Optics and Lasers in Engineering 67 (2015) 176–181

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Shape reconstruction from gradient data in an arbitrarily-shaped aperture by iterative discrete cosine transforms in Southwell configuration

Lei Huang [a,*], Mourad Idir [a], Chao Zuo [b], Konstantine Kaznatcheev [a,b,c,d], Lin Zhou [a,c], Anand Asundi [d]

[a] Brookhaven National Laboratory - NSLS II, 50 Rutherford Dr., Upton, NY 11973-5000, USA
[b] Jiangsu Key Laboratory of Spectral Imaging & Intelligence Sense, Nanjing University of Science and Technology, Nanjing 210094, China
[c] College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China
[d] School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore

ARTICLE INFO

Article history:
Received 29 September 2014
Received in revised form
18 November 2014
Accepted 20 November 2014
Available online 10 December 2014

Keywords:
Shape reconstruction from gradient
Wavefront reconstruction
Arbitrarily-shaped aperture
Discrete cosine transform
Southwell configuration

ABSTRACT

The shape reconstruction from gradient data is a common problem in many slope-based metrology applications. In practice, the gradient data may not be ideally available for the whole field of view as expected, due to the aperture or the unmeasurable part of sample. An iterative method by using discrete cosine transforms is addressed in this work to deal with the integration problem with incomplete gradient dataset in Southwell configuration. Simulation indicates that the discrete cosine transform provides better initial values than discrete Fourier transform does, and it converges to a more accurate level by updating with spectrum-based slopes comparing to the slope updates from finite difference in classical method. Experimental results show the feasibility of the proposed approach in a practical measurement.

1. Introduction

It is an integration problem to reconstruct a shape from its gradient data (the first derivatives) measured by slope sensors. The essence of this integration problem is to solve a Poisson equation with Neumann boundary conditions [1]. With decades of study in various fields, many "integrators" has been proposed [2–8]. These problem solvers are very popular and have been widely applied in wavefront metrology [9], phase measuring deflectometry [10–13], and even phase unwrapping [14], etcetera. After the proposal of finite-difference solvers with various configurations in late 1970 s [2–4], the Discrete Fourier Transform (DFT) has been used to reconstruct the wavefront in 1980 s [5], whereas the DFT algorithm essentially operates in rectangular domains. Instead of using DFT method which only provides a suboptimal solution with its implicit periodic boundary conditions, it is suggested to expand the estimation into cosine series, and as a result the reconstruction error is significantly reduced [8]. By simply padding slope matrices with flipped and positive or negative slope values accordingly, Discrete Cosine Transform (DCT) integration method can be implemented using the Fast Fourier Transform (FFT) algorithm [15,16].

However, since optical pupils are commonly circular, not in rectangle, the pixels outside the aperture do not provide available data from the measurement. In addition, the central region of the aperture in real applications, such as astronomy and X-ray metrology, may be blocked for a certain purpose. Moreover, the sample under test is not always ideal to measure and some regions are not measurable by a particular technique. Due to the reasons above, some pixels in sensor fail to offer valid measures as a result [see Fig. 1]. This brings troubles to the integration process, especially to the transform-based integration approaches. To solve this issue, missing slopes outside an arbitrarily-shaped aperture are extrapolated with an iterative DFT method [6,7], which is basically an application of the Gerchberg iteration [17]. Taking the loop continuity into account, another idea of slope extrapolation is proposed by Poyneer et al. with boundary and extension methods for Hudgin or Fried configuration [18]. Zou and Rolland employ the finite-difference solvers into the Gerchberg iteration to estimate wavefront for general-shaped pupils in Southwell configuration [19]. Unfortunately, the DCT integration method [8,15,16] is not

* Corresponding author.
E-mail address: huanglei0114@gmail.com (L. Huang).

