Direct phase extraction from interferograms with random phase shifts

Jiancheng Xu,[1,2,3] Qiao Xu,[2,4] Liqun Chai,[2] Yong Li,[1] and Hui Wang[1]

[1]*Institute of Information Optics, Zhejiang Normal University, Jinhua, Zhejiang, 321004, China*
[2]*Chengdu Fine Optical Engineering Research Center, Chengdu, Sichuan, 610041, China*
[3]*xujiancheng@zjnu.cn*
[4]*xuqiao@vip.sina.com*

Abstract: A new method that extracts phase directly from randomly phase-shifted interferograms is proposed. In this method, the background intensity and the modulation amplitude are determined by a temporal method and used as the system parameters and then the phase is extracted from one single interferogram by arc cosine function with range of $(0, \pi)$. By partitioning the extracted phase distribution into watershed regions and determining the sign of phase in each region, the extracted phase is recovered to its principle phase with range of $(-\pi, \pi)$. Numerical simulation and experiment are implemented to verify the effectiveness of this method. The method has applications in both static and dynamic phase measurement.

© 2010 Optical Society of America

OCIS codes: (120.3180) Interferometry; (120.2650) Fringe analysis; (120.5050) Phase measurement.

References and links

1. J. E. Greivenkamp, and J. H. Bruning, "Phase shifting interferometers, " in *Optical Shop Testing,* D. Malacara, ed., John Wiley and Sons, 1992, pp. 501–598.
2. K. Creath, "Temporal phase measurement method," in *Interferogram Analysis,* D.W. Robinson and G. T. Reid, eds. Institute of Physics, Bristol, UK, 1993, pp. 94–140.
3. C. J. Stolz, "The National Ignition Facility: The world's largest optical system," Proc. SPIE **6834**, 683402 (2007).
4. R. Smythe, and R. Moore, "Instantaneous phase measuring interferometry," Opt. Eng. **23**, 361–364 (1984).
5. 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).
6. G. Rodriguez-Zurita, C. Meneses-Fabian, N.-I. Toto-Arellano, J. F. Vázquez-Castillo, and C. Robledo-Sánchez, "One-shot phase-shifting phase-grating interferometry with modulation of polarization: case of four interferograms," Opt. Express **16**(11), 7806–7817 (2008), http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-16-11-7806.
7. K. Okada, A. Sato, and J. Tsujiuchi, "Simultaneous calculation of phase distribution and scanning phase shift in phase shifting interferometry," Opt. Commun. **84**(3-4), 118–124 (1991).
8. Z. Wang, and B. Han, "Advanced iterative algorithm for phase extraction of randomly phase-shifted interferograms," Opt. Lett. **29**(14), 1671–1673 (2004).
9. X. F. Xu, L. Z. Cai, Y. R. Wang, X. F. Meng, W. J. Sun, H. Zhang, X. C. Cheng, G. Y. Dong, and X. X. Shen, "Simple direct extraction of unknown phase shift and wavefront reconstruction in generalized phase-shifting interferometry: algorithm and experiments," Opt. Lett. **33**(8), 776–778 (2008).
10. P. Gao, B. Yao, N. Lindlein, J. Schwider, K. Mantel, I. Harder, and E. Geist, "Phase-shift extraction for generalized phase-shifting interferometry," Opt. Lett. **34**(22), 3553–3555 (2009).
11. Q. Hao, Q. Zhu, and Y. Hu, "Random phase-shifting interferometry without accurately controlling or calibrating the phase shifts," Opt. Lett. **34**(8), 1288–1290 (2009).
12. J. Park, and S.-W. Kim, "Vibration-desensitized interferometer by continuous phase shifting with high-speed fringe capturing," Opt. Lett. **35**(1), 19–21 (2010).
13. Q. Zhang, and X. Su, "High-speed optical measurement for the drumhead vibration," Opt. Express **13**(8), 3110–3116 (2005), http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-13-8-3110.
14. S. Zhang, D. Van Der Weide, and J. Oliver, "Superfast phase-shifting method for 3-D shape measurement," Opt. Express **18**(9), 9684–9689 (2010), http://www.opticsinfobase.org/abstract.cfm?URI=oe-18-9-9684.
15. L. Vincent, and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," IEEE Trans. Pattern Anal. Mach. Intell. **13**(6), 583–598 (1991).

