# Phase-shifting interferometry with uncalibrated phase shifts

Xin Chen, Maureen Gramaglia, and John A. Yeazell

A computationally efficient algorithm for phase-shifting interferometry with imprecise phase shifts is developed. It permits the use of an uncalibrated phase shifter and is also insensitive to spatial intensity variations. The measurement has both spatial and temporal aspects. Comparisons are made between pixels within the same interferogram, and these comparisons are extended across a set of interferograms by a maximum–minimum procedure. A test experiment is performed and confirms the theoretical results. An additional advantage of the algorithm is that an error measure can be developed. This error measure is used to implement an error correction scheme. © 2000 Optical Society of America

*OCIS codes:* 120.3180, 120.5050.

## 1. Introduction

Phase-shifting interferometry techniques provide phase information about an object beam by interfering it with a reference beam of known phase.[1,2] Several methods for extracting the phase distribution of the object beam have been developed; they may be broadly grouped as temporal and spatial techniques. In the temporal algorithms, several interferograms at different phase shifts of the reference beam provide the information with which the phase distribution can be determined. In the standard algorithms[1,2] (the three-, four-, and five-step techniques) the phase shifts of the reference beam must be known to a high degree of accuracy, so special efforts must be made to calibrate the phase shifter.[3] Changes in the environment and aging of the shifter can affect this calibration and so lead to erroneous results.

Spatial algorithms have been proposed that do not explicitly require a phase shifter.[4] However, these Fourier-transform-based methods do not lead to a unique solution unless additional information is available. For example, some assumptions about the nature of the phase distribution may be needed, or additional interferograms may be used, or a spatial carrier frequency may be introduced into the interferogram. A difficulty of the single-interferogram spatial algorithm is that a spatial variation in intensity of the object beam is interpreted as a phase variation. Temporal schemes avoid this difficulty by comparing each pixel with the phase-shifted versions of itself in the other interferograms. Hybrids of these spatial and temporal algorithms have been developed to overcome this problem.[5] They are less sensitive to noise, but the use of several interferograms reintroduces the need for a calibrated phase shifter.

A generalized method for dealing with arbitrary phase shifts was developed by Grievenkamp.[6] The amount of phase shift still must be known precisely, but no special values are needed. Fitting schemes to the sinusoidal variation in intensity with phase have been used to measure these phase shifts.[7] If these phase shifts are known, then with a collection of distinct interferograms it is possible to extract the phase distribution of the object beam. Farrell and Player[8] have also addressed this problem of unknown phase shifts. They used the Bookstein method[9] to measure an unknown phase step between two interferograms. This method uses the intensity at each pixel as independent variables that map out an ellipse. This ellipse is fitted, and the fit variables give the unknown phase step. Farrell and Player incorporated this basic method for measuring the unknown phase steps into several different algorithms[10] for retrieving the phase distribution. The fitting involved in these methods can represent a significant computational effort. Another approach is the use of feedback or self-calibrating methods, which generate an error measure. The minimization of this measure gives the desired phase distribution.[11,12] These methods accurately reproduce the desired phase distribution, provided that certain phase shifts are not used. These methods can also have substantial computational loads. An experimental solution to the prob-

The authors are with the Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802.

Received 19 July 1999; revised manuscript received 2 November 1999.

0003-6935/00/040585-07$15.00/0

lem of unknown phase shifts was demonstrated by Lai and Yatagai.[13] There an additional tilted mirror is added to the test surface to allow the phase shifts to be measured directly from the resultant Fizeau fringes.

In Section 2 we develop a spatial and temporal algorithm that does not require precise knowledge of the phase shifts for extraction of the phase information. It makes use of a maximum–minimum (max–min) technique that spans a collection of interferograms. First, a comparison of each pixel with itself in a collection of different interferograms removes the sensitivity to spatial intensity variation. Then the algorithm makes spatial comparisons of each pixel's intensity with the intensities of two reference pixels within the same interferogram. The temporal variation of these spatial comparisons from interferogram to interferogram gives the necessary information to enable one to find the phase distribution of the object beam. Note that the spatial location of the reference pixels remains the same for the collection of interferograms. In Section 3 this phase-shifting algorithm is implemented and tested in a Twyman–Green interferometer setup. The algorithm is also shown to be computationally efficient: The computational load of the algorithm grows only linearly with the size of the interferogram (i.e., the number of pixels). In Section 4 an analysis of the error is presented. In this section we also develop an error measure for this algorithm and use it to implement an error-correction scheme. Simulated interferograms are used to analyze the algorithm's sensitivity to noise and the capabilities of the error-correction scheme.