http://dx.doi.org/10.1016/j.optlaseng.2014.11.011
0143-8166/

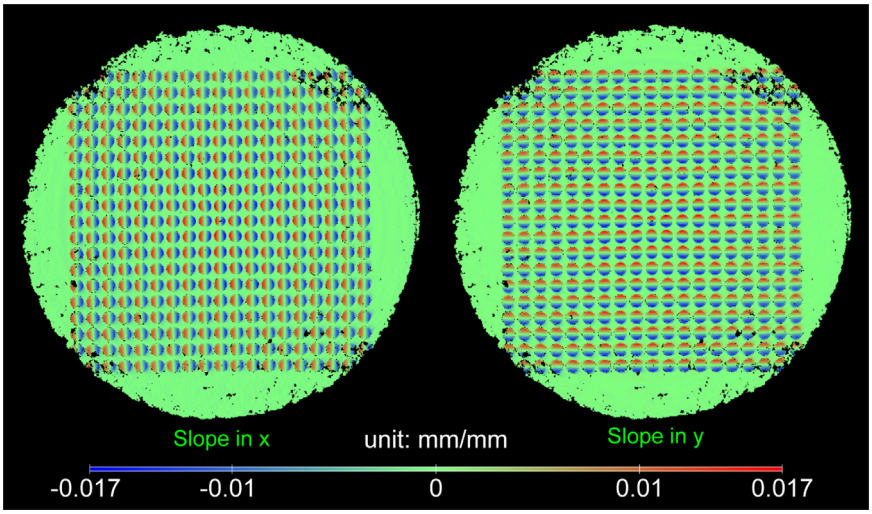

Fig. 1. Imperfections on samples or the shape of aperture may result in incomplete gradient data, which do not fulfill the whole sensing array and hence bring troubles to the implementation of DCT or FFT algorithm.

able to directly apply to incomplete gradient data for shape reconstruction [20].

In this work, a fast and accurate method for shape reconstruction from gradient in an arbitrarily-shaped aperture, or say incomplete gradient data, is addressed by employing iterative DCT in Southwell configuration. Of course, the available region of gradient data should be fully connected as a single piece, that is, there are no separate regions which have no connections with each other. The proposed method enjoys the high accuracy of the DCT integration, the high accuracy of extrapolation outside aperture with DCT-delivered derivatives through the Gerchberg iteration, and the high efficiency of the widely available FFT algorithm.

2. Method

In Southwell configuration, the shape is reconstructed where the slopes are measured as shown in Fig. 2, which is a beauty by itself.

Unlike the other configurations, there is no spatial shift between slope and shape grids, and as a result, the FFT-based slope-shape relation in Southwell configuration is really simple and elegant as [15,21].

$$z = \mathrm{FFT}^{-1}\{F_z\}, \tag{1}$$

where

$$F_z = \begin{cases} 0, & (u,v) = (0,0) \\ \frac{FFT\{s_x + i\cdot s_y\}}{i\cdot 2\pi(u+i\cdot v)}, & \text{elsewhere} \end{cases}, \tag{2}$$

s_x and s_y stand for the slope values in x and y direction, respectively. z denotes the estimated shape with the piston term ignored. The coordinates in spatial frequency domain (u, v) can be calculated through the size of matrix $(N_x \times N_y)$ and the x- and y-sampling intervals (h_x, h_y).

$$(u,v) = \left(\frac{n_x - \frac{N_x}{2}}{h_x N_x}, \frac{n_y - \frac{N_x}{2}}{h_y N_y}\right), \quad \begin{matrix} n_x = 0,1,\cdots,N_x - 1 \\ n_y = 0,1,\cdots,N_y - 1 \end{matrix}. \tag{3}$$

According to reference [16], the DCT integration can be implemented with the use of FFT algorithm by flipping slopes with positive or negative values, accordingly. Here, simpler expressions

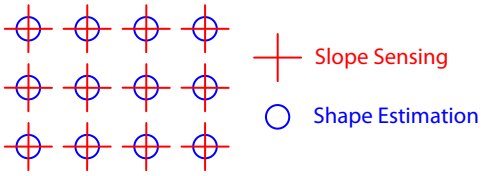

Fig. 2. In Southwell configuration, shape estimations happen at the same rectangular grid where the slopes are measured.

