Improvement of least-squares integration method with iterative compensations in fringe reflectometry

Huang, Lei; Asundi, Anand Krishna

2012

Huang, L., & Asundi, A. K. (2012). Improvement of least-squares integration method with iterative compensations in fringe reflectometry. Applied optics, 51(31), 7459-7465.

https://hdl.handle.net/10356/95887

https://doi.org/10.1364/AO.51.007459

Improvement of least-squares integration method with iterative compensations in fringe reflectometry

Lei Huang* and Anand Asundi

School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore

*Corresponding author: huanglei0114@gmail.com

Received 30 July 2012; accepted 31 August 2012;
posted 28 September 2012 (Doc. ID 173476); published 22 October 2012

Least-squares integration is one of the most effective and widely used methods for shape reconstruction from gradient data, which result from gradient measurement techniques. However, its reconstruction accuracy is limited due to the imperfection of the Southwell grid model, which is commonly applied in the least-squares integration method. An operation with iterative compensations is therefore proposed, especially for the traditional least-squares integration method, to improve its integration accuracy. Simulation and experiment are carried out to verify the feasibility and superiority of the proposed operation. This compensatory operation with iterations is suggested, and its good performance on integration accuracy improvement is shown. © 2012 Optical Society of America

OCIS codes: 150.6910, 120.3940, 110.3010.

1. Introduction

There are a variety of optical techniques in three-dimensional (3D) shape metrology for industrial inspection applications [1]. Among these measurement techniques, some are referred to as direct methods, as they directly get 3D results from the received optical information, e.g., the locations of laser spots or the distortion of pattern images. Other techniques deduce the surface shape by integration of the gradient information. The fringe reflection technique [2,3] and Shack–Hartmann methods [4] are two representatives of these indirect 3D shape measurement methods. A process to integrate the gradient data is consequently required to reconstruct the surface shape or the wavefront.

This integration process, commonly with a two-dimensional (2D) data set, does affect the profile results, and hence research efforts are being put to explore how to make this process more accurate and faster. Fried [5], and Hudgin [6,7] studied the wavefront reconstruction from phase difference or slope using least-squares fitting and evaluated the error propagation in reconstructions. Southwell [8] provided the wavefront estimation from slope with the least-squares method in more detail including its grid models, solutions, and errors. Li *et al.* [9] reviewed several evaluation methods, including the path-guided integration method, the Fourier-transform-based integration method, and the least-squares integration method, for gradient measurement techniques and made a comparison of these integration methods showing the superiority of the least-squares integration method in terms of accuracy, robustness, and flexibility. Legarda-Saenz and Espinosa-Romero [10] presented a Fourier-transform-based method to reconstruct the wavefront by integrating the multiple directional derivatives for deflectometric measurement. Koskulics *et al.* [11] further developed the least-squares integration methods with Southwell grid model by simultaneously considering the surface elevation, slope, and curvature for surface retrieval when the surface elevation, slope, and curvature are available from the measurement.

The least-squares integration method with Southwell grid model [8] is basically good at handling large

1559-128X/12/317459-07$15.00/0

gradient data sets due to its low memory cost by using sparse matrices. However, the Southwell grid model assumes the surface could be expressed as biquadratic functions. Obviously, this assumption is not always be satisfied in practical situations, in which the surface cannot be described with biquadratic functions, and finally it results in integration errors.

This work consequently aims to reduce the integration errors from the imperfection of the Southwell grid model in the traditional least-squares integration method. Section 2 first recalls the least-squares integration with Southwell model and then describes the operation of iterative compensations step by step. Simulations are demonstrated in Section 3 to show the feasibility and effectiveness of the proposed iterative compensations in improving the integration accuracy. Section 4 shows the experimental results by using the proposed method with the comparison of traditional method. Section 5 discusses the merits and limitations of the proposed method, and Section 6 concludes the work.