1. Introduction

Phase-shifting interferometry(PSI) [1,2], has been widely used in surface testing of optical elements, especially the large-aperture mirrors in Inertial Confinement Fusion (ICF) [3] or astronomical telescopes. Owing to the large size of measured surface, the optical path length is so long that mechanical vibration and air turbulence will change the preset phase shifts and cause inevitable errors to the measurement. In order to reduce the environmentally induced phase errors, simultaneous [4] or dynamic [5,6] phase-shifting interferometers have been developed to overcome the problem of vibration by capturing four images on four detectors or four images on one detector. However, complicated phase modulators are needed, and the polarization components or micropolarizer arrays in the interferometer must be carefully chosen and aligned to produce π/2 phase shifts, equal intensity and uniform images.

Meanwhile, many algorithms [7–10] of phase retrieval with unknown phase shifts have also been developed. Okada *et al* [7] and Wang *et al* [8] proposed the least-squares-based iterative algorithms which can determine phase shifts and phase distribution simultaneously. Xu *et al* [9] and Gao *et al* [10] introduced some algorithms to directly extract the unknown phase shifts on the assumption that the measured surface has an equal distribution in [0, 2π] over the whole interferogram. If the assumption is not fully met, iterative calculation is also needed to improve the accuracy of the extracted phase shifts. Therefore, all these algorithms cited above take phase shifts as parameters and thus need accurate phase shifts.

Recently Hao *et al* [11] proposed a random phase-shifting interferometry. In the method, the temporal intensity maximum and minimum in each pixel are taken as parameters, and then the phase is obtained by arc cosine function. Hao's method is accurate but time-consuming (20 min for calculation). In addition, it will have problems if the local phase shift between two adjacent interferograms is not small enough.

To deal with this problem, in this paper, we propose a new method to extract phase directly from randomly phase-shifted interferograms. Different from Hao's method, our method partitions the phase distribution into disjoint regions, determines the sign of the phase in each region, and then extracts the phase from its corresponding interferogram. Our method requires no substantial hardware modification in an existing phase-shifting interferometer, offering a cost-effective solution to vibration desensitization [12].Therefore; it can be used for dynamic measurement [13,14]. We will first discuss the principles of the method, and then give its verification by computer simulation and experiment.

2. Principle

Suppose that N frames of phase-shifted interferograms are collected, the intensity of an arbitrary pixel (x, y) in the nth interferogram is expressed as

$$I(x, y, n) = A(x, y) + B(x, y)\cos[\varphi(x, y) + \theta(x, y, n)] \tag{1}$$

where $A(x, y)$ and $B(x, y)$ represent the background intensity and the modulation amplitude in pixel (x, y), respectively. $\varphi(x, y)$ is the measured phase and $\theta(x, y, n)$ is the random phase shift. Here $\theta(x, y, n)$ usually contains three parts: $\theta(x, y, n) = \theta_1(n) + \theta_2(n) + \theta_3(x, y, n)$, where $\theta_1(n)$ is the preset phase shift; $\theta_2(n)$ is a random phase shift brought by mechanical vibration and should be the same for all pixels; and $\theta_3(x, y, n)$ is a random distribution over pixels which is brought by air turbulence. As long as the phase shift between two adjacent frames, i.e., $\theta(x, y, n) - \theta(x, y, n-1)$, is small enough, the intensity at (x, y) can be ergodic over N frames if N is large enough [11]. So the maximum and the minimum intensity, denoted as $I_{\max}(x, y)$ and $I_{\min}(x, y)$, can be easily found from N frames of interferograms [11]. Then $A(x, y)$ and $B(x, y)$ can be derived from

