Optics and Lasers in Engineering 64 (2015) 1–11

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

Comparison of two-dimensional integration methods for shape reconstruction from gradient data

Lei Huang [a,*], Mourad Idir [a], Chao Zuo [b], Konstantine Kaznatcheev [a], 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 30 May 2014
Accepted 4 July 2014
Available online 31 July 2014

Keywords:
Integration
Shape reconstruction
Wavefront reconstruction
Shape from gradient
Least squares estimation

ABSTRACT

As a requisite and key step in some gradient-based measurement techniques, the reconstruction of the shape, more generally the scalar potential, from the measured gradient data has been studied for many years. In this work, three types of two-dimensional integration methods are compared under various conditions. The merits and drawbacks of each integration method are consequently revealed to provide suggestions in selection of a proper integration method for a particular application.

1. Introduction

In metrology, some physical quantities from measurement may not be our desired objective directly, but they may have certain relationship with our objective. The measurements are therefore still very useful and can be employed to get our desired quantities. Many optical metrology techniques belong to this indirect measurement. For instance, wavefront measurement techniques (Hartmann-based wavefront sensing [1–3], lateral shearing interferometry [4–6], *etc.*) reconstruct the wavefront from the slopes measured by optical sensors. Moreover, the technique of shape from shading [7–9] estimates the surface profile by integrating the calculated gradient data. Similarly, three-dimensional shape measurement for specular surfaces, *e.g.* phase measuring deflectometry [10–18], integrates gradient data from metrology to get the surface shape as shown in Fig. 1. All these techniques above only measure the derivatives of the wanted quantity. In order to achieve our final goal, a two-dimensional (2D) integration procedure is necessary to reconstruct the shape from the measured derivatives.

Due to its wide application, 2D integration methods are investigated by many researchers and there are lots of articles in literatures [7,19–29]. The 2D integration problem can be considered as solving a Poisson equation with Neumann boundary conditions [30]. Research on 2D integration methods can be found since 1970s for wavefront reconstruction [19,20,22,30]. Generally, finite difference approaches were employed in those methods to connect the measured slope and desired shape, and least squares estimations are made for shape reconstruction. Fourier transform has been introduced into 2D integration in 1980s [7,31]. By applying the properties of Fourier transform, the integration operation can be implemented easily and quickly due to the well-known Fast Fourier Transform (FFT) algorithm. In 2004, Li *et al.* [23] compared the finite difference based least squares integration method and Fourier transform integration method with showing the finite difference based least squares integration method has higher accuracy at that time. By considering the boundary conditions, Talmi and Ribak [24] pointed out the best solution of gradient integration could be expressed in a Fourier cosine series, not the periodic Fourier series. The cosine transform integration method is therefore proposed with providing the integration result at half-integer positions in 2006. Coming to 2008, Ettl *et al.* [25] introduced an integration method by employing the radial basis functions which is flexible and robust. In 2012, Bon *et al.* [26] proposed a boundary-artifact-free Fourier integration method by simply padding slope matrices with accordingly flipped and positive or negative slope values. By noticing the accuracy of the traditional finite difference based least squares integration method is limited by its biquadratic shape assumption, Huang and Asundi [27,28] proposed an iterative compensation approach to obtain more accurate integration results. Recently, Li *et al.* [29] improved the finite difference based least squares integration method by applying higher-order numerical differentiation formats.

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

http://dx.doi.org/10.1016/j.optlaseng.2014.07.002
0143-8166/

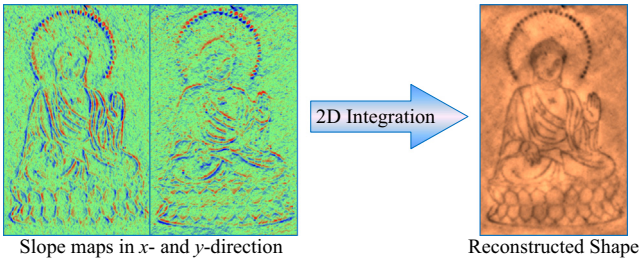

Fig. 1. The 2D integration is a vital process for the reconstruction of shape from the measured gradient data.

In this work, three families of integration methods are selected into our comparison, since they have been widely used in numerous applications. Several 2D integration methods are chosen as the representative of their corresponding families to make a comparison in both reconstruction accuracy and processing speed. These integration methods are invented in different application fields, and developed along their own paths. It is very interesting to know their merits and drawbacks after the improvement in recent years. Their "abilities" and "tempers" are revealed in order to help the selection of a proper 2D integration method for a specific application.

2. The 2D integration methods in comparison

There are three families of 2D integration methods to be compared in this work. The first family is the Finite-difference-based Least-squares Integration (FLI) methods. The second big family of integration methods is the Transform-based Integration (TI) methods. The third one is the Radial Basis Function based Integration (RBFI) method. Because we are interested in reconstruction of arbitrary shapes, the well-known modal wavefront reconstruction methods with the Zernike or other polynomials [32] are not compared here, which are more suitable for symmetrical optical components. During the writing up of this work, some other integration methods are noticed, such as the spline-based methods [33], showing convincing results and they may be strong potential competitors as well.

2.1. Finite-difference-based least-squares integration methods with Southwell configuration

The Traditional Finite-difference-based Least-squares Integration (TFLI) method in Southwell configuration [19] is well known and widely used not only due to its consistency between shape and slope locations (where you measure the slope, where you get the shape), but more significantly, because of its biquadratic spline shape in nature. The shape model in the Southwell configuration is essentially a biquadratic spline with an algorithm error of $O(h^3)$ where h stands for the interval of the sampling grids, whereas a bilinear curve in other similar configurations [20,22]. Due to its simple implementation and reliable performance, the TFLI method has been widely applied for shape reconstruction. The relations of the slope and shape in a M-by-N matrix are set locally in Eq. (1),

$$\begin{cases} \dfrac{z_{m,n+1}-z_{m,n}}{x_{m,n+1}-x_{m,n}} \triangleq \dfrac{s^x_{m,n+1}+s^x_{m,n}}{2}, & m=1,2,\ldots,M, \\ & n=1,2,\ldots,N-1 \\ \dfrac{z_{m,n+1}-z_{m,n}}{y_{m,n+1}-y_{m,n}} \triangleq \dfrac{s^y_{m+1,n}+s^y_{m,n}}{2}, & m=1,2,\ldots,M-1, \\ & n=1,2,\ldots,N \end{cases} \quad (1)$$

where x, y, z are the world coordinates, "$\triangleq$" stands for "equal in the least squares sense", s^x and s^y denote the measured slope values in x- and y-directions, respectively. The subscripts m and n are the matrix indices. The estimation can be handled globally as presented in Eq. (2).