2. Method

In the gradient-measuring optical metrology methods, numerical integration converts gradient data into surface shape or profile by using least-squares fitting. At first, it is very necessary to recall the traditional least-squares integration with Southwell grid model [8]. Following the Southwell grid model with a grid size of $M \times N$ as shown in Fig. 1, the relationship between slope and shape can be expressed as

$$\begin{cases} \frac{p_{i,j+1}+p_{i,j}}{2} = \frac{z_{i,j+1}-z_{i,j}}{x_{i,j+1}-x_{i,j}}, & i = 1,\ldots,M, \quad j = 1,\ldots,N-1 \\ \frac{q_{i+1,j}+q_{i,j}}{2} = \frac{z_{i+1,j}-z_{i,j}}{y_{i+1,j}-y_{i,j}}, & i = 1,\ldots,M-1, \quad j = 1,\ldots,N \end{cases}, \tag{1}$$

where x, y, z are the world coordinates. p and q denote the slopes in x and y directions, respectively, i.e., $p = dz/dx$ and $q = dz/dy$.

According to Eq. (1), the height information can be acquired as

$$\begin{cases} z_{i,j+1} - z_{i,j} = \frac{1}{2}(p_{i,j+1} + p_{i,j})(x_{i,j+1} - x_{i,j}), \\ \quad i = 1,\ldots,M, \quad j = 1,\ldots,N-1 \\ z_{i+1,j} - z_{i,j} = \frac{1}{2}(q_{i+1,j} + q_{i,j})(y_{i+1,j} - y_{i,j}), \\ \quad i = 1,\ldots,M-1, \quad j = 1,\ldots,N \end{cases}, \tag{2}$$

Furthermore, Eq. (2) can be rewritten in terms of matrices as

$$\mathbf{DZ} = \mathbf{G}, \tag{3}$$

where

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & -1 & 0 & \cdots & 0 & 1 \\ \hline -1 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix}_{[(M-1)N+M(N-1)]\times MN}, \tag{4}$$

$$\mathbf{Z} = \begin{bmatrix} z_{1,1} \\ z_{2,1} \\ \vdots \\ z_{M,N} \end{bmatrix}_{MN\times 1}, \tag{5}$$

$$\mathbf{G} = \frac{1}{2}\begin{bmatrix} (p_{1,2}+p_{1,1})(x_{1,2}-x_{1,1}) \\ (p_{1,3}+p_{1,2})(x_{1,3}-x_{1,2}) \\ \vdots \\ (p_{M,N}+p_{M,N-1})(x_{M,N}-x_{M,N-1}) \\ (q_{2,1}+q_{1,1})(y_{2,1}-y_{1,1}) \\ (q_{3,1}+q_{2,1})(y_{3,1}-y_{2,1}) \\ \vdots \\ (q_{M,N}+q_{M-1,N})(y_{M,N}-y_{M-1,N}) \end{bmatrix}_{[(M-1)N+M(N-1)]\times 1}. \tag{6}$$

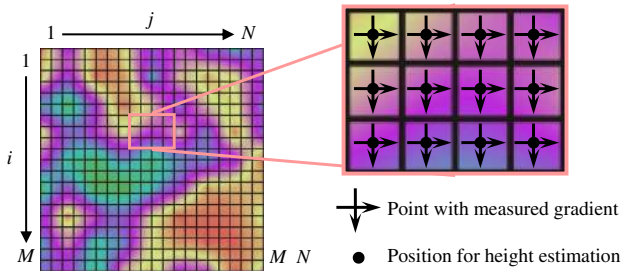

Fig. 1. (Color online) In Southwell grid model, the points for height estimation are at the same locations of those whose gradient data have been measured out.

In a practical situation ($MN > M + N$), Eq. (3) is commonly overdetermined, i.e., $(M-1)N+$

$M(N-1) > MN$. Hence, a least-squares estimation can be carried out to provide the height distribution as an optimized solution. By recalling the traditional least-squares integration, it can be found that there is an assumption implied in Eq. (1) that the slope distribution is bilinear within each tiny quadrilateral, i.e., the variation of surface height in a single quadrilateral is biquadratic. As mentioned above, this assumption is not always satisfied in actual measurements, and hence it is one of the major error sources in integration.

With noticing the existing imperfection of this assumption, an iterative compensation procedure is thus suggested to improve the integration accuracy. Basically, higher-order terms are contained in these residual gradient data, which cannot be fit with the biquadratic surface. It is possible to compensate the errors due to higher-order terms by integrating these residual gradient data. With several iterative compensations, the final shape can be accurately reconstructed free from the influence of the imperfect assumption. The specific steps of the algorithm are described as follows.