used in reference [15] are adopted as Eqs. (4)–(6).

$$\begin{cases} S_x := \begin{bmatrix} -s_x(-x,-y), & -s_x(-x,y) \\ s_x(x,-y), & s_x(x,y) \end{bmatrix} \\ S_y := \begin{bmatrix} -s_y(-x,-y), & s_y(-x,y) \\ -s_y(x,-y), & s_y(x,y) \end{bmatrix} \end{cases}, \tag{4}$$

$$F_Z = \begin{cases} 0, & (u,v) = (0,0) \\ \frac{FFT\{S_x + i\cdot S_y\}}{i\cdot 2\pi(u+i\cdot v)}, & \text{elsewhere} \end{cases}, \tag{5}$$

$$\begin{bmatrix} z(-x,-y), & z(x,-y) \\ z(-x,y), & z(x,y) \end{bmatrix} =: Z = \mathrm{FFT}^{-1}\{F_Z\}, \tag{6}$$

The resultant height distribution $z = z(x, y)$ is desired and can be retrieved by simply cropping it from matrix Z. By the way, the slopes estimated by transforms can be presented as

$$S_x = FFT^{-1}\{i\cdot 2\pi u\cdot F_Z\}, \tag{7}$$

$$S_y = FFT^{-1}\{i\cdot 2\pi v\cdot F_Z\}. \tag{8}$$

However, because of the blocking by the aperture or the "poor" quality of samples in practice, some of the sensor elements may not provide available slopes [e.g. Fig. 1]. The Gerchberg-type iteration is employed to extrapolate the missing data. The algorithm is illustrated in Fig. 3.

The implementation details can be described as the following steps.

Step 1: Flip the measured slopes by Eq. (4).
Step 2: Fill those unavailable slope values with 0 to get $S_{x,0}$ and $S_{y,0}$. Set the iterative number $n = 1$.
Step 3: Take the FFT according to Eq. (5) to get F_Z.
Step 4: Take inverse FFTs by Eqs. (6)–(8) to estimate the Z_n, $S_{x,n}$, and $S_{y,n}$. Update $n = n + 1$.

Step 5: Check the difference updated estimation and with the previous estimation. $\Delta S=(S_{x,n}, S_{y,n})-(S_{x,n-1}, S_{y,n-1})$ and $\Delta Z=Z_n-Z_{n-1}$. If both the standard deviations of slope differences and shape differences are smaller than the preset thresholds and the iteration number n is smaller the preset maximum number, overwrite the slope in the available region with the measured slopes and go back to Step 3. Otherwise, end the loop.

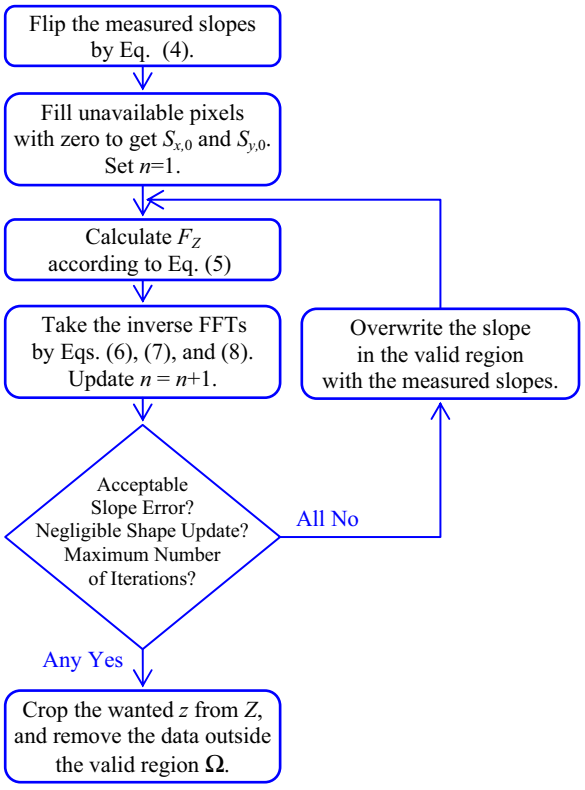

Fig. 3. Flow chart of the iterative DCT integration algorithm.

