Wave-front reconstruction from defocused images and the testing of ground-based optical telescopes

Claude Roddier and François Roddier

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822

Received November 19, 1992; revised manuscript received May 17, 1993; accepted June 2, 1993

A new method has been developed for testing the optical quality of ground-based telescopes. Aberrations are estimated from wideband long-exposure defocused stellar images recorded with current astronomical CCD cameras. An iterative algorithm is used that simulates closed-loop wave-front compensation in adaptive optics. Compared with the conventional Hartmann test, the new method is easier to implement, has similar accuracy, and provides a higher spatial resolution on the reconstructed wave front. It has been applied to several astronomical telescopes and has been found to be a powerful diagnostic tool for improving image quality.

1. INTRODUCTION

During the past 25 years new astronomical observatories have been built on high-altitude mountain sites that have excellent seeing conditions. In addition, progress has been made in reducing dome seeing effects by proper control of the telescope thermal environment. As a result, it is now found that the aberrations of optical telescopes often limit the angular resolution of the telescopes. Controlling large optics is not an easy task in the shop and is even less easy on a mountain site under observing conditions. The Hartmann test is the traditionally preferred technique. However, for telescope quality to be controlled on a permanent basis, there is a definite need for a simpler, but still highly accurate, optical testing method. Such a method is needed not only for maintaining proper telescope alignment but also for actively controlling the thin primary mirrors of large telescopes now under construction. It is shown here that one can quantitatively analyze the optical quality of a telescope by simply recording a small set of properly defocused stellar images with a CCD camera of good photometric quality. Image-processing algorithms have been developed for obtaining accurate estimates of the aberration terms as well as high-resolution maps of the wave-front errors.

Two different techniques must be distinguished. One technique, known as phase retrieval, is already widely used to control millimetric telescopes.[1] It has recently been successfully applied in the visible by several groups, including ourselves, in analyzing the aberrations of the Hubble Space Telescope.[2] The method works in the diffraction regime and requires taking monochromatic images of point sources either in focus or with a small amount of defocus. Like any other interferometric technique, it is sensitive to vibrations (the jitter of the Space Telescope was the main limitation) and to turbulence, which limits its application to ground-based telescopes at long wavelengths. On a good site such as Mauna Kea, we recently applied this technique to short exposures taken at 4 μm. We retrieved permanent telescope aberrations by averaging several reconstructed wave fronts.[3]

In this paper we describe a different technique that works with wideband long exposures taken in the visible with ground-based telescopes. Like the Hartmann test, it works in the geometrical optics regime and relies on long exposures for averaging out the effects of atmospheric turbulence. Best results are obtained with a large amount of defocus, well outside the so-called caustic zone. In this regime, intensity variations over the extrafocal image essentially reflect local changes in the wave-front total curvature (Laplacian). The observation of such images has long been known as a sensitive test for telescope alignment or mirror figure errors. It has sometimes been referred to as the eye-piece test or the inside-and-outside test. Surprisingly, there has been little attempt to extract quantitative information from such observations. In 1973 Behr[4] described the effect of coma and proposed the test as a means of aligning telescope optics. In a technical memorandum dated June 1980, Wilson[5] described also the effect of astigmatism and spherical aberration and gave simple formulas based on geometrical optics to estimate these aberrations from the location of the shadows produced by the edge of the pupil and its central obstruction. In 1988 one of us[6] showed that the defocused images contain information on both the wave-front Laplacian and the wave-front radial tilt at the edges. As a result, one can reconstruct the wave-front surface by solving a Poisson equation, using the edge tilt as a Neumann-type boundary condition. Compared with the Hartmann test, the technique has the advantages of simplicity, high throughput, and avoidance of calibration difficulties. In this paper we describe the latest algorithms that we developed to reconstruct the wave front accurately, together with the results of several different tests performed on different astronomical telescopes.

2. THEORY

The technique consists of recording the illuminations I_1 and I_2 in two out-of-focus beam cross sections on each side of the focal plane (Fig. 1, top). One records the illumination in plane P_1 at a distance l before the focal plane F and the illumination in plane P_2 at a similar distance l' after F. In the object space (Fig. 1, bottom), the recorded illumina-

0740-3232/93/112277-11$06.00

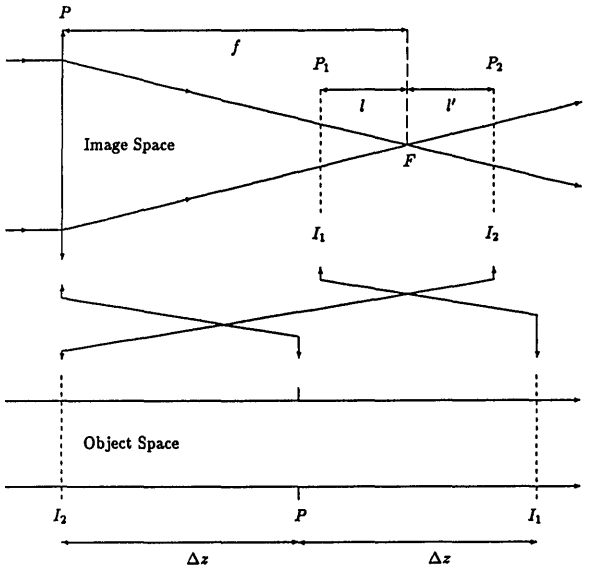

Fig. 1. In the image space (top), the recorded illuminations I_1 and I_2 appear as defocused stellar images. In the conjugate object space (bottom), they appear as defocused pupil images.

tions are conjugates of two cross sections of the incoming beam, one before the entrance pupil and one beyond the pupil. Hence I_1 and I_2 also can be considered defocused pupil images. In what follows, we assume that they are symmetrically defocused; that is, that the distances from the two beam cross sections to the pupil plane are the same and are equal to Δz. The theory is best described in terms of the irradiance transport equation,[7] which relates the illuminations I_1 and I_2 along the propagation path.[8-10] Assuming a paraxial beam propagating along the z axis, the irradiance transport equation states that

$$\partial I/\partial z = -(\nabla I \cdot \nabla W + I\nabla^2 W), \tag{1}$$

where $I(x, y, z)$ is the distribution of the illumination along the beam, $W(x, y)$ is the wave-front surface, and ∇ is the $\partial/\partial x$, $\partial/\partial y$ operator.