$$A(x, y) = [I_{\max}(x, y) + I_{\min}(x, y)]/2 \tag{2}$$

$$B(x, y) = [I_{max}(x, y) - I_{min}(x, y)]/2 \qquad (3)$$

Since $A(x, y)$ and $B(x, y)$ do not have frame-to-frame variation, i.e., they are functions of pixels only, they can be measured and used as the system parameters for dynamic phase measurement. With $A(x, y)$ and $B(x, y)$ already known, the phase of any frame, denoted as $\Phi(x, y, n) = \varphi(x, y) + \theta(x, y, n)$, can be solved from the intensity $I(x, y, n)$ with arc cosine function [11]

$$\Phi_{0,\pi}(x, y, n) = arc\cos[\frac{I(x, y, n) - A(x, y)}{B(x, y)}] \qquad (4)$$

Since the range of arc cosine function goes from 0 to π, the obtained $\Phi_{0,\pi}(x, y, n)$ by Eq. (4) also goes from 0 to π, as shown in Fig. 1(a).Before phase unwrapping, it is necessary to recover $\Phi_{0,\pi}(x, y, n)$ to its principle phase $\Phi_{-\pi,\pi}(x, y, n)$,whose range goes from $-\pi$ to π. According to the property of arc cosine function, $\Phi_{-\pi,\pi}(x, y, n)$ equals $\Phi_{0,\pi}(x, y, n)$ for those pixels where the principle phase is in $(0, \pi)$ and equals $-\Phi_{0,\pi}(x, y, n)$ for those pixels where the principle phase is in $(-\pi, 0)$.Thus we should partition $\Phi_{0,\pi}(x, y, n)$ into disjoint regions with the same sign of the principle phase and then determine the sign in each region.
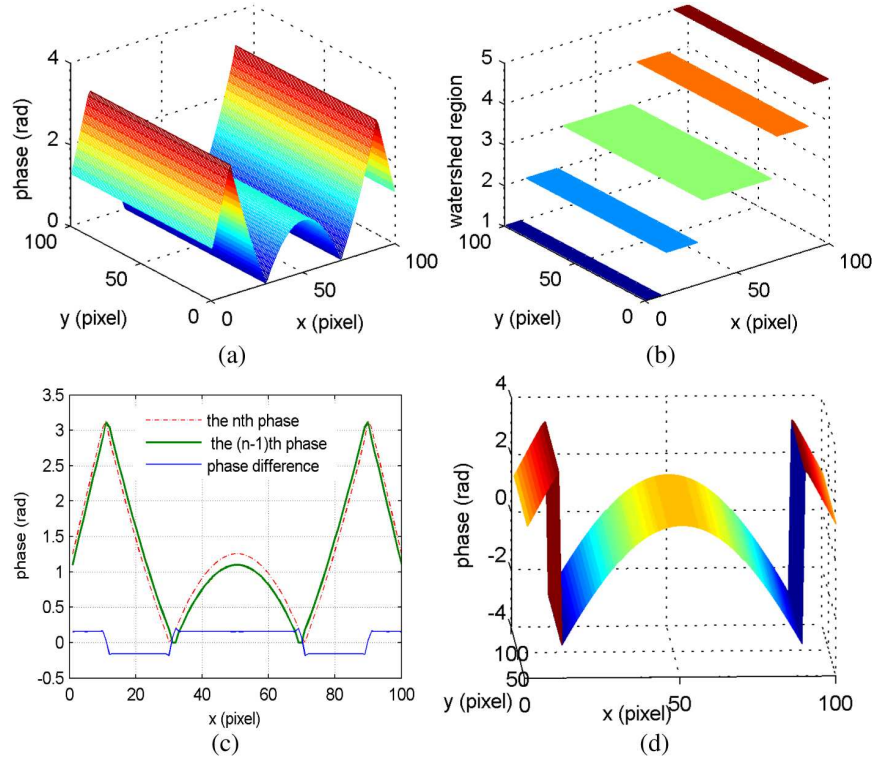

Fig. 1. The process of recovering the extracted phase to its principle phase. (a) The extracted phase by Eq. (4) with range of $(0, \pi)$; (b) the watershed regions of (a), (c) the phase difference between two adjacent frames, and (d) the recovered principle phase of (a) with range of $(-\pi, \pi)$

