Appl. Phys. B (2014) 115:355–364
DOI 10.1007/s00340-013-5610-3

Applied Physics B
Lasers and Optics

Error analysis of the principal component analysis demodulation algorithm

J. Vargas · J. M. Carazo · C. O. S. Sorzano

Received: 24 February 2013 / Accepted: 7 August 2013 / Published online: 14 September 2013

Abstract In this work, we present suitable phase accuracy indicators, which are obtained from the first three obtained eigenvalues of the principal component analysis (PCA) demodulation algorithm. These indicators can be used in the measuring process to determine a blind phase goodness assessment, without the need of using any ground truth phase information. Therefore, it is possible to perform further actions if required, as obtaining more interferograms or repeat the measure. Additionally, we present simulated and experimental results that support our mathematical analysis and conclusions. A complete MATLAB software package reproducing any result and figure shown in this work is provided in (http://goo.gl/fy5EC).

1 Introduction

Optical metrology may use optical interferometers [1] to measure with high accuracy a wide range of physical quantities as depth, strain analysis, temperature gradients and surface deformation. These physical quantities are determined studying the wavefront deformation when a beam of light passes through, or reflects in the sample of interest. Among wavefront reconstruction techniques based on automated interferogram analysis methods, temporal phase-shifting interferometry (TPI) is accepted as the most precise and accurate one [1]. TPI is an optical metrology technique for measuring the modulating phase of interferograms, using a controlled phase change or piston term between successive interferograms [1]. Other phase-shifting algorithms allow the determination of the modulating phase without the need of any prior knowledge about phase-shifts (asynchronous detection) [2, 3]. These asynchronous methodologies [2, 3] are usually iterative and require an initial guess to start the minimization process, typically by a least-squares minimization. As these methods are based on a minimization process, they require high computational load and processing time. Additionally, these methods require the background and contrast terms to be approximately spatially constant, and the number of interferograms has to be high enough to assure the solution convergence toward the global minimum, independently of the starting guess used.

Recently, a new asynchronous phase-shifting demodulation method based on the principal component analysis (PCA) algorithm, that can extract the phase distribution from unknown randomly phase-shifted interferograms, has been proposed [4, 5]. Different variants and applications have appeared since its publication [6–12]. The PCA demodulation approach is very fast—approximately two orders of magnitude faster than the advanced iterative algorithm (AIA) [2]—as it is not iterative, and it does not require performing a nonlinear optimization. Additionally, the PCA method does not need the background illumination and contrast to be spatially constant.

The different proposed phase-shifting demodulation algorithms (synchronous or asynchronous) cannot determine the accuracy of the obtained phase without the use of a reference or ground truth phase. Therefore, in practical experiments, we typically do not know how accurate our determined phase is. Note that this is an important issue as the experimental determination of the modulating phase is always affected by external disturbances. On one hand, synchronous methodologies, which acquire the interferograms with known phase-shifts between successive

J. Vargas (✉) · J. M. Carazo · C. O. S. Sorzano
Biocomputing Unit, Centro Nacional de Biotecnología-CSIC, C/Darwin 3, 28049 Cantoblanco, Madrid, Spain
e-mail: jvargas@cnb.csic.es

interferograms [1], can present a detuning error in the obtained phase [13, 14]. This is because the "known" phase-shifts may have some mismatch with respect to the actual ones, caused typically by the phase-shifter. On the other hand, asynchronous methods usually require phase-shifts to be well distributed in the [0, 2π] (rad) range. If phase-shifts are randomly distributed in the [0, 2π] range, typically it will be necessary to use a large number of interferograms to meet this requirement. The problem here is the lack of quantitative information about the required number of interferograms to obtain an accurate phase measure. Note that in some cases, with appropriate phase-shifts, five interferograms are sufficient. However, in other cases, it is required to use a much larger interferogram set. Additionally, the modulating phase obtained by any synchronous or asynchronous approach is always affected by the presence of noise in the acquired interferograms and by other external disturbances. Therefore, quantitative information about the accuracy of the retrieved phase is of great importance in practical applications. Note that if the obtained phase accuracy indicator is not acceptable, it is possible to further define automated actions as repeating the measure or acquiring additional interferograms.

In this work, we present a deep mathematical error analysis of the principal component analysis (PCA) demodulation algorithm [4, 5]. This analysis allows us to define suitable phase accuracy indicators for the PCA demodulation algorithm and a retrieved phase goodness criterion. These indicators can be used in practical applications to determine a blind estimate of the phase goodness, without the need of using any ground truth phase information, and to perform further actions if required.

In Sect. 2, we present the error analysis of the PCA method and we define the phase accuracy indicators and a phase goodness criterion. Section 3 includes some simulations and in Sect. 4, we show the experimental results. Finally, in Sect. 5, conclusions are drawn.

2 Theoretical foundations

In phase-shifting interferometry, an interferogram sequence of N samples can be described using the following expression

$$I_n = a + b\cos[\Phi + \delta_n] + \eta_n \quad n = [1, N] \tag{1}$$

where $I_n(x, y)$ is the nth phase-shifted interferogram, with size $N_x \times N_y$, $a = a(x, y)$ is the background component, $b = b(x, y)$ is the modulation or contrast term, $\Phi = \Phi(x, y)$ is the modulating phase map, δ_n are the phase-steps, that in our case can be randomly distributed in the [0, 2π] (rad) range, and $\eta_n = \eta_n(x, y)$ is the noise term that usually has a Gaussian distribution with zero mean. Note that the spatial dependence has been omitted for the sake of clarity. Expression (1) can be rewritten as

$$I_n = a + b(\cos[\Phi]\cos[\delta_n] - \sin[\delta_n]\sin[\Phi]) + \eta_n. \tag{2}$$