We apply this equation to the pupil plane ($z = 0$), where we assume the illumination to be fairly uniform and equal to I_0 inside the pupil and 0 outside. In this plane $\nabla I = 0$ everywhere but at the pupil edge, where

$$\nabla I = -I_0 \hat{n} \delta_c \tag{2}$$

Here δ_c is a linear Dirac distribution around the pupil edge and $\hat{n}$ is a unit vector perpendicular to the edge and pointing outward. Putting Eq. (2) into Eq. (1) yields

$$\frac{\partial I}{\partial z} = I_0 \frac{\partial W}{\partial n} \delta_c - I_0 P \nabla^2 W, \tag{3}$$

where $\partial W/\partial n = \hat{n} \cdot \nabla W$ is the wave-front derivative in the outward direction perpendicular to the pupil edge. $P(x, y)$ is a function equal to 1 inside the pupil and 0 outside. At the near-field, or geometrical optics, approximation the recorded illumination I_1 and I_2 are

$$I_1 = I_0 - \frac{\partial I}{\partial z}\Delta z, \tag{4}$$

$$I_2 = I_0 + \frac{\partial I}{\partial z}\Delta z. \tag{5}$$

The following quantity S, called the sensor signal, is computed:

$$S = \frac{I_1 - I_2}{I_1 + I_2} = \frac{1}{I_0}\frac{\partial I}{\partial z}\Delta z. \tag{6}$$

It should be noted that, since in practice images are recorded in the image space, one has to invert (rotate by 180 deg) the outside-focus image before computing S. Putting Eq. (3) into Eq. (6) gives

$$S = \left(\frac{\partial W}{\partial n}\delta_c - P\nabla^2 W\right)\Delta z. \tag{7}$$

The telescope objective reimages the beam cross section that is beyond the pupil plane at a distance l before the focal plane. According to Newton's law,

$$(\Delta z + f)l = f^2. \tag{8}$$

Hence

$$\Delta z = \frac{f(f - l)}{l}. \tag{9}$$

Putting Eq. (9) into Eq. (7) yields

$$S = \frac{f(f - l)}{l}\left(\frac{\partial W}{\partial n}\delta_c - P\nabla^2 W\right). \tag{10}$$

Equation (10) shows that the sensor signal consists of two terms. The first term is proportional to the wave-front radial slope at the pupil edge and is localized at the beam edge. The second term maps the wave-front Laplacian across the beam. Since these two terms do not overlap, one can measure them separately and reconstruct the wave-front surface by solving a Poisson equation, using the wave-front derivative normal to the edge as a Neumann-type boundary condition. However, Eq. (10) is only a first-order approximation valid for small Δz values, that is, highly defocused stellar images. The algorithm described in the following section uses the solution of the Poisson equation as a first-order solution that is further refined in an iterative process.

3. CLOSED-LOOP WAVE-FRONT RECONSTRUCTION TECHNIQUE

Earlier attempts to reconstruct the wave front from defocused images consisted of simply solving Eq. (10) with either direct integration[11] or fast Fourier transforms.[12] However, in the presence of large aberrations the reconstruction becomes inaccurate. We found that we could improve the accuracy of the wave-front reconstruction by iteratively compensating the effect of the estimated aberrations on the defocused images as in an active optics control loop. Residual aberrations are again estimated and compensated until the noise level is reached. The algorithm simulates the use of the wave-front sensing method in active optics. It generalizes a method, which we have described previously, for removing the effect of defocus and spherical aberration in the recorded images.[13] Compensation is done by geometrically distorting the images, as discussed below.

Let D, R, and f be the aperture diameter, the radius, and the focal length, respectively, of the telescope. We

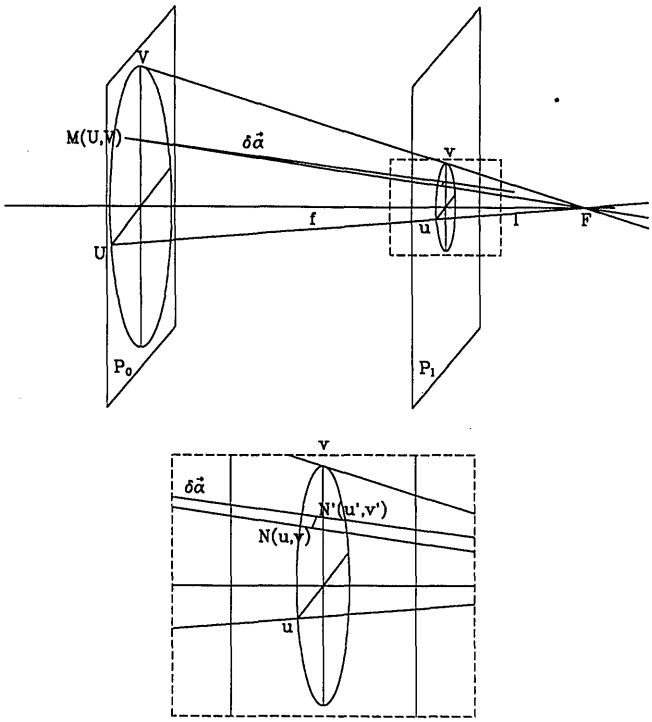

Fig. 2. Geometrical scheme showing the distortion introduced in plane P_l by a wave-front slope error in pupil plane P_0. The ray that would otherwise go through focus F and cross the plane P_l at point N(u,v) will in fact cross the plane at point N'(u,v).

denote by U,V the Cartesian coordinates in the pupil plane P_0 and by u,v the coordinates in a plane P_l normal to the optical axis, at a distance l from the focal plane (Fig. 2). We define reduced coordinates as $x = U/R$ and $y = V/R$ in the pupil plane P_0 and $x = u/r, y = v/r$ in plane P_l, where $r = lR/f$ is the radius of the beam cross section.

An aberration $W(x, y)$ at point $M(U,V)$ in the pupil plane produces a deviation of the optical ray $\delta\boldsymbol{\alpha}$ whose components are $-\partial W/\partial U$ and $-\partial W/\partial V$. The ray that would otherwise converge toward focus F and cross the plane P_l at point $N(u,v)$ will cross the plane at point $N'(u',v')$. For the inside-focus image, vector $\mathbf{NN'}$ is equal to $(f - l)\delta\boldsymbol{\alpha}$ with components

$$\mathbf{NN'}: -\frac{f-l}{R}\begin{vmatrix}\partial W/\partial x\\ \partial W/\partial y\end{vmatrix}.$$

In reduced coordinates, the displacement is given by vector $\delta\mathbf{N} = \mathbf{NN'}/r$ with components

$$\delta\mathbf{N}: -\frac{f(f-l)}{l}\frac{1}{R^2}\begin{vmatrix}\partial W/\partial x\\ \partial W/\partial y\end{vmatrix};$$

hence

$$\begin{cases}x' = x + C\partial W(x,y)/\partial x\\ y' = y + C\partial W(x,y)/\partial y\end{cases}, \tag{11}$$

with

$$C = -\frac{f(f-l)}{l}\frac{1}{R^2}. \tag{12}$$

A similar expression can be found for the outside-focus image. In practice, distance l is negligible compared with the focal distance f, and within a good approximation the displacements are the same at the same distance to focus. The signs are opposite because beyond the focal plane the coordinates are inverted. One achieves compensation by moving point $N'(x', y')$ back to location $N(x, y)$ according to Eq. (11).