$$\begin{bmatrix} z_{1,1} \\ z_{2,1} \\ \vdots \\ z_{M,N} \end{bmatrix} = (\mathbf{D}^{\mathbf{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathbf{T}}\mathbf{G}, \quad (2)$$

where $\mathbf{D}^{\mathbf{T}}$ stands for the transpose of $\mathbf{D}$. The matrix $\mathbf{D}$ (usually in sparse matrix format) and the vector $\mathbf{G}$ are

$$\mathbf{D} := \begin{bmatrix} \mathbf{D^x} \\ \mathbf{D^y} \end{bmatrix} = \begin{bmatrix} -1 & 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & -1 & 0 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix}, \quad (3)$$

$$\mathbf{G} := \begin{bmatrix} \mathbf{G^x} \\ \mathbf{G^y} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (s^x_{1,2}+s^x_{1,1})(x_{1,2}-x_{1,1}) \\ (s^x_{1,3}+s^x_{1,2})(x_{1,3}-x_{1,2}) \\ \vdots \\ (s^x_{M,N}+s^x_{M,N-1})(x_{M,N}-x_{M,N-1}) \\ \cdots\cdots\cdots\cdots \\ (s^y_{2,1}+s^y_{1,1})(y_{2,1}-y_{1,1}) \\ (s^y_{3,1}+s^y_{2,1})(y_{3,1}-y_{2,1}) \\ \vdots \\ (s^y_{M,N}+s^y_{M-1,N})(y_{M,N}-y_{M-1,N}) \end{bmatrix}. \quad (4)$$

With the development of metrology, an algorithm error of $O(h^3)$ at the integration stage may no longer be acceptable. Further investigation makes improvement on this path for a better accuracy during recent years. Iterations can be carried out to enhance the accuracy with Iterative Finite-difference-based Least-squares Integration (IFLI) method [27], which integrates the gradient residuals to implement iterative compensation onto the final result. Moreover, instead of using iterations, Li *et al.* [29] propose a more direct approach by considering higher order terms in Taylor expansion into the least squares estimation. Here we call it Higher-order Finite-difference-based Least-squares Integration (HFLI) method in this work. The expression of $\mathbf{G}$ is selected as Eq. (5) to maintain the same sparse matrix $\mathbf{D}$ as Eq. (3) which is usually the major concern of memory cost.

$$\mathbf{G} := \begin{bmatrix} \mathbf{G^x} \\ \mathbf{G^y} \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 12(s^x_{1,2}+s^x_{1,1})(x_{1,2}-x_{1,1}) \\ (-s^x_{1,4}+13s^x_{1,3}+13s^x_{1,2}-s^x_{1,1})(x_{1,3}-x_{1,2}) \\ (-s^x_{1,5}+13s^x_{1,4}+13s^x_{1,3}-s^x_{1,2})(x_{1,4}-x_{1,3}) \\ \vdots \\ 12(s^x_{M,N}+s^x_{M,N-1})(x_{M,N}-x_{M,N-1}) \\ \cdots\cdots\cdots\cdots \\ 12(s^y_{2,1}+s^y_{1,1})(y_{2,1}-y_{1,1}) \\ (-s^y_{4,1}+13s^y_{3,1}+13s^y_{2,1}-s^y_{1,1})(y_{3,1}-y_{2,1}) \\ (-s^y_{5,1}+13s^y_{4,1}+13s^y_{3,1}-s^y_{2,1})(y_{4,1}-y_{3,1}) \\ \vdots \\ 12(s^y_{M,N}+s^y_{M-1,N})(y_{M,N}-y_{M-1,N}) \end{bmatrix} \quad (5)$$

To disclose the improvement made in FLI family, a simple comparison between methods of TFLI, IFLI, and HFLI is carried out by integrating a gradient dataset with size of 256 (pixel) × 256 (pixel) × 2 (direction).

$$z(x,y) = 0.2 \times \Big\{ 3(1-x)^2\exp[-x^2-(y+1)^2]. \\ -10\Big(\frac{x}{5}-x^3-y^5\Big)\exp(-x^2-y^2)$$

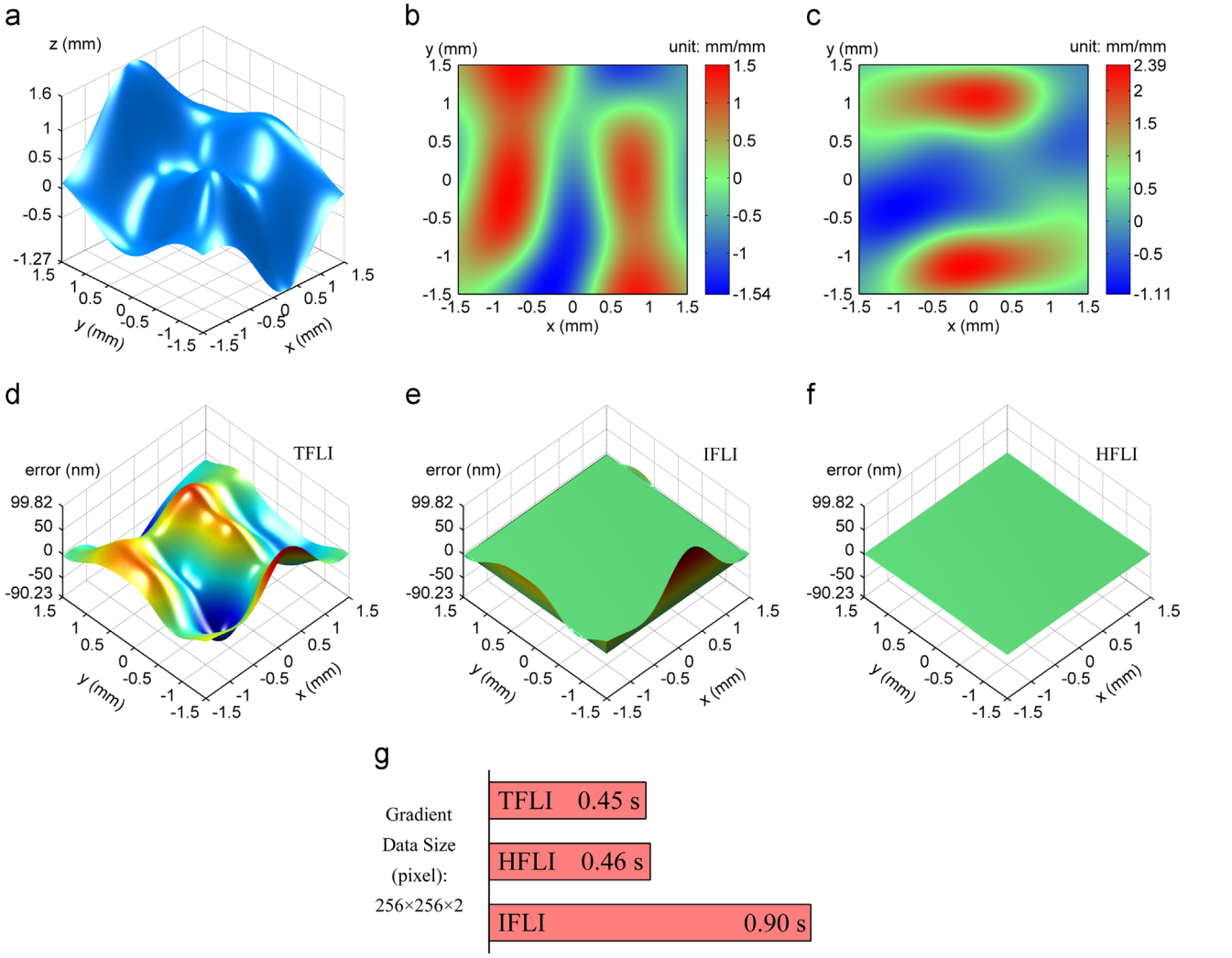

Fig. 2. The HFLI method outperforms the others in the FLI family. (a) Ground truth of shape, slope in x-direction (b) and y-direction (c), reconstruction error with the TFLI method (e), the IFLI method (d), and the HFLI method (f), and (g) comparison of running time.