Step 1. Integrate the gradient data with the traditional least-squares integration method recalled above to get initial height $z_0(x,y)$, and set the current height $z(x,y) = z_0(x,y)$.

Step 2. Calculate the residual slopes in x and y directions $dp(x,y)$ and $dq(x,y)$ with current height distribution $z(x,y)$.

Step 3. Integrate the residual slopes $dp(x,y)$ and $dq(x,y)$ with the traditional least-squares integration method as well to get $z_c(x,y)$ for compensation.

Step 4. Compensate the height $z(x,y) = z(x,y) + z_c(x,y)/n_k$, where the parameter for the kth compensation n_k is an empirical constant, which can be determined through simulation with the ideal surface containing a certain higher-order polynomial component, and k stands for the number of compensation times.

Step 5. Repeat the compensation as a loop from Step 2 to Step 4 until the compensating term $z_c(x,y)/n_k$ is less than a preset threshold z_{thr}.

Step 6. Record the current $z(x,y)$ as the final height.

The empirical parameters n_k are determined by simulation with surfaces containing different higher-order components. Through our calculation, these empirical parameters n_k are $n_1 \approx 3.0000$, $n_2 \approx 4.0909$, $n_1 \approx 3.0000$, $n_3 \approx 4.9476$, and $n_4 \approx 5.6768$ when $k = 1, 2, 3$, and 4, respectively. From the simulation study, the traditional integration is intrinsically able to perfectly integrate the surface with polynomial components up to the second order as analyzed above. Furthermore, the encountered error when reconstructing a surface with polynomial components up to the $2(n+1)$th order could be nicely compensated by taking the nth compensation.

3. Simulation

It is the objective of the proposed method to make improvement on integration accuracy. Since benchmarks are easy to find in simulation for assessment of the algorithm performance, a series of simulations are conducted to investigate the feasibility and efficiency of the iterative compensation.

The first example in simulation is to show the iterative compensation works efficiently to improve the integration accuracy when dealing with a case that the assumption of biquadratic form is not satisfied. Both x and y coordinates range from −5 to 5 mm with an interval of 0.02 mm; i.e., the size of matrices is 500 by 500, which yields 500,000 x- and y-slope values in total. As shown in Fig. 2(a), the true out-of-plane dimension is predefined as

$$z = 0.3\cos(0.4x^2 + 2x)\cos(0.4y^2 + 2y) + 0.7\cos[(x^3 + y^2)/4\pi], \quad (7)$$

and its corresponding slopes in x and y directions p and q are shown in Figs. 2(b) and 2(c).

The surface shape of the tested sample is specially chosen on purpose to investigate the iterative performance of the proposed method. In Fig. 2(b), the x slope varies relatively large when the absolute value of x is relatively big, and the y slope varies slightly at the lower left corner and acutely at the upper right corner in Fig. 2(c).

As a comparison, the result of the traditional method is shown at first to provide a rough idea of the conventional integration work. The result from the traditional least-squares integration method with no compensation is demonstrated in Fig. 3(a), which shows large integration errors as the true surface does not satisfy the biquadratic surface assumption in the traditional method described in Section 2. With a strategy of compensating these errors in Fig. 3(a), the proposed method with iterations could accurately reconstruct the out-of-plane height distribution.

As shown in Figs. 3(b)–3(d), with the proposed compensation process after the first compensation, the peak-to-valley value of height error significantly reduces to around 0.4% [from about 2000 nm in Fig. 3(a) down to about 8 nm in Fig. 3(b)], and even down to around 0.15% after the third compensation with iterations [from about 2000 nm in Fig. 3(a) down to about 3 nm in Fig. 3(d)].

The majority of compensated amount is completed after one or two compensations, which is also shown in Fig. 3(e). The standard deviation (Std.) of the height errors varies from about 250 nm by using the traditional method with no error compensation to less than 0.7 nm after the first compensation,