From Eq. (2) and grouping terms, we obtain

$$I_n = a + \alpha_{1n} I_c + \alpha_{2n} I_s + \eta_n \tag{3}$$

where $\alpha_{1n} = \cos[\delta_n]$, $\alpha_{2n} = -\sin[\delta_n]$ and $I_c = b\cos[\Phi]$, $I_s = b\sin[\Phi]$ correspond to the quadrature signals with size $N_x \times N_y$. From the set of phase-shifted interferograms, the background can be estimated by a temporal average as

$$a \cong \sum_{n=1}^{N} I_n \Big/ N \tag{4}$$

and we may define the background free interferogram as

$$\tilde{I}_n = I_n - a = \alpha_{1n} I_1 + \alpha_{2n} I_2 + \eta_n \tag{5}$$

Expression (5) shows that a background filtered interferogram set can be expressed as a linear combination of two signals and a noise term. Using Eq. (5), we can express the whole set of measurements as

$$\tilde{I} = Q\Lambda + N \tag{6}$$

with

$$\Lambda = [\alpha_1, \alpha_2]^T \tag{7}$$

and

$$Q = [q_1, q_2] \tag{8}$$

where N is the noise matrix of size $N_xN_y \times N$, where the nth column is taken columnwise from η_n; Λ is a $2 \times N$ matrix, where α_1 and α_2 are column vectors of size $N \times 1$ and $[\cdot]^T$ denotes the transposing operation. Note that the nth element of α_1 and α_2 corresponds to α_{1n} and α_{2n}, respectively. Q is a matrix of size $N_xN_y \times 2$ formed by the quadrature components, q_1 and q_2, that are column vectors with size N_xN_y whose elements are taken columnwise from I_c and I_s respectively. Finally, $\tilde{I}$ is a matrix with size $N_xN_y \times N$ where the nth column is taken columnwise from $\tilde{I}_n$.

The covariance matrix C of $\tilde{I}$ is given by $C = \tilde{I}^T\tilde{I}$ and using Eq. (6) corresponds to

$$C = \Lambda^T Q^T Q\Lambda + \Lambda^T Q^T N + N^T Q\Lambda + N^T N \tag{9}$$

Observe that matrices $(\Lambda^T Q^T)N$ and $N^T(Q\Lambda)$ are significantly smaller than $\Lambda^T Q^T Q\Lambda$ and $N^T N$ in a Frobenius norm sense as they correspond to the product of two uncorrelated matrices. Therefore, we can approximate Eq. (9) as

$$C \cong \Lambda^T Q^T Q\Lambda + N^T N \tag{10}$$

Observe that $\Lambda\Lambda^T$ corresponds to

$$\Lambda\Lambda^T = \begin{pmatrix} \|\alpha_1\|^2 & \langle\alpha_1,\alpha_2\rangle \\ \langle\alpha_1,\alpha_2\rangle & \|\alpha_2\|^2 \end{pmatrix} \tag{11}$$

with $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the norm and inner product operators given by $\|\alpha_1\|^2 = \sum_{n=1}^{N}\cos^2(\delta_n)$, $\|\alpha_2\|^2 = \sum_{n=1}^{N}\sin^2(\delta_n)$ and $\langle\alpha_1,\alpha_2\rangle = \sum_{n=1}^{N}\cos(\delta_n)\sin(\delta_n)$. In general, the two columns of Λ are not orthogonal and $\Lambda\Lambda^T$ is not diagonal. Note that because $\Lambda\Lambda^T$ is a real and symmetric matrix, it can always be diagonalized as $\Lambda\Lambda^T = P_\Lambda^T D_\Lambda P_\Lambda$, where D_Λ and P_Λ are 2×2 diagonal and orthogonal matrices, respectively. We can transform Λ and generate a new matrix $\hat{\Lambda}$ that verifies $\hat{\Lambda}\hat{\Lambda}^T = I$ as

$$\hat{\Lambda} = D_\Lambda^{-1/2} P_\Lambda \Lambda. \tag{12}$$

The diagonal elements of D_Λ are real and can be analytically calculated to be

$$\lambda_\pm = \frac{\left(\|\alpha_1\|^2+\|\alpha_2\|^2\right) \pm \sqrt{\left(\|\alpha_1\|^2-\|\alpha_2\|^2\right)^2+4\langle\alpha_1,\alpha_2\rangle^2}}{2} \tag{13}$$

with

$$D_\Lambda = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \tag{14}$$

and P_Λ is given by

$$P_\Lambda = \begin{pmatrix} \frac{\left(\lambda_+-\|\alpha_1\|^2\right)}{\sqrt{\langle\alpha_1,\alpha_2\rangle^2+\left(\lambda_+-\|\alpha_1\|^2\right)^2}} & \frac{\langle\alpha_1,\alpha_2\rangle}{\sqrt{\langle\alpha_1,\alpha_2\rangle^2+\left(\lambda_--\|\alpha_2\|^2\right)^2}} \\ \frac{\langle\alpha_1,\alpha_2\rangle}{\sqrt{\langle\alpha_1,\alpha_2\rangle^2+\left(\lambda_+-\|\alpha_1\|^2\right)^2}} & \frac{\left(\lambda_--\|\alpha_2\|^2\right)}{\sqrt{\langle\alpha_1,\alpha_2\rangle^2+\left(\lambda_--\|\alpha_2\|^2\right)^2}} \end{pmatrix} \tag{15}$$

Observe that if the phase-sifts δ_n are equally distributed in the [0, 2π] (rad) range, $\Lambda\Lambda^T$ is diagonal as $\|\alpha_1\|^2 = \|\alpha_2\|^2$ and $\langle\alpha_1,\alpha_2\rangle = 0$. In this case, D_Λ is diagonal and P_Λ is the 2×2 identity matrix.

Additionally, Q^TQ corresponds to

$$\begin{aligned} Q^TQ &= \begin{pmatrix} \|q_1\|^2 & \langle q_1,q_2\rangle \\ \langle q_2,q_1\rangle & \|q_2\|^2 \end{pmatrix} \\ &= \sum_{x,y}\begin{pmatrix} \cos^2(\Phi) & \cos(\Phi)\sin(\Phi) \\ \sin(\Phi)\cos(\Phi) & \sin^2(\Phi) \end{pmatrix} \end{aligned} \tag{16}$$

If there is more than one fringe in the resultant interferograms, we can approximate Q^TQ as

$$Q^TQ = D_Q \cong \kappa\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{17}$$

with $\kappa = \sum_{x,y}\cos^2(\Phi) \cong \sum_{x,y}\sin^2(\Phi)$. Using Eqs. (10), (12) and (17), we have

$$C \cong \hat{\Lambda}^T D_\Lambda^{1/2} P_\Lambda D_Q P_\Lambda^T D_\Lambda^{1/2}\hat{\Lambda} + N^TN \tag{18}$$