$$-\frac{1}{3}\exp[-(x+1)^2-y^2]\Big\}. \tag{6}$$

Fig. 2 shows a ground truth of shape to reconstruct (a) (see Eq. (6)) and its slope maps in x-direction (b) and y-direction (c), while the integration errors indicate the obvious accuracy improvement by the IFLI method (e) and the HFLI method (f) compared with the TFLI method (d). Here, the number of iterations in IFLI is set as $k=1$ only and the form for HFLI is selected with a truncation error of $O(h^5)$. For speed comparison, HFLI needs slightly longer time than TFLI as shown in Fig. 2(g), and it takes k times longer than TFLI for IFLI to get the final result, where k is the iteration number.

Considering accuracy and speed as well as the edge effect by IFLI (the same values as TFLI at the edge), the HFLI method should be the best one in the family of FLI methods with Southwell configuration. The HFLI method is therefore selected as the representative of FLI methods with Southwell configuration to make further comparisons with other integration approaches.

2.2. Transform-based integration methods

Another big family of integration approaches is composed of transform-based methods, which employ the property of the derivative in the spatial frequency domain to achieve the integration goal. Generally, these transform-based methods link all the slopes and shape globally and optimize them globally as well. Fourier Transform, the most famous transform in the scientific world, is well studied in the application of numerical integration [7,21,31,34]. Eqs. (7) and (8) are two different ways to reconstruct the shape in Fourier Transform Integration (FTI) method [21,35].

$$z(x,y)=F^{-1}\left\{\frac{i\cdot 2\pi[f_x F\{s^x(x,y)\}+f_y F\{s^y(x,y)\}]}{-4\pi^2(f_x^2+f_y^2)}\right\}. \tag{7}$$

$$z(x,y)=F^{-1}\left\{\frac{F\{s^x(x,y)+i\cdot s^y(x,y)\}}{i\cdot 2\pi(f_x+i\cdot f_y)}\right\}. \tag{8}$$

where F and F^{-1} stand for forward and inverse Fourier transform.

Talmi and Ribak [24] point out that Fourier transform may provide a suboptimal solution only due to the implicit periodic boundary conditions, whereas the optimal solution should be provided by applying the homogeneous Neumann boundary conditions to the residual field with use of the discrete cosine transform instead. The widely available FFT algorithm can be used to calculate the cosine transform, but the shape mesh is shifted to half-integer locations in their work. Bon *et al.* [26] present a simple way to utilize FFT to complete the reconstruction from slopes with no boundary artifact and keeping the shape mesh the same as the slope mesh (the Southwell configuration). By easily padding the slope matrices with

flipped and positive or negative slope values accordingly as Eqs. (9) and (10), Cosine Transform Integration (CTI) on the same data mesh is actually realized [36–38].

$$S^x(x,y) := \begin{bmatrix} -s^x(-x,-y), & -s^x(-x,y) \\ s^x(x,-y), & s^x(x,y) \end{bmatrix}. \tag{9}$$

$$S^y(x,y) := \begin{bmatrix} -s^y(-x,-y), & s^y(-x,y) \\ -s^y(x,-y), & s^y(x,y) \end{bmatrix}. \tag{10}$$

The same as FTI method, the CTI method could be implemented by applying Eq. (11) (CTI-R) or Eq. (12) (CTI-C).

$$Z(x,y) = F^{-1}\left\{\frac{i\cdot 2\pi[f_x F\{S^x(x,y)\}+f_y F\{S^y(x,y)\}]}{-4\pi^2(f_x^2+f_y^2)}\right\}. \tag{11}$$

$$Z(x,y) = F^{-1}\left\{\frac{F\{S^x(x,y)+i\cdot S^y(x,y)\}}{i\cdot 2\pi(f_x+i\cdot f_y)}\right\}. \tag{12}$$

The relation between padded $Z(x, y)$ and originally desired $z(x, y)$ are

$$Z(x,y) := \begin{bmatrix} z(-x,-y), & z(-x,y) \\ z(x,-y), & z(x,y) \end{bmatrix}. \tag{13}$$

Finally, $z(x, y)$ is obtained by cropping data from the inverse transform result $Z(x, y)$.

$$z(x,y) = Crop(Z(x,y)). \tag{14}$$

Fig. 3 demonstrates an example of non-iterative FTI method fails to correctly reconstruct the shape when the background just simply changed from flat (Eq. (15)) to a tilt surface (Eq. (16)), whereas the CTI method succeeds in both cases.

$$z(x,y) = \exp\{-(6x)^2-[6(y-0.7)]^2\}-\exp\{-(3x)^2-[3(y+0.7)]^2\}. \tag{15}$$

$$z(x,y) = 0.3x+\exp\{-(6x)^2-[6(y-0.7)]^2\} -\exp\{-(3x)^2-[3(y+0.7)]^2\}. \tag{16}$$

In fact, the Gershberg-type iteration [39] on extended slope matrices is really a mandatory "algorithm mate" for the FTI method to get much more accurate results (see Fig. 3(c)) in general cases. As shown in Fig. 3(d), the CTI method is faster than the iterative FTI method when the gradient data is complete, *i.e.* no missing gradient data in slope matrices.

The CTI-C algorithm is slightly faster than the CTI-R algorithm (see Fig. 3(d)) with the same accuracy of CTI-R algorithm, the CTI-C algorithm is therefore chosen as the representative of CTI method as well as transform-based global integration methods in later comparisons.

