Reprinted from APPLIED OPTICS, Vol. 14, page 2622, November 1975

Use of an ac heterodyne lateral shear interferometer with real-time wavefront correction systems

J. C. Wyant

An analysis is performed to determine the accuracy with which an ac heterodyne lateral shear interferometer can measure wavefront aberrations if a white light extended source is used with the interferometer, and shot noise is the predominate noise source. The analysis shows that for uniform circular or square sources larger than a derived minimum size, the wavefront measurement accuracy depends only upon the radiance of the source and not upon the angular subtense of the source. For a 1-msec integration time, a 25-cm^2 collecting area, and a source radiance of 10 W/m^2-sr the rms wavefront error is approximately 1/30 wave, assuming the signal is shot noise limited. It is shown that for both uniform circular and square sources an optimum shear distance is approximately 1/2 the aperture diameter required to resolve the light source. Comments are made on the optimum shear for nonuniform radiance distributions.

Introduction

Recently much interest has been generated in so-called active or adaptive optics systems that measure and correct in real time the wavefront aberrations in an optical system.[1-5] The wavefront aberrations may result from either surface deformations in the optics or atmospherically introduced deformations in the wavefront. Active optics systems consist of three basic components: (1) a wavefront sensor; (2) a wavefront corrector, such as a flexible mirror; and (3) electronics that convert the wavefront sensor signal into the appropriate signal for the wavefront corrector. This paper is concerned with the wavefront sensor and is specifically directed toward the use of a wavefront sensor for measuring phase deformations introduced by the atmosphere in a plane near the telescope pupil; however, the ideas expressed should be considered general and not restricted to this one application.

Several wavefront sensors for use with active optics systems have been described previously.[2-6]. The wavefront sensor discussed in this paper, a lateral shear interferometer that uses ac heterodyne phase measurement techniques[7,8] is believed to have particular merit because of the loose requirements on both the temporal and spatial coherence of the light source for its operation. As shown previously,[9] if the lateral shear is proportional to wavelength, the effect of temporal coherence is minimized. As long as the object size is smaller than the isoplanatic patch for the atmosphere, the only effect of reduced spatial coherence of the light source is to reduce the fringe visibility as given by the Van Cittert Zernike theorem. Although reduced fringe visibility can reduce the signal to noise to a point where the wavefront cannot be determined with sufficient accuracy, for any given light source the shear can be adjusted to give good visibility fringes. Small shear always gives good visibility; however, as shown later in this paper, if the shear is too small it becomes difficult to determine accurately the wavefront. Clearly a tradeoff must be made between fringe visibility and shear.

Being able to use white light extended sources with a wavefront sensor is very important in atmospheric correction systems since the isoplanatic patch for the atmosphere is generally only a few sec of arc, and both the object of interest and the light source for the wavefront sensor must be within this isoplanatic patch. Any reduction in the light source requirements increases the number of situations where use can be made of active optics.

Wavefront Measurement Accuracy

The phase information obtained directly from a shearing interferogram is called the wavefront difference function. The wavefront is calculated from the wavefront difference functions obtained from two shearing interferograms in two different (often orthogonal) directions. Recently improved techniques have been derived for obtaining the wavefront from the wavefront difference function.[10-12] Below, an approximate relationship between the rms error in

The author is with the Optical Sciences Center, University of Arizona, Tucson, Arizona 85721.

Received 10 May 1975.

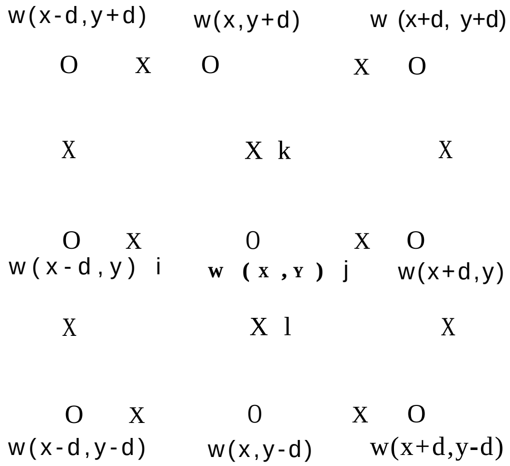

Fig. 1. Sample wavefront grid. X denotes points at which the wavefront difference function is measured, and O denotes points at which the wavefront is calculated.