## 2. Max–Min Algorithm

An interferogram created by two coherent beams is the key datum in phase-shifting interferometry. The fields of these two beams, reference and object beams, are, respectively, $E_r(x, y) = [I_r(x, y)]^{1/2}\exp(i\phi)$ and $E_o(x, y) = [I_o(x, y)]^{1/2}\exp[i\delta(x, y)]$. We assume that the phase of the reference beam, $\phi$, is uniform across the beam. The intensity of the reference beam, $I_r(x, y)$, is typically also uniform across the beam, but for the present algorithm this is not a requirement. The intensity of the object beam, $I_o(x, y)$, may also vary spatially. It is the phase distribution of the object beam, $\delta(x, y)$, that is the desired quantity. One finds this phase distribution by varying the phase of the reference beam, $\phi$. The resultant interferogram is

$$I(x, y) = I_b(x, y) + I_m(x, y)\cos[\delta(x, y) + \phi], \quad (1)$$

where $I_b(x, y)$ is the background intensity and $I_m(x, y)$ is the modulation intensity. Note that the phase difference between the reference beam and the object beam appears solely as an argument of the cosine term. This cosine term is independent of the intensity profiles of the two beams. It is advantageous to normalize the intensity and so eliminate the dependence of Eq. (1) on the background intensity and on the modulation intensity. This simplifies the analysis and removes the sensitivity of the algorithm to spatial intensity variations.

The data needed for this algorithm is a set of $M$ interferograms ($M \sim 15$) taken with various phase shifts of the reference beam that range roughly over $2\pi$. Nothing more needs to be known about these phase shifts. One use of these measurements is to normalize the intensity of Eq. (1). One does this by finding the maximum and minimum intensities for each point in this set of interferograms. From these data, $I_b(x, y)$ and $I_m(x, y)$ can in turn be found, which gives a normalized intensity:

$$I_n(x, y) = \cos[\delta(x, y) + \phi]. \quad (2)$$

The goal of the algorithm is to find $\delta(x, y)$ when the values of $\phi$ are not known precisely. In this case it is natural to consider the phase difference between different pixels. The Fourier-transform-based techniques described above are such an approach. Another approach that is closer in character to the one developed here is to build up a correlation[14] between the intensities of two of the pixels:

$$\cos[\delta(x, y) - \delta(x', y')] = 2\langle\cos[\delta(x, y) + \phi]\cos[\delta(x', y') + \phi]\rangle, \quad (3)$$

where the right-hand side is the ensemble average of a set of measurements associated with a uniform random distribution of the phase, $\phi$. To build such a correlation[15] typically requires the sampling of hundreds of different interferograms. The algorithm developed in this paper finds this phase difference without building such a correlation and without making an assumption about the distribution of the phases. The algorithm is based on the observation that the sum or difference of two cosine functions reveals the phase difference between them. When two cosine functions are in phase, adding them will yield maximum modulation; when two cosine functions are out of phase, adding them will yield minimum modulation.

Consider any two different points on a particular interferogram, $X = (x, y)$ and $X' = (x', y')$. First, we define $Q^{+}_{XX'}$, ($Q^{-}_{XX'}$) as the sum (difference) of the normalized intensities at these two points, $I_n(x, y)$ and $I_n(x', y')$:

$$Q^{\pm}_{XX'} = [\cos\delta(x, y) \pm \cos\delta(x', y')]\cos\phi - [\sin\delta(x, y) \pm \sin\delta(x', y')]\sin\phi. \quad (4)$$

Next we examine how the quantities $Q^{\pm}_{XX'}$ vary from interferogram to interferogram for the same points. Our goal is to find the extreme values of these quantities across this set of $M$ interferograms. We define a variable $R^{+}_{XX'}$ to be the difference between the maximum and the minimum values of $Q^{+}_{XX'}$. Similarly, $R^{-}_{XX'}$ is the difference between the maximum and the minimum values of $Q^{-}_{XX'}$. It is then straightforward

to show the relationship between $R^{\pm}_{XX'}$ and the phase difference $\delta(x, y) - \delta(x', y')$:

$$R^{\pm}_{XX'} = \max(Q^{\pm}_{XX'}) - \min(Q^{\pm}_{XX'}) = 2\{2 \pm 2\cos[\delta(x, y) - \delta(x', y')]\}^{1/2}. \quad (5)$$

