Curvature sensing from a single defocused image in partially coherent light

Ervin Goldfain

Welch Allyn Inc., Skaneateles Falls, NY 13153

ABSTRACT

Curvature sensing is an intensity-based technique for wavefront reconstruction using two defocused images located on the opposite sides of the focal plane. It requires either one detector placed at two consecutive axial locations or a dual path with a pair of detectors from which the sensor signal is obtained. The method yields a sensitivity comparable to that of the Hartmann test in the adjustment and evaluation of ground-based optical telescopes. We introduce the analytical framework underlying the function of a curvature sensor which operates from a single defocused image. A series of twin images is computed from the propagation law of the mutual intensity along the optical axis. The polynomial decomposition of the wavefront allows retrieval of Zernike coefficients by means of the standard least-squares algorithm. The paper concludes with a review of image sampling requirements and a discussion on the signal-to-noise ratio.

Keywords: wavefront sensing, telescope optics, adaptive optics, passive ranging, optical aberrations, Zernike polynomials.

1. INTRODUCTION

Wavefront retrieval based on curvature sensing has recently gained attention in telescope optics and optical testing applications as an effective alternative to the Hartmann-based methods [1,2] and iterative phase-retrieval methods [3,4]. Curvature sensing estimates the wavefront from two defocused images, typically located on opposite sides of the focal plane, by making use of the irradiance-transport theory [5,6]. The wavefront Laplacian ($\nabla^2 W$) and its normal gradient along the pupil edge are computed from:

$$S = d.(d - L).(1/L).[(\partial W/\partial n).\delta_c - P.\nabla^2 W] \tag{1}$$

where W (Xp,Yp) represents the wavefront in pupil coordinates (fig. 1):

$$Xp = (d/L).x, \quad Yp = (d/L).y \tag{2}$$

x,y are image coordinates in the two out-of-focus planes, δ_c is the Dirac delta function along the pupil edge, d is the paraxial image conjugate , L the axial defocus, P the pupil function (1 inside the pupil and 0 outside), n represents the outward normal direction to the pupil edge and S is the sensor signal given by:

$$S = [I_1(x,y) - I_2(x,y)]/[I_1(x,y) + I_2(x,y)] \tag{3}$$

in which $I_1(x,y)$ and $I_2(x,y)$ stand for the image intensity distributions[6,7]. Recording the two images requires either a single detector positioned at consecutive locations along the optical axis or a pair of detectors arranged in a dual path. We propose employing a single-detector/single-image setup and taking advantage of the propagation law for the mutual intensity along the optical axis[8]. To make the derivation universal, we assume that the wavefront emerges from a quasihomogeneous partially coherent source with a slowly varying amplitude distribution.

A number of approaches for solving the curvature sensing equation (1) have been developed and implemented[6,9,10]. We use herein the standard wavefront expansion in Zernike polynomials because: a) it leads to a solution that is tractable in terms of primary and higher order aberrations and b) enables retrieval of the expansion coefficients via well known routines of linear analysis, in particular, the least-squares fitting algorithm.

The outline of the paper is as follows: in the first section we obtain the digitized sensor signal (3) over a discrete (x,y) pixel array. The next section deals with the Zernike representation of the wavefront and the Gramm-Schmidt orthogonalization procedure which is applied to remove undesired numerical artifacts[11,12]. The least-squares solution of (1) is obtained in the last section, where image sampling requirements and the signal-to-noise ratio are also reviewed.

The described method may have potential benefits for a wide range of applications in optical metrology (such as testing of smooth aspherical refractive or reflective surfaces with large deviations from a reference profile[13]) and real-time passive ranging[14,15] (such as autofocus and machine vision systems).

2. DIGITAL REPRESENTATION OF THE SENSOR SIGNAL

For the sake of clarity, we consider only the one-dimensional imaging case. Referring to the setup depicted in fig.1, let $\{X\}$ $\{u\}$, $\{x_1\}$ and $\{x_2\}$ represent the coordinate set of the partially coherent source, its paraxial image and the two observation planes. Let the position of the detector coincide with plane $\{1\}$. Take "i" and "j" to denote two arbitrary points lying in the plane of the source. If the source is quasihomogeneous, its mutual intensity may be expressed as[16]:

$$J(X_i , X_j) = I(X_a).m(\Delta X) \tag{4}$$

in which $m(\Delta X)$ represents the complex coherence factor, $I(X_a)$ is the slowly varying intensity distribution and:

$$\Delta X = X_j - X_i$$
$$X_a = (X_i + X_j)/2 \tag{5}$$

Similar expressions hold for all three image planes, under the assumption that imaging does not change the quasihomogeneous property of the mutual intensity described by (4). Let $\{Xp\}$ denote the coordinate set of the exit pupil:

$$Xp = d.x_1/L \tag{6}$$

and let the source, exit pupil and all image planes be digitized using pixel arrays with the following spacing constants, respectively:

$$\delta X, \delta Xp, \delta u, \delta x_1, \delta x_2 \tag{7}$$

Take H to represent the linear extent of the source, D the lens diameter and λ the average wavelength. We assume that the coherence area of the light incident on the lens is much smaller than the lens area, that is[17]:

$$H.D \gg \lambda.d_0 \tag{8}$$

The mutual intensity reaching the paraxial image plane is defined by[18]:

$$J_i(u_i , u_j) = J_i(i.\delta u, j.\delta u) = \sum_i \sum_j J(i.\delta X, j.\delta X).K(i.\delta u , i.\delta X).K^*(j.\delta u , j.\delta X) \tag{9}$$

where "i" and "j" are point locators (i = 1,2...N ; j= 1,2...M) and K represents the amplitude spread function of the system:

$$K(i.\delta u , i.\delta X) = \exp\{j.\pi.(\lambda.d_0)^{-1} .(i.\delta u)^2 \}.\exp\{j.\pi.(\lambda.d)^{-1}.(i.\delta X)^2\}.[1/(\lambda^2.d_0.d)] .P(i.\delta Xp) \tag{10}$$

in which:

$$P(i.\delta Xp) = \sum_i P(i.\delta Xp).\exp\{-j.2.\pi.(\lambda.d)^{-1}.[i^2.(\delta u +d.d_0^{-1}.\delta X).\delta Xp] \tag{11}$$

with $P(i.\delta Xp)$ describing the complex pupil function:

$$P(i.\delta Xp) = P.\exp[-j.W(i.\delta Xp)] \tag{12}$$

By assumption (8), we can write:

$$P(i.\delta Xp).P^*(j.\delta Xp) \approx P^2 \tag{13}$$

which implies that the image mutual intensity is independent of pupil aberrations. The image intensity distribution corresponds to setting $u_i = u_j$ in (9), i.e.:

$$I(u_i) = J_i(u_i, u_i) \tag{14}$$

It follows from (4) and (5) that the image complex coherence factor is given by:

$$m(\Delta u) = J_i(u_i, u_j)/I(u_i) \tag{15}$$

in which:

$$\Delta u = u_j - u_i = (j - i).\,\delta u \tag{16}$$

According to the generalized van Cittert-Zernike theorem, the mutual intensity reaching plane {2} by forward radiation from the image plane , is found to be[19]:

$$J_2(x_{2,i}, x_{2,j}) = J_2(i.\delta x_2, j.\delta x_2) = k(x_{2a}).\exp(-j.\psi_{ij}).(\lambda.L)^{-2}.\ \sum_i \sum_j I(u_a).\exp[j.2.\pi.(\lambda.L)^{-1}.(\Delta x_2.u_a)] \tag{17}$$

if the following condition holds true:

$$L > (2.H.l_c/\lambda) \tag{18}$$

where l_c is the coherence length of the source. In (17) we have used the following notations:

$$x_{2a} = [(i + j).\delta x_2]/2$$

$$u_a = [(i + j).\delta u]/2$$

$$\Delta x_2 = (j - i).\,\delta x_2 \tag{19}$$

$$k(x_{2a}) = \sum_i \sum_j m(\Delta u).\exp[j.2.\pi.(\lambda.L)^{-1}.x_{2a}.\,\Delta u]$$

$$\psi_{ij} = \pi.(\lambda.L)^{-1}.\delta x_2^2.(j^2 - i^2)$$

From (17) we derive the intensity distribution at plane {2}:

$$I_2(x_{2i}) = J_2(x_{2,i}, x_{2,i}) \tag{20}$$

Replacing (20) and the detected intensity :

$$I_1(x_{1i}) = I_1(i.\delta x_1) \tag{21}$$

in (3), we obtain the digitized sensor signal as:

116

$$S(i) = [I_1(x_{1i}) - I_2(x_{2i})]/[I_1(x_{1i}) + I_2(x_{2i})]$$ (22)