2.1 Phase segmentation by watershed transformation

According to the property of arc cosine function, we know that the sign of $\Phi_{-\pi,\pi}(x, y, n)$ only changes at the place where $\Phi_{0,\pi}(x, y, n)$ approaches zero or π. Thus the sign of $\Phi_{-\pi,\pi}(x, y, n)$

is the same for every continuous region where the range of $\Phi_{0,\pi}(x,y,n)$ is limited to $(\varepsilon, \pi\text{-}\varepsilon)$ (here ε is a given small value). To partition $\Phi_{0,\pi}(x,y,n)$ into disjoint regions with the same sign of the principle phase, watershed transformation [15], a method of choice for image segmentation, is used according to the local maximum and minimum points of $\Phi_{0,\pi}(x,y,n)$. The process of watershed segmentation has the following two steps. (1)Converting $\Phi_{0,\pi}(x,y,n)$ into binary image. In each pixel of the binary image, the value is 1 if $\Phi_{0,\pi}(x,y,n) \leq \varepsilon$ or $\Phi_{0,\pi}(x,y,n) \geq \pi - \varepsilon$, otherwise it is 0.So in those continuous 0-value regions, the range of $\Phi_{0,\pi}(x,y,n)$ is limited to $(\varepsilon, \pi\text{-}\varepsilon)$ and the sign of $\Phi_{-\pi,\pi}(x,y,n)$ is the same in each continuous 0-value region. (2)Applying conventional watershed algorithm [15] to the binary image. We use the function of Watershed in Matlab to obtain the watershed regions from the binary image. In the function of Watershed, the watershed ridge lines are defined as those local points where the value of the binary image is 1. The watershed ridge lines partition $\Phi_{0,\pi}(x,y,n)$ into watershed regions where the sign of the principle phase is the same. Figure 1(b) is the watershed regions of Fig. 1(a). It shows that there are five watershed regions in Fig. 1(a) and the watershed ridge lines locate at the place where x approaches the 10 [th],30 [th], 70th and 90th pixels. It is easy to identify the watershed regions since each watershed region has its corresponding number. As shown in Fig. 1(b), 1-value pixels belong to the first watershed region, 2-value pixels belong to the second watershed region, and so on.

If there are very dense fringes in practical phase measurement, the obtained $\Phi_{0,\pi}(x,y,n)$ by Eq. (4) will have a high-frequency oscillation between zero and π. Since the lateral resolution of the CCD is limited, $\Phi_{0,\pi}(x,y,n)$ fails to get the values of 0 and π in some periods of the oscillation. In this case, to obtain accurate watershed regions, the value of ε in step 1 should be chosen larger, such as $\varepsilon = 0.3$ rad. However, a larger ε will lead to more 1-value pixels in the binary image. Those pixels do not belong to any watershed region, so it is hard to determine the sign of its principle phase for those pixels. Thus we should thin the 1-value pixels of the binary image to lines by morphologic method and make most of the pixels belong to the watershed regions.

2.2 Sign determination by phase difference

To determine the sign of phase in each watershed region, we define the phase difference $\Delta\Phi_{0,\pi}(x,y,n)$ between the nth and $(n\text{-}1)$th frames as

$$\Delta\Phi_{0,\pi}(x,y,n) = \Phi_{0,\pi}(x,y,n) - \Phi_{0,\pi}(x,y,n-1) \qquad (5)$$

Figure 1(c) shows the relation among the nth phase, the $(n\text{-}1)$th phase and their phase difference. Since $\Phi(x,y,n) = \varphi(x,y) + \theta(x,y,n)$, $\Delta\Phi_{0,\pi}(x,y,n)$ equals $\Delta\theta(x,y,n)$ for the place where both the nth and the $(n\text{-}1)$th principle phases are in $(0, \pi)$, and equals $-\Delta\theta(x,y,n)$ for the place where both the nth and the $(n\text{-}1)$th principle phase are in $(-\pi, 0)$.Here $\Delta\theta(x,y,n) = \theta(x,y,n) - \theta(x,y,n-1)$. If the average of $\Delta\theta(x,y,n)$ over pixels is a small positive number (this requirement can be achieved by choosing a proper preset phase shift $\theta_1(n)$ and incorporating a high-speed camera), then the sign of the principle phase in the mth watershed region can be determined as