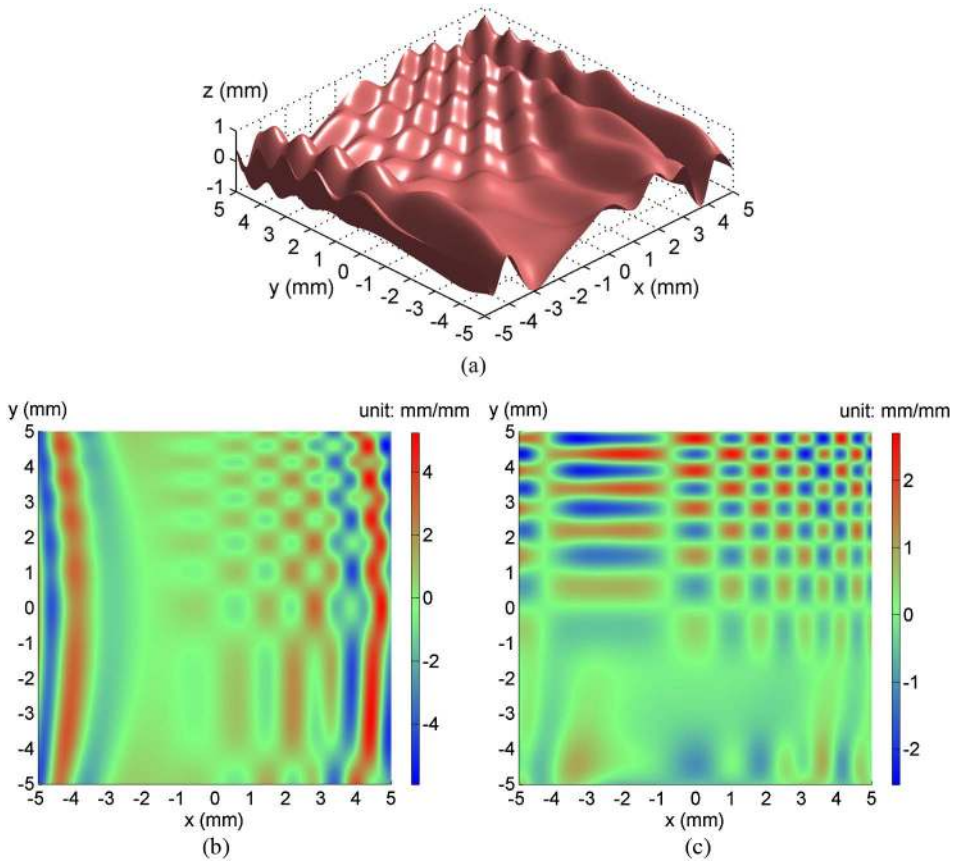

Fig. 2. (Color online) True shape is designed as the ground truth, and its corresponding gradient data are analytically calculated in simulation to test the performance of the iterative compensation. (a) The true shape, and the corresponding true slopes in x-direction $p = dz/dx$ (b) and in y-direction $q = dz/dy$ (c).

and even to about 0.15 nm after the third compensation with iterations. A one-line profile of the error is shown in Fig. 3(f) as well to see the capability of the proposed method to enhance the integration accuracy. As shown in Fig. 3(g), after the third iterative compensation, the out-of-plane height is effectively and accurately reconstructed. Since there is no noise existing on the gradient data in this simulation, the residual errors after iteration can be considered the so-called algorithm error. The next step is therefore to investigate the performance of the proposed method under noisy condition.

The second simulation demonstrates the ability of the proposed method to operate in the presence of noise. The true height shown in Fig. 4(a) is defined by

$$z = 3(1-x)^2 \cdot e^{-x^2-(y+1)^2} - 10\left(\frac{x}{5} - x^3 - y^5\right) \cdot e^{-x^2-y^2} - \frac{1}{3}e^{-(x+1)^2-y^2}, \tag{8}$$

where the in-plane dimensions x and y are limited within a range from -1 to 1 with sampling points of 400 by 400 and the corresponding out-of-plane height varies by more than 5 mm.

In a practical reflectometric measurement with fringe phase, noise exists on retrieved fringe phase, and further it can be approximately considered as normally distributed on surface-normal directions. Normally distributed angular noise with a standard deviation of 8 arc sec is added on the normal directions, which is a typical noise level in practical measurement with fringe reflection technique. The associated noisy slopes in x and y directions are shown in Figs. 4(b) and 4(c).