The sequence of steps leading to (22) is summarized below:

$$\text{source } [\mathbf{J}(X_i, X_j), \mathbf{m}(\Delta X)] \Rightarrow \text{ paraxial image plane } [I(u_i)] \Rightarrow \text{ plane } \{2\}[I_2(x_{2i})] \Rightarrow \text{ sensor signal } [S(i)]$$ (23)

If the source characteristics are unknown (which is often the case in practice), a similar approach may be developed by first computing the mutual intensity in the paraxial image plane from the recorded mutual intensity and complex coherence factor in plane {1}. The procedure is greatly simplified in incoherent light because the complex coherence factor reduces to a Dirac delta function[8]. The sequence of steps leading to (22) then becomes:

$$\text{plane } \{1\}[I_1(x_{1i})] \Rightarrow \text{ paraxial image plane } [I(u_i)] \Rightarrow \text{ plane } \{2\}[I_2(x_{2i})] \Rightarrow \text{ sensor signal } [S(i)]$$ (24)

3. WAVEFRONT EXPANSION IN ZERNIKE MODES

Zernike polynomials provide a convenient representation of the wavefront which is customarily used in optical testing and interferometry[11,20]. The standard wavefront expansion over the one-dimensional array is:

$$W(i.\delta Xp) = W(i) = \sum_k C_k.Z_k(i)$$ (25)

where $k = 1,2...K$ is the mode index. Zernike polynomials form a complete set of orthogonal functions over the continuous unit circle, but fail to maintain orthogonality over a discrete array of points[11], i.e.:

$$\sum_i Z_k(i).Z_{k'}(i) \neq \delta_{kk'}.$$ (26)

It can be shown that this condition may lead to undesired crosstalk between modes and noise amplification in applications involving wavefront fitting by the least-squares method[11,21]. The Gramm-Schmidt procedure is applied to restore the orthogonality property by constructing a new set of functions G_u such that:

$$\sum_i G_u(i).G_{u'}(i) = \delta_{uu'}.$$ (27)

$$W(i) = \sum_u B_u.G_u(i)$$ (28)

in which $u = 1,2...K$. The relationship between the original Zernike polynomials and the new ones is determined by[11]:

$$Z_k(i) = \sum_u A_{ku}.G_u(i)$$ (29)

where:

$$A_{ku} = \sum_i Z_k(i).G_u(i) \qquad u \neq k$$

$$A_{kk}^2 = \sum_i [Z_k(i)]^2 - \sum_u A_{ku}^2$$ (30)

It can be shown that the original expansion coefficients (C_k) and the new ones (B_u) are related through the following expressions:

$$C_K = B_K/A_{KK} \tag{31}$$

$$C_k = (B_k - \sum u \; A_{uk}.C_u)/ \; C_{kk} \tag{32}$$

in which summation is taken from u = k+1 to K and k = 1,2....(K-1). By differentiating (28) we obtain:

$$\partial W/\partial n = \partial W(i)/\partial i = \sum u \; B_u.\partial G_u(i)/\partial i \tag{33}$$

$$\nabla^2 W(i) = \sum u \; B_u.\nabla^2 G_u(i) \tag{34}$$

with u = 1,2...K. Direct substitution of (33) and (34) in (1) yields the linear system of equations :

$$S(i) = \sum u \; B_u \; Q_u(i) \tag{35}$$

where B_u are the unknowns expansion coefficients and:

$$Q_u(i) = d.(d - L).(1/L).[(\partial G_u(i)/\partial i).\delta_c - P.\nabla^2 G_u(i)] \tag{36}$$

The linear system may be cast in matrix form as:

$$\mathbf{S = Q.B} \tag{37}$$

where $\mathbf{Q}$ denotes a rectangular matrix with N rows and K columns (NxK), $\mathbf{S}$ is a N-dimensional vector and $\mathbf{B}$ a K-dimensional vector. Note that $Q_u(i)$ represents the sum of two components, the first one assuming nonzero values inside the pupil (i.e. $\nabla^2 G_u(i) \neq 0$) and the second one assuming nonzero values on the pupil edge ($\partial G_u(i)/\partial i \neq 0$). For applications involving retrieval of only wavefront tilt, defocus and astigmatism, the sensor signal $\mathbf{S}$ is exclusively collected from the pupil boundary due to the vanishing Laplacian associated with these modes[10].