From $R^{\pm}_{XX'}$ we can obtain the cosine of the phase difference:

$$\cos[\delta(x, y) - \delta(x', y')] = \frac{(R^{+}_{XX'})^2 - (R^{-}_{XX'})^2}{16}. \quad (6)$$

Thus this max–min procedure takes spatial information from within an interferogram and measures how that spatial information changes from interferogram to interferogram. The result is that the phase difference between the two points X and X′ is isolated from $\phi$, the unknown phase delays between these interferograms. Note that either $R^{+}_{XX'}$ or $R^{-}_{XX'}$ alone is sufficient to yield the cosine phase difference. We shall see that the above choice for expressing this cosine phase difference minimizes the error in the measurement. For example, the limited number of interferograms sample the $2\pi$ phase range, so it is likely that the true maxima and minima are missed. However, the missing of extrema tends to make both $R^{+}_{XX'}$ and $R^{-}_{XX'}$ smaller than their true values. It is also worth noting that an identity exists that relates $R^{+}_{XX'}$ and $R^{-}_{XX'}$:

$$(R^{+}_{XX'})^2 + (R^{-}_{XX'})^2 = 16. \quad (7)$$

This identity will be used in the error-correction scheme described in Section 4 of this paper.

Equation (6) is not sufficient to define precisely the phase difference between the two points (there is an undetermined sign; see below). Furthermore, it is not this phase difference that we desire but the value of $\delta(x, y)$ at each point.

To proceed, we use the above method as a basic building block. The general max–min algorithm is summarized in Fig. 1. We can apply Eq. (6) to the phase differences among three different points $(x, y)$, $(x_1, y_1)$, and $(x_2, y_2)$ that lie upon the same interferogram. The differences and the sums of the normalized intensities are formed in each possible pair of points in the manner indicated above. The extrema of these sums and differences are then found in the set of $M$ interferograms. Two of the points, $(x_1, y_1)$ and $(x_2, y_2)$, will act as references and so are fixed points. The third point consists of the independent variables for the desired phase distribution. Inasmuch as there is a global phase shift that need not be determined, we may set the phase at one of the reference points to zero. Specifically, we choose the phase at $(x_1, y_1)$, $\delta(x_1, y_1)$ equal to zero. This makes $\delta(x_2, y_2)$ the phase difference between the two reference points. Thus we use the following relations to determine $\delta(x, y)$:

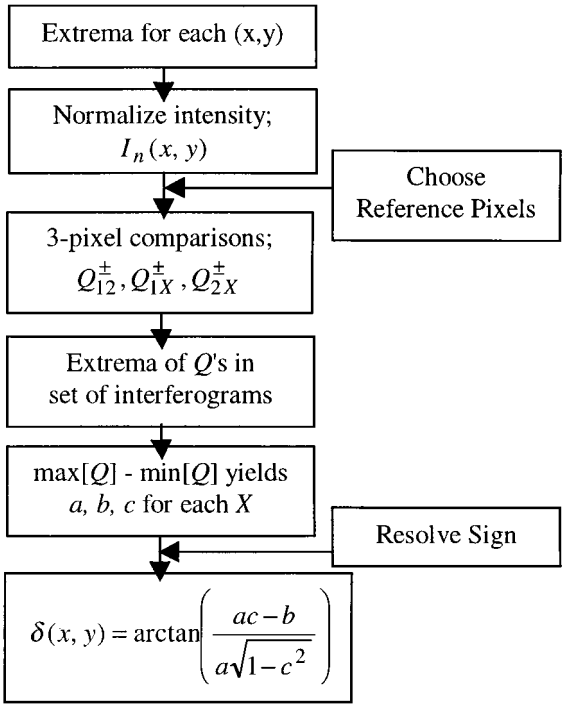


Fig. 1. Flow chart of the max–min algorithm. The reference pixels are denoted 1 and 2, and the test pixel is denoted $X = (x, y)$.

$$\cos[\delta(x, y)] = a, \quad (8)$$

$$\cos[\delta(x, y) - \delta(x_2, y_2)] = b, \quad (9)$$

$$\cos[\delta(x_2, y_2)] = c. \quad (10)$$

We have introduced the quantities $a$, $b$, and $c$ as a shorthand for the measured values defined by the right-hand of Eq. (6). Precisely,

$$a = \frac{(R^{+}_{X1})^2 - (R^{-}_{X1})^2}{16}, \quad (11)$$

$$b = \frac{(R^{+}_{X2})^2 - (R^{-}_{X2})^2}{16}, \quad (12)$$

$$c = \frac{(R^{+}_{12})^2 - (R^{-}_{12})^2}{16}, \quad (13)$$

where $X$ represents $(x, y)$ and the two reference points are represented simply by their numbers.