2.3. Integration with radial basis functions

The RBFI method is proposed by Ettl *et al.* [25] by applying Radial Basis Functions (RBFs) into the integration job in phase measuring deflectometry [10–15]. An interpolation function, especially tailored for gradient data, is expressed with a weighted combination of analytical derivatives of the selected RBFs as Eq. (17).

$$z(x,y) = \sum_i \alpha_i \frac{\partial\Phi(x-x_i,y-y_i)}{\partial x}+\beta_i\frac{\partial\Phi(x-x_i,y-y_i)}{\partial y}. \tag{17}$$

The same as Ref. [25], a Wendland's function is selected as Eq. (18).

$$\Phi(x,y) = \begin{cases} \frac{1}{3}\left(1-\frac{\sqrt{x^2+y^2}}{\rho}\right)^6\left(35\frac{x^2+y^2}{\rho^2}+18\frac{\sqrt{x^2+y^2}}{\rho}+3\right), & \sqrt{x^2+y^2}\le\rho, \quad \rho>0 \\ 0 & \text{otherwise} \end{cases}, \tag{18}$$

where ρ is a scaling factor which is able to adjust the region in effect.

By matching the analytical derivatives of $z(x, y)$ with the measured slope values $s^x(x, y)$ and $s^y(x, y)$ as Eq. (19), the shape estimation can be completed through the determination of α_i and β_i in a least squares sense.

$$\begin{bmatrix} \frac{\partial z}{\partial x}(x_j,y_j) \\ \frac{\partial z}{\partial y}(x_j,y_j) \end{bmatrix} = \begin{bmatrix} \frac{\partial^2\Phi}{\partial x^2}(x_j-x_i,y_j-y_i) & \frac{\partial^2\Phi}{\partial x\partial y}(x_j-x_i,y_j-y_i) \\ \frac{\partial^2\Phi}{\partial x\partial y}(x_j-x_i,y_j-y_i) & \frac{\partial^2\Phi}{\partial y^2}(x_j-x_i,y_j-y_i) \end{bmatrix}\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} \triangleq \begin{bmatrix} s^x(x_j,y_j) \\ s^y(x_j,y_j) \end{bmatrix} \tag{19}$$

Due to its relatively large memory consumption (matrices are not sparse), the size of gradient data is usually set as about 40 (pixel) × 40 (pixel) × 2 (direction) with a common computer by hand. Therefore, we classify this method as a regional approach. A stitching process of the divided subsets is needed to reconstruct the whole shape of a large dataset.

3. Comparison

Although the integration process with the HFLI, CTI, and RBFI methods is essentially shape estimation with optimization based on the measured gradient data, the strategies these three integration methods apply are different as shown in Fig. 4.

The HFLI method with Southwell configuration uses the relations of slopes and dimensions locally by finite difference to link the whole map together and make estimation with an optimization globally. The CTI method employs the transforms and their derivatives properties to create global relations and make global optimization as well. Due to the high memory requirement, the RBFI method applies RBFs to set a slope-shape relation regionally (based on radial distances) and makes regional optimization (based on the subset size), and then globally stitches the regional results with overlaps.

3.1. Accuracy under different grids and conditions

First of all, the reconstruction accuracy is the primary concern for an integration method. The accuracy of these integration methods is compared under several data grids and conditions (*e.g.* noise, missing data), which are actually encountered in real applications. Comparisons are carried out to reveal how good or how bad they can perform, or even whether they are applicable, when data are in rectangular, quadrilateral, or triangular grids. As the most common data grids in optical metrology, gradient data in rectangular grids under different conditions are studied in detail.

3.1.1. Rectangular grids

It commonly happens that measured slope data are located in a rectangular mesh from optical metrology, due to its rectangular positioned optical sensors (*e.g.* CCD sensors), or pinhole arrays (*e.g.* Hartmann wavefront sensor), or gratings (*e.g.* grating shearing interferometry), or scanning steps (*e.g.* point-wise deflectometry with scanning). Detailed investigation is carried out specifically for such grids under the conditions of higher order components, rotation, noise, and incompletion.

3.1.1.1. Higher order components. When reconstructing a discrete representation of shape which has the most-similar gradient values with the measured ones, there is an implied assumption that all of the ignored higher order terms (no matter in what kind of forms, *e.g.* polynomials, Fourier cosine series, *etc.*) are really negligible.

A free-form surface $[z_1(x, y)]$ with a sharp peak and valley $[z_2(x, y)]$ expressed in Eq. (20) is sampled with a low resolution [32 (pixel) × 32 (pixel) × 2 (direction)]as show in Fig. 5(a).

