Fringe-print-through error analysis and correction in snapshot phase-shifting interference microscope

YU ZHANG,1,2,3 XIAOBO TIAN,2 AND RONGGUANG LIANG2,4

1Institute of Materials Physics, College of Science, Northeast Electric Power University, Jilin, Jilin 132012, China
2College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA
321zhangyu2008@163.com
4rliang@optics.arizona.edu

Abstract: To reduce the environmental errors, a snapshot phase-shifting interference microscope (SPSIM) has been developed for surface roughness measurement. However, fringe-print-through (FPT) error widely exists in the phase-shifting interferometry (PSI). To ensure the measurement accuracy, we analyze the sources which introduce the FPT error in the SPSIM. We also develop a FPT error correction algorithm which can be used in the different intensity distribution conditions. The simulation and experiment verify the correctness and feasibility of the FPT error correction algorithm.

References and links
1. D. Malacara, Optical Shop Testing, 3rd ed. (John Wiley & Sons, Inc., 2007), Chap. 1–7.
2. B. Kimbrough, N. Brock, and J. Millerd, “Dynamic surface roughness profiler,” Proc. SPIE 8501, 85010D (2012).
3. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, “Digital wavefront measuring interferometer for testing optical surfaces and lenses,” Appl. Opt. 13(11), 2693–2703 (1974).
4. C. Tian and S. Liu, “Two-frame phase-shifting interferometry for testing optical surfaces,” Opt. Express 24(16), 18695–18708 (2016).
5. K. Kinnstaetter, A. W. Lohmann, J. Schwider, and N. Streibl, “Accuracy of phase shifting interferometry,” Appl. Opt. 27(24), 5082–5089 (1988).
6. K. Hibino, B. F. Oreb, D. I. Farrant, and K. G. Larkin, “Phase-shifting algorithms for nonlinear and spatially nonuniform phase shifts,” J. Opt. Soc. Am. A 14(4), 918–930 (1997).
7. M. Servin, J. C. Estrada, and J. A. Quiroga, “The general theory of phase shifting algorithms,” Opt. Express 17(24), 21867–21881 (2009).
8. D. Malacara, Optical Shop Testing, 3rd ed. (John Wiley & Sons, Inc., 2007), Chap. 14.
9. P. J. de Groot, “Vibration in phase-shifting interferometry,” J. Opt. Soc. Am. A 12(2), 354–365 (1995).
10. L. L. Deck, “ Suppressing phase errors from vibration in phase-shifting interferometry,” Appl. Opt. 48(20), 3948–3960 (2009).
11. M. Novak, J. Millerd, N. Brock, M. North-Morris, J. Hayes, and J. Wyant, “Analysis of a micropolarizer array-based simultaneous phase-shifting interferometer,” Appl. Opt. 44(32), 6861–6868 (2005).
12. C. L. Koliopoulos, “Simultaneous phase-shift interferometer,” Proc. SPIE 1531, 119–127 (1991).
13. Z. Chen, X. Wang, and R. Liang, “Snapshot phase shift fringe projection 3D surface measurement,” Opt. Express 23(2), 667–673 (2015).
14. D. Wang and R. Liang, “Simultaneous polarization Mirau interferometer based on pixelated polarization camera,” Opt. Lett. 41(1), 41–44 (2016).
15. P. de Groot, “Generating fringe-free images from phase-shifted interferometry data,” Appl. Opt. 44(33), 7062–7069 (2005).
16. C. T. Farrell and M. A. Player, “Phase-step insensitive algorithms for phase-shifting interferometry,” Meas. Sci. Technol. 5(6), 648–652 (1994).
17. K. Hibino, B. F. Oreb, D. I. Farrant, and K. G. Larkin, “Phase-shifting algorithms for nonlinear and spatially nonuniform phase shifts,” J. Opt. Soc. Am. A 14(4), 918–930 (1997).
18. J. Xu, Q. Xu, and L. Chai, “Iterative algorithm for phase extraction from interferograms with random and spatially nonuniform phase shifts,” Appl. Opt. 47(3), 480–485 (2008).
1. Introduction

