained phase residues that gave rise to discontinuities in the calculated unwrapped phase $\phi_{i,j}$ within the region of interest. It was therefore necessary to mask the values of $\psi_{i,j}$ outside the aperture. To confine any remaining residues to the square boundary of the phase function, away from the circular aperture edge, we filled the masked region with values taken from pixels at the edge of the aperture. Pixel values were simply repeated along a row (or column) from the pixel at the edge of the aperture to the closest boundary. This process did not, of course, affect the values of the calculated phase within the aperture.

This research was partially funded by the British Council (Japan) and the Research Councils UK in the Basic Technology scheme. M. Booth is a Royal Acad-

emy of Engineering/Engineering and Physical Sciences Research Council research fellow.

References

1. R. P. Grosso and M. Yellin, "The membrane mirror as an adaptive optical element," J. Opt. Soc. Am. **67,** 399–406 (1977).
2. G. V. Vdovin and P. M. Sarro, "Flexible mirror micromachined in silicon," Appl. Opt. **34,** 2968–2972 (1995).
3. G. V. Vdovin, P. M. Sarro, and S. Middelhoek, "Technology and applications of micromachined adaptive mirrors," J. Micromech. Microeng. **9,** R8–R20 (1999).
4. M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "Adaptive aberration correction in a confocal microscope," Proc. Natl. Acad. Sci. USA **99,** 5788–5792 (2002).
5. L. Sherman, J. Y. Ye, O. Albert, and T. B. Norris, "Adaptive correction of depth-induced aberrations in multiphoton scanning microscopy using a deformable mirror," J. Microsc. **206,** 65–71 (2002).
6. P. N. Marsh, D. Burns, and J. M. Girkin, "Practical implementation of adaptive optics in multiphoton microscopy," Opt. Express **11,** 1123–1130 (2003).
7. D. Dayton, J. Gonglewski, S. Restaino, J. Martin, J. Phillips, M. Hartman, S. Browne, P. Kervin, J. Snodgrass, N. Heimann, M. Shilko, R. Pohle, B. Carrion, C. Smith, and D. Thiel, "Demonstration of new technology MEMS and liquid crystal adaptive optics on bright astronomical objects and satellites," Opt. Express **10,** 1508–1519 (2002).
8. T. Ota, S. Kawata, T. Sugiura, M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "Dynamic axial-position control of a laser-trapped particle by wave-front modification," Opt. Lett. **28,** 465–467 (2003).
9. T. Ota, T. Sugiura, S. Kawata, M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "Enhancement of laser trapping force by spherical aberration correction using a deformable mirror," Jpn. J. Appl. Phys. **42,** L701–L703 (2003).
10. E. Theofanidou, L. Wilson, W. J. Hossack, and J. Arlt, "Spherical aberration correction for optical tweezers," Opt. Commun. **236,** 145–150 (2004).
11. F. Gonte, A. Courteville, and R. Dandliker, "Optimization of single-mode fiber coupling efficiency with an adaptive membrane mirror," Opt. Eng. **41,** 1073–1076 (2002).
12. E. J. Fernández and P. Artal, "Membrane deformable mirror for adaptive optics: performance limits in visual optics," Opt. Express **11,** 1056–1069 (2003).
13. L. Zhu, P. C. Sun, D. U. Bartsch, W. R. Freeman, and Y. Fainman, "Wave-front generation of Zernike polynomial modes with a micromachined membrane deformable mirror," Appl. Opt. **38,** 6019–6026 (1999).
14. C. Paterson, I. Munro, and J. C. Dainty, "A low cost adaptive optics system using a membrane mirror," Opt. Expr. **6,** 175–185 (2000).
15. R. K. Tyson and B. W. Frazier, "Microelectromechanical system programmable aberration generator for adaptive optics," Appl. Opt. **40,** 2063–2067 (2001).
16. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge U. Press, 1992).
17. M. A. A. Neil, M. J. Booth, and T. Wilson, "Dynamic wave-front generation for the characterization and testing of optical systems," Opt. Lett. **23,** 1849–1851 (1998).
18. M. A. A. Neil, M. J. Booth, and T. Wilson, "New modal wavefront sensor: a theoretical analysis," J. Opt. Soc. Am. A **17,** 1098–1107 (2000).
19. M. J. Booth, "Direct measurement of Zernike aberration modes with a modal wave front sensor," in *Advanced Wavefront Control: Methods, Devices, and Applications*, J. D. Gonglewski, M. A. Vorontsov, and M. T. Gruneisen, eds., Proc. SPIE **5162,** 79–90 (2003).
20. M. A. Vorontsov, "Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion," J. Opt. Soc. Am. A **19,** 356–368 (2002).
21. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Cambridge U. Press, 1975).
22. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," J. Opt. Soc. Am. **72,** 156–160 (1982).
23. W. M. Macy, "Two-dimensional fringe-pattern analysis," Appl. Opt. **22,** 3898–3901 (1983).
24. C. Roddier and F. Roddier, "Interferogram analysis using Fourier transform techniques," Appl. Opt. **26,** 1668–1773 (1987).
25. D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping* (Wiley, 1998).
26. E. S. Claflin and N. Bareket, "Configuring an electrostatic membrane mirror by leastsquares fitting with analytically derived influence functions," J. Opt. Soc. Am. A **3,** 1833–1839 (1986).
27. D. Dayton, S. Restaino, J. Gonglewski, J. Gallegos, S. McDermott, S. Browne, S. Rogers, M. Vaidyanathan, and M. Shilko, "Laboratory and field demonstration of a low cost membrane mirror adaptive optics system," Opt. Commun. **176,** 339–345 (2000).
28. M. Schwertner, M. J. Booth, M. A. A. Neil, and T. Wilson, "Measurement of specimen induced aberrations of biological samples using phase stepping interferometry," J. Microsc. **213,** 11–19 (2004).
29. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66,** 207–211 (1976).