Note that as D_Q is approximately diagonal, with its diagonal elements almost equal to each other, it commutes with any matrix. Taking into account that P_Λ verifies that $P_\Lambda^T P_\Lambda = P_\Lambda P_\Lambda^T = I$, as it is an orthogonal matrix, we can rewrite Eq. (18) as

$$C \cong \hat{\Lambda}^T(D_\Lambda D_Q + D_N)\hat{\Lambda} \tag{19}$$

where we have taken into account that $N^TN \cong D_N$, with D_N a diagonal $N\times N$ matrix with all its diagonal elements approximately equal, and equal to the noise variance. Observe that in Eq. (18), we have increased the dimensions of $D_\Lambda D_Q$ from 2×2 to $N\times N$ adding zeros. Therefore, $D_\Lambda D_Q$ only has values different than zero in the first and second diagonal elements. Additionally, we have also increased the dimensions of $\hat{\Lambda}$ from $2\times N$ to $N\times N$ adding conveniently orthonormal rows with respect to the first and second row of $\hat{\Lambda}$. Taking into account that $C = \tilde{I}^T\tilde{I}$, and using Eqs. (10) and (12), we obtain the relationship between the background subtracted intensity matrix, $\tilde{I}$, and the quadrature components in presence of noise as

$$\tilde{I} \cong Q\left(P_\Lambda^T D_\Lambda^{1/2}\hat{\Lambda}\right) + N = \left(QP_\Lambda^T\right)D_\Lambda^{1/2}\hat{\Lambda} + N \tag{20}$$

PCA is a technique from statistics for reducing an image or dataset that transforms a number of possibly correlated images into the smallest number of uncorrelated images called the principal components. PCA is based on the search of orthogonal directions explaining as much variability of the data as possible. The first step of the PCA demodulation algorithm [4–6] consists of obtaining the covariance matrix C of the background subtracted dataset, as $C = \tilde{I}^T\tilde{I}$. Note that because C is real and symmetric, it is always possible to diagonalize this matrix as

$$C = A^TDA \tag{21}$$

where A and D are orthogonal and diagonal matrices, respectively, both with size $N\times N$ (it is assumed that the eigenvalues in D are sorted in descendent order). Observe that this factorization is unique if all eigenvalues are distinct. If the multiplicity of any of the eigenvalues is larger than one, then the factorization is unique up to a permutation of the corresponding columns of A. The principal components of the interferogram set, which correspond to images containing the highest variance in the data, are given by

$$\begin{aligned} Y &= \tilde{I}A^T \\ \tilde{I} &= YA \end{aligned} \tag{22}$$

where Y is a matrix of size $N_xN_y \times N$ and its column vectors (principal components) y_n are orthogonal and uncorrelated. The principal component analysis demodulation approach [4–6] reconstructs the modulating phase from the first and second principal components (y_1 and y_2), that corresponds to the highest eigenvalues, as

$$\Phi = \arctan(I_s/I_c) = \pm\arctan(y_2/y_1) \tag{23}$$

Note that the PCA demodulation algorithm also obtains the diagonal D matrix that contains the different eigenvalues. Taking into account that the matrix factorizations shown in Eqs. (19) and (21) are unique up to possible interchanges in some columns of A and $\hat{\Lambda}$, we can state that A and D correspond to $\hat{\Lambda}^T$ and $D_\Lambda D_Q + D_N$ respectively, up to possible interchanges in some columns. Additionally, from Eqs. (20) and (22), we have that the first two principal components are related to the actual quadrature components as

$$\begin{aligned} y_i &= \tilde{I}\hat{\Lambda}^T = \left((QP_\Lambda^T)D_\Lambda^{1/2} + N\hat{\Lambda}^T\right)_i, \quad i = 1,2 \\ y_i &= \left(N\hat{\Lambda}^T\right)_i, \quad i \in [3,N] \end{aligned} \tag{24}$$

where $(\cdot)_i$ refers to the ith column. Observe that P_Λ^T is in general an unknown 2×2 orthogonal matrix, if we have no information about phase-shifts and its only effect is an unknown phase-shift, or piston term, of θ rad in the obtained modulating phase, which is irrelevant. We can mathematically show this point performing the following transformation to Q

$$\hat{Q} = QP_\Lambda^T \tag{25}$$

In the following, we make use of the fact that P_Λ^T is a 2×2 orthogonal matrix and, consequently, it can be written as $P_\Lambda^T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. After this transformation, Eq. (24) is rewritten as

$$y_i = \tilde{I}\hat{\Lambda}^T = \left(\hat{Q}D_\Lambda^{1/2} + N\hat{\Lambda}^T\right)_i, \quad i = 1,2 \tag{26}$$

The first and second columns of Q, that corresponds to $q_1 = b\cos(\Phi)$ and $q_2 = b\sin(\Phi)$, after the rotation corresponds to $\hat{q}_1 = \cos(\theta)q_1 + \sin(\theta)q_2 = b\cos(\Phi + \theta)$ and $\hat{q}_2 = \cos(\theta)q_1 - \sin(\theta)q_2 = b\sin(\Phi + \theta)$. Therefore, the only effect of P_Λ^T is an unknown phase-shift of θ rad in the modulating phase Φ and, without loss of generality, we can assume that $\theta = 0$ rad and P_Λ^T corresponds to the 2×2 identity matrix. If we suppose that the noise is smaller than the signal, we can write from Eqs. (23) and (24)

$$\Phi = \arctan\left(\frac{q_1\lambda_+^{1/2} + \left(N\hat{\Lambda}^T\right)_1}{q_2\lambda_-^{1/2} + \left(N\hat{\Lambda}^T\right)_2}\right) \cong \arctan\left(\alpha\frac{q_1}{q_2}\right) \tag{27}$$

with $\alpha = (\lambda_+/\lambda_-)^{1/2}$. Observe from Eq. (27) that we will obtain only an accurate phase reconstruction, using the PCA demodulation algorithm, with at least one fringe in the interferograms, if $\alpha \cong 1$ and the noise term is smaller than the signal term. This provides us with a method to quantify the obtained error by the PCA demodulation method. Using Eqs. (19) and (21), we have that the eigenvalues computed by the PCA demodulation method corresponds to $D = D_\Lambda D_Q + D_N$. Taking into account that D_N is approximately a diagonal $N \times N$ matrix, with all its diagonal elements equal to σ^2, we have that the eigenvalues given by the PCA demodulation method have the following mathematical expression,

$$\begin{aligned} D_{11} &= \kappa\lambda_+ + \sigma^2 \\ D_{22} &= \kappa\lambda_- + \sigma^2 \\ D_{ii} &= \sigma^2 \quad i \in [3,n] \end{aligned} \tag{28}$$