Interferometry is the industry standard metrology method for optical surface [1] and roughness measurement [2]. The phase-shifting interferometer (PSI) was introduced by Branning [3] to achieve accurate metrology in 1974, PSI and its variations have been widely used in optical measurement [1, 4, 5]. However the accuracy of the standard phase-shifting algorithm (PSA) depends on the accuracy of phase shift [5–7]. Since the phase-shifted interferograms are collected sequentially, the instabilities of light source intensity and frequency, vibration, and air turbulence in the working environment [8–10] will lead to the unavoidable and unknown phase error. To overcome these problems, two simultaneous PSIs were developed [11]. The first configuration uses three and more cameras to acquire phase-shifted interferograms simultaneously [12]. While it is insensitive to the environmental errors, it is complex and expensive. The second configuration uses only one polarization camera to acquire four phase-shifted interferograms [11, 13, 14]. Although the simultaneous PSI based on polarization camera can avoid the instability error of the light source and environmental errors, the fringe-print-through (FPT) error [15] can be introduced by the defects in polarized components, such as polarization beam splitter (PBS), quarter-wave plate (QWP), and polarizer array (PA) in front of the pixels.

To address the issues caused by the inaccurate phase shift, a number of error correction algorithms have been developed [16–22]. One simple and effective algorithm is Lissajous ellipse fitting algorithm. In 1992 Farrell and Player [20] utilized Lissajous figures and ellipse fitting to calculate the phase difference between two interferograms, but the correction result is not accurate if the intensity distribution is non-uniform. In 2014 Kimbrough [21] proposed a correction algorithm of systematic polarization aberrations in polarization based dynamic PSIs, also using Lissajous ellipse, while this method is not sensitive to non-uniform intensity distribution because every pixel is corrected separately, it is time-consuming because at least 6 measurements are necessary and two interferograms are needed from each measurement. Recently, Liu et al. [22] proposed a general algorithm to correct the phase error in PSI, where the points of a circle on the interferograms are used to create the Lissajous ellipse.

To achieve the high measurement accuracy, in addition to ensure the quality of the components (e.g. light source, PBS, QWP, PA), the error correction algorithm is critical as well. For the optical surface roughness measurement, Zernike fitting can’t be used to eliminate the high-frequency component since it is what we need, hence the FPT error due to the PBS, QWP and PA can’t be reduced or eliminated by the fitting process.

In this paper, we will discuss the FPT error analysis and correction in the snapshot phase-shifting interference microscope (SPSIM) which was developed for measuring surface roughness. Section 2 presents the principle of the SPSIM, system analysis using Jones vectors, and error analysis. The algorithms of correcting FPT error and generating Lissajous ellipse are introduced in this section as well. In Section 3 the simulation of the error analysis
and correction algorithm is discussed, Section 4 evaluates the correction algorithm with the experimental data. The conclusion is finally drawn in Section 5.

2. Principles

2.1 Principle of snapshot phase-shifting interference microscope

A schematic SPSIM layout is depicted in Fig. 1. The light from a LED with the wavelength of 625 nm and the minimum power of 700 mw (Thorlabs M625L3) is attenuated by the neutral density filter (NDF) and its polarization direction is controlled by a polarizer (P) to adjust the fringe contrast of the interferogram. The $p$-polarization and $s$-polarization of the light are transmitted and reflected from the PBS respectively, serving as the test beam (red line) and the reference beam (green line). The test arm includes the first QWP (QWP1), an objective with 10X magnification and 0.28 NA (Edmund M Plan Apo 10X), and the optical surface under test, while the reference arm consists of the second QWP (QWP2), an objective, and a reference mirror M with the surface quality of 15-5 scratch-dig and the surface flatness of $\lambda/10$ at 632.8 nm (Newport 10D20ER.2). The reflected test beam and reference beam are transformed to opposite-circular polarization beams by the third QWP (QWP3) and interfere after passing the polarizer array (PA). Without the physical phase shift, four phase-shifted interferograms can be extracted from a single image snapshotted by the 12 bit polarization camera PolarCam with the pixel number of 1208x1348 and the pixel size of 7.4 µm from 4D Technology, Inc [11, 14, 23].