Manipulation of Eqs. (8)–(10) produces the expression $ac \pm \sin\delta(x, y)\sqrt{1 - c^2} = b$. Here we digress for the moment and discuss the two possible choices of sign in this expression. The undetermined sign has its origin in the term $\sin[\delta(x_1, y_1) - \delta(x_2, y_2)]$. If nothing is known about the phase shifts associated with the interferograms, there is no way to choose between the two possible solutions for the phase distribution. However, if we do know that the phase is progressively advanced between successive interfero-

grams, the sign ambiguity can be resolved by a logical comparison of the relative order of the extrema of the normalized intensities at the reference points $(x_1, y_1)$ and $(x_2, y_2)$. Consider the case in which the maximum at point 2 occurs before the minimum in the sequence of interferograms. Between these two interferograms, reference point 1 must go through either a maximum or a minimum. If it is a maximum, then $\delta(x_1, y_1) < \delta(x_2, y_2)$; if it is a minimum, then $\delta(x_1, y_1) > \delta(x_2, y_2)$. When the minimum at point 2 occurs first, the relative sizes of the phases are interchanged. These relative sizes precisely determine the sign of $\sin[\delta(x_1, y_1) - \delta(x_2, y_2)]$. Note that we define the difference $\delta(x_1, y_1) - \delta(x_2, y_2)$ such that it lies in the range $(-\pi, \pi)$. This analysis determines the sign for all points of the interferogram.

With this mechanism for resolving the sign in place, we return to further manipulation of Eqs. (8)–(10) to obtain

$$\sin \delta(x, y) = \pm \frac{ac - b}{\sqrt{1 - c^2}}. \tag{14}$$

Finally, after dividing Eq. (14) by Eq. (8), we obtain the resultant equation for $\delta(x, y)$:

$$\delta(x, y) = \arctan\left(\pm \frac{ac - b}{a\sqrt{1 - c^2}}\right). \tag{15}$$

The study of Farrell and Player[8,10] has some similarity to this proposed algorithm. It also uses a combination of spatial and temporal information to extract the phase distribution from the interferograms. The retrieval of the phase distribution in both cases relies intrinsically on the arccosine function. Also, points in the same interferogram with $n\pi$ phase differences are of little use to either method. However, the ellipse fitting technique used in Refs. 8 and 10 and the max–min technique presented here are quite different. The primary difference is that no fitting is used in the max–min technique. Also, the generation of the ellipse requires a variation of at least a full fringe across the interferogram. The max–min technique will work well on an object beam whose phase varies as little as 8° across the interferogram. Of course a tilt in the object beam can introduce a variation of a fringe or more across the interferogram. Such tilts have been used to create a spatial carrier frequency for sinusoidal fitting algorithms, such as those of Lassahn *et al.*[7] However, user intervention to introduce the tilt is not needed in the max–min algorithm.

### 3. Experiment

The algorithm was tested experimentally with a Twyman–Green interferometer. The mirrors used in this interferometer are flat to $\lambda/10$. We obtained the object beam by slightly tilting one of the mirrors. A piezoelectric transducer was used to translate the mirror in the reference arm of the interferometer and so advance the phase delay of the reference beam. Fifteen steps were taken over a range of more than $2\pi$. These steps progressively advanced the phase, but no attempt was made to make these steps in a uniform or calibrated fashion. At each step an interferogram was recorded with an 8-bit CCD array. As prescribed by the above algorithm, the intensity of each pixel was normalized. Two reference pixels were chosen that have different normalized intensities or different phases. If their phases were too similar, they would have given essentially the same information about the test points. We used the logical comparison of the ordering of the extrema for these two reference pixels to resolve the undetermined sign. Next, the extrema of the sums and differences of the normalized intensities at these two reference pixels were found and used to yield the quantity $c$ in Eq. (10). Similarly, the extrema of the sum and difference of each pixel with respect to each of the reference points was found, and these in turn gave the quantities $a$ and $b$. These data were then inserted into Eq. (15) and produced the phase distribution of the object beam shown in Fig. 2. The phase increased linearly, as expected. Note that over the 2.5 mm × 2.5 mm region shown in the figure the phase distribution indicates that the optical surfaces are flatter than their specifications. That is, the fluctuations in phase are much smaller than $\lambda/10 \sim 0.06$ rad. This experiment demonstrates that each of the steps of the algorithm is straightforward to implement experimentally and computationally.