Step 6: Crop the wanted z from Z, and keep the data inside the valid region Ω only.

3. Simulation

Simulation is carried out to verify the proposed algorithm is effective and accurate. A surface expressed as Eq. (9) with arbitrarily-shaped aperture [see Fig. 4(a)]. 128 × 128 points are sampled in an area of 3 unit × 3 unit and here the unit could be any unit, such as mm, μm, or nm, which depends on the used measuring technique.

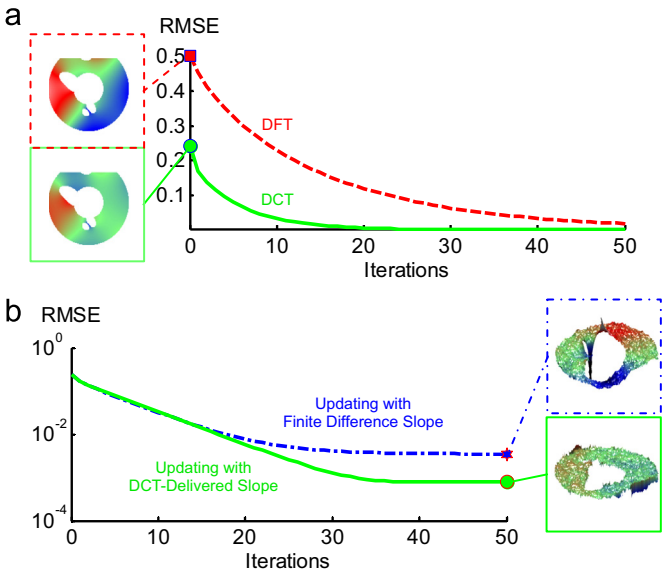

Fig. 5. The DCT method offers a better initial value than the DFT method does (a), and comparing to the slopes from finite difference, updates with DCT-delivered slopes leads to a better convergence (b).