Using Eq. (28), we can define two-phase accuracy indicators. One of these indicators (γ_{ps}) takes into account possible phase errors because the phase-shifts are not well distributed in the [0, 2π] (rad) range and then, i.e., we do not have that $\alpha \cong 1$ in Eq. (27). The other indicator (γ_n) takes into account phase reconstruction errors because the presence of noise in the interferograms. These indicators are obtained from the first three eigenvalues by

$$\begin{aligned} \gamma_{ps} &= \left(\frac{(D_{11}/D_{22}) - 1}{10}\right) \\ \gamma_n &= \left(\frac{D_{33}}{D_{11} + D_{22} - 2D_{33}}\right) \end{aligned} \tag{29}$$

Note that in Eq. (29), we have defined $\gamma_{ps} = 0$ if $D_{11} = D_{22}$. Finally, we can combine both indicators in order to define a unique phase error indicator as

$$\begin{aligned} \gamma &= \sqrt{I_{ps}^2 + I_n^2} \\ \gamma &= \sqrt{\left(\frac{(D_{11}/D_{22}) - 1}{10}\right)^2 + \left(\frac{D_{33}}{D_{11} + D_{22} - 2D_{33}}\right)^2} \end{aligned} \tag{30}$$

This phase error indicator can be used to define a phase goodness criterion for the PCA demodulation algorithm. For instance, we may establish that a retrieved phase is reliable if γ is smaller than a certain threshold. We have seen heuristically that a good criterion to establish if a retrieved phase is reliable consists in verifying whether γ is smaller than 10 %.

3 Numerical analysis

In order to verify our mathematical analysis, we have performed some simulations. In the first experiment, we

use an interferogram set composed by $N = 25$ fringe patterns with equidistant phase-shifts. The phase-shift of the nth interferogram corresponds to $\delta_n = 2\pi \times (n/N)$ with $n \in [0, 24]$. The interferograms have a size of 800×800 px and are noiseless. The modulating phase, background and modulation signals are Gaussian shaped, and their mathematical expressions correspond to

$$\begin{aligned} a(x,y) &= 0.5 + 0.3\exp(-5((x-10)^2 + (y-15)^2)/10^5) \\ b(x,y) &= \exp(-(x^2+y^2)/10^5) \\ \Phi(x,y) &= 80\pi\exp(-(x^2+y^2)/10^5) \end{aligned} \tag{31}$$

where (x, y) corresponds to pixel coordinates and the origin of coordinates is placed in the image center. The first two interferograms of this dataset are shown in Fig. 1. In order to verify the presented mathematical analysis, the phase error indicators and the phase goodness criterion presented above, we have obtained the modulating phase using the PCA demodulating approach and using different number of interferograms belonging to the dataset introduced above. First, we compute the modulating phase using the first ten phase-shifted interferograms. Next, we obtain the phase using the first eleven patterns and so on until we use the N phase-shifted interferograms. In every case, we obtain the accuracy of the reconstructed phase computing the root-mean-square (*rms*) error between the computed phase and the reference theoretical phase given in Eq. (27). Additionally, we also calculate for each case γ_{ps}, γ_n and γ phase error indicators. Observe that if we compute the phase using the $N = 25$ available fringe patterns, it is verified that $\alpha = 1$ in Eq. (27) as $\lambda_+ = \lambda_-$ and $\|\alpha_1\|^2 = \|\alpha_2\|^2$, $\langle \alpha_1, \alpha_2 \rangle = 0$, in Eq. (11). In this case, the phase-shifts are exactly equally distributed in the $[0, 2\pi]$ (rad) range. However, if we use fewer patterns, it is not verified that $\alpha = 1$ and it will appear a phase error. This error must be larger as fewer patterns are used. Additionally, observe that in this first experiment, because the interferograms are noiseless, the only source of phase error comes because phase-shifts are not well distributed in the $[0, 2\pi]$ (rad) range and then, $\alpha \neq 1$ in Eq. (27). In Fig. 2(a), we plot the obtained phase error indicators with respect to the accuracy of the reconstructed phases, when different number of interferograms are used. As can be seen from Fig. 2(a), γ_n is approximately zero in every case, as the interferograms used are noiseless. On the other hand, γ_{ps} is highly correlated with the *rms* of the obtained modulating phase. Observe that the curve that relates γ_{ps} with the *rms* values is nonlinear because the nonlinear nature of Eq. (28). Additionally, in Fig. 2b, we show a plot between the obtained phase error indicators and the number of interferograms used to reconstruct the phase. As can be seen from Fig. 2b, both magnitudes are related by a smooth mathematical expression in this case. In Fig. 3, we show the theoretical reference wrapped phase and the obtained wrapped phases, when we use the first ten, nineteen and twenty-five phase-shifted interferograms. Note that in Fig. 3, we present a zoomed version of the wrapped phases to show clearly phase details. From Fig. 3b, we see that the obtained wrapped phase, when the first ten interferograms are used, clearly presents phase errors. In this case, the phase error indicator γ is equals to 80 % and the *rms* corresponds to 0.46 rad. Additionally, we can see that the recovered wrapped phase, when we use the twenty-five phase-shifted interferograms (Fig. 3d), is very similar to the reference phase (Fig. 3a). In this case, γ is equal to 0.08 % and the *rms* corresponds to 0.0034 rad. Finally, when we use nineteen interferograms, the obtained phase presents some errors but it is still reliable. Note that in this case, γ is approximately equal to 9 % and the *rms* is of 0.19 rad. The obtained processing times are 0.32, 0.93 rad and 1.4 s, when are used ten, nineteen and twenty-five phase-shifted interferograms, respectively, with a 2.67 GHz laptop and using MATLAB. Regarding this