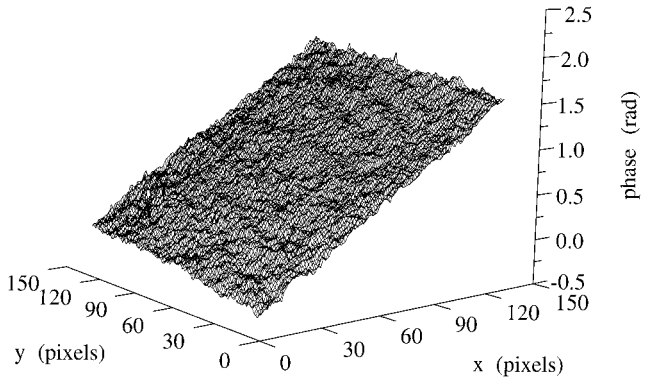


Fig. 2. Experimentally measured phase distribution of the object beam. The object beam is created by introduction of a tilt along the $x$ axis (note that the construction of the mirror mount also introduces a slight tilt along the $y$ axis).

The algorithm is computationally efficient. The computation time increases linearly with the number of pixels ($N$) and with the number of interferograms ($M$). The total number of operations is given by a multiple ($\simeq$15) of the product of $M \times N$. An operation is defined as either a logical comparison or a simple arithmetic operation. Such linear performance in $N$ is particularly important when the image size is large. Also important is the convergence of the algorithm to the correct answer. The error analysis performed below indicates that the maximum error in the phase distribution decreases quadrati-

cally with increasing $M$, whereas the computational load increases only linearly with $M$.

We can estimate the computation time by dividing the number of operations, $\simeq 15MN$, by the number of operations per second that can be performed by a standard desktop computer. For example, a 400-MHz Pentium computer will perform roughly $50 \times 10^6$ operations/s. For an array of $100 \times 100$ pixels and 15 interferograms, the number of operations required is $2.25 \times 10^6$. This leads to a computation time of 0.045 s.

#### 4. Error Analysis and Error Correction

Errors in the retrieved phase distribution can have several sources. The problems introduced by intensity noise and digitization errors are common to all phase-shifting interferometries. The sensitivity of this algorithm to these noise sources is discussed below. Of special interest to the present algorithm is the error introduced by the use of a relatively small number of interferograms. This leads to the possibility that the measured extrema might not be the true extrema. For example, when we measure the two-point sum we find a sinusoidal signal whose measured amplitude is given by $R^{+}_{XX'}/2$. In sampling these signals it is likely that the true value, $R^{+(0)}_{XX'}$, will be missed. The measured value must be somewhat smaller than this true value, $R^{+}_{XX'} = R^{+(0)}_{XX'}(1 - \varepsilon)$, $\varepsilon > 0$.

We define the largest difference in phase shift between interferograms as $\delta\alpha$; then the extrema may be missed by only half of this value. Therefore the error is $\varepsilon \leq 1/2(1/2\delta\alpha)^2$, where we have expanded the sinusoidal function to second order. The measured value of $R^{-}_{XX'}$ has a similar error that we denote $\varepsilon'$. These two errors combine to give the following error in the cosine phase difference:

$$\begin{aligned} &|\cos[\delta(x, y) - \delta(x', y')] - \cos[\delta^0(x, y) - \delta^0(x', y')]| \\ &\quad = |(\varepsilon - \varepsilon') + \cos[\delta^0(x, y) - \delta^0(x', y')](\varepsilon + \varepsilon')| \\ &\quad \leq 2 \max[\varepsilon, \varepsilon'] \leq 1/4(\delta\alpha)^2. \end{aligned} \quad (16)$$

As the number of interferograms is increased, $\delta\alpha$ should decrease, and the error falls as $(\delta\alpha)^2$.