For the 45 degree linear polarization incident beam, the Jones vector can be expressed as $E_i = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. The Jones vectors of the test beam and reference beam before they interfere through the PA are

$$E_r = Q_3 * T_{phs} * Q'_{r} * M_r * Q_2 * R_{phs} * E_i * \exp(i\varphi).$$  \hfill (1)

$$E_r = Q_3 * R_{phs} * Q'_{r} * M_r * Q_2 * T_{phs} * E_i.$$

![](image1.png)

Fig. 1. Snapshot phase-shifting interference microscope. NDF - neutral density filter, P - polarizer, PBS - polarization beam splitter, QWP - quarter-wave plate, L - lens, M - mirror.
where \( Q_1, Q_1', Q_2, Q_2', \text{ and } Q_3 \) are the Jones matrix of QWP1, QWP2 and QWP3, defined as
\[
Q = \begin{bmatrix}
\cos^2 \beta + i \sin^2 \beta & \cos \beta \sin \beta (1-i) \\
\cos \beta \sin \beta (1-i) & i \cos^2 \beta + \sin^2 \beta
\end{bmatrix}
\]
where \( \beta \) refers to the angle that the fast axis of QWP makes with respect to the \( x \)-axis. The transmittance and reflectance Jones matrices of the PBS can be written as \( T_{pbs} = \begin{bmatrix} T_p & 0 \\ 0 & T_r \end{bmatrix} \), \( R_{pbs} = \begin{bmatrix} R_p & 0 \\ 0 & R_r \end{bmatrix} \), \( T_p \) and \( T_r \) are the transmittance of the \( p \) and \( s \) polarized light, and \( R_p \) and \( R_s \) denote the reflectance of the \( p \) and \( s \) light. \( M_T \) and \( M_R \) are the Jones matrices of optical surface under test and reference mirror, \( M_T = M_R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \). Ideally the test and reference beams after passing QWP3 should be left and right circularly polarized.

As shown in Fig. 1, the PA inside the polarization camera consists of linear polarizers with four different polarizations (0, \( \pi/4 \), \( \pi/2 \), and \( 3\pi/4 \)). After passing through the linear polarizer at an angle \( \gamma \) with respect to the \( x \)-axis, the reference and test beams interfere. The electric field emerging from the interference is derived and expressed as
\[
E = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = P^* E_r + P^* E_g.
\]
where \( P = \begin{bmatrix} \cos^2 \gamma & \sin \gamma \cos \gamma \\ \sin \gamma \cos \gamma & \sin^2 \gamma \end{bmatrix} \) is the Jones matrix of the polarizer, \( \gamma = 0, \pi/4, \pi/2, 3\pi/4 \) for 4 different polarizers.

The intensity of the phase-shifted interferogram is
\[
I = E^2 = E^2_x + E^2_y.
\]
With four phase-shifted interferograms extracted from the snapshotted interferogram, the phase under test can be calculated by the 4-step PSA
\[
\phi = \arctan \left( \frac{I_3 - I_2}{I_1 - I_3} \right).
\]

**2.2 Error analysis**

While SPSIM is much less sensitive to the environmental errors than the standard PSI since it can obtain four phase-shifted interferograms from one single image without the physical phase shift, it also suffers the other errors introduced by the non-ideal PBS, the defect in QWP and its orientation error, the defect in the PA and its misalignment with camera pixels. In addition, the camera response to different light level and polarization will introduce errors as well.

For an ideal PBS, the transmittance and reflectance matrices are
\[
T_{pbs} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad R_{pbs} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.
\]
However the \( T_p \) and \( R_s \) are not equal to 1 practically, and \( T_r \) and \( R_p \) are not zero either. As a result, the test beam and reference beam are non-ideal \( s \) and \( p \) polarized light before passing through QWP3.
To generate the ideal orthogonal circularly polarized light after passing through the QWP3, the orientation of the QWP1, QWP2 and QWP3 should be oriented at 45° to the x-axis. However, the alignment is never perfect. After going through the QWP3 the Jones vectors of the non-ideal orthogonal circularly polarized light are

\[
E_r = \begin{bmatrix}
\cos^2(\beta_i) + i\sin^2(\beta_i) & \cos(\beta_i)\sin(\beta_i)(1-i) \\
\cos(\beta_i)\sin(\beta_i)(1-i) & i\cos^2(\beta_i) + \sin^2(\beta_i)
\end{bmatrix} \begin{bmatrix} T_p & 0 \\ 0 & T_i \end{bmatrix}
\]