$$\Phi_{-\pi,\pi}(m,n) = \begin{cases} \Phi_{0,\pi}(m,n) & if\ \overline{\Delta\Phi_{0,\pi}(m,n)} \geq 0 \\ -\Phi_{0,\pi}(m,n) & otherwise \end{cases} \qquad (6)$$

where $\Delta\overline{\Phi_{0,\pi}}(m,n)$ means the average of the phase difference in the mth watershed region. With the sign of the principle phase in one watershed region (e.g. the mth watershed region) already known, the principle phase in all watershed regions can be easily recovered because the signs of the principle phase in the adjacent watershed regions (i.e., the mth and the $(m + 1)$th watershed regions) have opposite signs. Figure 1(d) shows the recovered principle phase from Fig. 1(a) by the proposed method. After phase unwrapping, the phase $\Phi(x, y, n)$ can be finally obtained from its corresponding interferogram $I(x, y, n)$. In practical phase measurement, it requires that $A(x, y)$ and $B(x, y)$ do not have frame-to-frame variation and must be measured as the system parameters before phase measurement.
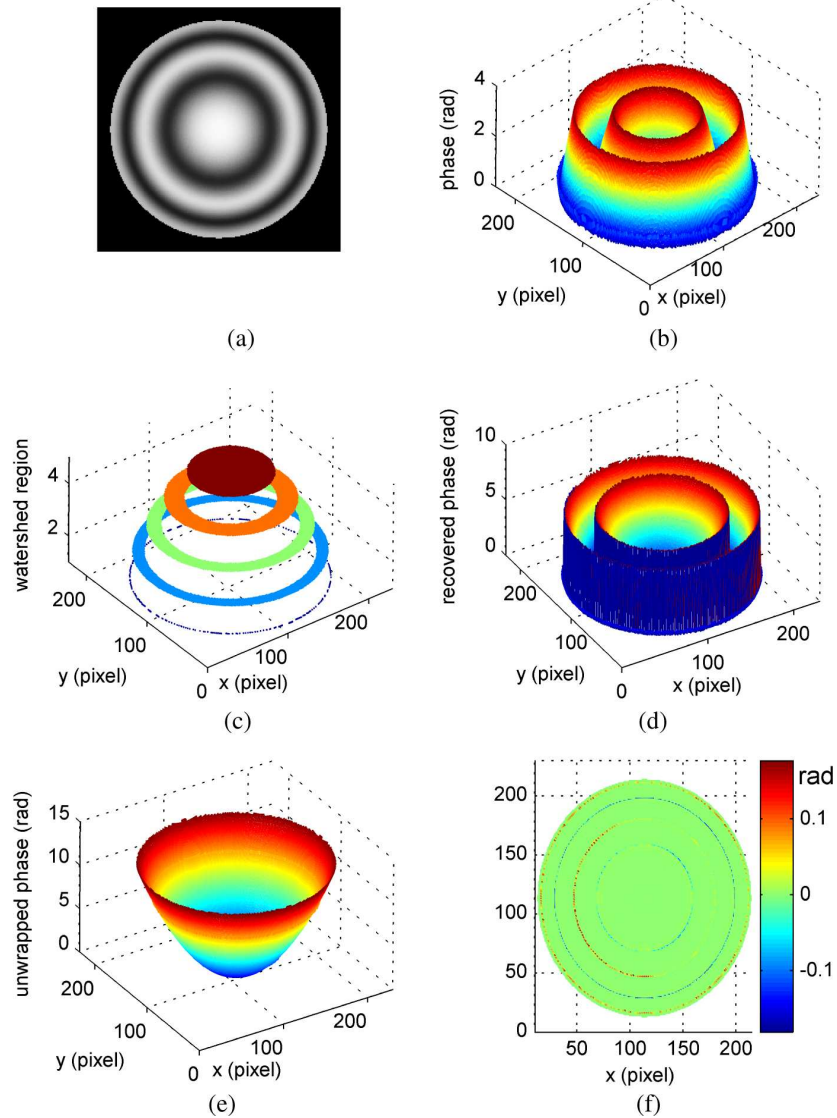