In this paper three points are used to resolve the phase of a particular pixel, so we must propagate this error in the cosine phase difference to the final result, i.e., the phase of a pixel. As was shown above,

$$\tan[\delta(x, y)] = \pm \frac{ac - b}{a\sqrt{1 - c^2}}. \quad (17)$$

Thus to find the error $\delta\Phi[\Phi = \delta(x, y)]$ that is due tothe errors of $a$, $b$, and $c$, namely, $\delta a$, $\delta b$, and $\delta c$, we use the following equation:

$$\frac{1}{\cos^2(\Phi)}\,\delta\Phi = \left|\left(\delta a \frac{\partial}{\partial a} + \delta b \frac{\partial}{\partial b} + \delta c \frac{\partial}{\partial c}\right)\left(\pm \frac{ac - b}{a\sqrt{1 - c^2}}\right)\right|. \quad (18)$$

**Table 1. Variation of Maximum Error (in radians) in the Phase Distribution with the Phase Difference ($\phi_{12}$) between the Reference Pixels[a]**

| $\phi_{12}$ (°) | $\mathrm{Err}_{\max}{}^{u}$ | $\mathrm{Err}_{\max}{}^{c}$ |
|---|---|---|
| 90 | 0.036 | 0.015 |
| 45 | 0.073 | 0.023 |
| 15 | 0.149 | 0.026 |
| 8 | 0.295 | 0.031 |

[a]The uncorrected maximum error is denoted $\mathrm{Err}_{\max}{}^{u}$, and the corrected maximum error is denoted $\mathrm{Err}_{\max}{}^{c}$.

Evaluating the right-hand side of Eq. (18) and using $\cos^2(\Phi) = a^2$ give

$$\delta\Phi = \left|\frac{b}{\sqrt{1 - c^2}}\,\delta a - \frac{a}{\sqrt{1 - c^2}}\,\delta b + \frac{a^2(a - bc)}{(1 - c^2)^{3/2}}\,\delta c\right|. \quad (19)$$

As we see, the errors $\delta a$, $\delta b$, and $\delta c$ have an upper boundary $(1/4)\delta\alpha^2$. If the cosine phase difference between the two fixed reference pixels ($c$) is not too close to 1, the error in the reconstructed phase will be of the same order, i.e., $(1/4)\delta\alpha^2$. When $c \simeq 0$, we have

$$\delta\Phi = |b\delta a - a\delta b + a^2(a - bc)\delta c| \leq 1/2(\delta\alpha)^2. \quad (20)$$

This error will increase as the phase difference between the two reference pixels approaches 0 (or $n\pi$). However, even with a phase difference of just 30° between the two reference pixels, the overall error increases only by approximately a factor of 2 from the optimum at 90°. Thus this error is relatively insensitive to the choice of reference pixels, which also indicates that the algorithm will work well on rather flat phase objects as well as on those that have strong variations in phase.

Sets of numerically generated interferograms were used to test this error analysis. A range of values for the phase difference between the two reference points was studied. The phase step between interferograms is generated randomly with a uniform distribution over a range from $2\pi/15$ to $2\pi/10$. The set of interferograms generated more than spanned a phase shift of $2\pi$. Digitization error of the intensities of the numerically generated interferograms was also included in these numerical tests (a resolution of 8 bits was used). The generated interferograms have a phase distribution that has the form $\delta(x, y) = 3x + 0.5y$, and it varies by a maximum of $\pi$ across the surface. Table 1 summarizes the results. As expected, when the phase difference between the reference pixels is 90° the maximum phase error of any point is small, only 0.036 rad. The error grows to 0.295 as this phase difference decreases to 8°. Table 1 also shows that the error-correction scheme described below can further reduce the sensitivity of the error to the phase difference between the two reference pixels.

For the phase reconstruction algorithm suggested

Table 2. Variation of Maximum Error (in radians) in the Phase Distribution with the Phase Difference ($\phi_{12}$) between the Reference Pixels[a]

| $\phi_{12}$ (°) | $Err_{max}{}^{u}$ | $Err_{max}{}^{c}$ |
|---|---|---|
| 90 | 0.054 | 0.027 |
| 45 | 0.069 | 0.037 |
| 15 | 0.210 | 0.057 |
| 8 | 0.327 | 0.209 |

[a]Intensity noise with a maximum of 5% of the modulation intensity is added. The uncorrected maximum error is denoted $Err_{max}{}^{u}$, and the corrected maximum error is denoted $Err_{max}{}^{c}$.