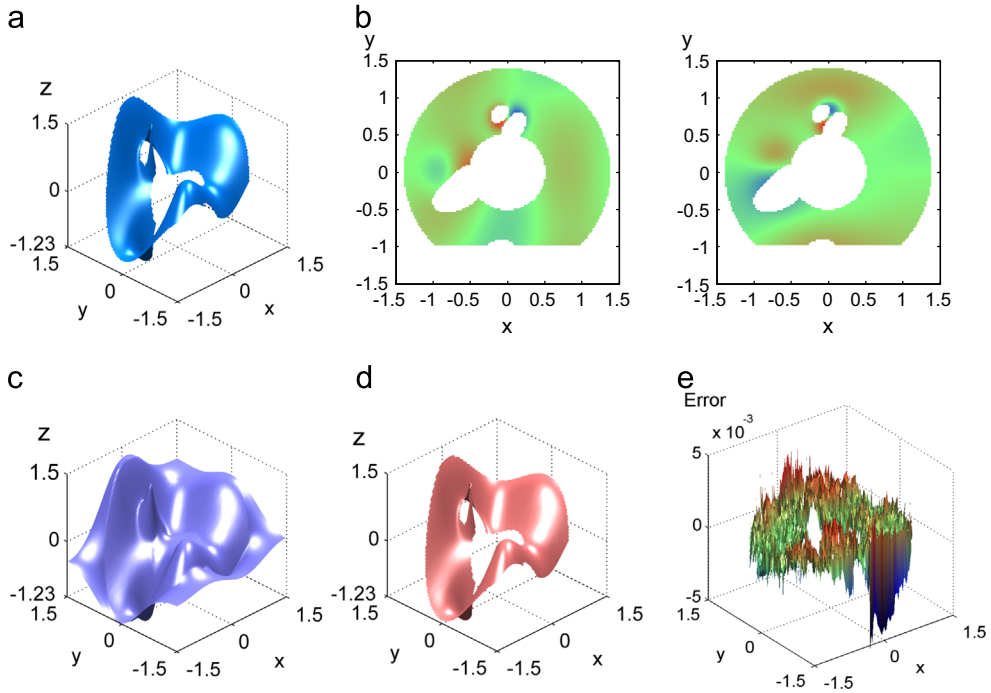

Fig. 4. Reconstruction error (e) indicates the proposed iterative DCT can successfully reconstruct a shape [true shape (a) and the reconstructed ones before (c) and after (d) cropping] from its slopes with additive noise and an arbitrarily-shaped aperture (b).

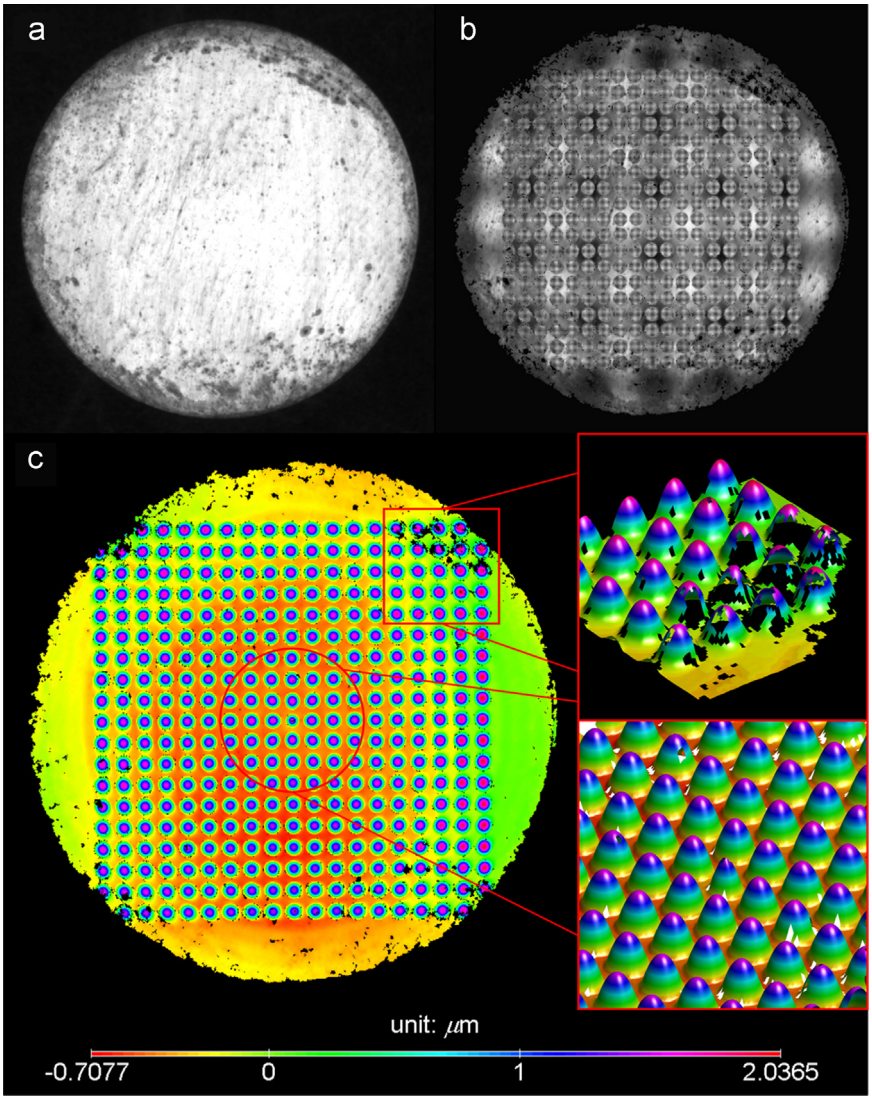

Fig. 6. The reconstructed shape (c) from incomplete gradient data [Fig. 1] in real measurement [see sample under (a) uniform illumination and (b) structured-light illumination] shows the feasibility of the proposed method.