If $I(x, y)$ is the intensity at point N and $I'(x', y')$ the intensity at point N', flux conservation requires that

$$I(x,y)d^2\mathbf{N} = I'(x',y')d^2\mathbf{N'} = I'(x',y')Jd^2\mathbf{N},$$

where $d^2\mathbf{N}$ is the elementary area and J the Jacobian of the transformation:

$$J = \begin{vmatrix}\partial x'/\partial x & \partial x'/\partial y\\ \partial y'/\partial x & \partial y'/\partial y\end{vmatrix};$$

that is,

$$I(x,y)/I'(x',y') = J = \begin{vmatrix}1 + C\partial^2W/\partial x^2 & C\partial^2W/\partial xy\\ C\partial^2W/\partial xy & 1 + C\partial^2W/\partial y^2\end{vmatrix}.$$

Hence image compensation also requires changing the intensity $I'(x, y)$ into

$$I(x,y) = I'(x',y')\left\{1 + C\left(\frac{\partial^2W}{\partial x^2} + \frac{\partial^2W}{\partial y^2}\right) + C^2\left[\frac{\partial^2W}{\partial x^2}\frac{\partial^2W}{\partial y^2} - \left(\frac{\partial^2W}{\partial xy}\right)^2\right]\right\}. \tag{13}$$

Because in our program the optical aberrations are expressed in terms of Zernike polynomials, we have computed the derivatives and Jacobians for the first 15 polynomials. The coefficients of the first- and second-order terms of the Jacobians are given in Table 1. The quantity ρ^2 is defined as $\rho^2 = x^2 + y^2$. A wave-front tilt correction is tantamount to a recentering of the images. A correction of defocus is tantamount to a rescaling. At each iteration the coefficients of a Zernike expansion are estimated and a given number of terms are compensated in the recorded images with use of the tabulated analytic expressions. The reconstructed wave front is obtained by addition of the compensated Zernike terms to the residuals.

4. PRACTICAL IMPLEMENTATION

First one must determine quite accurately the geometrical radii of the observed beam cross sections. This could be done qualitatively on the screen of a workstation or in a program that uses different kinds of threshold on the intensity. However, if the f ratio of the telescope and the positions of the images are sufficiently well known, it is preferable to infer the radii from the image distances to the estimated focal position. Before any computation, we approximately cocenter our images and occasionally rescale them to make the estimated geometrical radii equal (tip/tilt and defocus compensation). Then we subtract the two images and compute the sensor signal S [Eq. (6)]. We solve the Poisson equation numerically and get our first estimate of the aberrations by least-squares fitting Zernike polynomials to the reconstructed wave front.

Let Z_n be the coefficient of polynomial W_n. Coefficients Z_2, Z_3, and Z_4 express tip, tilt, and defocus errors, respectively. Compensation of these terms deserves

special attention because the error can be large and detrimental to the accuracy of the wave-front reconstruction process. First the defocus error can be translated into a new focus position that differs from the previous estimate by an amount

$$\Delta F = 16(f/D)^2 Z_4. \tag{14}$$

From this new focus position we determine new geometrical radii for the images. Similarly, from the tip/tilt error we determine new centers for the images. Using these new values, we then compensate the effects of the higher-order aberrations. First we select a vector $\mathbf{r}$ with integer coordinates (i, j), and we calculate the coordinates of vector $\mathbf{r}'$ as given by Eq. (11). The value of the illumination $I(\mathbf{r}')$ at point $\mathbf{r}'$ is interpolated as the weighted sum of the illuminations of the four neighboring points with use of a standard bilinear interpolation routine.[14] The value of the new illumination at point $\mathbf{r}(i, j)$ is given by Eq. (13). Images are then again recentered, rescaled, and subtracted to produce a new sensor signal S from which a new wave front is estimated.

From now on, two different algorithms are implemented. In one algorithm, called A, we correct 15 Zernike polynomials, and the whole set of aberration coefficients is determined at each iteration. In another algorithm, called B, used when aberrations (mainly coma and spherical) are large, we correct 22 polynomials, and the coefficients are determined sequentially in a specific order. The first set of coefficients to be determined consists of the terms in sin/cos $\theta(Z_{2,3}, Z_{7,8}, Z_{16,17})$. Only these terms are compensated at each iteration. When the algorithm has converged for these terms, their value is considered the final value. The second set of coefficients to be determined consists of the terms independent of $\theta(Z_4, Z_{11}, Z_{22})$. At each iteration both these terms and the sin/cos θ terms are compensated until convergence is observed. The third set consists of the terms in sin/cos 2θ $(Z_{5,6}, Z_{12,13})$. Only these terms and again the sin/cos θ terms are compensated. The final set consists of both the sin/cos 3θ and the sin/cos 4θ terms. The cos/sin θ terms are again compensated with them. We found that this procedure minimizes error propagation, especially the centering of (tip/tilt) errors. Simulations show that, when the aberrations are large, algorithm B gives the best results. When aberrations are small, the algorithms give the same results, but algorithm A runs faster.

A major source of uncertainty was found to come from the estimation of the coefficients Z_n by a least-squares fit over the reconstructed wave front. We believe that this problem is general and independent of the wave-front sensing method. If the wave front is quite smooth, the estimate of Z_n is fairly independent of the domain over which the least-squares fitting is done. But this is seldom the case: Most optical surfaces have large errors near the edges. Taking more or fewer of the edges into account may give dramatically different results. Cross talk occurs mainly between aberration terms with the same azimuthal frequency. For instance, the value of the third-order spherical aberration (which we call just spherical aberration) may differ appreciably, depending on whether fifth-order spherical aberration is also estimated. In our applications, although we sometimes correct only 15 Zernike terms, we have always fitted 22 terms, thus taking fifth-order spherical aberration into account. We found it difficult to estimate *a priori* the uncertainty in the reconstructed wave front. It depends on the position and size of the images and on the quality of seeing when the images were taken. In what follows, we call error the dispersion of our measurements on independent sets of defocused images. The accuracy of the method was investigated by means of both computer-simulated images and real images, with independent measurements of the telescope aberrations.

5. TESTS ON SIMULATED IMAGES

The linear mapping technique used to compensate image distortions produced by optical aberrations can also be used to simulate distorted images. However, the approximation is valid only outside the caustic zone. Close to focus, it breaks down. A more accurate way of simulating defocused images consists of computing diffraction patterns by means of fast Fourier transforms. We have used this method extensively to analyze data from the Space Telescope and have discussed the sampling conditions in the literature.[2] A way of checking our linear mapping algorithm was in fact to compare the results with those of a diffraction calculation. Within the range of validity of

Table 1. Analytic Expression of the First- and Second-Order Terms of the Jacobian of the Transformation

Polynomial Number	First Order	Second Order
$W2$ (x tilt)	0	0
$W3$ (y tilt)	0	0
$W4$ (defocus)	8	16
$W5$ (x astigmatism)	0	-4
$W6$ (y astigmatism)	0	-4
$W7$ (x coma)	$24x$	$108x^2 - 36y^2$
$W8$ (y coma)	$24y$	$108y^2 - 36x^2$
$W9$ (x triangular coma)	0	$-36\rho^2$
$W10$ (y triangular coma)	0	$-36\rho^2$
$W11$ (spherical)	$96\rho^2 - 24$	$144(12\rho^4 - 8\rho^2 + 1)$
$W12$ (x spherical astigmatism)	$48(x^2 - y^2)$	$36(8x^2 - 1)(1 - 8y^2)$
$W13$ (y spherical astigmatism)	$96xy$	$36[8\rho^2 - 16(x^2 - y^2)^2 - 1]$
$W14$ (x quadratic astigmatism)	0	$-144\rho^4$
$W15$ (y quadratic astigmatism)	0	$-144\rho^4$

Table 2. Example of Aberrations Retrieved from Simulated Data

	Zernike Aberration Terms									
Iteration	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	Z_{10}	Z_{11}
Introduced										
	0.5	0	0	0	0.8	0	1	0	0	0.3
Retrieved										
1	0.20	−0.25	−0.010	0.04	0.50	0.01	0.96	0.02	−0.02	0.270
2	0.29	0.13	−0.005	0.03	0.55	0.00	0.93	0.01	−0.01	0.290
3	0.30	0.14	−0.009	0.05	0.69	−0.02	0.93	0.01	−0.01	0.298
4	0.33	0.08	0.004	0.05	0.74	−0.03	0.94	0.02	−0.02	0.299
5	0.42	0.07	0.009	0.05	0.75	0.00	0.95	0.02	0.02	0.300
6	0.46	0.07	0.010	0.05	0.76	0.01	0.94	−0.01	0.02	0.301

the linear mapping, a good agreement was observed. The only difference is that, with the geometrical transformation, all the diffraction patterns (rings, fringes,...) disappear. A drawback of the diffraction calculation is that it applies only to monochromatic images. With white light, diffraction patterns also disappear. We have not attempted to extend our diffraction calculations to wideband sources.