\[
\begin{bmatrix}
\cos^2(-\beta_i) + i\sin^2(-\beta_i) & \cos(-\beta_i)\sin(-\beta_i)(1-1-i)(-1) \\
\cos(-\beta_i)\sin(-\beta_i)(1-1-i) & i\cos^2(-\beta_i) + \sin^2(-\beta_i)
\end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\]

\[
\begin{bmatrix}
\cos^2(\beta_i) + i\sin^2(\beta_i) & \cos(\beta_i)\sin(\beta_i)(1-1-i) \\
\cos(\beta_i)\sin(\beta_i)(1-1-i) & i\cos^2(\beta_i) + \sin^2(\beta_i)
\end{bmatrix} \begin{bmatrix} R_p & 0 \\ 0 & R_i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 1 \\ 1 & \frac{1}{\sqrt{2}} \end{bmatrix} \exp(i\varphi)
\]

After going through the PA, the two beams interfere and the 4-frame phase-shifted interferograms can be obtained. The linear polarizers with different directions in PA may have the polarized angle errors, the actual $\gamma$ are $\delta_1$, $\pi/4 + \delta_2$, $\pi/2 + \delta_3$, $3\pi/4 + \delta_4$, $\delta_5$, $\delta_6$, and $\delta_7$ are the polarized angle errors.

The last, but the not least, error we need to consider is the misalignment between the PA and the camera sensor, as shown in Fig. 2. The misalignment error of the PA and the camera sensor will lead to the phase-shift error because each sub-pixel in one super pixel may receive light from four linear polarizers. Provided that, for each sub-pixel in the super pixel, the contribution of the light from four different polarizers are $m$, $n$, $p$, and $q$ separately, the electrical field of the phase-shifted interferograms are

\[
E_1 = m^* P_1^* (E_r + E_g) + n^* P_2^* (E_r + E_g) + p^* P_3^* (E_r + E_g) + q^* P_4^* (E_r + E_g)
\]

\[
E_2 = m^* P_1^* (E_r + E_g) + n^* P_2^* (E_r + E_g) + p^* P_3^* (E_r + E_g) + q^* P_4^* (E_r + E_g)
\]

\[
E_3 = m^* P_1^* (E_r + E_g) + n^* P_2^* (E_r + E_g) + p^* P_3^* (E_r + E_g) + q^* P_4^* (E_r + E_g)
\]

\[
E_4 = m^* P_1^* (E_r + E_g) + n^* P_2^* (E_r + E_g) + p^* P_3^* (E_r + E_g) + q^* P_4^* (E_r + E_g)
\]

where $E_1$, $E_2$, $E_3$ and $E_4$ represent the electrical fields of the 4-frame phase-shifted interferograms, $P_1$, $P_2$, $P_3$ and $P_4$ represent the Jones matrixes for four different polarizers of a super-pixel in the PA. In this paper, we don’t analyze the camera response error, we directly use the calibration method designed by Z. Chen et al [24].
2.3 Correction algorithm of the FPT error

In SPSIM the intensity expressions of 4-frame interferograms are

\[
I_1 = A + B \cos(\varphi + \varepsilon_1) \\
I_2 = A + B \cos\left(\varphi + \frac{\pi}{2} + \varepsilon_2\right) = A - B \sin(\varphi + \varepsilon_2) \\
I_3 = A + B \cos(\varphi + \pi + \varepsilon_3) = A - B \cos(\varphi + \varepsilon_3). \\
I_4 = A + B \cos\left(\varphi + \frac{3}{2} \pi + \varepsilon_4\right) = A + B \sin(\varphi + \varepsilon_4)
\]  

(10)

where \( A \) and \( B \) are the background intensity and the modulation amplitude respectively, \( \varphi \) is the real phase, \( \varepsilon_1, \varepsilon_2, \varepsilon_3 \) and \( \varepsilon_4 \) are the phase-shift errors due to the above error analysis in Section 2.2. Note that \( \varepsilon_1, \varepsilon_2, \varepsilon_3 \) and \( \varepsilon_4 \) are approximate constants for every pixel, the reasons are: (1) the errors introduced by the non-ideal PBS, the defect in QWP and its orientation error, the defect in the PA and its misalignment with camera pixels are the same for the different pixels, and these errors are the main errors which can introduce the phase errors; (2) the effect of the environmental errors for each pixel is also approximately equal because four phase-shifted interferograms are extracted from the same raw image; and (3) with the camera nonlinear response calibrated by the method in Ref (23), the phase error due to the camera nonlinear response error can be ignored.

