Phase determination of an amplitude modulated complex wavefront

Demetri Psaltis and David Casasent

An electronic–optical method is described for obtaining the phase of an arbitrary complex 2-D wavefront. The method consists of intensity detection, limiting, pulse width normalization, differentiation, dc filtering, and integration. Theoretical analyses and experimental confirmation are included. Applications of this scheme that are discussed include space variant processing, adaptive optics, data transmission, kinoforms, holography, and optical data processing.

I. Introduction

An optical wavefront is in general complex and is completely described only by both its amplitude and phase. However, only the time-averaged intensity of the wavefront can be detected rather than its instantaneous values. Holography provides a method by which the total amplitude and phase of the wavefront can be recorded, but extracting the phase content alone of the wavefront from this hologram is our present concern. Heterodyne detection (as used in acoustic–optic signal processing and elsewhere) likewise preserves the phase of the wavefront, but is limited to use with 1-D signals.

As a result, numerous indirect methods have been developed by which the phase of an optical wavefront can be estimated. Most of these prior methods can only be used to determine the phase of a pure phase function and include the knife edge test[1] and the Zernike phase-contrast test.[2] These involve differentiation and phase detection whereby the phase variations in the wavefront are converted to intensity modulation.[3,4] Refinements such as spatial filtering[5] have been included to provide an output proportional to the phase rather than its modulus squared.

These methods thus cannot be applied to an arbitrary, amplitude-modulated wavefront. The Hartman test[2] can be used to obtain the phase of a general wavefront. However, it requires an array of lenses, each of which focuses the light incident on its aperture to a position proportional to the local phase slope across the portion of the input wavefront. Thus, this method requires very large apertures and an array of lenses.

The final and most used methods by which data on the relative phase of two wavefronts can be found are the interferometric techniques.[6] In this case the relative phase data are encoded in the location of the interference fringes. The intensity of these interference fringes are proportional to the magnitude of the interfering waves at a point in space. In these interference techniques, the shape of the fringes are used to estimate the phase variations, and quantitative data are not provided. By curve fitting and computer analysis using Zernike polynomials, point-by-point optical path differences data are possible. But such data are only available after considerable computer time and only on a limited point-by-point basis, e.g., 20 × 20 points.

The method of phase detection we propose is a modified interferometric technique in which the intensity variations of the fringes are removed by a hybrid electronic/optical system. In Sec. II, we describe interference patterns in the notation to be used and then provide the theoretical considerations of our proposed method in Sec. III. Experimental confirmation of the technique and the design of the optical system used are then presented in Sec. V. This is followed by a discussion of several uses for this method in kinoforms, adaptive optics, space, variant processing, data transmission, holography, and optical data processing.

II. Interference Patterns

We consider a general amplitude modulated wavefront $u(x) = a(x) \exp[j\phi(x)]$ with amplitude $a(x)$ and with a phase $\phi(x)$ which we wish to determine. We use 1-D functions to simplify notation only, as the proposed method is directly usable on 2-D wavefronts, and in fact

The authors are with Carnegie-Mellon University, Department of Electrical Engineering, Pittsburgh, Pennsylvania 15213.

experimental confirmation will be shown for such 2-D cases.

An interference pattern is first created. It is well known from the basic principles of holography that such a recording contains both the amplitude and phase information about the input wavefront. We consider three different methods of producing the interference pattern below. The one selected depends on the specific application and origin of the wavefront whose phase is desired.

If the intensity of the interference pattern of $u(x)$ and $u^*(x)$ with an angle of inclination 2θ between the two wavefronts is recorded, we obtain

$$|u(x) \exp(j\alpha x) + u^*(x) \exp(-j\alpha x)|^2 = |u(x)|^2 \cos^2[\phi(x) + \alpha x], \quad (1)$$

where $\alpha = (\sin\theta)/\lambda$ is the carrier spatial frequency due to the off-axis angle θ. If $u^*(x)$ is the reference, the nulls of the interference pattern described by Eq. (1) will be zero when

$$\phi(x) + \alpha x = (2n + 1)\pi/2, \quad (2)$$

where n is an integer. Thus, the desired phase $\phi(x)$ of $u(x)$ is encoded in the locations of the zeroes of the interference pattern.