Table 3. Variation of Maximum Error (in radians) in the Phase Distribution with Added Intensity Noise[a]

| Noise (%) | $Err_{max}{}^{u}$ | $Err_{max}{}^{c}$ |
|---|---|---|
| 0 | 0.036 | 0.015 |
| 2.5 | 0.036 | 0.024 |
| 5.0 | 0.054 | 0.027 |
| 7.5 | 0.054 | 0.064 |
| 10 | 0.067 | 0.068 |

[a]The uncorrected maximum error is denoted $Err_{max}{}^{u}$, and the corrected maximum error is denoted $Err_{max}{}^{c}$.

in this paper, one of the major sources of error is the noise in the measured intensities of the pixels. However, recall the identity, Eq. (7), found in Section 2. The presence of errors in the measured intensity will cause a deviation from this identity. Thus it is possible to define a measure to gauge the error in a particular measurement:

$$\eta = 16 - [(R^{+}_{XX'})^2 + (R^{-}_{XX'})^2]. \quad (21)$$

In addition to quantifying the error associated with intensity noise, this measure can be used to partially correct the error. Note that a low-noise environment induces only small deviations from the identity, and the measured values of $R^{+}_{XX'}$ and $R^{-}_{XX'}$ are only slightly away from their true value. The goal is to find the minimal corrections to $R^{+}_{XX'}$ and $R^{-}_{XX'}$ that will satisfy the identity. For notational convenience we suppress the subscripts $X$ and $X'$ in the following analysis. Thus we denote the corrections $\Delta R^{+}$ and $\Delta R^{-}$ such that

$$(R^{+} + \Delta R^{+})^2 + (R^{-} + \Delta R^{-})^2 = 16, \quad (22)$$

which gives

$$2R^{+}\Delta R^{+} + 2R^{-}\Delta R^{-} = 16 - [(R^{+})^2 + (R^{-})^2] = \eta. \quad (23)$$

We define the shortest distances from the true values of $R^{+}$ and $R^{-}$ by minimizing $(\Delta R^{+})^2 + (\Delta R^{-})^2$. After some straightforward algebra, we find the corrections for $R^{+}$ and $R^{-}$:

$$\Delta R^{+} = \frac{\eta R^{+}}{2[(R^{+})^2 + (R^{-})^2]}, \quad (24)$$

$$\Delta R^{-} = \frac{\eta R^{-}}{2[(R^{+})^2 + (R^{-})^2]}. \quad (25)$$

The effect of the intensity noise on this algorithm has been tested numerically. Tables 2 and 3 show the results for uncorrected and corrected versions of the algorithm. The background intensity and the modulation intensity of the interferogram were set so significant digitization errors would be generated in the intensities. This choice leads to a maximum digitization error in the intensity of 0.5%. In Table 2 a uniformly distributed random noise term is added to the unnormalized intensity. The noise has a maximum value of 5% of the modulation intensity, $I_m(x, y)$. The generated interferograms have the same phase distribution as described for Table 1. The phase shifts were also chosen in the same manner. The effect of different choices for the phase difference between reference pixels is studied. The sensitivity of the error to this choice is greatly reduced, as in Table 1. However, the added noise does reduce the effectiveness of the error correction when the reference phase difference is small.

It is interesting to make a comparison of this technique with the Grievenkamp algorithm.[6] These two techniques have quite different assumptions about what prior knowledge exists concerning these interferograms. The Grievenkamp algorithm assumes that the interferograms are separated by a known phase shift. However, with the use of a relatively large number of interferograms this algorithm can also retrieve $\delta(x, y)$, even if the phase shifts are not precisely known. The test makes use of 15 simulated interferograms (with digitization error) whose phase steps vary randomly about the nominal shift of $2\pi/15$ rad. The phase errors have a uniform distribution about these nominal shifts. We examined three different distributions whose maximum widths are $\pm 3°$, $\pm 5°$, and $\pm 10°$. The maximum errors associated with these distribution are 0.038, 0.064, and 0.124 rad, respectively. The maximum errors quoted in Table 1 for the uncorrected and especially for the corrected algorithm are generally significantly smaller these values. Note that it is typically straightforward to choose two reference pixels that are separated by roughly 90°. If intensity noise on the level of 5% is added to this simulation of the Grievenkamp algorithm, the maximum errors for the same three phase-error distributions are 0.055, 0.071, and 0.127 rad. Again, the maximum errors quoted in Table 2 are generally significantly smaller these values. This is by no means a complete comparison of the two algorithms. It should be noted that the two algorithms were designed from fundamentally different points of view, i.e., known versus unknown phase shifts.