the calculated wavefront and the rms error in the wavefront difference function is derived. For simplicity in the derivation it is assumed that the two shearing interferograms from which the wavefront difference is determined have the shear in orthogonal directions and that the wavefront is calculated and the wavefront difference is measured on a square grid. It is also assumed that the shear is much smaller than the smallest period of disturbance being measured.

The wavefront difference function is equal to the phase difference between two points on the wavefront separated by the shear distance. Equivalently, the wavefront difference function can be thought of as the average slope of the wavefront (in the direction of shear) over the shear distance times the shear distance.

Figure 1 shows a sample set of points in the wavefront being measured. The W's represent the wavefront values to be determined for the point designated by the symbol 0. The wavefront difference function ***V*** is measured at the points designated by the symbol X. The following four equations can be written

$$W(x,y) = W(x-d,y) + a_1, \quad (1)$$

$$W(x,y) = W(x+d,y) - a_2, \quad (2)$$

$$W(x,y) = W(x,y+d) + b_1, \quad (3)$$

$$W(x,y) = W(x,y-d) - b_2. \quad (4)$$

The *a*'s and *b*'s are determined from the wavefront difference function and would be exactly equal to the wavefront difference function if the shear distance were equal to the detector spacing. In this derivation the *a*'s and *b*'s can be taken to be equal to the wavefront difference function scaled by the ratio of the detector spacing *d* to the shear. If S_X and S_Y are the shear in the x and y directions, respectively,

$$a_1 = V_i(d/S_x), \quad (5)$$

$$a_2 = V_j(d/S_x), \quad (6)$$

$$b_1 = V_k(d/S_y), \quad (7)$$

$$b_2 = V_l(d/S_y). \quad (8)$$

The above four equations are precisely correct for any quadratic form of wavefront aberration such as defocus or astigmatism. For other aberrations the above equations are only approximately correct; however, as long as the wavefront is sampled at a sufficiently high spatial rate and also the shear is sufficiently small, Eqs. (5)-(8) can be used in the following analysis. Substituting Eqs. (5)-(8) into Eqs. (1)-(4) and adding the four resulting equations yield

$$W(x,y) = \tfrac{1}{4}\Big[W(x-d,y) + W(x+d,y) + W(x,y-d) + W(x,y+d) + \frac{d}{S_x}\{V_i - V_j\} + \frac{d}{S_y}\{V_k - V_l\}\Big]. \quad (9)$$

If $\Delta\theta$ is the rms error in determining the wavefront at each point on the grid, and it is assumed that the wavefront is determined to the same accuracy at each point, and $\Delta\phi$ is the rms error in measuring the wavefront difference function at each point, it follows from Eq. (9) that

$$\Delta\theta = (M\Delta\phi)/\sqrt{3}, \quad (10)$$

where

$$M = \frac{d}{\sqrt{2}}\left(\frac{1}{S_x^2} + \frac{1}{S_y^2}\right)^{1/2}. \quad (11)$$

If S_x is equal to S_y, ***M*** reduces to

$$M = d/S. \quad (12)$$

Equation (10) gives the relationship between the rms error in the calculated wavefront and the rms error in the measured wavefront difference function for the situation where the point at which the wavefront is determined is connected by four paths as shown in Fig. 1. It follows easily from a more general analysis that if a point is connected by N paths, the $\sqrt{3}$ in Eq. (10) would be replaced with a $(N-1)^{1/2}$. Therefore, as long as each point at which the wavefront is determined is connected by two or more integration paths, the rms error in the wavefront is equal to or less than the rms error in the measurement of the wavefront difference function multiplied by the factor *M* given in Eq. (11). A compromise must be made between the amount of shear, the number of measurements made of the wavefront difference function, and the accuracy of the wavefront determination.