The object of the next section is to derive the generic least-squares solution of (37) for the expansion coefficients B_u and to evaluate the contribution of noise to the wavefront retrieval process.

4. THE LEAST-SQUARES SOLUTION AND WAVEFRONT ESTIMATION ERRORS

The system described by (37) may be solved using a variety of numerical methods such as direct ones (Gauss elimination, the method of band matrices) or iterative matrix algorithms[22,23]. If N>K, the linear system is overdetermined (the number of sampling points is larger than the number of Zernike modes) and it is suitable for the least-squares algorithm. Enhanced resolution may be gained by computing the sensor signal in multiple sampling planes along the optical axis and averaging the individual outcomes.

Multiplying (37) from the left by $\mathbf{Q}^T$ ($\mathbf{Q}$ transpose) we obtain the normal form:

$$(\mathbf{Q}^T.\mathbf{Q}).\mathbf{B} = \mathbf{Q}^T.\mathbf{S} \tag{38}$$

To extract the least-squares solution from (38), one needs to multiply the right-hand side by the inverse of $(\mathbf{Q}^T.\mathbf{Q})$. This standard procedure is valid only if the normal equations are not ill-conditioned, that is, if the square matrix $(\mathbf{Q}^T.\mathbf{Q})$ is not singular. It can be shown that (38) is well-conditioned if $G_u(i)$ are normalized to have zero mean and the wavefront expansion is taken with the piston term removed (minimum norm solutions)[23,24].

Assuming that the normal equations are or have been brought to a well-conditioned form, the standard least-squares solution is given by:

$$B = (Q^T.Q)^{-1}. Q^T.S \tag{39}$$

There are two basic sources of error in the wavefront retrieval process.The first one relates to the finite sampling density of the continuous image received at plane {1}. For an aberration-free lens, the image is bandwidth limited by diffraction, i.e. contains spatial features with frequencies less than the cutoff value:

$$f_c = 2.NA/\lambda \tag{40}$$

where NA is the lens numerical aperture in image space. Undersampling occurs when the sampling frequency is less than the Nyquist frequency ($2.f_c$). Thus , to avoid undersampling, δu must satisfy:

$$\delta u \leq \lambda/(4.NA) \tag{41}$$

which defines the resolution limit in the paraxial image plane. The corresponding object and image resolution limits in {1} and {2} may be obtained from (41) by appropiate scaling. Aliasing may occur if:

$$\delta u \geq \lambda/(4.NA) \tag{42}$$

The second source of error is created by noise in the detection and computation of the sensor signal. We assume that the sensor signal contains randomly distributed additive noise:

$$S = S_0 + \delta S$$

$$\delta S = \delta S_d + \delta S_c \tag{43}$$

where δS_d and δS_c are the detection and computation noise contributions and S_0 the noise-free signal. The first order variation of the sensor signal (3) in vector form is:

$$\delta S = (\partial S/\partial I_1).\delta I_1 + (\partial S/\partial I_2).\delta I_2 = [2/(I_1 + I_2)^2].(I_2.\delta I_1 - I_1.\delta I_2) \tag{44}$$

in which $\delta I_{1(2)}$ are the intensity noise vectors. Let the resulting estimation error in the expansion coefficients matrix be δB, that is:

$$B = B_0 + \delta B \tag{45}$$

with B_0 representing the noise-free coefficient matrix. Substituting (43) and (45) in (38) yields the normal set of equations :

$$(Q^T.Q).\delta B = Q^T.\delta S \tag{46}$$

and the standard least-squares solution:

$$\delta B = (Q^T.Q)^{-1}. Q^T.\delta S \tag{47}$$

From (28), the wavefront error induced by δB is:

$$\delta W(i) = \sum u \; \delta B_u.G_u(i) \tag{48}$$

It can be shown[23] that for signal variations δS that are equal in magnitude (δS_0) and uncorrelated, the mean-square wavefront error is given by:

$$E = (1/N^2) . \sum i \langle \delta W(i)^2 \rangle = (\delta S_0)^2 . tr \, (\mathbf{Q}^T.\mathbf{Q})^{-1} \tag{49}$$

where tr $(\mathbf{Q}^T.\mathbf{Q})^{-1}$ stands for the sum of the diagonal elements. On the other hand, the noise-free wavefront is:

$$W_0(i) = \sum u \, B_{o.u}.G_u(i) \tag{50}$$

with $B_{o.u}$ representing the noise-free components of matrix $\mathbf{B_0}$. From (44), (49) and (50) one can define the "signal-to-noise" ratio for the wavefront retrieval as:

$$SNR(i) = W_0(i)/E = W_0(i)/\{[2/(I_1 + I_2)^2.(I_2.\delta I_1 - I_1.\delta I_2)]^2. \, tr \, (\mathbf{Q}^T.\mathbf{Q})^{-1}\} \tag{51}$$

or:

$$SNR(i) = W_0(i)/\{[2/(I_1 + I_2)^2.I_1.I_2.(\delta I_1/I_1 - \delta I_2/I_2)]^2. \, tr \, (\mathbf{Q}^T.\mathbf{Q})^{-1}\} \tag{52}$$

The above formula relates the signal-to-noise ratio of the wavefront retrieval to the signal-to-noise ratio of the detection and computation processes:

$$SNR_d = I_1/\delta I_1 \tag{53}$$

$$SNR_c = I_2/\delta I_2 \tag{54}$$

We may use (52) to determine the combination of values for SNR_c and SNR_d that maximizes SNR. The general conditions defining the SNR extremum are:

$$\partial SNR/\partial SNR_d = 0$$

$$\partial SNR/\partial SNR_c = 0$$

$$[\partial^2 SNR/\partial(SNR_d)^2].[\partial^2 SNR/\partial(SNR_c)^2] - [\partial^2 SNR/\partial(SNR_d).\partial(SNR_c)]^2 > 0 \tag{55}$$

$$\partial^2 SNR/\partial(SNR_d)^2 < 0$$

Next we illustrate the above procedure with a simple numerical example. Assume that the ideal detected image is a unit step function and let it be degraded by the following range of uniform noise values:

$$\delta I_{1,q} = .1 + .01.q \tag{56}$$

where the range variable (q) is specified by :

$$q = 1,2,...50 \tag{57}$$

The corresponding detection signal-to-noise ratio is :

$$SNR_{d,q} = (1 - \delta I_{1,q})/\delta I_{1,q} \tag{58}$$

Assume that the computation noise is also uniform and described by the range:

$$\delta I_{2,q} = .05 + .002.q \tag{59}$$

120

$$SNR_{c,q} = [(1-\delta I_{1,q})-\delta I_{2,q}]/\delta I_{2,q} \tag{60}$$

If the noise-free wavefront W_0 and the sum of diagonal elements tr $(Q^T.Q)^{-1}$ are taken as multiplicative constants, the wavefront signal-to-noise ratio (52) becomes:

$$SNR_q = (const)/\{4.[2.(1-\delta I_{1,q})-\delta I_{2,q}]^2.(1-\delta I_{1,q})^2.(1-\delta I_{1,q}-\delta I_{2,q})^2.[(1/SNR_{d,q})-(1/SNR_{d,q})]^2\} \tag{61}$$

The variations of $SNR_{c,q}$, SNR_q as functions of the range variable (q) and the variation of $SNR_{d,q}$ as function of the detection noise range $\delta I_{1,q}$ are plotted in figs. 2 to 4. The SNR_q maximum occurs for:

$$\delta I_{1,q} = .110$$

$$\delta I_{2,q} = .098 \tag{62}$$

5. SUMMARY

The operation of a single-detector curvature sensor in partially coherent illumination has been analyzed. Using the Zernike decomposition of the wavefront, we have derived the least-squares solution of the sensor signal equation (1). We have also investigated how the overall signal-to-noise ratio is degraded by generic detection and computation errors.