The aperture is shaped with a circle boundary which is common in optics. Furthermore, the central region is blocked to simulate conditions in X-ray or astronomy applications. Other irregular edges are added to make aperture shape even more complex. Normally distributed random noise (σ=0.03 unit/unit) is added to the analytical slopes [see Fig. 4(b)].

As an iterative algorithm, the initial values and the variable updates are critical for its performance. The flipped slopes before zero padding can be considered as setting up an initial guess from DCT point of view, which is a better initial value comparing to the directly zero padding from DFT point of view as shown in Fig. 5(a). The Root Mean Square Error (RMSE) of initial estimation

$$z=\begin{cases} 0.2\left\{3(1-x)^2\exp\left[-x^2-(y+1)^2\right]-10\left(\frac{x}{5}-x^3-y^5\right)\exp\left(-x^2-y^2\right)-\frac{1}{3}\exp\left[-(x+1)^2-y^2\right]\right\} \\ +\exp\left\{-(6x)^2-[6(y-0.7)]^2\right\}-\exp\left\{-[3(x+0.7)]^2-(3y)^2\right\}+0.3x & ,\ (x,y)\in\Omega \\ \text{no available values} & ,\ (x,y)\notin\Omega \end{cases}. \tag{9}$$

After about 40 iterations, the surface shape in and out of the aperture is estimated as Fig. 4(c), whereas only the shape in slope-available region Ω is of interest which can be cropped out [Fig. 4(d)]. The error distribution is shown in Fig. 4(e) with a root mean square of 1.1×10^{-3} unit and a peak-to-valley value of 9.9×10^{-3} unit when the slope noise σ=0.03 unit/unit and the sampling interval is 3/128 units.

from DCT is smaller than that from DFT, which will lead a faster convergence.

On the other hand, for the updates of slopes, the transform-delivered slopes from Eqs. (7) and (8) are updated instead of using the finite differences of the estimated shape traditionally [6,7]. The slopes from spatial frequency domain are inherent with the FFT algorithm, which induces the iterations to converge at more