In the absence of noise the output signal *i* from the ac heterodyne interferometer can be written as

$$i = i_s[1 + \gamma \sin(\omega t + \phi)], \quad (13)$$

where ω is the modulation frequency, ϕ is the phase being measured, and γ is the fringe visibility, which for an interferometer in which the two interfering beams have the same intensity is determined by the coherence of the light source. A common technique for measuring ϕ is to measure the time difference between the zero crossing of the sinusoidal signal given in Eq. (13) and the zero crossing of a reference signal. Thus, the error in the phase measurement results from error in measuring the correct time of the zero crossing. It can be shown that for zero crossing measurement of phase the rms phase error $\Delta\theta$ is given by

$$\Delta\theta = 1/(S/N), \quad (14)$$

where S/N is the current signal to noise.[13,14]

For low light level conditions, a better method for measuring phase is to divide the period of the sinusoidal signal into three or more segments and count the number of photons detected for each segment. If the period is divided into four segments the integrated current (or photon count) can be converted into phase information as shown below.

Let P equal the number of photoelectrons detected during one period T. Therefore P, which is equal to $i_s T/e$, is given by

$$P = \int_{-T/8}^{T/8} \frac{idt}{e} + \int_{T/8}^{3T/8} \frac{idt}{e} + \int_{3T/8}^{5T/8} \frac{idt}{e} + \int_{5T/8}^{7T/8} \frac{idt}{e} \quad (15)$$

$$= A + B + C + D, \quad (16)$$

where e is the electronic charge.

The integration limits were selected as shown to simplify the derivation. If Eq. (13) is substituted into Eq. (15) it can be shown that

$$A = P\left(\frac{1}{4} + \frac{\sqrt{2}\gamma}{2\pi}\sin\phi\right), \quad (17)$$

$$B = P\left(\frac{1}{4} + \frac{\sqrt{2}\gamma}{2\pi}\cos\phi\right), \quad (18)$$

$$C = P\left(\frac{1}{4} - \frac{\sqrt{2}\gamma}{2\pi}\sin\phi\right), \quad (19)$$

$$D = P\left(\frac{1}{4} - \frac{\sqrt{2}\gamma}{2\pi}\cos\phi\right). \quad (20)$$

It then follows that

$$\tan\phi = \frac{A - C}{B - D}. \quad (21)$$

Now that the phase has been determined by integrating the current (counting detected photons), the error in the phase measurement can be determined as follows[15]:

From Eq. (21)

$$\phi = \tan^{-1}\left(\frac{A - C}{B - D}\right). \quad (22)$$

The variance of ϕ is given by[16]

$$(\Delta\phi)^2 = \left[\frac{B - D}{(B - D)^2 + (A - C)^2}\right]^2 [(\Delta A)^2 + (\Delta C)^2] + \left[\frac{A - C}{(B - D)^2 + (A - C)^2}\right]^2 [(\Delta B)^2 + (\Delta D)^2]. \quad (23)$$

Using the fact that

$$(\Delta A)^2 + (\Delta C)^2 = A + C = P/2 \quad (24)$$

and

$$(\Delta B)^2 + (\Delta D)^2 + B + D = P/2, \quad (25)$$

the rms wavefront error becomes

$$\Delta\phi = \pi/(2\gamma\sqrt{P}). \quad (26)$$

If the measurement is performed for several periods, P becomes the total number of photons detected.

From Eqs. (10) and (26) it follows that

$$\Delta\theta = \frac{M\pi}{2\sqrt{3}\gamma\sqrt{P}} = \frac{(0.9)M}{\gamma\sqrt{P}} \quad (27)$$

Since in the derivation of Eq. (10) it was assumed that the wavefront difference function has the same error in x and y, the shear should be selected such that y is the same for both x and y.