6. REFERENCES

1. R. H. Hudgin, "Wave-front reconstruction for compensated imaging", J. Opt. Soc. Amer. **67** , 375-378, 1977.
2. D. L. Fried, "Least-square fitting a wave-front distorsion estimate to an array of phase-difference measurements", J. Opt. Soc. Amer. **67**, 370-375, 1977.
3. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform", Opt. Lett. **3**, 27-29, 1978.
4. J. C. Dainty, J. R. Fienup, "Phase retrieval and image reconstruction for astronomy", *Image Recovery*, pp. 231-272, H. Stark (ed.), Academic, New York, 1987.
5. F. Roddier, "Wavefront sensing and the irradiance transport equation", Appl. Opt. **29**, 1402-1403, 1990.
6. C. Roddier, F. Roddier, "Wave-front reconstruction from defocused images and the testing of ground-based optical telescopes", J. Opt. Soc. Amer. **10**, 2277-2287, 1993.
7. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics", Appl. Opt. **27**, 1223-1225, 1988.
8. J. W. Goodman, *Statistical Optics*, pp. 195-202, John Wiley & Sons, New York, 1985.
9. F. Roddier, C. Roddier, "Wavefront reconstruction using iterative Fourier transforms", Appl. Opt. **30**, 1325-1327, 1991.
10. I. Han, "New method for estimating wavefront from curvature signal by curve fitting", Opt. Eng. **34**, 1232-1237, 1995.
11. J. Y. Wang, D.E. Silva, "Wave-front interpretation with Zernike polynomials", Appl. Opt. **19**, 1510-1518, 1980.
12. J. Schwiegerling, J. E. Greivenkamp, J. M. Miller, "Representation of videokeratoscopic height data with Zernike polynomials", J. Opt. Soc. Amer. **12**, 2105-2113, 1995.
13. G. Vdovin, "Reconstruction of an object shape from the near-field intensity of a reflected paraxial beam", Appl. Opt. **36**, 5508-5513, 1997.
14. A. W. Lohmann, D. Mendlovic, Z. Zalevsky, "Digital method for measuring the focus error", Appl. Opt. **36**, 7204-7209, 1997.
15. E. R. Dowski, Jr., W. T. Cathey, "Single-lens single-image incoherent passive-ranging systems", Appl. Opt. **33**, 6762-6773, 1994.
16. Ref. 8, p. 219 and p. 224.
17. Ref. 8, p. 309.
18. Ref. 8, p. 298 and p. 299.
19. Ref. 8, p. 221.

20. D. Malacara, *Optical Shop Testing*, pp. 489-505, John Wiley & Sons, New York, 1978.

21. R. Cubalchini, "Modal wave-front estimation from phase derivative measurements", J. Opt. Soc. Amer. **69**, 972-977, 1979.

22. M. B. Allen III, E. L. Isaacson, *Numerical Analysis for Applied Science* , chapter 2, John Wiley & Sons, New York, 1998.

23. W. H. Southwell, "Wave-front estimation from wave-front slope measurements", J. Opt. Soc. Amer. **70**, 998-1006, 1980.