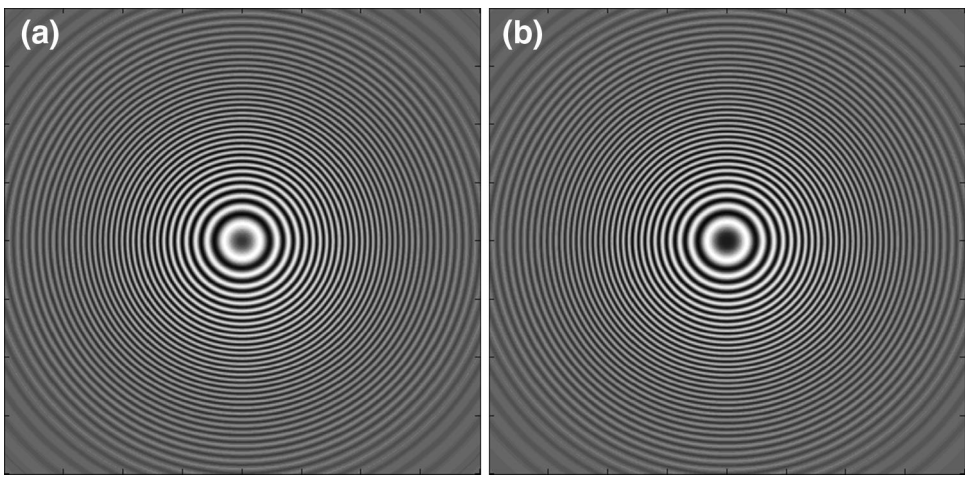

Fig. 1 First two interferograms used in the first simulation

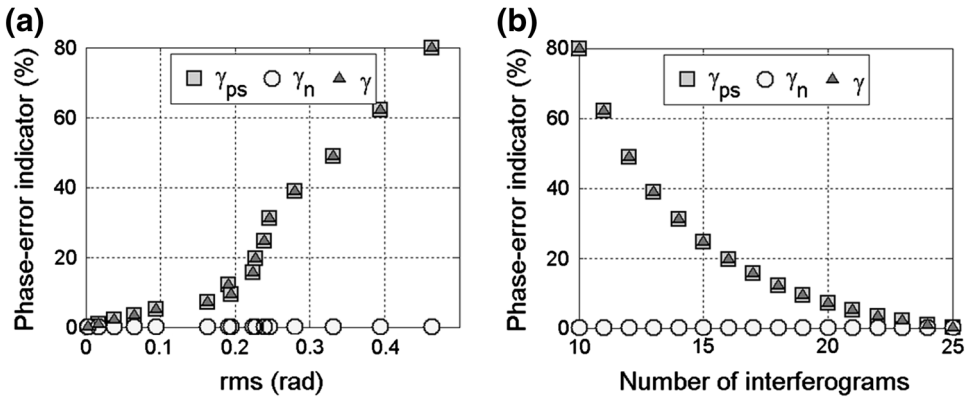

Fig. 2 Obtained phase error indicators with respect to the accuracy of the reconstructed phases (**a**) and the number of interferograms used (**b**) in the first simulation

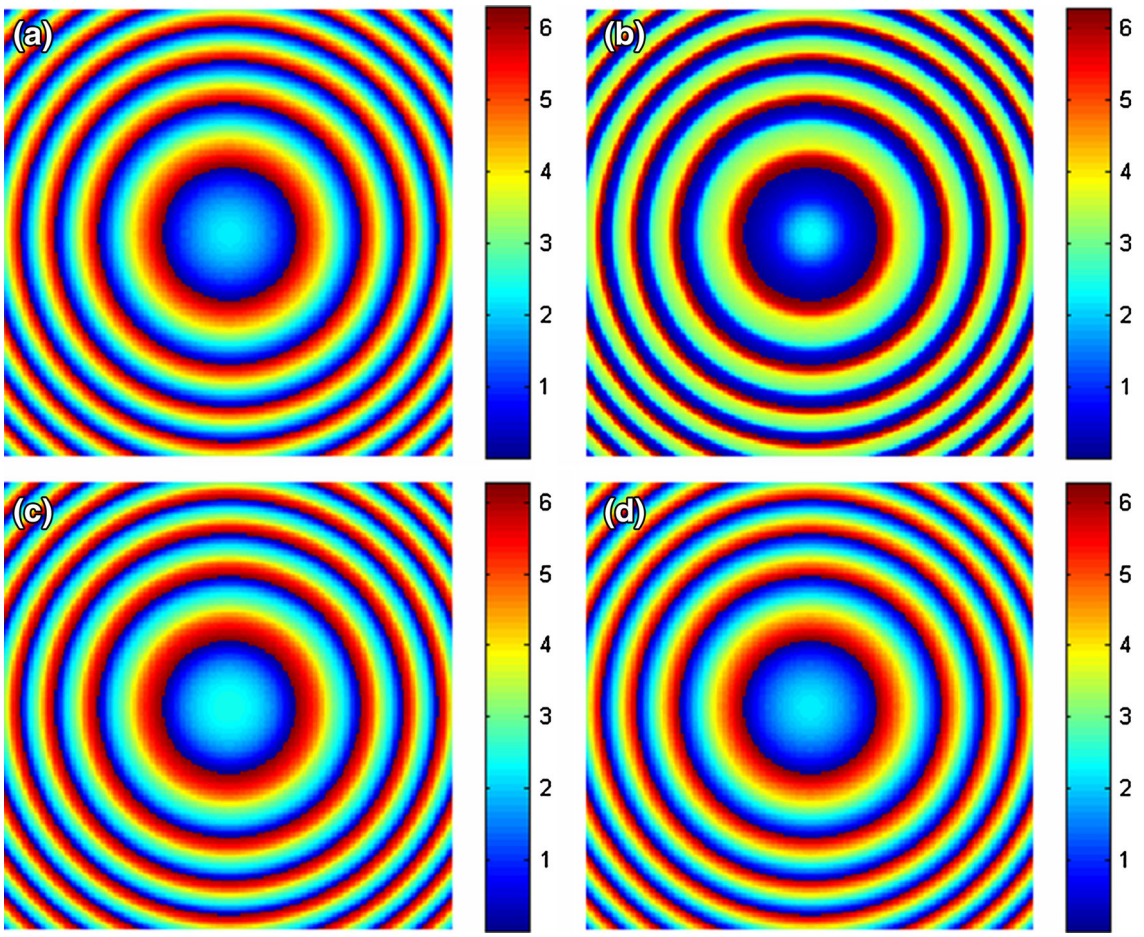

Fig. 3 Zoomed versions of the reference wrapped phase (**a**) and obtained wrapped phases when are used the first thirteen (**b**), nineteen (**c**) and twenty-five (**d**) phase-shifted interferograms in radians

example, we propose that a phase retrieval is considered as acceptable if its γ index is below 10 %.

In the second experiment, we used ten phase-shifted interferograms equally distributed in the $[0, 2\pi]$ (rad) range but with different levels of noise. The modulating phase, background and contrast signals are given in Eq. (31), and the noise is additive and Gaussian with zero mean. We performed eleven phase reconstructions using the PCA demodulation approach and adding different levels of noise to the ten fringe patterns. The patterns have a noise-to-signal ratio (NSR) that goes from 0 to 100 % and is defined as

$$\mathrm{NSR} = \frac{P_\mathrm{n}}{P_\mathrm{s}} \tag{32}$$

where P_n and P_s are the noise and signal power, respectively. Note that we have used this definition instead the

typical signal-to-noise ratio (SNR) to avoid the instabilities of the SNR for low error levels.