In Table 3 the effect of increasing levels of added noise is studied. Note that digitization noise is still included as described above. When the intensity noise reaches a large value ($\simeq$10%) the max–min algorithm leads to a large maximum error as noise spikes are interpreted as extrema. However, for smaller values of the noise ($\simeq$5%) the algorithm

works well and the error correction can further improve the retrieved phase distribution. The intensity noise in phase-shifting interferometry is often well below this value of 5%, so it does not pose a serious limitation on the use of the algorithm. The error-correction procedure is computationally straightforward. Therefore the inclusion of the error correction adds only slightly to the computational load.

## 5. Conclusions

The max–min algorithm presented here overcomes the common difficulty of calibrating and environmentally stabilizing the phase shifter of the apparatus. Its implementation does not require the addition of equipment for use on existing interferometers. The algorithm is also insensitive to spatial intensity variations. Computationally the algorithm is efficient. One disadvantage lies in the need to collect a relatively large number of interferograms. However, the computational efficiency should more than offset this disadvantage in a wide range of applications. The rapid decrease of the error with increasing number of interferograms allows the algorithm to gain accuracy with only modest increases in computational time because of the efficiency of the algorithm. The algorithm also permits the incorporation of an error-correction scheme to reduce the effect of intensity noise such as digitization error.

This study was supported by the National Science Foundation under grant PHY-9733643.

## References and Notes

1. P. K. Rastogi, *Holographic Interferometry* (Springer-Verlag, Berlin, 1994).
2. D. Malacara, M. Servin, and Zacarias Malacara, *Interferogram Analysis for Optical Testing* (Marcel Dekker, New York, 1998).
3. N. A. Ochoa and J. M. Huntley, "Convenient method for calibrating nonlinear phase modulators for use in phase-shifting interferometry," Opt. Eng. **37,** 2501–2505 (1998).
4. T. Kreis, "Digital holographic interference-phase measurement using the Fourier-transform method," J. Opt. Soc. Am. A **3,** 847–855 (1986).
5. R. Windecker and H. J. Tiziani, "Semispatial, robust, and accurate phase evaluation algorithm," Appl. Opt. **34,** 7321–7326 (1995).
6. J. E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng. **23,** 350–352 (1984).
7. G. D. Lassahn, J. K. Lassahn, P. L. Taylor, and V. A. Deason, "Multiphase fringe analysis with unknown phase shifts," Opt. Eng. **33,** 2039–2044 (1994).
8. C. T. Farrell and M. A. Player, "Phase-step insensitive algorithms for phase-shifting interferometry," Meas. Sci. Technol. **5,** 648–652 (1994).
9. F. Bookstein, "Fitting conic sections to scattered data," Comput. Graph. Image Process. **9,** 56–71 (1979).
10. C. T. Farrell and M. A. Player, "Phase step measurement and variable step algorithms in phase-shifting interferometry," Meas. Sci. Technol. **3,** 953–958 (1992).
11. G.-S. Han and S.-W. Kim, "Numerical correction of reference phases in phase-shifting interferometry by iterative least-squares fitting," Appl. Opt. **33,** 7321–732 (1994).
12. A. Dobroiu, D. Apostol, V. Nascov, and V. Damian, "Statistical self-calibrating algorithm for phase-shift interferometry based on a smoothness assessment of the intensity offset map," Meas. Sci. Technol. **9,** 1451–1455 (1998).
13. G. Lai and T. Yatagai, "Generalized phase-shifting interferometry," J. Opt. Soc. Am. A **8,** 822–827 (1991).
14. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge U. Press, Cambridge, 1995).
15. A more comprehensive description of the use of correlations to find the phase difference between two pixels $i$ and $j$ is given by $\cos(\delta_i - \delta_j) = \langle I_i I_j \rangle - \langle I_i \rangle \langle I_j \rangle / [\langle I_i^2 \rangle - \langle I_i \rangle^2 \langle I_i^2 \rangle - \langle I_i \rangle^2]^{1/2}$, where $I_i$ and $I_j$ are the intensities at the pixels $i$ and $j$, respectively. The angle brackets denote an ensemble average over the randomly varying phase $\phi$.