If N is the radiance of the light source that subtends a solid angle Ω from the telescope, T is the transmission of the atmosphere and optics, including the interferometer, ρ is the responsivity of the photodetector, A_0 is the telescope aperture seen by each detector, and τ is the total integration time, P is given by

$$P = N\Omega T A_0 \tau\rho/e, \quad (28)$$

where e is the electronic charge. Equation (27) then becomes

$$\Delta\theta = \frac{0.9M}{\gamma}\left(\frac{e}{N\Omega T A_0 \tau\rho}\right)^{1/2} \quad (29)$$

It should be stressed that while $\Delta\theta$ does depend upon the area of the telescope aperture seen by each detector, $\Delta\theta$ does not depend upon the total telescope aperture. Increasing the telescope aperture does not improve the wavefront measurement accuracy.

By substituting M given by Eq. (11) into Eq. (29), it follows that if the diameter of the effective aperture of each detector is equal to the detector spacing, in which case all the available light is being used, $\Delta\theta$ is independent of the detector spacing and effective detector area. However, care must be taken not to make the effective detector area so large as to have a sufficiently large wavefront difference function variation across the detector to reduce appreciably the signal. Likewise, the integration time τ must be short enough to temporarily resolve the atmospheric disturbance.

It is of interest to calculate $\Delta\theta$ for a couple of examples. The first example will be for the two planets Venus and Mars, and the second example will be for rectangular sunlit diffuse objects of various values of angular subtense. A sunlit diffuse object is of particular interest since if sunlight reflected by existing satellites can be used as the light source for a wavefront sensor in an atmospheric compensation system, large portions of the sky can be viewed with resolution never before achieved from the earth. In all calculations d the detector spacing will be taken as 0.05 m, A_0 the telescope aperture area seen by each detector will be taken as $(0.05\ \text{m})^2$, the transmittance T will be taken as 0.15, the integration time τ will be taken as 1 msec, and ρ the spectral responsivity of the photodetector will be taken as 0.05 A/W.

In calculating $\Delta\theta$ for Venus and Mars it will be assumed that the planets are uniform disks. Therefore, the coherence or visibility function γ is found from the Van Cittert Zernike theorem to be given by

$$\gamma = \frac{2J_1(\pi S\alpha/\lambda)}{\pi S\alpha/\lambda}, \quad (30)$$

where α is the angular subtense of the planet as seen from the earth. Equation (29) then becomes

$$\Delta\theta = \frac{0.9d}{\lambda J_1(\pi S\alpha/\lambda)}\left(\frac{e\pi}{NTA_0\rho\tau}\right)^{1/2}, \quad (31)$$

where use has been made of the fact that $\Omega = \pi\alpha^2/4$, and since for a circular source the shear can be the same for both the x and y directions, $M = d/S$. It is seen that the shear S should be selected to maximize $J_1(\pi S\alpha/\lambda)$. The maximum value of $J_1(x)$ is approximately 0.58, which occurs for $x \cong 1.8$. Therefore, the shear should be selected so that

$$S = 1.8\lambda/\pi\alpha \quad (32)$$

That is, the shear should be approximately one-half of the aperture width required to resolve the source. Equation (31) then becomes

$$\Delta\theta = (1.1 \times 10^{-9})\frac{d}{\lambda(NTA_0\rho\tau)^{1/2}}. \quad (33)$$

It is interesting to note that $\Delta\theta$, the rms error in the wavefront determination, depends only upon the source radiance and not upon the solid angle subtended by the circular source. However, it must be realized that for this to be true, Eq. (32) must be satisfied. This is equivalent to saying that the source must be large enough to be resolved with an aperture approximately twice a reasonable shear distance. The smallest subtense of Venus and Mars is on the order of 4.8×10^{-5} rad and 1.8×10^{-5} rad,[17] respectively, and hence Eq. (32) shows that at a wavelength of 0.6 µm the optimum shear is 0.7 cm for Venus and 1.9 cm for Mars. Therefore, for Venus and Mars there is no problem satisfying Eq. (32).