24. J.Herrmann, "Least-squares wave-front errors of minimum norm", J. Opt. Soc. Amer. **70**, 28-35, 1980.
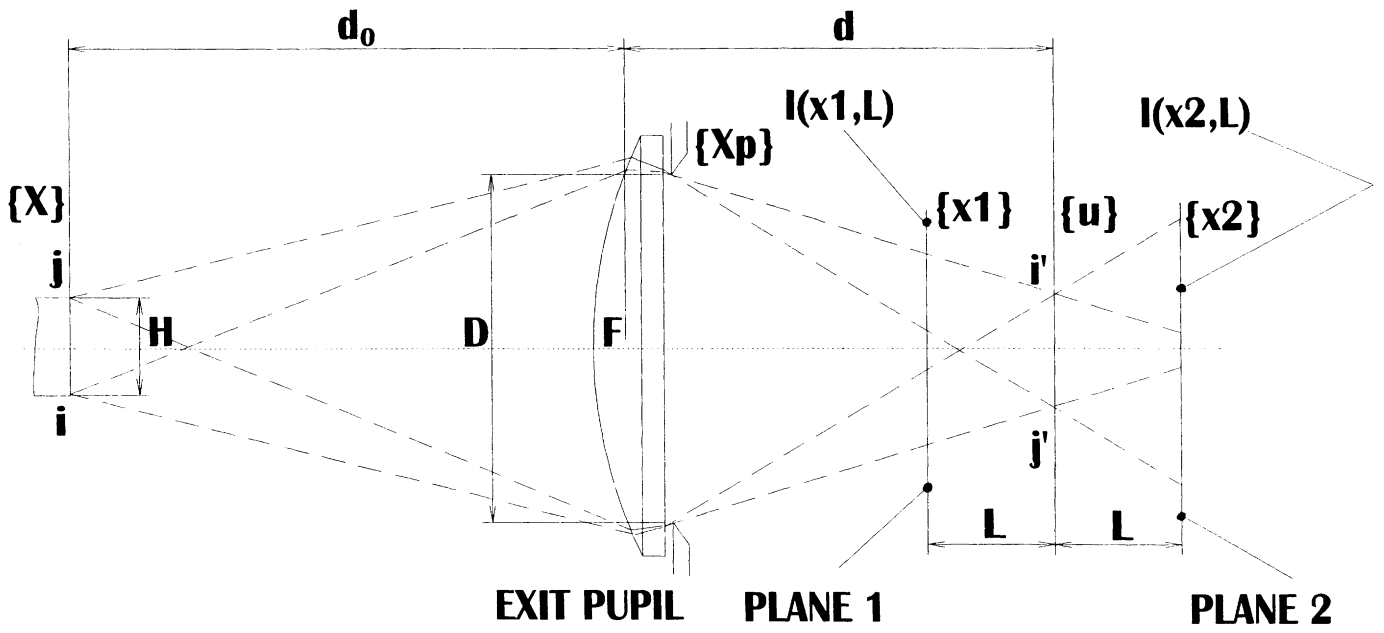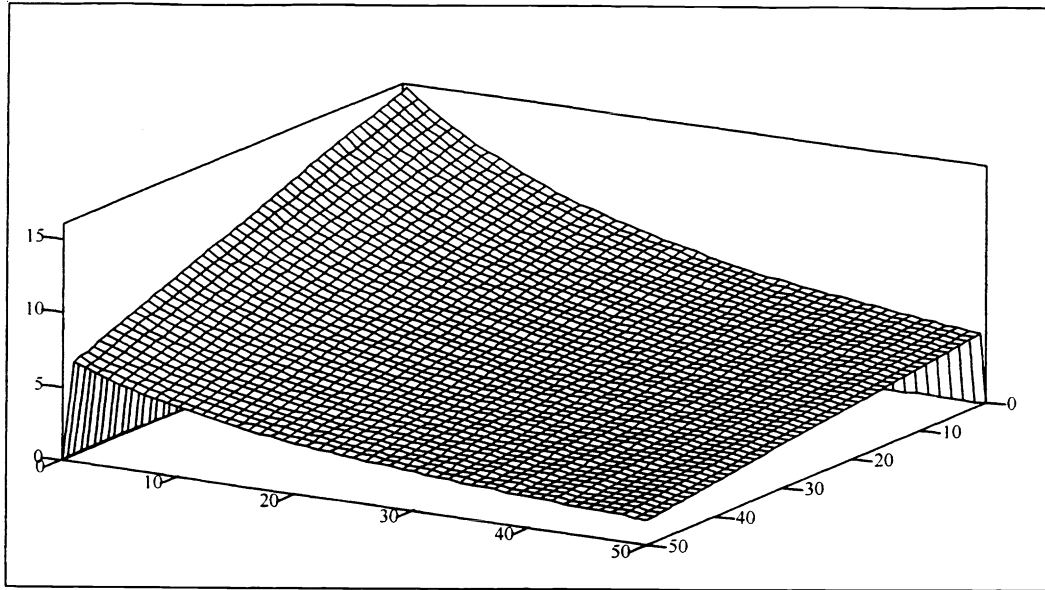