$$z(x,y) = z_1(x,y)+z_2(x,y), \tag{20}$$

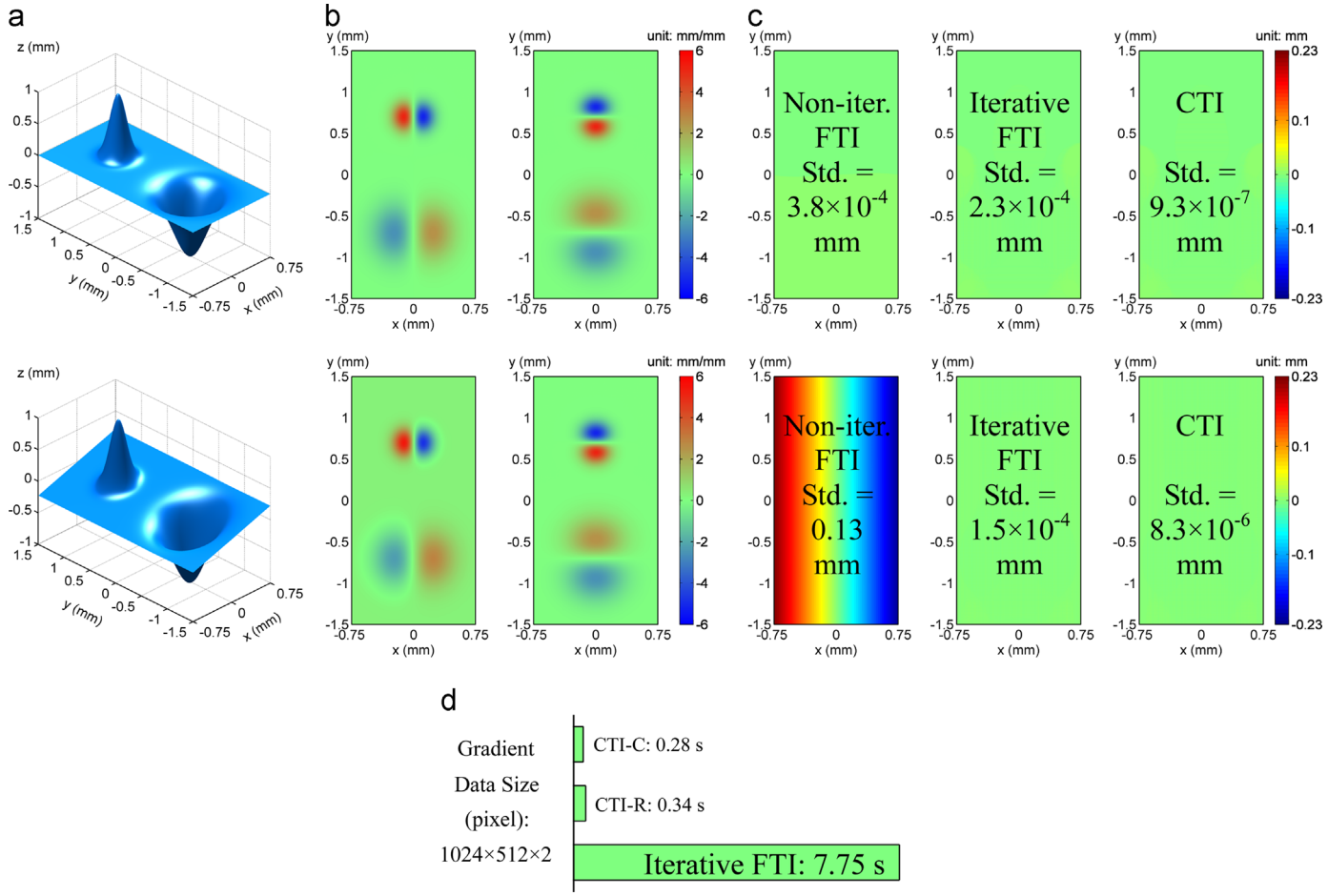

Fig. 3. The CTI method outperforms the FTI method. The ground truth of the shapes (a) and (b) their corresponding slope maps (b), (c) reconstruction errors from methods of non-iterative FTI, iterative FTI, and CT, (d) the CTI-C algorithm is faster than the CTI-R algorithm and the iterative FTI method.

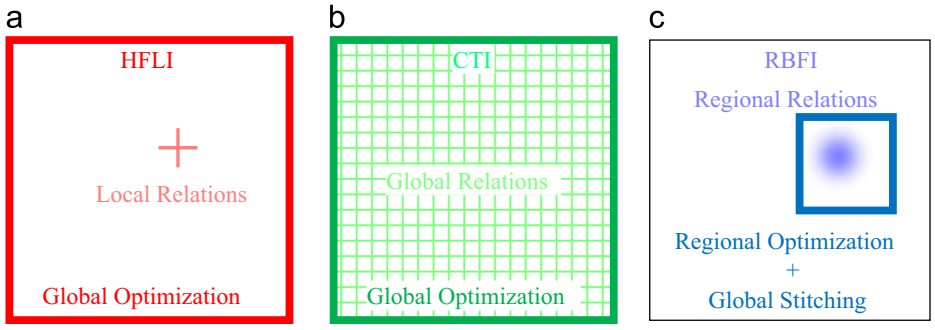

Fig. 4. Reconstruction strategies in methods of HFLI (a), CTI (b), and RBFI (c) are different.

where

$$z_1(x,y) = 0.2 \times \left\{ 3(1-x)^2 \exp[-x^2-(y+1)^2] - 10\left(\frac{x}{5} - x^3 - y^5\right)\exp(-x^2-y^2) - \frac{1}{3}\exp[-(x+1)^2-y^2] \right\}, \quad (21)$$

$$z_2(x,y) = 0.3x + \exp\{-(6x)^2 - [6(y-0.7)]^2\} - \exp\{-(3x)^2 - [3(y+0.7)]^2\}. \quad (22)$$

The influence of higher order terms is more obvious in low sampling ratio for comparison purpose. The slope maps in Fig. 5 (b) shows the sudden changes on slope around peak and valley regions.

The shape errors after HFLI method (Fig. 5(c)) indicate that although it has an improvement in algorithm error from TFLI method by considering higher order terms in local slope-height relations, it still has obvious local shape error when handling really serve changing slopes. Different from HFLI method, CTI method generates slope-height relations globally and the higher order terms are actually higher frequency terms, therefore the shape errors spread into the whole map (Fig. 5(d)). The higher order components affect the result from RBFI method less and the influence appears regionally (Fig. 5(e)).

3.1.1.2. Rotational field. Theoretically, the gradient of a scalar potential is a conservative field, or say a curl-free field. However,

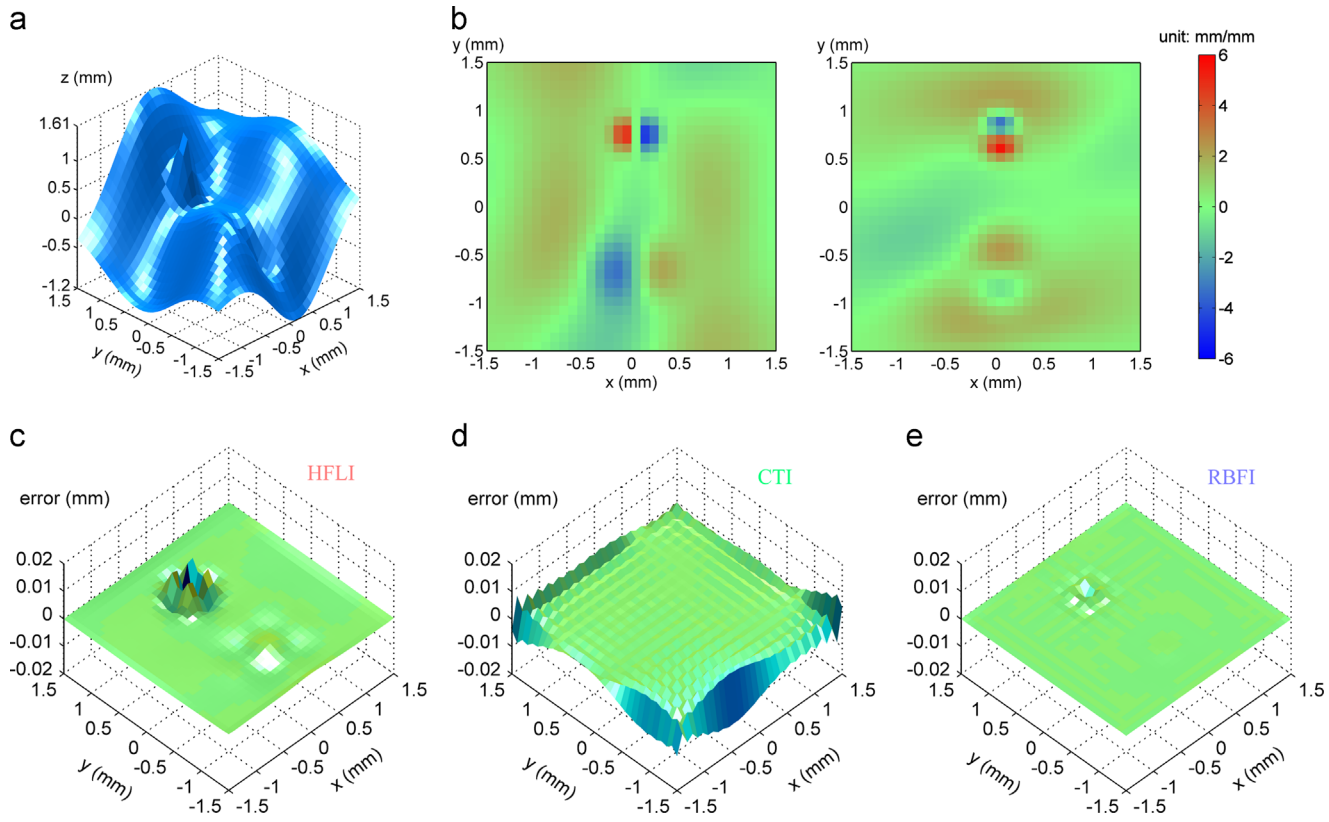

Fig. 5. The higher order components on shape with low resolution (a) and its slopes (b) affect results from the HFLI method (c), the CTI method (d), and the RBFI method (e).