According the 4-step PSA as shown in Eq. (5), the expressions of nominator \( N \) and denominator \( D \) are

\[
N = I_4 - I_2 = B \sin(\varphi + \varepsilon_4) + B \sin(\varphi + \varepsilon_2) = 2B \sin\left(\varphi + \frac{\varepsilon_4 + \varepsilon_2}{2}\right) \cos\left(\frac{\varepsilon_4 - \varepsilon_2}{2}\right)
\]  

(11)

\[
D = I_1 - I_3 = B \cos(\varphi + \varepsilon_1) + B \cos(\varphi + \varepsilon_3) = 2B \cos\left(\varphi + \frac{\varepsilon_1 + \varepsilon_3}{2}\right) \cos\left(\frac{\varepsilon_1 - \varepsilon_3}{2}\right)
\]  

(12)

After setting \( a_x = 2B \cos\frac{\varepsilon_4 - \varepsilon_2}{2}, \ a_y = 2B \cos\frac{\varepsilon_1 - \varepsilon_3}{2}, \ \Phi = \varphi + \frac{\varepsilon_1 + \varepsilon_3}{2}, \ \varepsilon = \frac{\varepsilon_1 + \varepsilon_3 - \varepsilon_2 - \varepsilon_4}{2} \), \( N \) and \( D \) can be rewritten as
\[ N = a_x \sin (\Phi). \] (13)
\[ D = a_y \cos (\Phi + \epsilon). \] (14)

A constant phase offset will not affect the whole phase distribution, hence, we can use \( \varphi \) instead of \( \Phi \), the final expressions of \( N \) and \( D \) are
\[ N = a_x \sin \varphi. \] (15)
\[ D = a_y \cos (\varphi + \epsilon). \] (16)

Equations (15) and (16) describe an ellipse with no center offset, and semi-major and semi-minor amplitudes are \( a_x \) and \( a_y \). The phase with the FPT error can be calculated using Eqs. (15) and (16)

\[ \varphi = \arctan \left( \frac{N}{D} \right) = \arctan \left( \frac{a_x \sin \varphi}{a_y \cos (\varphi + \epsilon)} \right). \] (17)

According to Eqs. (15) and (16), a general conic function can be obtained
\[ \frac{N}{a_x^2} + \frac{D}{a_y^2} + \frac{2N \cdot D}{a_x \cdot a_y} \sin \epsilon = \cos^2 \epsilon. \] (18)

A general conic function can be also expressed by the following second order polynomial:
\[ F = ax^2 + bxy + cy^2 + dx + fy + g. \] (19)

For an ellipse, Eq. (18) needs to meet the conditions of \( F = 0 \) and \( b^2 - 4ac < 0 \). If an ellipse doesn’t have center offset, then \( d = 0 \) and \( f = 0 \). According to Eqs. (20) and (21), the semi-major amplitude \( a_x \), semi-minor amplitude \( a_y \), and the ellipse orientation angle \( \theta \) with respect to the x-axis can be calculated by
\[ a_x = \sqrt{\frac{2g}{-(a-c)^2 + b^2 - (a+c)}}, \quad a_y = \sqrt{\frac{2g}{-(a-c)^2 + b^2 - (a+c)}}. \] (20)

\[ \theta = \frac{1}{2} \arctan \frac{b}{a-c} \quad \text{for } a < c. \]
\[ \theta = \frac{\pi}{2} + \frac{1}{2} \arctan \frac{b}{a-c} \quad \text{for } a > c \] (21)

From Eqs. (17) and (18), we can see that the phase \( \varphi \) calculated from Eq. (17) is exactly the real phase \( \varphi \) when \( a_x = a_y, \epsilon = 0 \). In other words, when \( N \) and \( D \) lie on a circle centered at the origin, the real phase \( \varphi \) can be obtained. Hence, to correct the FPT error, the ellipse needs to be transformed to a perfect circle centered at the origin.