When the complex conjugate wavefront $u^*(x)$ is not available, as is often the case, Eq. (1) can be realized by interfering $u(x)$ with a reference beam $\exp(-j\alpha x)$ and high pass filtering as in conventional holographic matched spatial filtering.

Quite often the wavefront $u(x)$ whose phase $\phi(x)$ is desired is the Fourier transform of a spatial input function, e.g., $f(x)$. In this case Fourier transform hologram techniques can be used. However, a modified joint-transform correlator in which $f(x - x_0) + f(-x + x_0)$ is recorded in the input plane [i.e., the function $f(x)$ and its mirror image separated by a distance $2x_0$ is often most useful since the interference pattern in this case has full modulation at all spatial frequencies independent of the spatial frequency chosen[7] etc.]. The intensity of the Fourier transform of this input pattern is

$$|F(u) \exp(-jux_0) + F^*(u) \exp(jux_0)|^2$$
$$= 2|F(u)|^2 \{1 + \cos[2ux_0 - 2\phi(u)]\}, \quad (3)$$

where $F(u) = |F(u)| \exp[j\phi(u)]$ is the Fourier transform of $f(x)$, $\phi(u)$ is the phase of this transform that we desire, and u is the spatial frequency variable in the transform plane.

III. Theoretical Considerations

We now consider the proposed method for obtaining the phase $\phi(x)$ from an interference pattern of the form described by Eqs. (1) or (3). For clarity of understanding, the process used is divided into steps, and each is discussed separately.

A. Limiter

The intensity of the fringes in the interference pattern described by Eq. (1) depends only on $|A(x)|^2$, and

their position depends on $\phi(x)$. To obtain only the phase from the phase and amplitude modulated wavefront, we first remove the amplitude variations present on the cosine. We detect the received signal described by Eq. (1) on a vidicon or other array detector and pass this received signal through a nonlinear limiter. The threshold level R of the limiter is chosen so that the output of the limiter $T(x)$ is high if

$$\cos[2\phi(x) + 2\alpha x] > R/2|A(x)|^2 - 1 = c(x), \quad (4)$$

i.e.,

$$T(x) = \begin{cases} 1 & \text{if } |\alpha x + \phi(x) + n\pi| < q(x)/2 \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

where

$$q = q(x) = \cos^{-1}[c(x)], \quad (6)$$

with $n = 0, \pm 1, \pm 2, \ldots$. For later use in step 2, we rewrite Eq. (5) as a Fourier series expansion

$$T(x) = \sum_n \left[\frac{\sin(nq)}{n} \right] \exp\{jn[\phi(x) + \alpha x]\}, \quad (7)$$

where the dependence of $T(x)$ on $|A(x)|$ is contained in q.

B. Pulsewidth Normalization

The function $T(x)$ is a binary signal. However, the duration of a 1 state still depends on $|A(x)|$, and this dependence must also be removed. To visualize the effect of $|A(x)|$ on $T(x)$, we consider the case when the threshold level R in Eq. (4) is chosen to be small such that $|A(x)|^2 \gg R/2$; then $c \simeq -1$, and $T(x)$ is low only when $\cos[2\alpha x + 2\phi(x)] \simeq -1$. If $|A(x)|^2 = R$, $T(x)$ will be low for half of each cycle. Thus, $|A(x)|$ has the effect of varying the relative duration of the low and high states of $T(x)$. To remove this pulsewidth modulation effect of $|A(x)|$ on $T(x)$, we use $T(x)$ to trigger a monostable whose output pulse is of constant width each time $T(x)$ goes low. The monostable's output $T'(x)$ we write as

$$T'(x) = \sum_n \left[\frac{\sin(nq_0)}{n} \right] \exp\{jn[\phi(x) + \alpha x]\}, \quad (8)$$

where q_0 is a constant independent of x, and $T'(x)$ is now independent of $|A(x)|$.