Using either algorithm, we have simulated defocused images with known low-order aberrations ($Z_n = 0$ if $n >$ 15) and applied to them our wave-front-reconstruction technique. In all cases the aberrations were retrieved with an error less than 5%, the largest error being the error on the astigmatism. Spherical aberration and coma are retrieved with an error of 2–3% at the second or third iteration, while it takes a minimum of five iterations to get the astigmatism within the 5% range. Zernike polynomials for which the azimuthal frequency equals the radial degree ($Z_{5,6}, Z_{9,10}, Z_{14,15}$) have zero Laplacian and are therefore more difficult to retrieve from Laplacian measurements since all the information comes from the boundary conditions only. The term that gives the most trouble is the astigmatism, which, in spite of great care, is sensitive to centering errors. Only when all other aberrations are removed can the astigmatism be precisely estimated as a null measurement.

Next we added low-order Zernike aberration terms to maps of the residual phase errors observed on real telescopes. The results depend on the amplitude of these residuals. For the best telescopes the errors are of the same order of magnitude as on true simulations. For less-good telescopes the errors can grow as high as 10%. Table 2 shows an example of results that we obtained by adding arbitrarily given Zernike terms to higher-order aberrations estimated at the Cassegrain focus of the Canada–France–Hawaii Telescope (CFHT) on Mauna Kea. The high-order-aberration phase map was obtained from real data after removal of the 15 first Zernike terms. The numbers are given in micrometers on the wave front. A contour plot of the residual phase map is given in Fig. 3. It shows the intrinsic quality of the telescope together with the above-mentioned edge errors.

6. EXAMPLE OF RECONSTRUCTION FROM REAL DATA

On May 16, 1992, defocused stellar images were recorded with the New Technology Telescope (NTT) at the European Southern Observatory (ESO) in Chile. The primary mirror of this telescope has active supports. One can introduce known aberrations either by changing the distribution of the forces applied by the support or by moving the secondary mirror by a known amount. Here, as an example, we describe the analysis of defocused images obtained with a known independently calibrated coma of 585 nm rms. The telescope was first perfectly aligned by use of the ESO Shack–Hartmann sensor. Then the secondary mirror was decentered by a known amount, in this case 2.1 mm, which according to ray tracing should produce the indicated coma value.

The different steps of the data-reduction procedure are illustrated in Fig. 4 and Table 3. Different rows in the figure show the evolution of image compensation as iterations proceed. The top row shows the raw data, the second row the data after one iteration, the third row after six iterations. The first two columns display the defocused images (inside-focus and outside-focus). The third column displays the sensor signal S: the outside gray level represents zero signal, black is negative, and white is positive. The fourth column shows the different domains used in the wave-front reconstruction process and discussed below. For convenience we refer to these images as i_{mn}, m being the row number and n the column number. i_{11} is the raw inside-focus image, and i_{34} shows the domains of integration at the sixth iteration.

A coma is easily detected in a defocused image: the illumination varies linearly across the image, and the

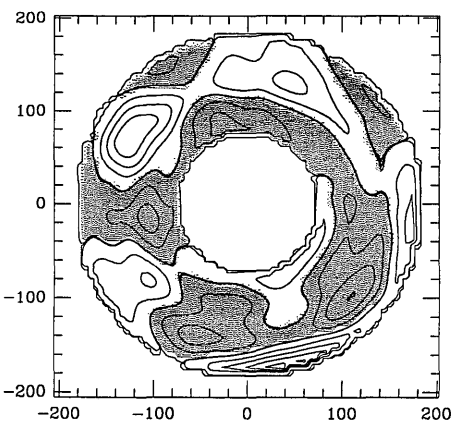

Fig. 3. CFHT Cassegrain focus phase map with 15 Zernike terms removed. The contour intervals are 0.02 μm. Gray areas are positive; white areas are negative.

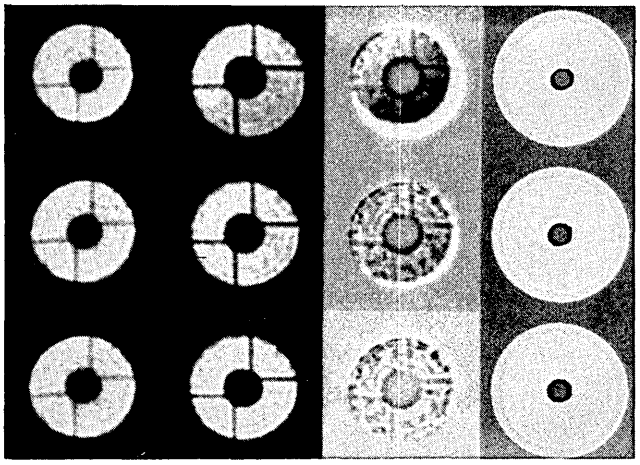

Fig. 4. Reduction of a coma term introduced in the ESO NTT. From left to right: inside-focus image, outside-focus image, normalized difference between the two images (sensor signal), and domain boundaries; from top to bottom: starting data, data after one iteration, and data after six iterations.

shadow of the secondary mirror is decentered toward the brightest part of the image. The effect is quite strong on images i_{11} and i_{12}, although the coma is only approximately 1 wave rms. This high sensitivity was achieved because the excellent quality of the NTT made it possible for images to be recorded quite close to focus. When these images were taken, the exact position of the telescope focus was not yet known, so their distances to focus, and therefore their sizes, are not identical, as one can see from both the outer and the inner boundaries. We obtained the signal i_{13} by subtracting these two images after arbitrarily centering them on the inner boundaries. It shows both the effect of defocus and coma.

The Poisson equation is solved with use of the iterative Fourier-transform method described in a previous paper.[12] On simulations it gives the most-accurate results. On the right-hand side of Fig. 4 are displayed the four domains used at each iteration. Inside the white domain the original signal is kept, in the light gray domain (outside ring) the outside normal derivative of the reconstructed aberration is put to zero, in the black domain (inner ring) the inside normal derivative is put to zero, and everywhere else (dark gray) the extrapolated signal is left. One obtains the boundaries of the white domain by putting a threshold at approximately one tenth of the sum of the intensities in the two images. The outer boundary is a circle 30% larger than the pupil size. The inner boundary is a circle 30% smaller than the diameter of the central obstruction.