The integration error from the traditional least-squares method is shown in Fig. 5(a) without any compensation. The height error is not only resulted from the influence from gradient noise but also from the mismatch between the assumed biquadratic surfaces and the actual surface to be reconstructed. Most of the errors due to this assumption can be compensated through the proposed iterative process with showing the height error in Fig. 5(b). In addition, a one-line profile ($y = 0$) of errors before and after compensation is compared as well in Fig. 5(c), from which it can be easily found that the iterative process improves the integration accuracy with successful elimination of the waviness, and the residual error is mainly from the gradient noise. The result shows the iterative compensation is still effective in the presence of noise.

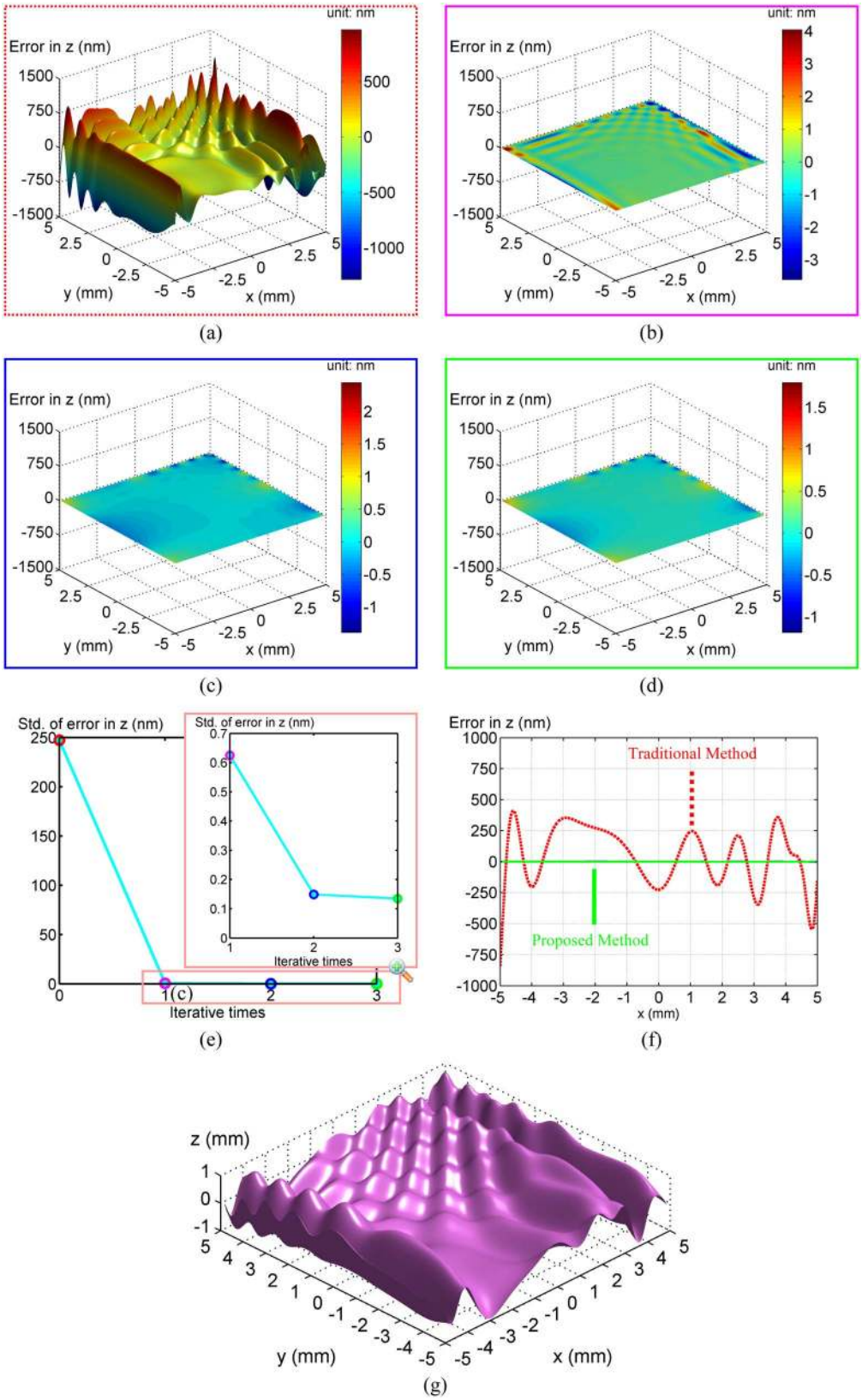

Fig. 3. (Color online) Integration results with the traditional method and the proposed iterative compensation are compared, showing the validity of the iterations. (a) Height errors with no compensation with the traditional method, (b) after the first compensation, (c) after the second compensation, (d) after the third compensation, (e) standard deviations of height error before and after compensations, (f) one-line profile of the errors with the traditional method and proposed compensation approach, and (g) the reconstructed result after the third compensation.