In Fig. 4a, we show a plot of the obtained phase error indicators with respect to the accuracy of the reconstructed phases. As can be seen from Fig. 4a, γ_{ps} is approximately zero in every case, as the interferograms are equally distributed in the [0, 2π] range. On the other hand, γ_n is highly correlated with respect to the *rms* of the obtained modulating phase. In Fig. 4b, we show a plot of the obtained phase errors indicators with respect to the NSR of the phase-shifted interferograms. As can be seen from Fig. 4b, in this case, there is a linear relationship between both magnitudes. Finally, in Fig. 5, we show the theoretical reference phase and the obtained unwrapped phases, for NSRs equal to 1, 10 and 81 %. Note that these values correspond to SNRs of 100, 10 and 1.23 %. The wrapped phases obtained by the PCA demodulation algorithm are unwrapped using the method presented in [15, 16] that corresponds to a fast unwrapping two-dimensional algorithm based on a linear recursive filter. This method is robust to noise thanks to its smoothing capabilities. As can be seen from Fig. 5b, the retrieved phase for NSR equals to 1 % is very similar to the reference phase. Note that in this case, the *rms* and the phase error indicator correspond to 0.09 (rad) and 0.8 %, respectively. When the NSR corresponds 81 %, the reconstructed phase shown in Fig. 5d is highly affected by artifacts and is not reliable. In this case, the *rms* and the phase error indicator correspond to 0.9 (rad) and 66 %. Finally, in Fig. 5c, we show the retrieved phase for SNR equals to 10 %. The obtained *rms* and phase error indicator (γ) correspond to 0.27 (rad) and 7.3 %, respectively. Observe that this reconstructed phase is inside the limit of our proposed phase goodness criterion and as can be seen from Fig. 5c, the unwrapped phase is very similar to the reference one. The obtained processing times to compute the results shown in Fig. 5b–d are 0.35, 0.38 and 0.38 s.

4 Experimental results

We have also checked our proposed phase error indicators with real interferograms. We have obtained a reference phase map from a large interferogram set composed by nineteen interferograms with phase-shifts randomly distributed in the [0, 2π] range and using the PCA method [4–6]. In Fig. 6, we show the first three interferograms of this dataset. The images have size of 600 × 800 px. In order to study the performance of the proposed phase error indicators, we have obtained the modulating phases when different numbers of phase-shifting interferograms are processed. This number goes from five to fifteen. For each phase reconstruction, we calculate the different phase error indicators and the *rms* between the obtained and reference phases. In Fig. 7, we show the results obtained. In Fig. 7a, we plot the obtained phase error indicators with respect to the accuracy of the reconstructed phases, and in Fig. 7b, c, we show the obtained phase errors indicators and *rms* with respect to the number of phase-shifted interferograms used. As can be seen from Fig. 7a, the recovered phases are slightly affected by noise and the most significant source of phase error is due to the use of a set of interferograms with phase-shifts not well distributed in the [0, 2π] range. Finally, in Fig. 8, we show the reference phase (a), and the obtained phases when we use three (b), ten (c) and seventeen (d) interferograms. The phase error indicators and *rms* values obtained in these cases are 105, 13, 4 % and 0.60, 0.14 and 0.04 rad, respectively. Observe that the phase map shown in Fig. 8c presents clear phase errors and γ index is larger than 10 %. In this case, the obtained processing times to compute the results shown in Fig. 8b–d are 0.10, 0.15 and 0.29 s, respectively. As can be seen from the results presented above, the proposed accuracy phase indicators, computed from the first three obtained eigenvalues of the principal component analysis algorithm, are very suitable to determine a blind estimate of phase

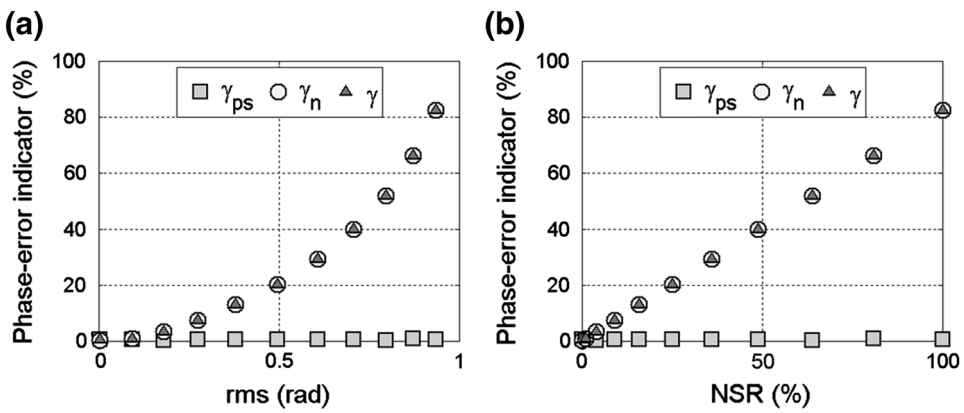

Fig. 4 Obtained phase error indicators with respect to the accuracy of the reconstructed phases (**a**) and the signal-to-noise ratio of the interferograms used (**b**) in the second simulation