We typically run four loops to obtain our first wave-front estimate. A least-squares fit gives a first set of Zernike coefficients Z_n. For comparison with independent estimates made simultaneously at ESO we use the rms coefficients as defined by Noll.[15]

Since we did not rescale the original images, in this stage of the reconstruction the encoder value for the focus position was assumed to be the half-sum of the encoder positions for the two images, in our case −3.4 encoder units, eu. From the Zernike defocus term we get a new estimate for the focus position, −3.283 eu. Aberrations are compensated accordingly, as described in Section 3. i_{21} and i_{22} are the compensated images. The effect of the coma is clearly smaller, and there is no evidence for defocus. However, both effects are still revealed on the signal i_{23}. One notices that some structures in the images get blurred when the difference is taken, because their locations do not match exactly. In i_{24} the domains are nearly circular.

The last row displays the same material after six iterations. One can see hardly any difference between images i_{21} and i_{31} or between i_{22} and i_{32}, and there is even less difference between i_{24} and i_{34}. However, one can clearly see a difference between i_{23} and i_{33}. The sensor signal becomes more contrasted, because now the small structures in the two defocused images match exactly.

Table 3 gives (in nanometers) the values of the aberrations estimated at each iteration together with the values measured by the ESO Shack–Hartmann sensor. Astigmatism clearly shows slower convergence. The agreement with the Shack–Hartmann sensor is certainly remarkable. Apart from coma, the largest absolute difference is 18 nm on astigmatism.

7. COMPARISON OF OUR RESULTS WITH OTHER INDEPENDENT ESTIMATES

In general it is difficult to compare results obtained from real data taken with astronomical telescopes with other independent measurements made on the same telescope, because very few telescopes have been tested on the sky.

Table 3. Aberrations Retrieved at the ESO NTT with a Known Coma Introduced, along with Results of the ESO Shack–Hartmann Sensor

Iteration	Focus (eu)	Spherical (nm)	Coma (nm)	Astigmatism (nm)	Triangular Coma (nm)	Quadratic Astigmatism (nm)
0	−3.40	37	454	217	43	10
1	−3.28	22	545	237	62	24
2	−3.25	27	556	244	73	29
3	−3.24	29	553	248	79	28
4	−3.24	29	543	232	77	28
5	−3.24	30	535	213	74	28
6	−3.24	30	534	205	70	29
7	−3.24	30	535	202	72	31
Shack–Hartmann		25	490	184	83	27
Coma Introduced			585			

Table 4. Mean Difference and rms Dispersion between the ESO Shack–Hartmann Sensor and the Out-of-Focus-Image Method

	Mean Difference over Six Measurements (nm)	rms (nm)
Spherical	10	10
Coma	72	21
Astigmatism	40	42
Triangular coma	25	29
Quadratic astigmatism	5	11
Total	30	35

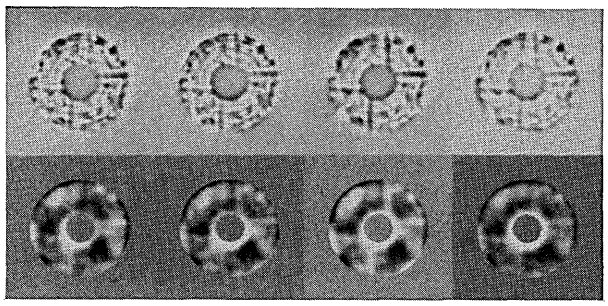

Fig. 5. Four independent estimates of the ESO NTT mirror figure. Top: residual signal; bottom: associated phase map (15 first Zernike terms removed).

In that regard, the ESO NTT is quite an exception. Here we present results obtained with various telescopes that confirmed the accuracy and the spatial resolution of the technique.

A. Comparison of Our Results with Those of the New Technology Telescope Shack–Hartmann Sensor

A detailed account of the May 1992 engineering run at the NTT will be given in another paper. Here we merely give the results of a comparison of our measurements with those of the ESO Shack–Hartmann sensor. The data consist of six independent sets of measurements, some with aberrations removed as effectively as possible, some others with independently known aberrations applied (spherical, coma, or triangular coma). Table 4 shows the mean difference between the two sets of measurements and the rms fluctuation, in nanometers. On the average, systematic differences are of the order of 30 nm and the dispersion of the order of 35 nm, which demonstrates that both methods are quite precise and give consistent results.

B. New Technology Telescope Residual Wave-front Error

As described in Section 4, our final phase map is obtained with 15 Zernike terms removed and represents what we call the residual wave-front errors. These small-scale errors are quite insensitive to optical misalignments and mirror-support problems. They essentially reflect mirror figure errors left by the polishing tool on either the primary or the secondary mirror. Figure 5 shows the final sensor signal and its associated phase map for four independent measurements made at the NTT. For images in the first column the telescope was perfectly tuned. For images in the second column a coma was introduced (the signal is the i_{33} image in Fig. 4). For images in the third column a spherical aberration, and in the fourth column a triangular coma, was introduced. Images in the last two columns are slightly more blurred, owing to poorer seeing conditions.

All these images are remarkably similar, giving us confidence that these wave-front errors are real. Two bright spots are visible near the edge in the first quadrant (upper right) of the phase maps. They appear to be associated with figure errors in the primary mirror. Similar errors are clearly seen in the primary-mirror figure obtained by Zeiss during the final optical shop tests.[16] A contour plot of the average phase map is given in Fig. 6 together with a cross section of the four independently obtained phase maps. The uncertainty in a wave-front reconstruction process is known to be at maximum at the pupil boundaries. Telescope mirrors often have poor edges, which makes the reconstruction even more difficult, because boundary conditions are affected by large edge-slope defects. For instance, the sensor signal displayed in image i_{33} (Fig. 4) shows an edge error in the second quadrant (upper left). A similar edge error would have appeared with a good mirror if the coma had been overestimated. The fact that we observe exactly the same bad edges with four different initial aberrations gives us confidence that the wave-front reconstruction process is accurate.

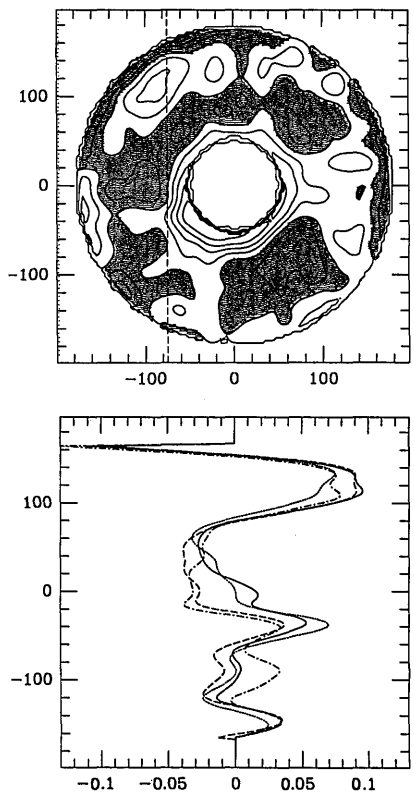

Fig. 6. ESO NTT mirror figure (15 first Zernike terms removed). Contour plot (top) and cross section (bottom). The cross section was taken on four independent phase maps, along the vertical dashed line shown on the contour plot. The contour plot shows the average phase map. The contour intervals are 0.04 μm. Gray areas are positive; white areas are negative.