The correction procedure is as follows: 1) calculate \( N \) and \( D \) according to Eqs. (11) and (12); 2) use the least-squares fitting method to calculate the coefficients of the best fit ellipse \( (a, b, c, d, f, g) \), \( d \) and \( f \) should be equal to zero since the center of the ellipse is the origin; 3) calculate the semi-major amplitude \( a_x \), semi-minor amplitude \( a_y \), and orientation angle \( \theta \) of the ellipse by Eqs. (20) and (21); 4) rotate the ellipse to align the major and minor axes.
with the x and y axes which represent \( N \) and \( D \); 5) transform the ellipse to a circle; and 6) rotate the circle back to the original angle. The correction steps can be represented by Eq. (22)

\[
\begin{bmatrix}
N_c \\
D_c
\end{bmatrix} = T * \begin{bmatrix}
N \\
D
\end{bmatrix},
\]

where \( N_c \) and \( D_c \) are numerator and denominator after the FPT error correction, the transformation matrix is

\[
T = \begin{bmatrix}
\cos(-\theta) & -\sin(-\theta) \\
\sin(-\theta) & \cos(-\theta)
\end{bmatrix} * \begin{bmatrix}
1 & 0 \\
0 & r
\end{bmatrix} * \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}, \quad r = a_x / a_y .
\]

After the correction of the FPT error, the real phase can be calculated by Eq. (23)

\[
\phi = \arctan \left( \frac{N_c}{D_c} \right).
\]

### 2.4 Method of generating Lissajous ellipse

The method of generating Lissajous ellipse may affect correction result and accuracy, so an appropriate method is essential to obtain optimal correction of FPT error. There are two approaches to generate Lissajous ellipse.

- **First method:** Takes a series of 6 measurements with actual phase shift (a circular phase shift period is needed, i.e., \( 2\pi \) ) \[21\], every measurement has 2-frame interferograms which can form a group of \( N \) and \( D \). For every pixel, there are 6 groups of \( N \) and \( D \). Using \( N \) and \( D \) as the x and y coordinates, an ellipse can be plotted by fitting 6 groups of \( N \) and \( D \) for each pixel. This method is insensitive to the quality of the interferograms and doesn’t need uniform intensity distribution. In addition, it can correct the phase error pixel by pixel, so it has high correction accuracy. However, this method needs actual phase shift between 6 measurements, not suitable for the SPSIM which achieves the phase shift through the PA in the camera.

- **Second method:** Only needs one measurement and some points in the field are chosen for fitting the ellipse \[20, 22\]. This method is relatively simple and doesn’t need phase-shift, but the chosen points must meet the requirement of uniform intensity distribution. For the uniform intensity distribution or the interferogram with relatively small noise, every point in the field can be used to generate an ellipse. For partially non-uniform intensity distribution, the points on a line with uniform intensity distribution in the test field can be used as the fitting points. For the non-uniform but with circularly symmetric intensity distribution, the points on a circle or a ring are suitable for generating Lissajous ellipse. In real measurement with irregular non-uniform intensity distribution, we need to develop effective method for choosing points with nearly uniform background intensity to generate Lissajous ellipse.

The intensity expression of the phase-shifted interferogram is

\[
I'_{i,j} = A_{i,j} + B_{i,j} \cos(\varphi_j + \Delta_i).
\]

where \( i \) is the number of the interferogram (\( i = 1,2,3,...,M \), \( M \) is the number of the interferograms), \( j \) is the number of the pixel in one interferogram (\( j = 1,2,3,...,N \), \( N \) is the number of pixels in one interferogram), \( \varphi_j \) is the phase of the pixel \( j \), and \( \Delta_i \) is the phase shift of the interferogram \( i \).

Provided that the background intensity \( A_{i,j} \) and modulation amplitude \( B_{i,j} \) are irrelevant to \( i \), only relevant to \( j \), then \( A_{i,j} = A_{i,j} = \ldots = A_{i,j} = A_j \), \( B_{i,j} = B_{i,j} = \ldots = B_{i,j} = B_j \). By setting \( a_j = A_j \), \( b_j = B_j \cos \varphi_j \), and \( c_j = -B_j \sin \varphi_j \), Eq. (24) is rewritten as

\[
I'_{i,j} = a_j + b_j \cos \Delta_i + c_j \sin \Delta_i .
\]
The sum of squared differences between the theoretical intensity and actual intensity of the interferogram can be expressed as

\[ S_j = \left( \sum_{i=1}^{M} (I'_{i,j} - I_{i,j}) \right)^2 = \sum_{i=1}^{M} \left( a_j + b_j \cos \Delta_i + c_j \sin \Delta_i - I_{i,j} \right)^2. \]  