C. First-Order Filter

Since methods exist for extracting the phase of a pure phase function, we next convert $T'(x)$ given by Eq. (8) into a pure phase function by passing only its first diffracted order. We thus obtain the pure phase function

$$T_1'(x) = [\sin(q_0)] \exp\{j[\phi(x) + \alpha x]\} = K \exp\{j[\phi(x) + \alpha x]\}, \quad (9)$$

where K is a constant. Equation (9) simply describes the phase $\phi(x)$ of $u(x)$ on a carrier α. Separation of the first diffracted order from the other terms present requires the proper choice of α or x_0 in Eqs. (1) and (3). The requirement on α and x_0 is the same as used in holography or matched spatial filtering.

D. Differentiation and High-Pass Filtering

The phase of the pure phase function in Eq. (9) can be found by forming the derivative of the function, i.e.,

$$\left|\frac{dT_1'(x)}{dx}\right| = K\left|\left[\frac{d\phi(x)}{dx} + \alpha\right]\exp\{j[\phi(x) + \alpha x]\}\right|$$

$$= K\left[\frac{d\phi(x)}{dx} + \alpha\right]. \qquad (10)$$

This pattern described by Eq. (10) is then integrated to produce $\phi(x)$, and a dc stop is inserted to remove the constant term α and any other constant terms present in the derivative of $\phi(x)$. We thus obtain the desired phase term $\phi(x)$ with no restrictions on the input wavefront. Let us now consider the hybrid electronic/optical realization of this phase determination scheme.

IV. System Realization

A schematic block diagram of the hybrid electronic/optical system to realize the operations described in Sec. III is shown in Fig. 1. The input function, e.g., $f(x - x_0) + f(-x + x_0)$, is placed at plane P_1 and the intensity of its Fourier transform recorded on the TV camera or other detector array. Alternatively the interference of the wavefronts $u(x)$ and $u^*(x)$, or $u(x)$ and a plane wave reference beam, is recorded on the detector array. The electronic output of the TV camera corresponds to Eqs. (1) or (3) in all cases. The output of the limiter is $T(x)$ described by Eqs. (5) or (7). The output of the monostable is $T'(x)$ described by Eq. (8). This is the output of the electronic section of the system. This $T'(x)$ pattern is recorded as a transparency with amplitude transmittance $T'(x)$ at plane P_2. A variety of devices for this purpose exist.[8,9]

Plane P_2 is now illuminated with parallel, collimated, coherent light. The Fourier transform of $T'(x)$ is formed at plane P_3 by lens L_2 where an aperture stop A passes only the first diffracted order term $T_1'(x)$ described by Eq. (9). The filter F placed at the location of this first-order term in plane P_3 has transmittance jx_3 (where x_3 is the spatial coordinate of plane P_3). The resultant pattern detected at plane P_4 is the derivative of $T_1'(x)$ described by Eq. (10).

Lenses L_5 and L_6 in Fig. 1 comprise an integrator with a dc stop placed at the frequency plane P_5. The dc stop removes the constant α term in Eq. (10) and any other constant term in $d\phi(x)/dx$, and at plane P_6 we obtain the desired phase function $\phi(x)$.