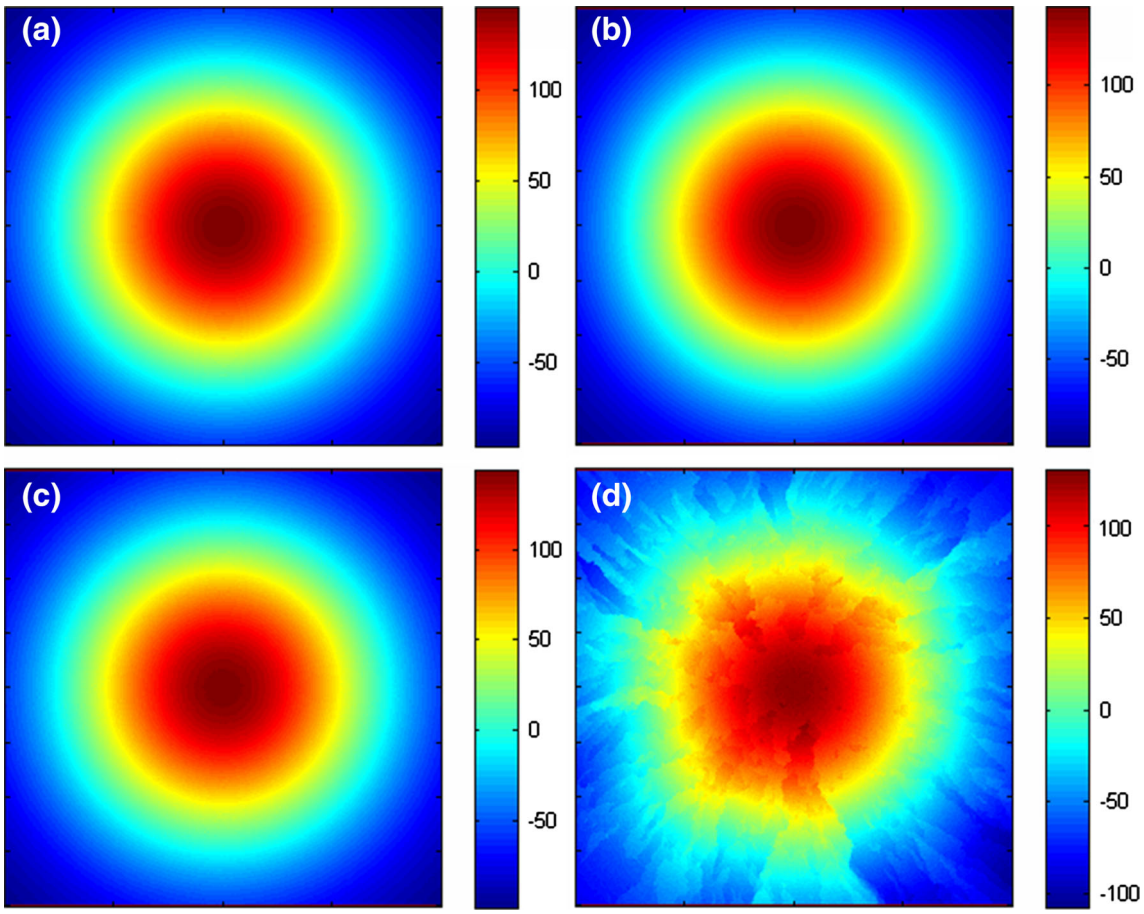

Fig. 5 Reference unwrapped phase (**a**) and obtained unwrapped phases when are used interferograms with noise-to-signal ratios of 9 % (**b**), 36 % (**c**) and 81 % (**d**) in radians

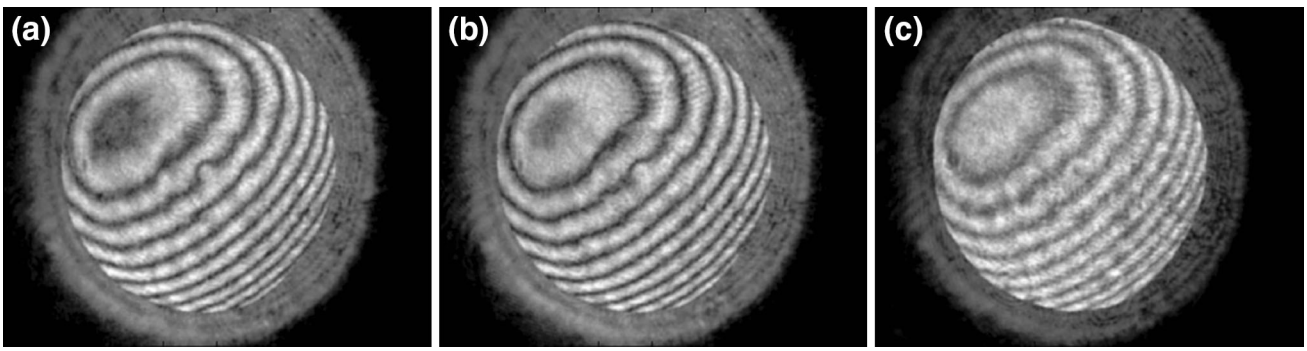

Fig. 6 First three real interferograms used in the Sect. 4

goodness, without the need of using any ground truth phase information, and perform further actions if the retrieved phase accuracy is not acceptable.

5 Conclusions

In this work, we have shown a thorough mathematical error analysis of the principal component analysis demodulation algorithm. In this analysis, we show that the possible phase errors are caused because the interferograms have phase-shifts not uniformly distributed in the [0, 2π] range and/or the presence of noise in the fringe patterns. Additionally, we present suitable phase accuracy indicators, computed from the first three obtained eigenvalues of the principal component analysis demodulation algorithm. These indicators can be used in the measuring process to determine a blind indicator of the phase

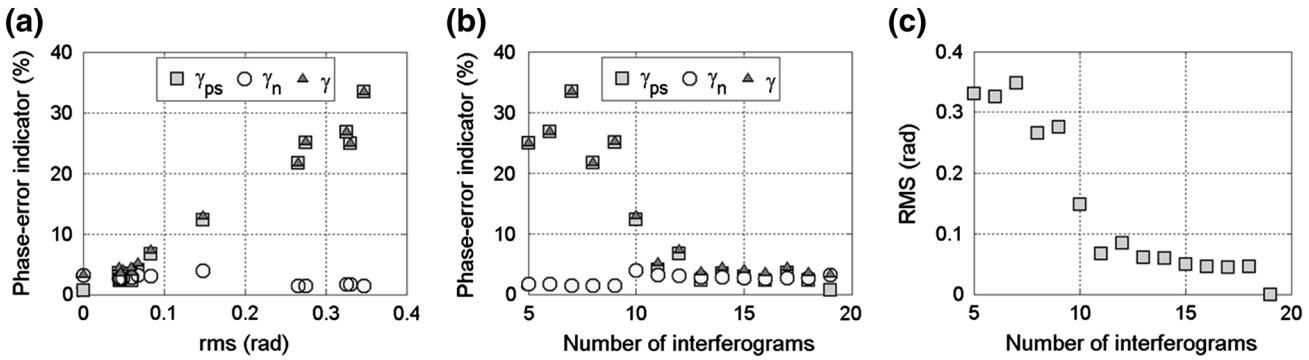

Fig. 7 Obtained phase error indicators with respect to the accuracy of the reconstructed phases (**a**) and the number of interferograms used (**b**). In **c**, it is presented the relation between the *rms* obtained and the number of interferograms processed