(26)

According to the least-squares theory [25, 26], \( S_j \) should be minimum, \( \partial S_j / \partial a_j = 0, \partial S_j / \partial b_j = 0, \partial S_j / \partial c_j = 0 \), so

\[ X_j = S^{-1} R_j. \]  

(27)

\[ S = \begin{bmatrix} \sum_{i=1}^{M} \cos \Delta_i & \sum_{i=1}^{M} \sin \Delta_i \\ \sum_{i=1}^{M} \sin \Delta_i & \sum_{i=1}^{M} \cos \Delta_i \end{bmatrix} \]  

(28)

\[ R_j = \begin{bmatrix} \sum_{i=1}^{M} I'_{i,j} \\ \sum_{i=1}^{M} I_{i,j} \cos \Delta_i \\ \sum_{i=1}^{M} I_{i,j} \sin \Delta_i \end{bmatrix}. \]  

(29)

If the phase shift is relatively accurate, we can calculate the background intensity and modulation from Eq. (27), \( A_j = a_j, B_j = \sqrt{b_j^2 + c_j^2} \). With \( A_j \) and \( B_j \), we can select the points with non-uniform ratio less than 15% to generate an ellipse, where the non-uniform ratio is the ratio of the difference between the max or min intensity and the mean of the intensity to the mean of the intensity.

**3. Simulation**

To validate the effectiveness of the error analysis and correction algorithm discussed in Section 2, we perform two simulations for the case of irregular non-uniform intensity distribution, one is the correction of the double-frequency FPT error, the other is the correction of the single frequency FPT error. A simulated test surface is shown in Fig. 3(a). For the double frequency FPT error, the errors discussed in Section 2 are used in this simulation and summarized as: \( T_p = 0.9, R_s = 0.095, T_5 = 0.1, R_p = 0.005, \beta_2 = 43.6^\circ, \beta_1 = 45.8^\circ, \beta_3 = 43^\circ, \delta_1 = -0.13 \text{ rad}, \delta_2 = -0.1 \text{ rad}, \delta_3 = 0.15 \text{ rad}, \text{ and } \delta_4 = -0.09 \text{ rad}. \) In addition, the primary polarized state accounts for 75% of the whole sub-pixel, other polarized states account for 10%, 5% and 10% separately. In order to generate the irregular non-uniform intensity distribution, the Signal-to-Noise Ratio (SNR) of interferogram is set as 20 dB.

The interferogram with the added errors and intensity distribution of the 4-frame phase-shifted interferograms are shown in Figs. 3(b) and 3(c), we can see that the amplitudes of 4-frame interferograms are different, and the intensity distribution is non-uniform, after the normalization of the intensity, we need to find relative uniform points from the calculated background intensity to apply the FPT error correction algorithm. Figure 4(a) is 2D map of the calculated surface with the FPT error, where the double frequency FPT is obvious. With the correction of FPT error by choosing the points with the similar background intensity (5238 pixels with non-uniform ratio of the intensity less than 15% are used), the FPT error is
removed as shown in Fig. 4(c). The phase error before and after FPT correction are shown in Fig. 4(b) (PV = 42.1 nm, RMS = 9.3 nm) and Fig. 4(d) (PV = 14.1 nm, RMS = 2.2 nm), most of the double frequency error has been removed, the residual error is due to the non-uniform intensity distribution. For the double frequency FPT error, before the correction we can see the obvious ellipse in the Lissajous figures (Fig. 5(a)), where the semi-major amplitude is 2.05, semi-minor amplitude is 1.67, and the center is the origin. After the correction, the Lissajous ellipse becomes a perfect circle (Fig. 5(b)). This simulation demonstrates that the correction algorithm for double FPT error is effective even for the most complicated case - irregular non-uniform intensity distribution.

In order to generate single frequency FPT error, we reset the simulation parameters as: \( \beta_3 = 35^\circ, \delta_1 = -0.19 \text{ rad}, \delta_2 = 0.1 \text{ rad}, \delta_3 = 0.14 \text{ rad}, \text{ and } \delta_4 = -0.2, \) other errors are same as the double frequency FPT error. From Figs. 6(a) and 6(b), we can clearly distinguish the
single frequency FPT error. After the correction by the choosing points (5609 pixels with non-uniform ratio of the intensity less than 15% are used), most of the single frequency FPT error is removed (Figs. 6(c) and 6(d)), although the difference between the Lissajous ellipse before correction and the circle after correction for single frequency FPT error is not obvious (Figs. 7(a) and 7(b)). The effect of single frequency FPT error correction is similar as the double frequency FPT error correction. Hence, we can conclude that, the correction algorithm is effective for both double and single frequency FPT error.