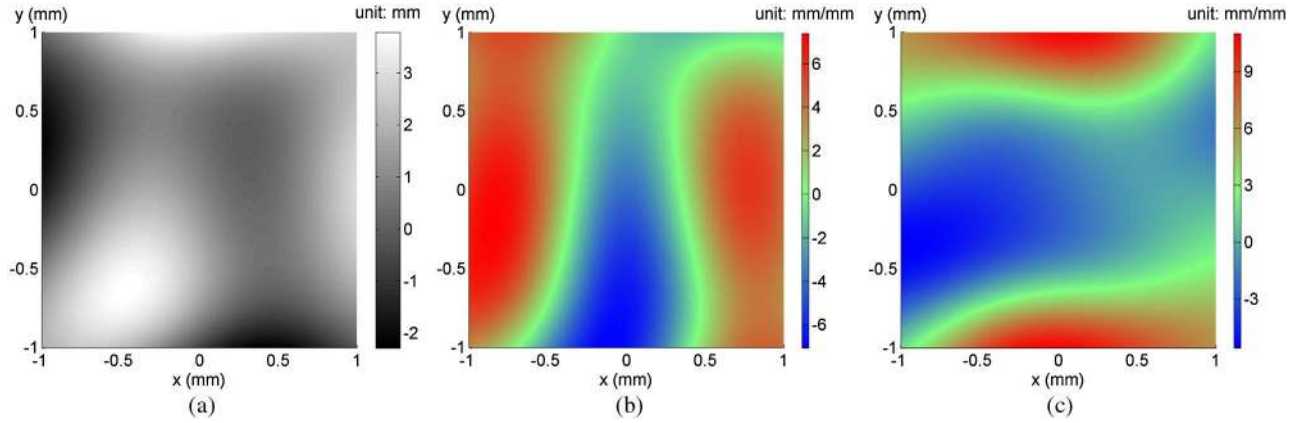

Fig. 4. (Color online) Integrate gradient data with noise. (a) True shape to be reconstructed, (b) slope $p = dz/dx$, and (c) slope $p = dz/dy$.

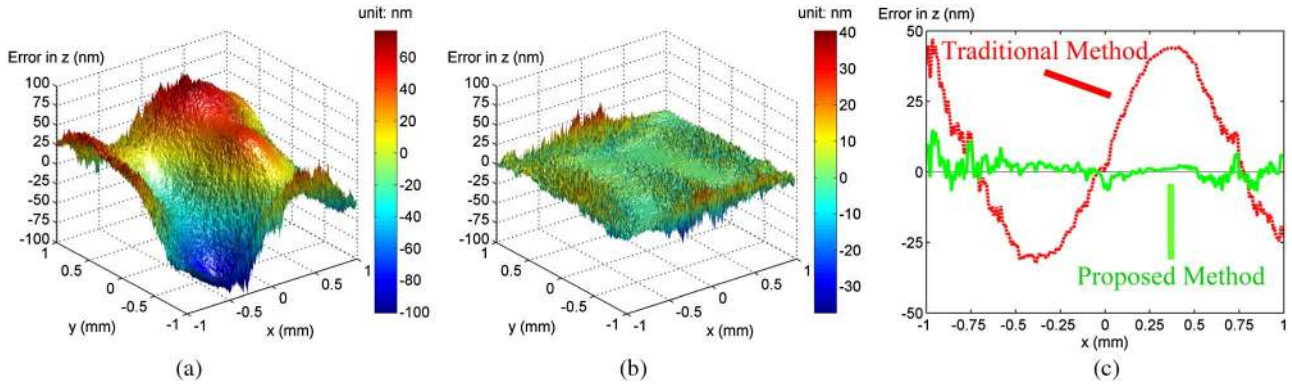

Fig. 5. (Color online) A comparison is carried out between the traditional method and the proposed iterative compensation with the existence of noise on gradient data. (a) Height error before compensation, (b) height error after iterative compensation, and (c) profiles of height error in one line ($y = 0$).