Fig. 2. Simulation results. (a) interferogram (Media 1), (b) the calculated phase by Eq. (4) with range of $(0, \pi)$, (c) the watershed regions of (b), (d) the recovered phase with range of $(0, 2\pi)$ (Media 2), (e) the unwrapped phase and (f) the residual phase error.

3. Numerical simulation and discussion

To prove the effectiveness of the proposed method, 100 frames of randomly phase-shifted interferograms are generated according to Eq. (1) by setting the parameters as follows. $A = 140\exp[0.4(x^2 + y^2)]$, $B = 110\exp[0.4(x^2 + y^2)]$ and $\varphi = 4\pi(x^2 + y^2)$ where $x^2 + y^2 \leq 1$. $\theta_1(n) = n\pi/10$, $\theta_2(n)$ is a random number ranging from -0.1 to 0.1 rad, and $\theta_3(x, y, n)$ is a random distribution over pixels and ranges from -0.04 to 0.04 rad. One of the generated interferograms with pixels of 200×200 is shown in Fig. 2(a). After $A(x, y)$ and $B(x, y)$ are determined from 100 frames of randomly phase-shifted interferograms, the phase of Fig. 2(a) is calculated by Eq. (4) and shown in Fig. 2(b). It shows that the range of Fig. 2(b) is $(0, \pi)$. By setting $\varepsilon = 0.2$ rad, the watershed regions of Fig. 2(b) are obtained and shown in Fig. 2(c).It shows that there are 5 watershed regions in Fig. 2(b). By determining the sign of the principle phase in each watershed region and adding 2π to those watershed regions whose principle phase is negative, the principle phase of Fig. 2(b) is recovered and shown in Fig. 2(d) with range of $(0, 2\pi)$. Finally, the measured phase φ is obtained by phase unwrapping and shown in Fig. 2(e), and the residual phase error is calculated by comparing the obtained phase with the preset phase and shown in Fig. 2(f). Figure 2(f) shows that the residual phase error mainly locates at the place where $\Phi_{0,\pi}(x, y, n)$ [Fig. 2(b)] approaches zero or π. The reason is that our method fails to determine the sign of the principle phase where $\Phi_{0,\pi}(x, y, n) \leq \varepsilon$ or $\Phi_{0,\pi}(x, y, n) \geq \pi - \varepsilon$. The result also shows that the peak-to-valley (PV) and RMS values of the residual phase error are 0.39 (near 2ε) and 0.03rad respectively.

In addition, 20 frames of the randomly phase-shifted interferograms are generated according to Eq. (1) and shown in Media 1. These corresponding phases (20 frames) are extracted one by one and shown in Media 2. It shows that our method can extract the phase from one single interferogram if $A(x, y)$ and $B(x, y)$ are determined, thus it has potential application in dynamic measurement.

Some factors that influence the accuracy of the proposed method should be analyzed and discussed. Our method assumes that the number of frames N is large enough and the intensity at (x, y) is ergodic over N frames. If the number of fringe patterns used is not large enough, the intensity at (x, y) is not ergodic, and the obtained $A(x, y)$ and $B(x, y)$ will have errors which will cause inevitable errors to phase measurement. Thus, the residual phase error of our method depends on the number of fringe patterns used. We assume that the number of pixels in interferogram is 200×200, $\varepsilon = 0.2$, and the other parameters of the interferograms are the same as the above simulation. Then the residual phase error is calculated for $N = 10$, 20,40,60,80 and 100 and the result is shown in Table 1. It shows that the residual phase error decreases with the number of fringe patterns used increasing. Since the preset phase shift $\theta_1(n) = n\pi/10$ and the random phase shift $\theta_2(n) < 0.2$, the intensity at (x, y) is near ergodic over more than 20 frames. Thus, when N is smaller than 20, the residual phase error is reduced effectively as N increases; and when N is larger than 20, the residual phase error is not reduced significantly as N increases. Therefore, in order to reduce the residual phase error, the number of fringe patterns used should be larger than $2\pi/\theta_1(1)$ for small random phase shift. In practical phase measurement, because of mechanical vibration and air turbulence, the number of fringe patterns used should be about 2~5 times as large as $2\pi/\theta_1(1)$.