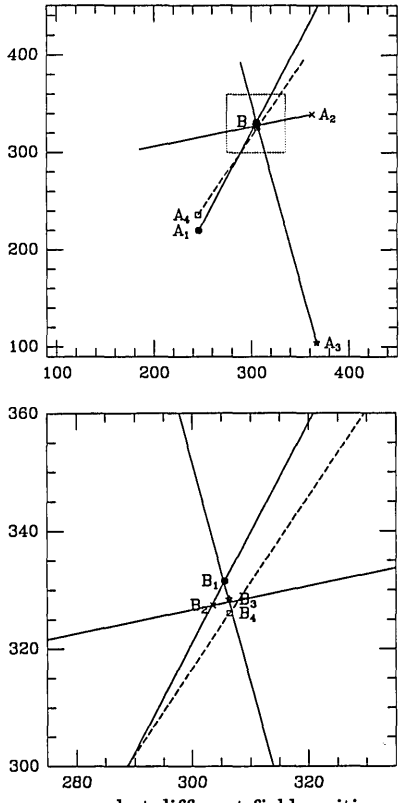

Fig. 7. Coma measured at different field positions at the prime foxus of the NASA/Infrared Telescope. Coordinates are given in CCD pixel units. One pixel corresponds to 0.728 arcsec on the sky. (a) For each field position A, a vector is drawn parallel to the coma, with its origin at point A and its length proportional to the amount of coma. For a perfect measurement all the extremities B should fall at the same point (telescope optical axis); (b) the enlarged portion of the field shows that all the extremities fall within 2.5 pixel of their center of gravity, which indicates a 45-nm peak uncertainty on the coma values.

C. Coma as a Function of Field Position

We now describe results obtained at the primary focus of the NASA Infrared Telescope Facility on Mauna Kea. Extrafocal images were taken at various field positions with a 512 × 512 pixel CCD camera. As is shown in Fig. 7(a), first we took pairs of defocused images centered at positions A_1, A_2, and A_3. Sometime later, after having moved the telescope and pointed it at several other stars, we did the same at position A_4. Figure 7 shows the result of the coma measurements. The coordinates are in CCD pixel units. One pixel corresponds to 0.728 arcsec. At each field position A, we determined a value of the coma and drew on Fig. 7(a) a vector parallel to the direction of the coma, with its origin at the observing position A and its length proportional to the amount of coma. All the vectors are expected to point toward the same point, the telescope optical axis. Since the amount of coma is expected to be proportional to the distance from this axis, one can theoretically scale the vectors so that all their extremities fall on the axis. Here, we arbitrarily scaled the vectors so that the extremity B_3 of the third vector falls at its intersection with the direction of the second vector. As shown in the enlarged portion of the field [Fig. 7(b)], the four vector extremities fall within 2.5 pixels of their center of gravity. The maximum deviation is 1% of the value of the coma at point A_3, that is, 45 nm. This is consistent with the 35-nm rms dispersion quoted above. The fact that the fourth coma measurement made later is consistent with the first three measurements is also a good indication of the telescope's mechanical stability. Such measurements could be made routinely for telescope alignment purposes.

D. Spherical Aberration as a Function of Focus Position

Moving the focus position away from its nominal position by moving the secondary mirror along the optical axis produces spherical aberration. For a Ritchey–Chrétien telescope, the amount of aberration introduced is given by the following expression[17]:

$$Z_{11} = \frac{m(m^2-1)}{6 \times 128 F^3 F_p}\left[1 + \frac{2}{(m-1)(m-\beta)}\right] dS, \qquad (15)$$

where

m is the secondary-mirror magnification,
F is the focal ratio at the final focus,
F_p is the focal ratio at the primary focus,
β is the backfocus from the primary-mirror vertex divided by the focal length of the primary mirror,
S is the distance from the focus of the primary mirror to the vertex of the secondary mirror.

This effect has been measured at the 4-m telescope of the Cerro–Tololo Interamerican Observatory (CTIO) in Chile. The results are shown in Fig. 8. We made two independent sets of measurements. For the first set the camera was set at the focus that corresponds to secondary-mirror position M (encoder value 108). The mirror was moved symmetrically at three different pairs of positions on each side of M. The image pairs were processed as usual. The resulting spherical-aberration values are indicated by stars above the letter M in Fig. 8. For the second

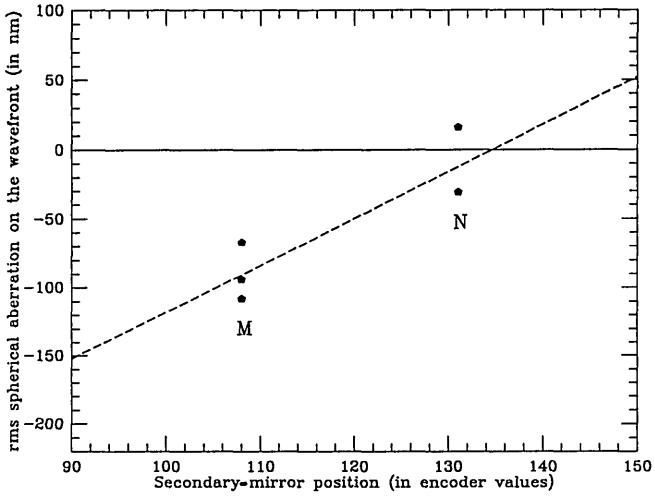

Fig. 8. Effect of the secondary-mirror position on spherical aberration. Spherical aberration was measured for two different focus position M and N on the CTIO 4-m telescope (stars). The horizontal scale shows the encoder value for the secondary-mirror position. The dashed line indicates the expected theoretical variation of the spherical aberration as a function of focus position. The best focus position (free from spherical aberration) is found to be at the encoder value 134.

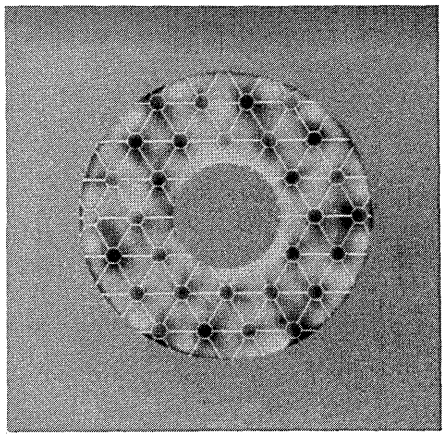

Fig. 9. Estimated primary-mirror figure of the Hale telescope after removal of 22 Zernike terms. White lines show the mirror honeycomb structure.

set the camera was set at the focus that corresponds to secondary-mirror position N (encoder value 131). The mirror was moved symmetrically at two different pairs of positions on each side of N. The spherical-aberration value for each pair is indicated by a star above the letter N. The dashed line is a linear fit with the theoretical slope given by Eq. (15). It shows that our results are consistent with theory. This measurement helped us to determine the best focus position on the CTIO 4-m telescope. Because of this effect, one may question results obtained when defocusing is done by moving the secondary mirror rather than the camera. At the primary focus of a telescope, defocusing can be achieved only by moving the camera. At a Cassegrain focus it is much easier to move the secondary mirror. As we have seen, this will introduce some amount of spherical aberration. However, the effect is a linear function of the mirror position and is opposite on each side of the focal plane. It is expected to cancel out when the difference between the two defocused images is taken. In our experiments we always were careful to take images with the secondary mirror at two positions as symmetrical as possible on each side of the focal plane. We did not find any systematic error that was due to the motion of the secondary mirror.