The radiance of Venus and Mars is approximately 80 W/m^2-sr and 10 W/m^2-sr,[17,18] respectively, and for these values of radiance Eq. (33) yields an rms wavefront measurement error of approximately 0.07 rad (or $\cong\lambda/90$) and 0.21 (or $\cong\lambda/30$), respectively.

If the light source is the diffuse reflection from a uniform rectangular sunlit object the visibility γ is given by

$$\gamma = \frac{\sin(\pi S_x\alpha_x/\lambda)}{\pi S\alpha_x/\lambda}, \quad (34)$$

where the shear is S_x, and in the shear direction the angular subtense is α_x. If the angular subtense of the object is α_x and α_y (i.e., $\Omega = \alpha_x\alpha_y$), and if the visibility is required to be the same for x shear and y shear, it follows that $S_x\alpha_x = S_y\alpha_y$. Using this fact Eq. (29) becomes

$$\Delta\theta = \left(\frac{\alpha_x}{\alpha_y} + \frac{\alpha_y}{\alpha_x}\right)^{1/2}\frac{0.9d\pi}{\lambda\sin(\pi S_x\alpha_x/\lambda)}\left(\frac{e}{2NTA_0\rho\tau}\right)^{1/2}. \quad (35)$$

It is noted that the optimum value of shear S, is such that $\sin(\pi S_x\alpha_x/\lambda) = 1$, i.e.

$$S_x = (n + \tfrac{1}{2})\lambda/\alpha_x, \quad n = 0, 1, 2, \cdots, \quad (36)$$

that is, the shear width should be 1/2, 3/2, 5/2, . . . , the aperture width required to resolve the source. It is seen that the wavefront phase error is the same for many different values of shear. The phase error is given by

$$\Delta\theta = (8 \times 10^{-10})\left(\frac{\alpha_x}{\alpha_y} + \frac{\alpha_y}{\alpha_x}\right)^{1/2}\frac{d}{\lambda(NTA_0\rho\tau)^{1/2}}. \quad (37)$$

For a square object Eq. (37) becomes

$$\Delta\theta = (1.1 \times 10^{-9})\frac{d}{\lambda(NTA_0\rho\tau)^{1/2}}. \quad (38)$$

It is interesting that for a square source the wavefront measurement accuracy is also independent of the angular subtense of the object, and in fact to within two significant figures the measurement accuracy is the same as for a circular source. For the wavefront measurement accuracy to be independent of angular subtense Eq. (36) must be satisfied, which means, as for the circular disk case, the source must be sufficiently large that it can be resolved with an aperture approximately twice a reasonable shear distance. If it is assumed that the maximum allowable shear is 10 cm, for visible light the object must subtend a minimum angle of approximately 0.5 sec of arc.

Using a radiance value of 100 W/m^2-sr, which is believed to be a reasonable radiance for a sunlit diffuse satellite, Eq. (38) gives a value of A8 equal to 6×10^{-2} rad $\cong \lambda/100$. Even if the source is 25 times larger in the shear direction than in the direction perpendicular to shear, the rms phase error is approximately 1/30 wave.

The fact that the measurement accuracy is independent of object size for either a square source or a circular source larger than a minimum size is not as surprising as it first appears. For a square or circular source the rms wavefront error A0 is given by

$$\Delta\theta = \frac{0.9d}{S\gamma\sqrt{P}}. \quad (39)$$

Let d, the detector spacing, remain constant. If the radiance of the source is kept constant as the source size increases, the number of detected photons is proportional to the solid angle subtended by the source, i.e., $\sqrt{P}$ is proportional to α, the angular subtense of the source. For constant fringe visibility Eqs. (31) and (35) show that the shear is inversely proportional to α. Therefore, $\Delta\theta$ is independent of angular subtense of the source for both square and circular light sources.