V. Experimental Results

To demonstrate this phase calculation method, a complex wavefront was produced by interfering three plane waves. The resultant pattern [Fig. 2(a)] was detected by a TV camera, displayed on a monitor, and photographed. As shown, it resembles a uniform grating with low-frequency sinusoidal amplitude modulation superimposed on it. The amplitude modulation in Fig. 2(a) appears primarily as a variation in the intensity of the fringes, although significant fringe
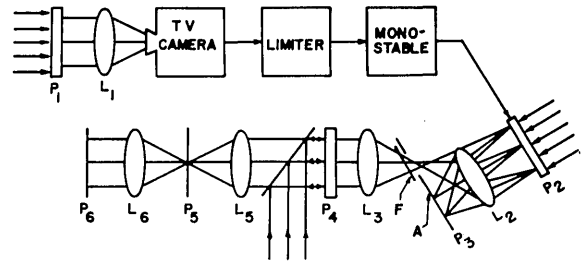
width modulation is also present. The electronically thresholded version of this pattern $T(x,y)$ is shown in Fig. 2(b). The effects of the amplitude modulation now appear primarily as a variation in the width of the fringes. This thresholded signal was then used to trigger a monostable. In this output pattern $T'(x,y)$ of Fig. 2(c), the variation in fringe width has been removed and hence all the effects of the amplitude modulation.

The pattern in Fig. 2(c) is thus the desired pure phase function, in this case a grating at a single frequency. To extract the phase from this function, a transparency of Fig. 2(c) was placed at the input plane P_2 of lens L_2 of Fig. 1. All the frequency components in the transform of $T'(x,y)$ were blocked at plane P_3 except those contained in one of the first-order terms. A high-pass filter was placed at this first-order term to approximate differentiation. The optical waveform transmitted through this filter was then Fourier transformed to produce the differentiated pattern described by Eq. (10) at plane P_4 of Fig. 1. The effective uniformity of the plane P_4 pattern is expected since the phase of the Fourier transform of the effective input was zero. This confirms the validity of the proposed scheme.

The high contrast and high resolution film used to photograph the image of Fig. 2(d) shows all minor nonuniformities severely. A plot of the cross section of the pattern is quite flat. The nonflat field of the lens used and noise in the system show up at low energies. This causes less monostable triggerings, and the fringes do not appear uniform.

To show more vividly the effectiveness of the proposed method of phase determination, the original interference pattern of Fig. 2(a) was recorded at Plane P_2 of Fig. 1. The resultant pattern at plane P_4 of Fig. 1 is shown in Fig. 2(e). It clearly shows the presence of the sinusoidal amplitude modulation and thus the need for the thresholding and pulse equalization steps to separate the phase and amplitude of the input pattern completely.

To insure that the pulses produced by the monostable occurred at the center of each fringe in Fig. 2(b), a peak detector circuit was used in conjunction with the threshold circuit.

There is no limitation on the wavefront $u(x) = a(x)\exp[j\phi(x)]$. However, if $|a(x)|^2 < R/4$, then $c(x) > 1$.

However, since the cosine is never greater than unity, the output of the limiter in Fig. 1 will be low for all values of $\phi(x)$ when $|a(x)|^2 < R/4$. This produces a set of false pulses at the monostable's output and an error in the calculation of the phase function. If the frequency of these false pulses is much less than the carrier frequency $\alpha = (\sin\theta)/\lambda$, their effect can be removed by filtering in plane P_3. This requirement simply requires selection of a large enough value for θ or α as is normally done in matched filtering and holography.

VI. Applications

The ability to detect the phase of an arbitrary, complex, amplitude modulated wavefront accurately in real time is of potential use in many optical data processing systems. We highlight some of these below.