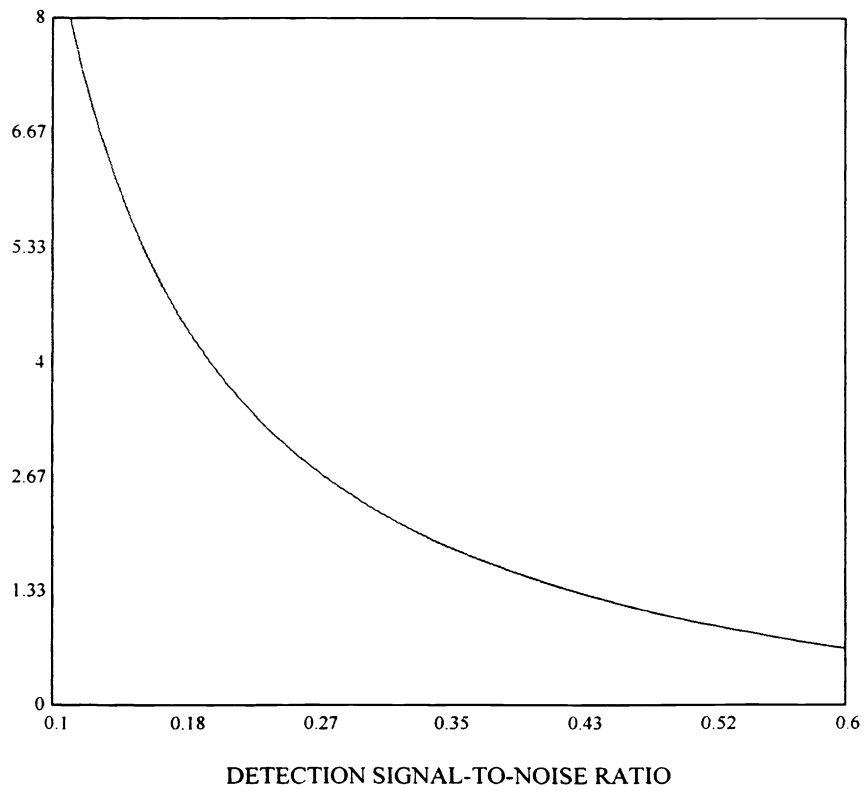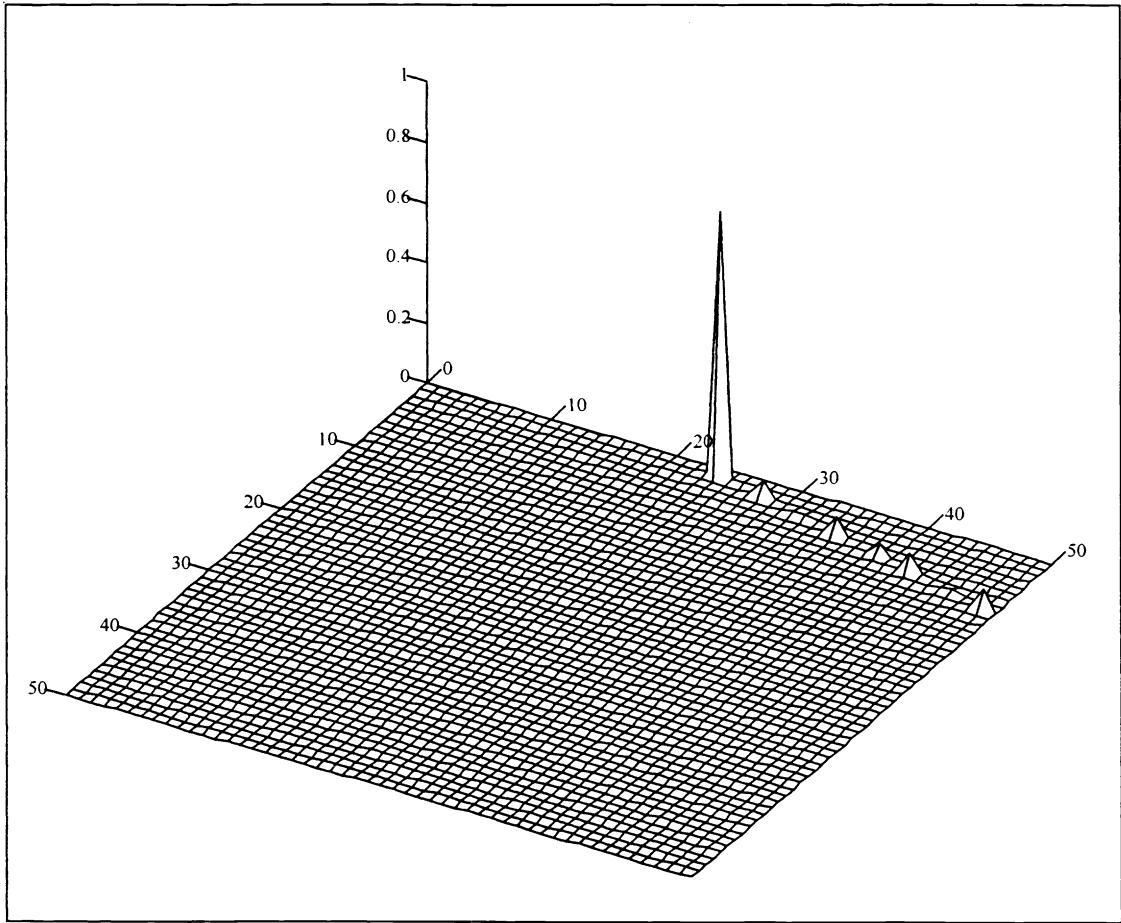
FIG. 1

COMPUTATION SIGNAL-TO-NOISE RATIO

FIG 2

DETECTION SIGNAL-TO-NOISE RATIO

FIG 3

WAVEFRONT SIGNAL-TO-NOISE RATIO

FIG 4