Fig. 6. Simulation results of the single frequency FPT error correction with irregular non-uniform intensity distribution. (a) the 2D map of the calculated surface with the FPT error, (b) the phase error before error correction (PV = 33.4 nm, RMS = 8.1 nm), (c) the 2D map of calculated surface after the FPT error correction, (d) the phase error after error correction (PV = 16.7 nm, RMS = 2.9 nm).

Fig. 7. Simulated Lissajous figures by plotting N against D for single frequency FPT error. (a) an ellipse: $a_1 = 1.81$, $a_2 = 1.72$ (with errors); (b) a circle (without errors).

4. Demonstration with Experimental Data

To demonstrate the proposed algorithm, we measured the roughness of a smooth surface after fine polishing with the SPSIM. The first frame phase-shifted interferogram and intensity distribution of a line are shown in Figs. 8 (a) and 8(b). After calibrating the camera nonlinear response [23], the amplitudes of 4-frame interferograms become similar and intensity distributions become more uniform, as shown in Fig. 8(c) and 8(d). However the intensity distribution is still not sufficiently uniform, we still need to calculate the background intensity to find some points with relatively uniform intensity distribution (125025 pixels with non-uniform ratio of the intensity less than 15% are used). Figures 9(a), 9(b) and 9(c) are the 3D, 2D map and line profile of the calculated phase with the FPT error, the single frequency FPT error is obvious in Fig. 9(b) and the information of the roughness cannot be distinguished from the profile in Fig. 9(c). After the correction of FPT error, the 3D, 2D map and line profile are plotted in Fig. 10 (the first 9 Zernike terms which are low-order aberrations have been removed since we measure the roughness). The FPT error is almost disappeared in Fig. 10(b) and the information of the roughness is clearly shown in Figs. 10(b) and 10(c). The difference between the Lissajous ellipse before correction (Fig. 11(a)), whose semi-major amplitude is 1.66 and semi-minor amplitude is 1.60, and the circle (Fig. 11(b)) after correction is not obvious since the FPT error is single frequency, consistent with the
simulation of single FPT error. However, the effect of the FPT error correction is very obvious, we can conclude that the error correction algorithm is also effective for the experiment data.

Fig. 8. Interferogram and intensity distribution before and after the response non-uniformity calibration of the polarization camera. (a) the first frame interferogram before the calibration, (b) the intensity distribution of a line in four phase-shifted interferograms before the calibration, (c) the first frame interferogram after the calibration, (d) the intensity distribution of a line in four phase-shifted interferograms after the calibration.

Fig. 9. Experimental results before the FPT error correction with irregular non-uniform intensity distribution. (a), (b) and (c) the 3D, 2D map and line profile of the calculated phase with the FPT error (PV = 39.7 nm, RMS = 4.7 nm).

Fig. 10. Experimental results after the FPT error correction with irregular non-uniform intensity distribution. (a), (b) and (c) the 3D, 2D map and line profile of the calculated phase after FPT error correction (PV = 39.7 nm, RMS = 3.9 nm).
5. Conclusion

In this paper, we analyze the errors which can introduce the FPT error in the SPSIM and present an effective error correction algorithm which uses all the phase-shifted interferograms and can be applied in the different intensity distribution conditions. This method can correct the error by only one measurement even the intensity distribution is non-uniform. We have demonstrated the proposed error correction method with the simulated data and experimental data of a diamond turned conic surface. This method has the potential applications for the FPT error correction in phase-shifting interferometry.

Funding

Y. Zhang is supported by National Natural Science Foundation of China (NSFC) (11304034), Program of China Scholarship Council (201508220104), Program of Jilin Provincial Educational Department (2015241), Doctoral Research Foundation of Northeast Electric Power University (BSJXM-201218), and Scientific Basic Research Foundation of Northeast Electric Power University. X. Tian is supported by Oculus LLC.