A. Space-Variant Processing

To increase the class of operations achievable in an optical processor and the flexibility and practicality of an optical processor, space variant optical systems have
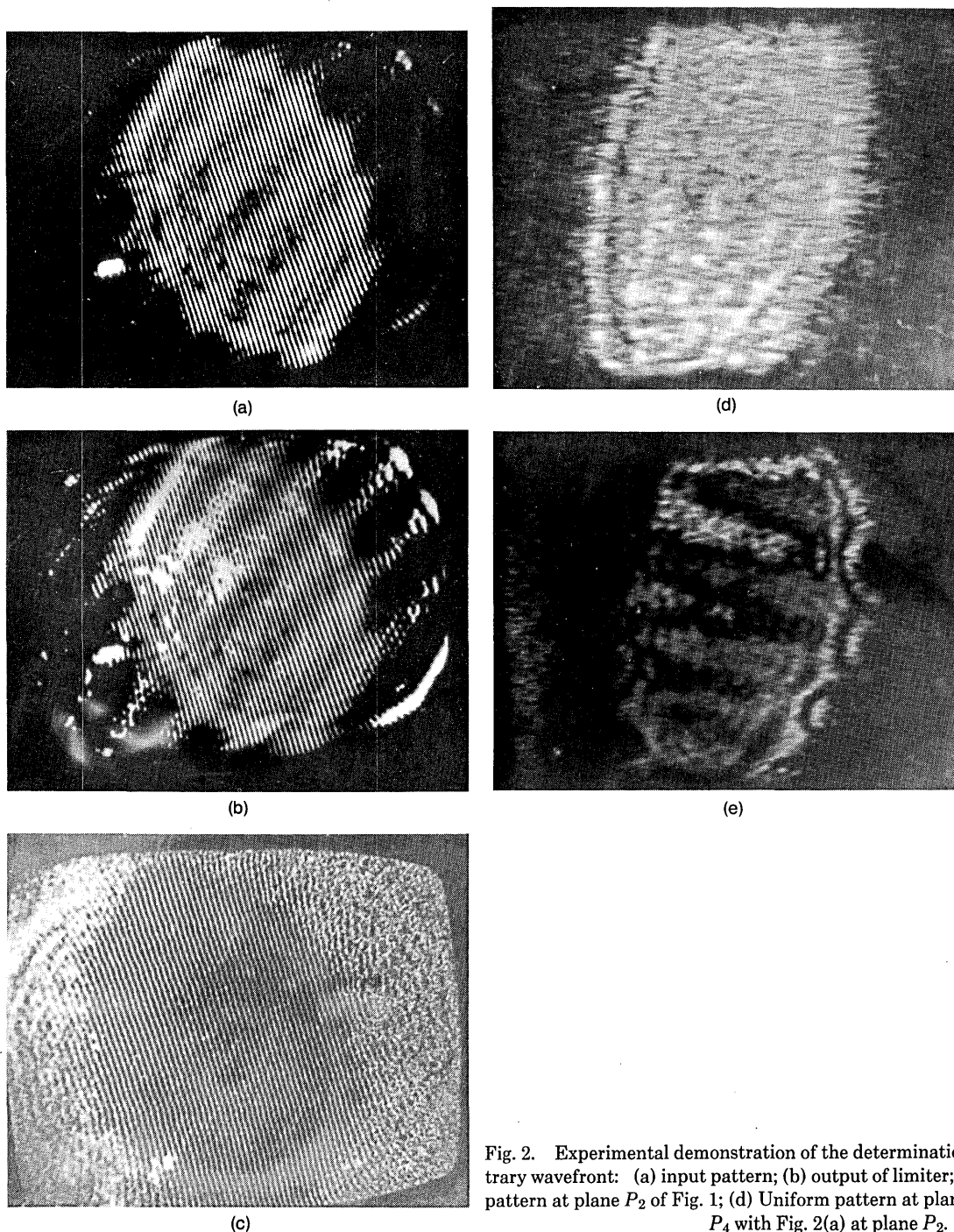

(a)

(b)

(c)

(d)

(e)

Fig. 2. Experimental demonstration of the determination of the phase of an arbitrary wavefront: (a) input pattern; (b) output of limiter; (c) output of monostable, pattern at plane P_2 of Fig. 1; (d) Uniform pattern at plane P_4; (e) pattern at plane P_4 with Fig. 2(a) at plane P_2.

recently been developed.[10] These systems are by definition quite sensitive to shifts in the input function and thus recessitate accurate positioning of the input function. Shifts in the input coordinates one equivalent to a linear phase factor in the Fourier transform of this input function. If we detect the phase of the Fourier transform as above and eliminate the linear terms in it, the shift in the input function will be eliminated automatically, and accurate positioning of the input can be achieved in real time. We can eliminate the linear phase term with no phase loss by taking the derivative of Eq. (10) and performing a double integration.