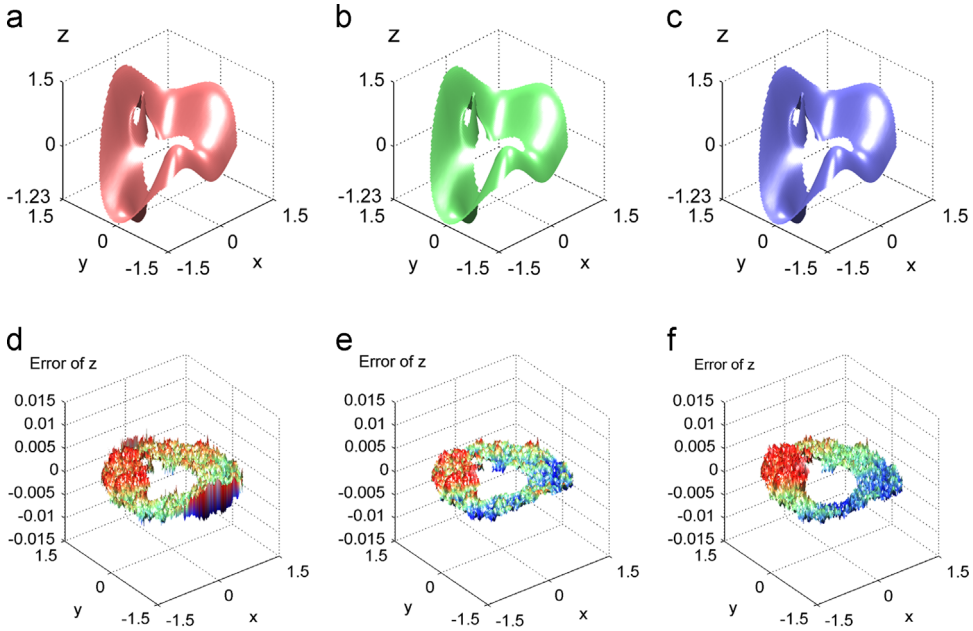

Fig. 7. The comparison of the integration results in an arbitrarily-shaped aperture. Result (a) and its error (d) of the proposed transform-based integration method. Result (b) and its error (e) of the finite-difference-based least-squares integration method. Result (c) and its error (f) of the Radial Basis Functions based integration method.

accurate results as shown in Fig. 5(b). Moreover, it is not costly to compute the slopes from Eqs. (7) and (8), because anyway we have to calculate F_Z for Z, and then S_x and S_y are obtained almost for free.

4. Experiment

The proposed method is also applied to a set of real data from measurements to validate its actual use. The input slopes shown in Fig. 1 are measured with phase measuring deflectometry [10]. The sample under test is an optical component with many reflective micro-mirrors positioned in an array. The height of each micro-mirror is about 2 μm. The area of the specular surface is about 14 mm × 14 mm. Two images (640 × 640 pixels) show the sample with "freckles" under uniform lighting [Fig. 6(a)] or structured-light illumination [Fig. 6(b)], respectively. With uniform illumination the micro-mirrors are very difficult to observe, but it becomes obvious when illuminated with structured light. Although many pixels fail to offer reliable slope values, the proposed method is still able to reconstruct the shape from incomplete gradient data straightforward as shown in Fig. 6(c) in details.

In fact, the change of the aperture shape or the sample may change the available region for slope measurement, and it is very common to happen in practice. Considering this, it is not hard to find an obvious advantage of the proposed method comparing to Poyneer's boundary methods in which matrices for least squares estimation have to be adjusted accordingly every time when the available region changes. Here we do not need to consider this problem if the proposed method is used, because the procedure of the proposed method is the same for any shape of available regions.

5. Discussion

The proposed iterative method is developed in the path of the transform-based integration methods. Its superiority to the other transform-based integration methods is demonstrated in Fig. 5. Here we discuss its performance comparing to other existing methods which do not use transforms to complete the integration. As shown in the Fig. 7, the finite-difference-based least-squares integration method [22,23] and radial basis functions based integration method [24,25] are applied to make the comparison.

Very similar reconstruction results, referring to Fig. 7(a–c), can be obtained by using these three methods. The these state-of-art integration methods are able to deal with the problem of shape reconstruction results indicate that from gradient data measured by slope sensors even with an arbitrarily-shaped aperture. At the boundary of the aperture, the transform-based methods usually suffer the Gibbs ringing effect as shown in Fig. 7(d) and thus the error is commonly larger than those of the non-transform based approaches Fig. 7(e–f) around the inner and outer boundary regions. More detailed comparisons in different circumstances can be found in our recent comparison work [20].

6. Conclusions

In this work, an iterative DCT integration method is proposed to reconstruct the surface or wavefront shape from gradient data in an arbitrarily-shaped aperture. The feasibility of the method is verified in both simulation and experiment. The DCT integration method is pushed to be applicable for handling incomplete gradient data with missing slopes either inside a "hole" or outside the aperture. This method is designed for the situation that the gradient data are in a single piece of available region. If there are disconnected sub-regions, the proposed method will have no capability of providing the relative depth/height information between sub-regions.

Acknowledgment

Authors would like to thank Prof. Peng Su and Dr. Tianquan Su in University of Arizona for the helpful discussions.

References

[1] Noll RJ. Phase estimates from slope-type wave-front sensors. J Opt Soc Am 1978;68:139–40.