4. Experiment

To verify the validity of the proposed iterative compensation process, a practical measurement is carried out with fringe reflection technique [2,3]. The target to measure in the experiment is a concave mirror. The fringe patterns with phase shifting are displayed on an LCD, screen and the typical images captured by a CCD camera are shown in Fig. 6(a). The slope data in x and y directions determined by the phase measuring reflectometric method are shown in Figs. 6(b)–6(c).

By applying the proposed integration method with iterative compensations on the measured gradient data, the surface shape of the concave mirror can be reconstructed as shown in Fig. 6(d). It indicates the proposed method is effective and feasible in a practical measurement. In addition, the shape difference between the proposed and the traditional methods is also investigated and shown in Fig. 6(e). Improvement in integration accuracy is achieved by using the iterative operation.

5. Discussion

Generally speaking, a fast, accurate, and easily implemented iterative compensation process is presented to improve the integration accuracy by compensating the errors due to imperfection of biquadratic surface assumption. The iterations with high efficiency do not sacrifice much computational time. The merits of the revised integration method with the proposed iterative compensations can be summarized as follows.

First, the proposed compensation process improves the accuracy of the traditional least-squares integration method. It is able to compensate the error from the imperfection of assumption implied in the traditional least-squares integration method. This is the major improvement.

Second, as improving the accuracy by iterations, the algorithm is still fast since the iterative compensation is very effective. Commonly, only one or two compensations are enough to get a satisfactory result.

Third, the same as the traditional least-squares method, the proposed method is able to cope with both a complete and incomplete data set. A data set with a size of 500,000 can be handled by the proposed method without stitching.

The limitation for this proposed method is also investigated. As the traditional least-squares integration method, the proposed method can only handle a gradient data set with rectangular mesh grid, and it