Table 1. The relation between the residual phase error and the number of fringe patterns used. (The number of pixels is (200 × 200), ε = 0.2, and unit is rad)

Number of frames	10	20	40	60	80	100
PV of phase error	0.6578	0.4051	0.3990	0.3950	0.3927	0.3900
RMS of phase error	0.2294	0.0377	0.0343	0.0327	0.0316	0.0305

In Table 1, it also shows the PV and RMS of residual phase error is bigger than 0.39 and 0.03rad even if $N = 100$. The reason is that our method fails to determine the sign of the principle phase where $\Phi_{0,\pi}(x,y,n) \le \varepsilon$ or $\Phi_{0,\pi}(x,y,n) \ge \pi - \varepsilon$ and ε cannot be set too small because the number of pixels ($M \times M$) in interferogram is limited. We calculate the residual phase error for different numbers of pixels and values of ε and the result is shown in Table 2. It shows that the residual phase error decreases with the number of pixels increasing and the value of ε decreasing. Thus in order to reduce the residual phase error of the proposed method, we should capture more interferograms, use high resolution CCD and choose a small value of ε.

Table 2. The relation between the residual phase error and the number of pixels ($M \times M$) and the preset small value of ε. (N is 40 and unit is rad)

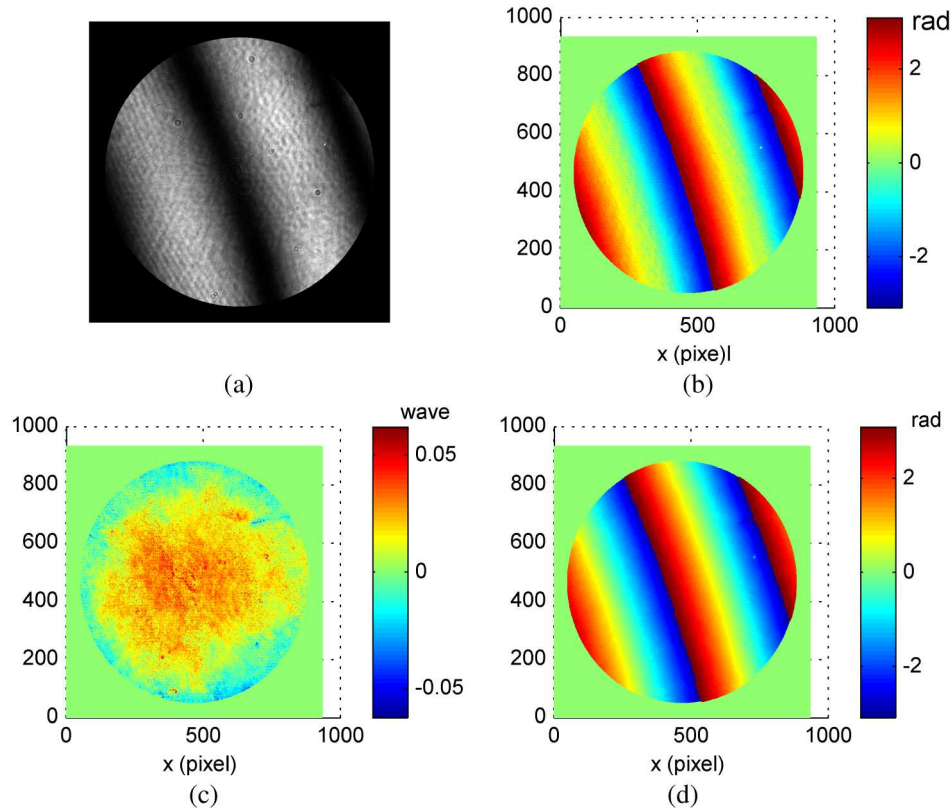|	M = 200, ε = 0.2	M = 400, ε = 0.2	M = 400, ε = 0.1	M = 600, ε = 0.1
PV of phase error	0.3990	0.2697	0.2577	0.1993
RMS of phase error	0.0343	0.0292	0.0234	0.0183