E. Mirror Honeycomb Structure of the Hale Telescope

Extrafocal images taken at the primary focus of the Hale telescope on Mount Palomar were given to us for analysis. The primary mirror of this telescope was the first telescope mirror with a honeycomb structure. This structure is represented by white lines in Fig. 9. Superimposed upon this pattern is the reconstructed mirror phase map after removal of the first 22 Zernike terms. The match is striking, giving us again confidence that small-scale wave-front errors are well retrieved in the wave-front reconstruction process. The amplitude of the bumps and dips on the wave front is typically 0.3 μm peak to valley. Figure 10 shows the associated point-spread function, that is, a stellar image that the telescope would produce at 0.5 μm if both seeing and the first 22 Zernike terms were removed by means of adaptive optics. The Strehl ratio is 0.3, and the intensities in the six spots are approximately one tenth of the central intensity.

F. Comparison of Our Estimates with Reported Encircled Energy Distribution

On several occasions we computed the point-spread function from our reconstructed wave fronts and compared the distribution of encircled energies with other independent estimates. Figure 11 displays the results of such comparisons made for the CFHT. First we compare our results with encircled energies given in the report on the primary-mirror acceptance tests.[18] The values in the report are derived from the results of Hartmann tests made at the optical shop before mirror delivery. Comparison is made with wave-front data obtained at the primary focus. To make a fair comparison, we removed from our reconstructed wave front three aberration terms that were not taken into account during the acceptance tests. These are coma, which is field dependent and reflects the distance from our images to the optical axis; spherical aberration, which was compensated by a null lens during the test; and astigmatism, which depends on the telescope orientation. Astigmatism was not seen during the tests

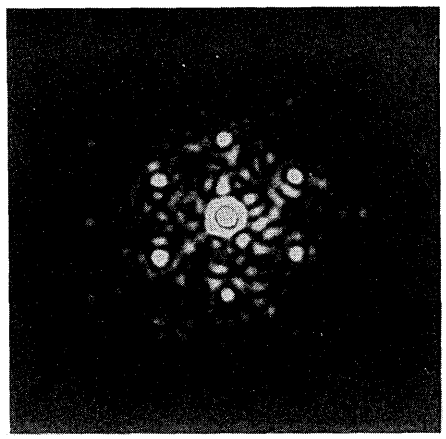

Fig. 10. Point-spread function associated with the phase map shown in Fig. 9.

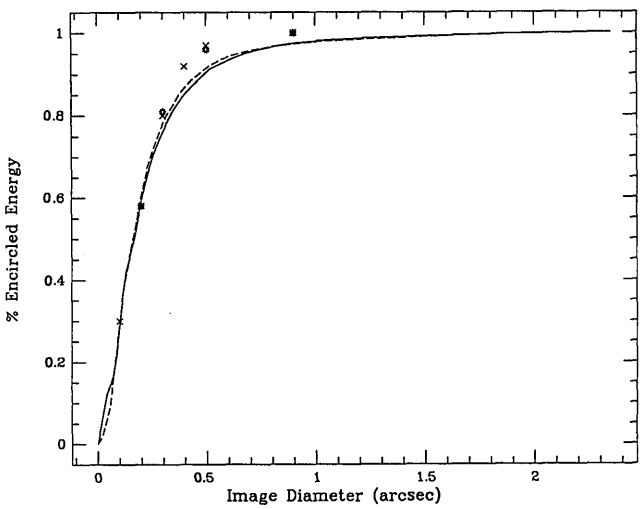

Fig. 11. Estimated encircled energies for the CFHT. The curves show our estimate from data taken at the prime focus with coma, astigmatism, and spherical aberrations removed (solid curve) and at the Cassegrain focus with coma removed (dashed curve). The experimental points are from a Shack–Hartmann spot diagram obtained during the primary-mirror acceptance test (asterisks) and later at the Cassegrain focus (crosses).

and is believed to be produced by the mirror supports. A second comparison was made with encircled energies derived from Hartmann tests made on the sky in 1983 at the Cassegrain focus.[19] Our data were also recorded at the Cassegrain focus. Only coma was removed from our reconstructed wave front, as it was for the Hartmann tests. It is clear from Fig. 11 that all these results are quite consistent. The spherical-aberration term observed at the primary focus appears to be removed effectively by the secondary mirror. A particularity of the CFHT is that this term can be adjusted by changing an air bag pressure in the back of the secondary mirror.[19] The largest discrepancy between the reported encircled energies and our estimates is in the wings. This is because the reported energies were obtained from geometrical spot diagrams, whereas ours were obtained from full diffraction calculations.

8. HOW TO TAKE OPTIMUM OUT-OF-FOCUS IMAGES

The technique of taking optimum out-of-focus images requires a science-grade CCD camera that is available on most astronomical telescopes. Since it works with broadband light, no filter is needed. The exposure time must be long enough to average out seeing effects but short enough to avoid any degradation that is due to telescope tracking errors. Experience shows that a 30-s exposure is usually a good choice. The stellar magnitude is dictated by the desire to obtain a good signal-to-noise ratio while staying well within the linear range of the CCD camera. On a 4-m telescope an 8-magnitude star taken from the Smithsonian Astrophysical Observatory (SAO) star catalog when it comes near zenith is appropriate.

The question then arises of how much defocus should be introduced for best results. To a first approximation, the defocused stellar image can be viewed as a blurred pupil image. The width of the blurring function is the width of the focal-plane image. Hence one can determine the width of the defocused image by multiplying the width of the focal plane image by the desired spatial resolution expressed in resolved wave-front elements per image diameter. For a given telescope, under a given seeing condition, the width of the required defocused image grows as the beam f ratio and may exceed the size of a standard CCD chip. For instance, at the $f/30$ Cassegrain focus of a 4-m telescope, a 1-arcsec seeing disk produces a 0.6-mm-diameter spot. Hence filling a 12-mm diameter CCD chip with a defocused image yields a maximum resolution of 20 independent wave-front elements across the pupil diameter. In the case of large f ratios, lenses can be used to reimage the beam cross sections onto the CCD camera with the desired magnification.

One must keep in mind that increasing the distance to focus increases the spatial resolution on the reconstructed wave front but decreases the sensitivity of the method. Hence the optimum distance depends also on the application. For telescope alignment, a smaller distance yields a higher sensitivity on the low-order terms such as coma. Moreover, there are fewer pixels to process in the image, which speeds up the computation. For example, the results on coma described in Subsection 7.C were obtained with images taken rather close to focus. Images shown in Fig. 4 were taken farther away to produce a good balance between sensitivity and resolution. Finally the primary-mirror figure of the Hale telescope shown in Fig. 9 was obtained with highly defocused images emphasizing spatial resolution. The images, which have 520 pixels across a diameter, have been smoothed. The mirror figure shown in Fig. 9 still has 340 pixels across a diameter, which corresponds to 1.5 cm/pixel on the mirror surface.