So far this paper has considered only uniform rectangular or circular light sources. What happens for more complicated sources? To obtain a precise answer, each source must be treated separately. However, a general answer is that objects with structure can be better than uniform objects. This can be seen as follows: The visibility function is equal to the Fourier transform of the radiance distribution of the source. Objects with structure will contain more high frequency information than objects without structure, i.e., the fringe visibility or coherence function will have higher peaks at reasonably large shears for structured objects than for uniform objects. For a simple example, consider a source that has a periodic radiance distribution similar to that of a Ronchi ruling. The coherence distribution is given by the Fourier transform of the Ronchi ruling, which is also a periodic function. If the shear is equal to the distance between the peaks of the coherence function, good contrast fringes are obtained as illustrated in Ref. 9.

Conclusions

In conclusion, the results in this paper indicate that at least for the situation where shot noise is the predominate noise source, an extended white light source can be used as the light source for an ac heterodyne lateral shear interferometer. The optimum shear depends upon the size and radiance distribution of the light source. For a uniform circular source the optimum shear distance is equal to approximately one-half of the minimum diameter aperture that will resolve the source. For uniform square sources, the optimum shear distance is 1/2, 3/2, 5/2, . . .) the minimum diameter aperture that will resolve the source. The wavefront measurement error is inversely proportional to the square root of the radiance of the light source, and for uniform circular or square sources the error is independent of the angular subtense of the source as long as the subtense is sufficiently large that the source can be resolved with an aperture twice the width of the shear distance. To within two significant figures, both square and circular uniform sources of the same radiance give the same measurement accuracy. Using a 1-msec integration time, a 25-cm^2 collecting area, and a planet such as Mars or Venus as the light source gives a measurement accuracy on the order of 1/30 to 1/90 wavelength, while square (or circular) objects having a radiance of 100 W/m^2-sr give an accuracy of approximately 1/100 wave. The accuracy obtained using nonuniform sources depends upon the radiance distribution. The accuracy is improved if the object has periodic structure.

The author acknowledges beneficial conversations with V. P. Bennett, J. Hardy, M. P. Rimmer, and R. Vyce. Portions of this work were performed while the author was at the Itek Corporation, 10 Maguire Road, Lexington, Massachusetts 02173. The work was partially supported by the U.S.A.F. under contract F33657-72-C-0605.

References

1. H. W. Babcock, Publ. Astron. Soc. Pac. 65,229 (1953).
2. R. A. Muller and A. Buffington, J. Opt. Soc. Am. 64, 1200 (1974).
3. J. Hardy, J. Feinleib, and J. C. Wyant, in Digest of *OSA Topical Meeting on Optical Propagation through Turbulence* (Optical Society of America, Washington, D.C., 1974).
4. W. N. Peters, R. A. Arnold, and S. Gowrinathan, Appl. Opt. 13,1785 (1974).
5. W. B. Bridges, P. T. Brunner, S. P. Lassera, T. A. Nussmeier, T. R. O'Meara, J. A. Sanguinet, and W. P. Brown, Jr., Appl. opt. 13,291(1974).
6. L. J. Golden, R. V. Shack, and P. N. Slater, NASA Final Report for contract NAS827863 (1974).
7. R. Crane, Appl. Opt. 8,538 (1969).
8. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, Appl. Opt. 13,2693 (1974).
9. J. C. Wyant, Appl. Opt. 13,200 (1974).
10. J. B. Saunders and R. J. Bruening, Astron. J. 73,415 (1968).
11. M. P. Rimmer, Appl. Opt. 13,623 (1974).
12. R. V. Shack, J. Opt. Sot. Am. 64,1362A (1974).
13. E. D. Sunde, *Communication Systems Engineering Theory* (Wiley, New York, 1969), p. 120
14. M. I. Skolnik, *Introduction to Radar Systems* (McGraw-Hill, New York, 1962), p. 463.
15. The author. acknowledges a beneficial conversation with V. P. Bennett on the derivation.
16. A. G. Worthing and J. Geffner, *Treatment of Experimental Data* (Wiley, New York, 1943), p. 208.
17. S. P. Wyatt, *Principles of Astronomy* (Allyn and Bacon, Boston, 1964), p. 205.
18. *RCA Electra-Optics Handbook Technical Series EOH-I1* (Harrison, N.J., 1974), p. 68.