B. Adaptive Optics

The proposed phase calculation method is also of immense importance in adaptive optical systems. These systems are designed to adjust themselves to correct for phase distortions in an optical wavefront usually due to atmospheric turbulence. In these systems, the phase of the optical wavefront must first be detected to determine the degree of deformation and the optical system then adapted accordingly. High speed scanning output photodiode or charge coupled device array detectors may eventually provide the requisite speed demanded for such an application.

C. Hologram Transmission

The bandwidth of a hologram must be adequate to accommodate both the frequency of the carrier and the bandwidth B of the function itself. This usually requires six to ten times the bandwidth of the function itself. Since we are able (by the above scheme) to separate the phase and magnitude of the holographic recording, we can transmit each separately at a bandwidth B and recombine them at the receiver. This would result in a large reduction in the required bandwidth of the transmission channel.

D. Data Normalization and Centering

These phase extraction techniques are also of use in a novel method whereby an image of an object can automatically be normalized and centered on its center of mass. Consider two objects (in 1-D for simplicity) $f(x)$ and $af(x - x_0)$ which differ in exposure and in the location of their centers. Normalized masses of these objects with each centered on the object's center of mass can be produced by forming the second derivative of the natural log of the Fourier transform of the functions. The problem in the optical realization of this scheme is the implementation of the complex log since it requires separation of the phase of the signals. The method noted in this paper appears to provide this property. Details of this normalization and center of mass system will be provided later after experimental details and confirmation have been completed.

E. Image Restoration and Correlation

The restoration of a blurred or distorted image requires the use of complex filters to alter the system's impulse response to correct for the input distortion. With the ability to produce phase functions as described above separate amplitude and phase filters can be fabricated and in-line image restoration performed with reduced requirements (less resolution) for the filter plane recording material, less lens aberration and aperture problems, etc. Similar advantages result when this phase recording scheme is used in optical correlation and pattern recognition.

F. Kinoforms

The computer writing of a kinoform is not easy if the object is real. The size of these computer generated holograms is also quite limited. By use of the proposed phase detection scheme, over 1000×1000 point on axis kinoforms of real objects can be produced using even a TV detector as in Fig. 1.

By these brief remarks, the potential use of this phase detection scheme should be apparent.

VII. Summary and Conclusion

An electronic/optical method has been presented for determining the phase alone from a complex amplitude modulated optical wavefront or for separating the phase portion from the entire pattern in an optical hologram or matched spatial filter. Experimental confirmation of the basic concepts has proven successful. From the numerous potential applications shown for this technique, its importance and usefulness in a wide range of areas should be apparent.

References
1. J. W. Goodman, Introduction to Fourier Optics, (McGraw-Hill, New York, 1968).
2. J. Wyant, Laser Forms, 38 (November 1976).
3. R. A. Sprague and B. J. Thompson, Appl. Opt. 11, 1469 (1972).
4. D. Kermisch, J. Opt. Soc. Am., 65, 887 (1975).
5. B. Horwitz, Opt. Commun. 17, 231 (1976).
6. J. Wyant, Appl. Opt. 14, 2622 (1975).
7. D. Casasent and A. Furman, Appl. Opt. 16, 1652, 1662 (1977).
8. D. Sasasent, Proc. IEEE 65, 143 (1977).
9. D. Casasent, J. Soc. Inf. Disp. 15, 131 (1974).
10. D. Casasent and D. Psaltis, Proc. IEEE 65, 77 (1977).