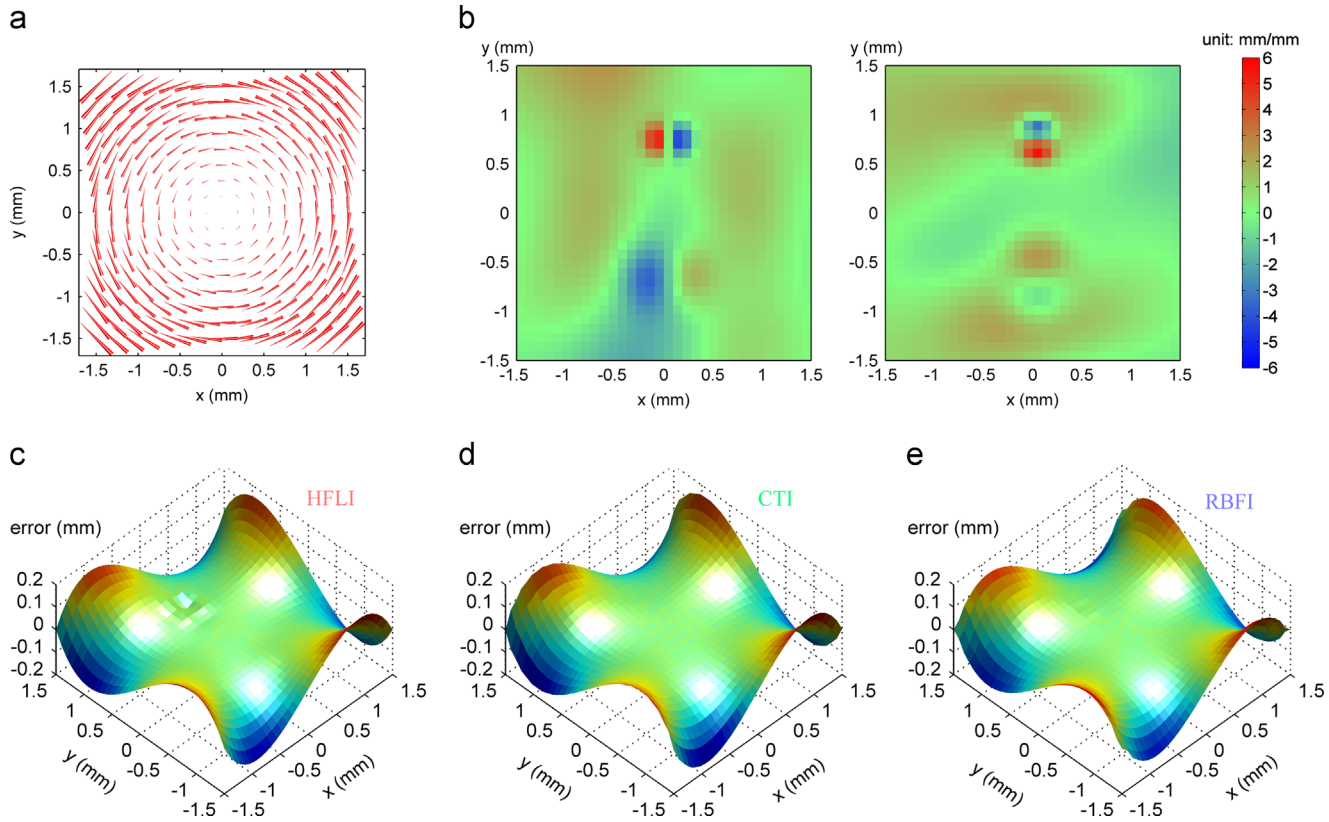

Fig. 6. Errors occur with the HFLI method (c), the CTI method (d), and RBFI method when a rotational field (a) is added on a gradient data (b).

gradient data from a real measurement may include a rotational vector field (*e.g.* a vortex) due to inaccurate system calibration, inaccurate slope calculation, or the measurement noise. To study the performance of these methods in a field with non-zero curl, a rotational field (Fig. 6(a)) is added onto the same slope maps in Fig. 5(b), and the resultant slope maps with a rotational field are shown in Fig. 6(b).

From the error distribution from the HFLI, CTI, and RBFI methods (see Fig. 6(c), (d), and (e)), it shows all the methods in our comparison are sensitive to rotational fields, whereas it is a reasonable result with the data in hand, because the derivative of the reconstructed shape is the most similar gradient to the vector field obtained from measurement. Additional information may be needed to determine a unique gradient field from Helmholtz decomposition of the rotational vector field. Further studies on how to diminish the affection from fields with non-zero curl may be necessary to perform a better integration.

3.1.1.3. Noise. In a practical measurement, noise always exists. With a consideration on the principle of slope measurement in optical metrology, additive angular noise is simulated on the surface normal. In order to investigate the performance of these integration methods under noise condition in either a large or a small dataset, gradient data of the same shape in Eq. (20) in a dense sampling of