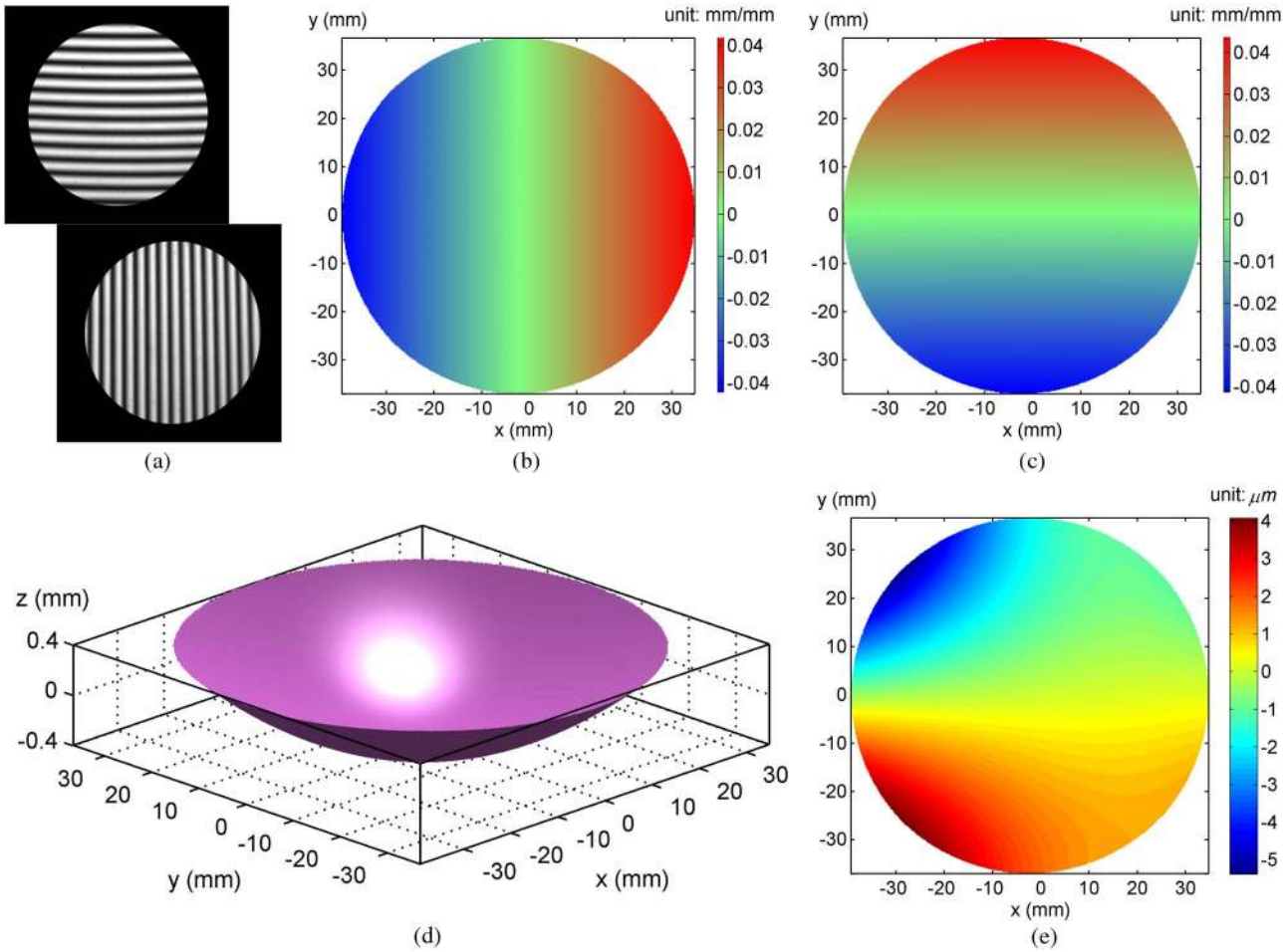

Fig. 6. (Color online) A comparison of the traditional and proposed methods is also carried out with experimental data from the fringe reflection technique. (a) Fringe patterns from fringe reflection technique, (b) slope in x-direction, (c) slope in y-direction, (d) integrated mirror surface by using proposed method, and (e) the shape difference between proposed method and the traditional least squares integration method.

is not able to integrate an arbitrarily distributed data set, but generally the data are in a rectangular mesh grid in many practical optical measurements.

6. Conclusion

This work aims to improve the least-squares integration method with iterative compensation to solve the issue of the incorrect biquadratic shape assumption. The proposed method is investigated by both simulation and experiment. Improvement in integration accuracy is verified by comparing with the traditional least-squares integration method. The merits of the proposed method are accurate, fast, and able to handle large data sets. In summary, this least-squares integration with iterative compensation method is an effective and accurate 2D integration tool to handle shape from slope problems in some gradient-measuring-based optical inspection applications.

References

1. F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," Opt. Eng. **39**, 10–22 (2000).
2. M. C. Knauer, J. Kaminski, and G. Häusler, "Phase measuring deflectometry: a new approach to measure specular free-form surfaces," Proc. SPIE **5457**, 366–376 (2004).
3. T. Bothe, W. Li, C. von Kopylow, and W. P. O. Jüptner, "High-resolution 3D shape measurement on specular surfaces by fringe reflection," Proc. SPIE **5457**, 411–422 (2004).
4. B. C. Platt and R. Shack, "History and principles of Shack–Hartmann wavefront sensing," J. Refractive Surg. **17**, S573–S577 (2001).
5. D. L. Fried, "Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements," J. Opt. Soc. Am. **67**, 370–375 (1977).
6. R. H. Hudgin, "Wave-front reconstruction for compensated imaging," J. Opt. Soc. Am. **67**, 375–378 (1977).
7. R. H. Hudgin, "Optimal wave-front estimation," J. Opt. Soc. Am. **67**, 378–382 (1977).
8. W. H. Southwell, "Wave-front estimation from wave-front slope measurements," J. Opt. Soc. Am. **70**, 998–1006 (1980).
9. W. Li, T. Bothe, C. von Kopylow, and W. P. O. Jüptner, "Evaluation methods for gradient measurement techniques," Proc. SPIE **5457**, 300–311 (2004).
10. R. Legarda-Saenz and A. Espinosa-Romero, "Wavefront reconstruction using multiple directional derivatives and Fourier transform," Opt. Eng. **50**, 040501 (2011).
11. J. Koskulics, S. Englehardt, S. Long, Y. Hu, and K. Stamnes, "Method of surface topography retrieval by direct solution of sparse weighted seminormal equations," Opt. Express **20**, 1714–1726 (2012).