[2] Hudgin RH. Wave-front reconstruction for compensated imaging. J Opt Soc Am 1977;67:375–8.
[3] Fried DL. Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements. J Opt Soc Am 1977;67:370–5.
[4] Southwell WH. Wave-front estimation from wave-front slope measurements. J Opt Soc Am 1980;70:998–1006.
[5] Freischlad KR, Koliopoulos CL. Modal estimation of a wave front from difference measurements using the discrete fourier transform. J Opt Soc Am A 1986;3:1852–61.
[6] Roddier F, Roddier C. Wavefront reconstruction using iterative fourier transforms. Appl Opt 1991;30:1325–7.
[7] Guo H, Wang Z. Wavefront reconstruction using iterative discrete fourier transforms with fried geometry. Optik Int J Light and Electron Opt 2006;117:77–81.
[8] Talmi A, Ribak EN. Wavefront reconstruction from its gradients. JOSA A 2006;23:288–97.
[9] Platt BC, Shack R. History and principles of Shack-Hartmann wavefront sensing. J Refract Surg 2001;17:S573–7.
[10] Knauer MC, Kaminski J, Häusler G. Phase measuring deflectometry: a new approach to measure specular free-form surfaces. In: Osten Wolfgang, Takeda Mitsuo, editors. Optical Metrology in Production Engineering. Strasbourg, France: SPIE; 2004. p. 366–76.
[11] Su P, Parks RE, Wang L, Angel RP, Burge JH. Software configurable optical test system: a computerized reverse Hartmann test. Appl Opt 2010;49:4404–12.
[12] Huang L, Seng C, Krishna Asundi Ng, A. Fast full-field out-of-plane deformation measurement using fringe reflectometry. Opt Laser Eng 2012;50:529–33.
[13] Su T, Wang S, Parks RE, Su P, Burge JH. Measuring rough optical surfaces using scanning long-wave optical test system. 1. principle and implementation. Appl Opt 2013;52:7117–26.
[14] Ghiglia DC, Pritt MD. Two-dimensional phase unwrapping: theory, Algorithms, and Software. New York: John Wiley and Sons; 1998.
[15] Bon P, Monneret S, Wattellier B. Noniterative boundary-artifact-free wavefront reconstruction from its derivatives. Appl Opt 2012;51:5698–704.
[16] Zuo C, Chen Q, Asundi A. Boundary-artifact-free phase retrieval with the transport of intensity equation: fast solution with use of discrete cosine transform. Opt Express 2014;22:9220–44.
[17] Gerchberg RW. Super-resolution through error energy reduction. Opt Acta: Int J Opt 1974;21:709–20.
[18] Poyneer LA, Gavel DT, Brase JM. Fast wave-front reconstruction in large adaptive optics systems with use of the fourier transform. J Opt Soc Am A 2002;19:2100–11.
[19] Zou W, Rolland JP. Iterative zonal wave-front estimation algorithm for optical testing with general-shaped pupils. J Opt Soc Am A 2005;22:938–51.
[20] Huang L, Idir M, Zuo C, Kaznatcheev K, Zhou L, Asundi A. Comparison of two-dimensional integration methods for shape reconstruction from gradient data. Opt Laser Eng 2015;64:1–11.
[21] Kottler C, David C, Pfeiffer F, Bunk O. A two-directional approach for grating based differential phase contrast imaging using hard x-rays. Opt Express 2007;15:1175–81.
[22] Huang L, Asundi A. Improvement of least-squares integration method with iterative compensations in fringe reflectometry. Appl Opt 2012;51:7459–65.
[23] Li G, Li Y, Liu K, Ma X, Wang H. Improving wavefront reconstruction accuracy by using integration equations with higher-order truncation errors in the southwell geometry. J Opt Soc Am A 2013;30:1448–59.
[24] Ettl S, Kaminski J, Knauer MC, Häusler G. Shape reconstruction from gradient data. Appl Opt 2008;47:2091–7.
[25] Huang L, Asundi AK. Framework for gradient integration by combining radial basis functions method and least-squares method. Appl Opt 2013;52:6016–21.