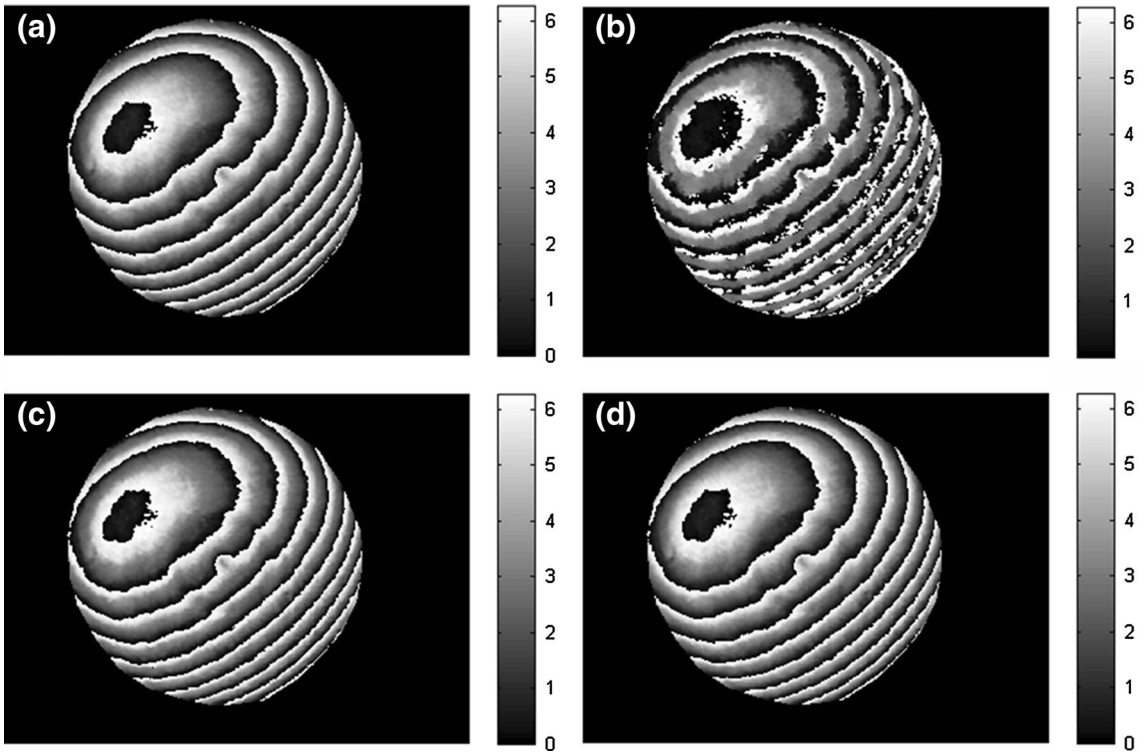

Fig. 8 Reference wrapped phase (**a**) and obtained wrapped phases when are used the first three (**b**), ten (**c**) and seventeen (**d**) phase-shifted interferograms in radians

accuracy, without the need of using any ground truth phase information. If the retrieved accuracy is not acceptable, we can perform further actions as obtaining additional interferograms or repeating the experiment. Additionally, we present simulated and experimental results that support our mathematical analysis and conclusions. A complete MATLAB software package where can be reproduced any result and figure shown in this work is provided in (http://goo.gl/fy5EC).

References

1. D. Malacara, M. Servín, Z. Malacara, *Interferogram Analysis for Optical Testing* (Marcel Dekker, Inc., New York, 1998)
2. Z.Y. Wang, B.T. Han, Advanced iterative algorithm for phase extraction of randomly phase-shifted interferograms. Opt. Lett. **29**(14), 1671–1673 (2004)
3. I.B. Kong, S.W. Kim, General algorithm of phase-shifting interferometry by iterative least-squares fitting. Opt. Eng. **34**(1), 183–188 (1995)

4. J. Vargas, J. Antonio Quiroga, T. Belenguer, Phase-shifting interferometry based on principal component analysis. Opt. Lett. **36**(8), 1326–1328 (2011)
5. J. Vargas, J. Antonio Quiroga, T. Belenguer, Analysis of the principal component algorithm in phase-shifting interferometry. Opt. Lett. **36**(12), 2215–2217 (2011)
6. J. Vargas, C.O.S. Sorzano, Quadrature Component Analysis for interferometry. Opt. Lasers Eng. (2013). doi:10.1016/j.optlaseng.2013.01.004
7. J. Xu, W. Jin, L. Chai, Q. Xu, Phase extraction from randomly phase-shifted interferograms by combining principal component analysis and least squares method. Opt. Express **19**(21), 20483–20492 (2011)
8. J. Vargas, C.O.S. Sorzano, J.C. Estrada, J.M. Carazo, Generalization of the principal component analysis method for interferometry. Opt. Commun. (2012). doi:10.1016/j.optcom.2012.09.017
9. J. Xu, W. Jin, L. Chai, Q. Xu, Principal component analysis of multiple-beam Fizeau interferograms. Opt. Express **19**(15), 14464–14472 (2011)
10. H. Du, H. Zhao. B. Li, S. Cao, Three frames phase-shifting shadow moiré using arbitrary unknown phase steps. Meas. Sci. Technol. **23**(10), 105201–7 (2012)
11. S. Wang, L. Xue, J. Lai, Z. Li, An improved phase retrieval method based on Hilbert transform in interferometric microscopy. Optik (2012). doi:10.1016/j.ijleo.2012.05.029
12. D.N. Schimpf, S. Ramachandran, Polarization-resolved imaging of an ensemble of waveguide modes. Opt. Lett. **37**(15), 3069–3071 (2012)
13. P. Hariharan, B.F. Oreb, T. Eiju, Digital phase-shifting interferometry: a simple error compensating phase calculation algorithm. Appl. Opt. **26**(13), 2504–2506 (1987). http://ao.osa.org/abstract.cfm?URI=ao-26-13-2504
14. J.F. Mosino, M. Servin, J.C. Estrada, J.A. Quiroga, Phasorial analysis of detuning error in temporal phase shifting algorithms. Opt. Express **17**(7), 5618–5623 (2009)
15. M.A. Navarro, J.C. Estrada, M. Servin, J.A. Quiroga, J. Vargas, Fast two-dimensional simultaneous phase unwrapping and low-pass filtering". Opt. Express **20**(3), 2556–2561 (2012)
16. J.C. Estrada, J. Vargas, J.M. Flores-Moreno, J.A. Quiroga, Windowed phase unwrapping using a first-order dynamic system following iso-phase contours. Appl. Opt. **51**(31), 7549–7553 (2012)