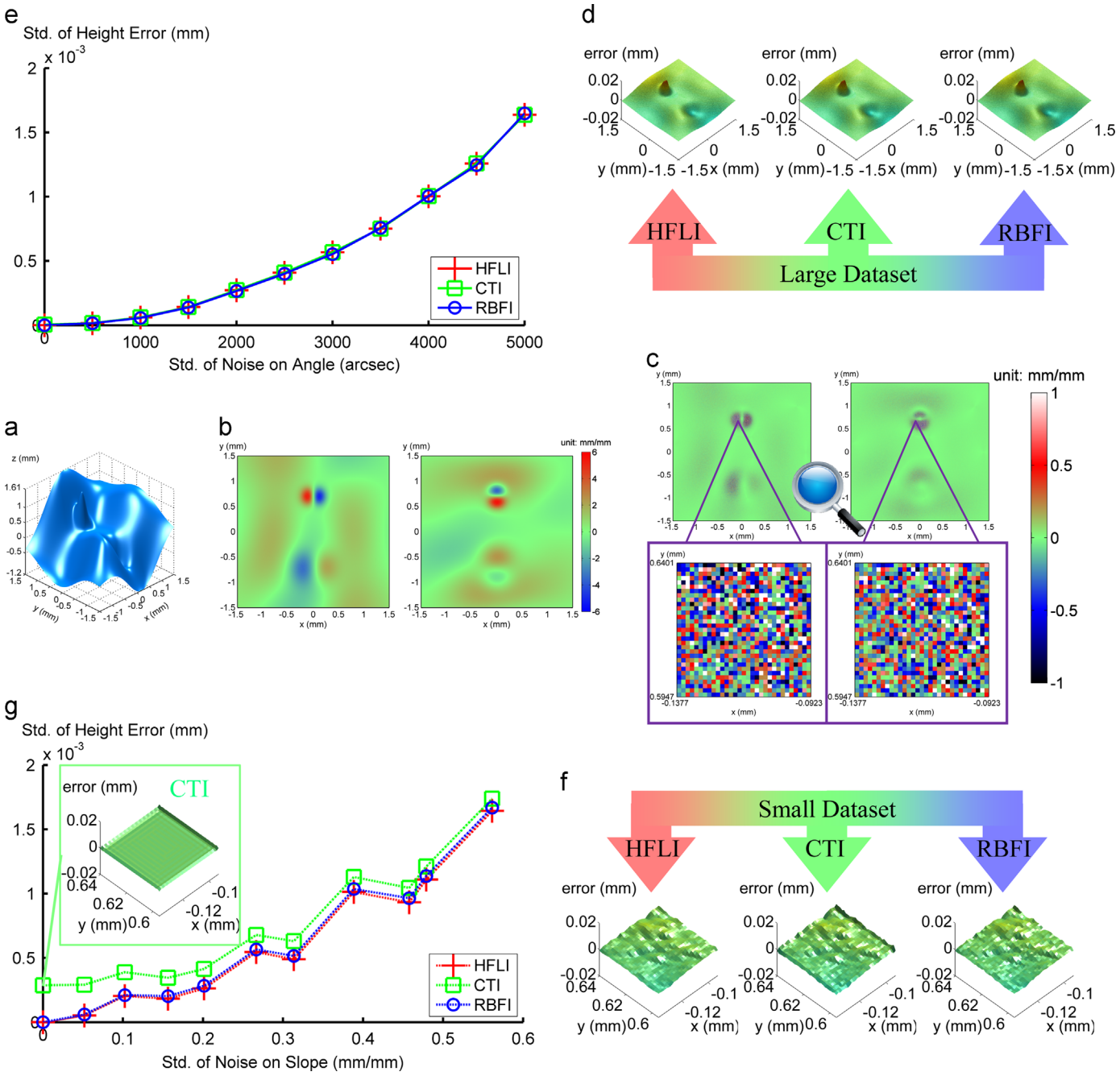

Fig. 7. Noise on angle influence the reconstruction accuracy. (a) The ground truth of shape is a large dataset with a dense sampling. (b) The slope maps is added with an angular noise with Std.=5000 arcs. (c) The whole maps of slope errors (upper) shows larger slope errors at steeper regions, and a zoomed-in small region with big slope errors is selected (lower). (d) Reconstruction errors are similar when using large datasets of slope in (b). (e) Three integration methods behave similarly under noise condition when handling large datasets. (f) Reconstruction errors are from integrating slope in the small region of (c) with big slope errors. (g) Reconstruction errors become bigger with the noise getting more severe, and the CTI method gets larger errors than the other two methods do when the dataset is small.

2048 (pixel) × 2048 (pixel) × 2 (direction) (see Fig. 7(a)) with Gaussian noise (Std. σ=5,000 arcs) are simulated on gradient (see Fig. 7(b)), from which a small piece of data [32 (pixel) × 32 (pixel) × 2(direction)] is cropped out. From the slope error maps in Fig. 7(c), it can be found, under the same angular noise, the slope error will be larger when the surface is steeper, *i.e.* the absolute value of the slope is larger.

Reasonably as shown in Fig. 7(d), the reconstruction error is larger at the region with larger slope errors. It needs to note that a

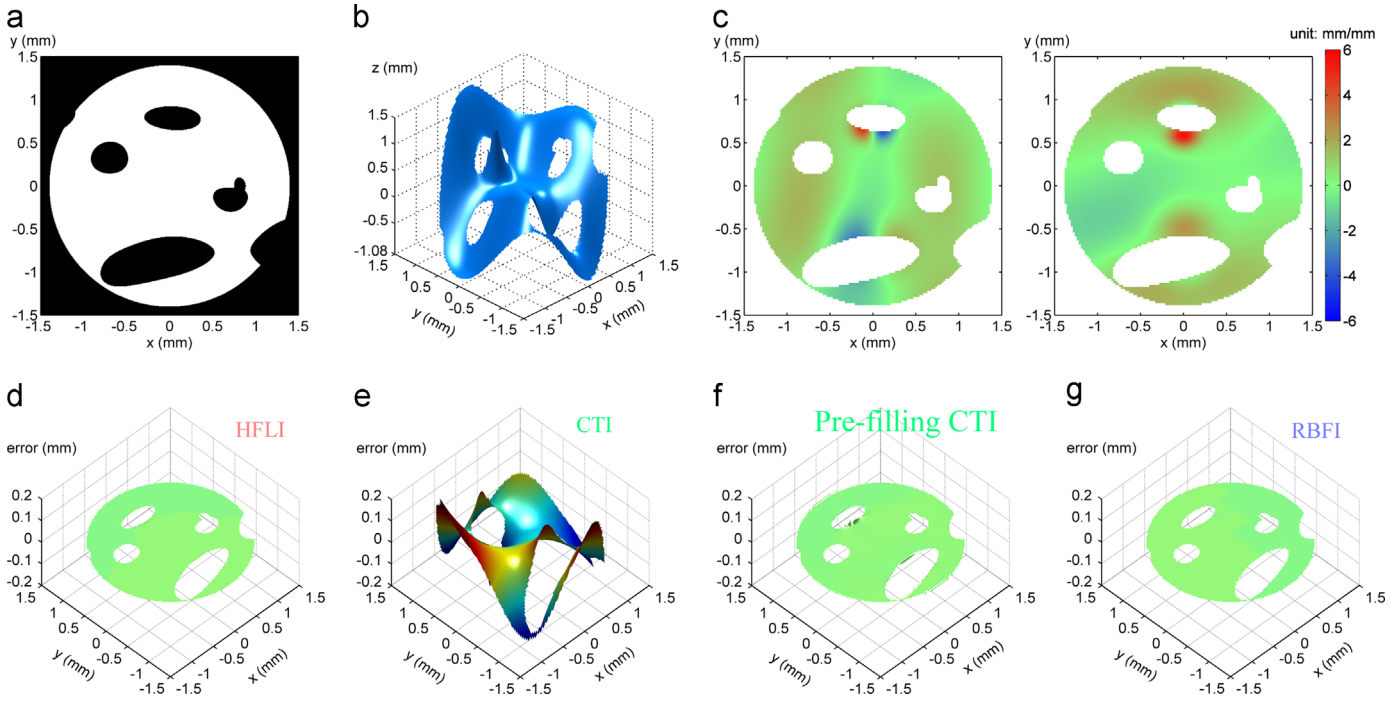

Fig. 8. Incomplete gradient data bring some troubles to the CTI method. (a) The mask indicates regions with available data. (b) The true 3D shape and (c) its slope data. Reconstruction errors of the methods of HFLI (d), CTI without pre-filling (e), CTI with pre-filling (f), and RBFI (g).