In choosing the distance to focus, one must also pay attention to another condition that has to be met. Our reduction process is valid only for images taken outside the caustic zone, that is, the zone inside which rays coming from different sampled pupil points intersect. One must take images far enough from the focal plane for this condition to apply. Unfortunately the size of the caustic zone depends on the aberrations of the telescope that we are supposed to measure, and no general rule can be given. The same situation arises for the classical Hartmann test. One can state that the distance to focus must be at least the same as that at which a Hartmann plate would be taken.

In some cases a particular telescope aberration dominates, and the dimensions of the caustic zone can be precisely stated. This is the case for extrafocal images taken at the (uncorrected) primary focus of a Ritchey–Chrétien telescope. In this case the primary mirror is hyperbolic and the primary focus is not stigmatic. It shows a strong negative spherical aberration. One can still record out-of-focus images to reconstruct the primary-mirror figure and measure its conical constant accurately.[13] We found that errors in the conical constant were a major source of aberration in most of the telescopes that we tested.[20] In the case of a negative spherical aberration, the caustic zone extends beyond the paraxial focus over a distance equal to three times the longitudinal aberration.[21] At this distance the diameter of the beam is eight times larger than the diameter of the circle of least confusion. Therefore the minimum defocus distance is the distance at which the diameter of the defocused image is eight times the diameter of the image at best focus. One must also allow for the seeing blur that must be added to the image diameter. In practice, a value at least twice as large will allow the algorithm to converge more easily. An even greater distance is desirable if one wishes to resolve any smaller feature on the reconstructed wave front.

Our experience is that in most cases satisfactory results are obtained when the telescope spider arms are clearly visible on the defocused images but the effect of the aberrations is only barely visible.

9. CONCLUSION

A new wave-front-sensing method was developed. It consists of reconstructing the wave front from defocused point-source images. The method works with broadband long-exposure stellar images taken by a ground-based optical telescope through the turbulent atmosphere and requires only a science-grade astronomical CCD camera. As originally proposed,[6] the wave-front reconstruction algorithm is based on the solution of a Poisson equation. The solution is further refined by means of an iterative

algorithm that simulates an adaptive optics control loop. This refinement considerably increases the dynamic range of the original method, allowing small aberrations to be retrieved in the presence of much larger ones.

The results were compared with those of a conventional Shack–Hartmann sensor. The accuracy appears to be similar. However, the new method is easier to implement, does not require the use of a flat reference wave-front, and generally provides a higher spatial resolution on the reconstructed wave front. It was successfully tested on several astronomical telescopes and was found to be a powerful diagnostic tool for telescope aberrations. The most frequently encountered aberrations are coma resulting from misalignment and spherical aberration resulting from inaccurate conical constants. Other aberration terms were often found to be related to mirror-support problems. Astigmatism and triangular coma (trefoil) were found to depend on the distance of the star to zenith and were related to primary-mirror-support adjustments. Higher-order terms were found to rotate with the secondary mirror and were related to the secondary-mirror support (mainly for infrared chopping secondaries). The information gathered in these tests is now being used on several telescopes to improve image quality.

A user-friendly interactive algorithm has been written, with instructions on how to use it. It is available on request to the authors. When the algorithm is automated on a SUN SPARC-2 work station, the computation time can be less than the time required for acquiring both images on a CCD camera. Application of this method to the active control of the primary-mirror supports and the alignment of large telescopes is now envisaged.

ACKNOWLEDGMENTS

This study was made possible with the help of several groups who provided us telescope time to take out-of-focus images or directly provided such images. We are particularly thankful to Guy Monnet and Derrick Salmon of the CFHT, Malcolm Northcott and Richard Baron of our institute, Mike Shao and Mark Colavita of the Jet Propulsion Laboratory, Lothar Noethe and Alain Gilliotte of the ESO, and Jack Baldwin and Brooke Gregory of the CTIO.

Note added in proof: Since this paper was accepted, a new version of the program has been written in which image distortions are compensated for by direct numerical differentiation rather than by use of the analytic expressions of Eqs. (11) and (13). The program, which now avoids the need for fitting Zernike polynomials, has been fully automated without any loss of accuracy.

REFERENCES

1. D. Morris, "Phase retrieval in the radio holography of reflector antennas and radio telescopes," IEEE Trans. Antennas Propag. **AP-33,** 749–755 (1985).
2. C. Roddier and F. Roddier, "Combined approach to Hubble Space Telescope wave-front distortion analysis," Appl. Opt. **32,** 2992–3008 (1993).
3. C. Roddier and F. Roddier, "New optical testing methods developed at the University of Hawaii: results on ground-based telescopes and Hubble Space Telescope," in *Advanced Optical Manufacturing and Testing II,* V. J. Doherty, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1531,** 37–43 (1991).
4. A. Behr, "A proposal for the alignment of large telescopes," Astron. Astrophys. **28,** 355–358 (1973).
5. R. Wilson, "Procedures and formulae for the adjustment of telescopes and analysis of their performance," memorandum, ESO Telescope Project Division (European Southern Observatory, Garching bei München, Germany, June 18, 1980).
6. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," Appl. Opt. **27,** 1223–1225 (1988).
7. F. Roddier, "Wavefront sensing and the irradiance transport equation," Appl. Opt. **29,** 1402–1403 (1990).
8. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am. **73,** 1434–1441 (1983).
9. N. Streibl, "Phase imaging by the transport equation of intensity," Opt. Commun. **49,** 6–10 (1984).
10. K. Ichikawa, A. Lohmann, and M. Takeda, "Phase retrieval based on the irradiance transport equation and the Fourier transform method: experiments," Appl. Opt. **27,** 3433–3436 (1988).
11. N. Roddier, "Algorithms for wave-front reconstruction out of curvature sensing data," in *Active and Adaptive Optical Systems,* M. A. Ealey, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1542,** 120–129 (1991).
12. F. Roddier and C. Roddier, "Wavefront reconstruction using iterative Fourier transforms," Appl. Opt. **30,** 1325–1327 (1991).
13. C. Roddier, F. Roddier, A. Stockton, A. Pickles, and N. Roddier, "Testing of telescope optics: a new approach," in *Advanced Technology Optical Telescopes IV,* D. L. Barr, ed., Proc. Soc. Photo.-Opt. Instrum. Eng. **1236,** 756–766 (1990).
14. W. H. Press, B. P. Flannery, S. A. Tenkolsky, and W. T. Vetterling, *Numerical Recipes in C* (Cambridge U. Press, Cambridge, 1988), p. 104.
15. R. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66,** 207–211 (1976).
16. R. N. Wilson, F. Franzia, P. Giordano, L. Noethe, and M. Tarenghi, "Active optics: the NTT and the future," Messenger (ESO) **53,** 1–7 (1988).
17. D. J. Schroeder, *Astronomical Optics* (Academic, San Diego, 1987), p. 113.
18. J. C. Fouéré and G. Ratier, "Report on the optical quality of the primary mirror," Rep. No. 78/222 (CFHT primary-mirror acceptance test report) (Canada–France–Hawaii Telescope Project Office, Meudon, France, 1978).
19. In "Modified F/8 secondary gives excellent images," CFHT Inform. Bull. 10 (Canada–France–Hawaii Telescope Office, Kamuela, Hawaii, 1984), p. 1.
20. F. Roddier, "Mirror aberration communication," Phys. Today **43**(11), 117 (1990).
21. D. Malacara, *Optical Shop Testing,* 2nd ed. (Wiley, New York, 1992), p. 750.