(a)

(b)

(c)

(d)

Fig. 3. Experiment results. (a) interferogram (Media 3), (b) the principle phase of (a) by our method (Media 4), (c) the unwrapped phase of (b) after tilt subtraction, and (d) the principle phase of (a) by Wang's iterative method.

4. Experiment

Optical experiment has also been carried out to investigate the performance of our method. An optical flat with aperture of 100mm is measured in a standard phase-shifting Fizeau interferometer with PZT uncalibrated and active vibration isolation workstation turned off. 100 frames of interferograms, whose preset phase shift is $n\pi/8$, are captured at a frame frequency of 20 fps. The background intensity and the modulation amplitude are determined

by Eq. (2) and Eq. (3) from these 100 frames of randomly phase-shifted interferograms. One typical interferogram is shown in Fig. 3(a). The noisy interferogram leads to noisy phase. Thus, before watershed transformation, a low-pass filter is used to remove the noise of $\Phi_{0,\pi}(x,y,n)$, and ε is set as 0.3 to obtain a binary image without breaks in the center of 1-value pixels. Then we thin the 1-value pixels to lines by morphologic method. Finally, the principle phase of Fig. 3(a) is recovered and shown in Fig. 3(b) with range of (-π, π). By phase unwrapping and tilt subtraction, the measured phase of optical flat is obtained and shown in Fig. 3(c). The PV and RMS of the measured phase is 0.1268 and 0.0126 λ (here λ is the wavelength of light and equals 632nm). Meanwhile, we extract the phase from 4 frames of interferograms [including Fig. 3(a)] by Wang's iterative algorithm [8] with initial phase shift of $n\pi/8$ .After 10 iteration cycles, the principle phase is recovered and shown in Fig. 3(d).

After phase unwrapping and tilt subtraction, the PV and RMS of the measured phase is 0.1115 and 0.0111λ. It shows that the principle phase from our method coincides well with that from Wang's iterative algorithm. However, our method extracts the phase directly from one single interferogram, while Wang's iterative algorithm extracts the phase from four phase-shifted interferograms. So our method is more sensitive to noise than Wang's iterative algorithm, which can be easily demonstrated by comparing Figs. 3(b) with 3(d). But if we extract the four phases from the four phase-shifted interferograms by our method, and then calculate the average of the four phase, we find that the PV and RMS of the averaged phase are 0.1154 and 0.0117 λ respectively, which is very close to that of Wang's iterative algorithm. Therefore, our method will become insensitive to local noise by use of average operation, which is useful for static phase measurement. On the other hand, our method has potential application in dynamic measurement. For example, Media 3 shows 16 frame of randomly phase-shifted interferograms, which are captured in dynamic environment. By use of our method, these corresponding phases (16 frames) are extracted one by one and shown in Media 4. It shows that our method extracts the phase from one interferogram without iteration, so it is suitable for dynamic measurement.

5. Conclusion

We have presented a method that directly extracts phase from randomly phase-shifted interferograms, and the performance of this method is tested numerically and experimentally.

In this method, the temporal intensity maximum and minimum in each pixel, instead of phase shifts, are taken as parameters, the background intensity and the modulation amplitude are determined and used as the system parameters, and the phase is extracted from one single interferogram by arc cosine function. Numerical simulation shows that we should capture more interferograms, use high resolution CCD and choose a small value of ε to reduce the phase error of the proposed method. Experiment shows that the result from our method coincides well with that from Wang's iterative algorithm. Although hundreds of interferograms are needed to determine the background intensity and the modulation amplitude, they can be measured and used as the system parameters before phase measurement. Finally, our method extracts the phase information from one single interferogram and thus it works well for dynamic measurement.