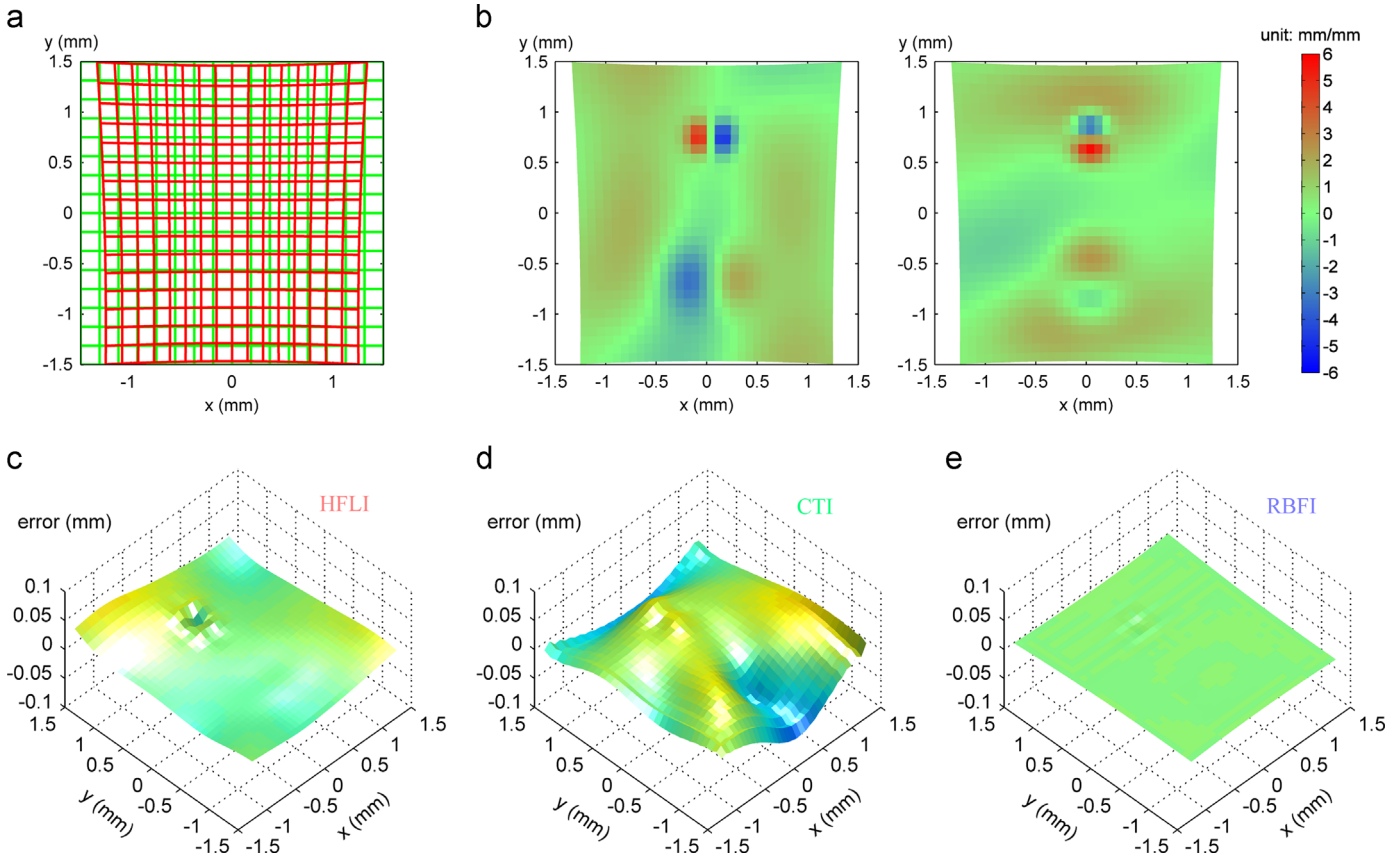

Fig. 9. The three methods are applicable to gradient data in quadrilateral grids, but only the RBFI method makes good reconstruction. (a) The rectangular grids becoming quadrilateral ones, (b) the corresponding slope maps, and the reconstruction errors from the methods of HFLI (c), CTI (d), and RBFI (e).

stitching process is required for RBFI method to deal with the large dataset. A stitching method [40] can be used stitch the subset results from RBFI method under noise condition. The integration errors under different noise levels shown in Fig. 7(e) indicate the three methods behave similarly under noise condition with large dataset.

The standard deviation (Std.) of slope error in that cropped small region can be estimated out as the constant. When the Std. of angular noise is 5000 arcs (the Std. of slope error is about 0.56 mm/mm in that region), the reconstruction errors with the three methods are almost similar (see Fig. 7(f)). However, when the noise is getting smaller as shown in Fig. 7(g), the influence of the higher order components, which the CTI method suffers while handling small data size, becomes more obvious.

3.1.1.4. Incomplete gradient data. In a practical measurement, the missing of data absolutely happens. It may result from the blocking of light by the aperture, "dead pixels" on detector, unmeasurable regions on samples, out of measuring range, or some other reasons. However, the good integration method should be able to deal with incomplete slope data from such imperfect measurement. Here, a mask (Fig. 8(a)) is made arbitrarily to crop the 3D shape (Fig. 8(b)) and its slope maps and the incomplete slope maps are shown in Fig. 8(c).

The error map from HFLI method in Fig. 8(d) indicates that missing of gradient data has little influence on the reconstructed shape and a similar observation can be found with RBFI method (Fig. 8(g)) as well. The connection of local (HFLI) or regional (RBFI) relations can naturally keep the none-slope pixels away from the optimization. Therefore, the implementation of HFLI and RBFI method is quite straightforward to handle incomplete gradient data. However, it is not straightforward to apply the CTI method to process gradient data with empty pixels. If those empty pixels in gradient data are just simply filled with "zero", large shape errors will show up as shown in Fig. 8(e). Additional preprocessing to the gradient data is necessary [41]. As shown in Fig. 8(f), if the missing gradient data is filled to maintain the loop continuity, shape error by using the CTI method can be reduced, but still larger than the other two methods. On the other hand, an iterative approach is suggested for the CTI method to deal with the incomplete dataset with good performance as well [42].

3.1.2. Other grids

In some applications, such as monoscopic phase measuring deflectometry [11–13], the imaging device is set up to observe the specimen in a certain perspective angle. Moreover, the lens distortion introduces additional dimensional changes of sensor grids in world coordinates. In this case, the rectangular grids certainly become quadrilateral ones as shown in Fig. 9(a). The corresponding slope values at the quadrilateral grids shown in Fig. 9(b) could be determined by some gradient measuring approach. Since all of these three integration methods are applicable in such a situation, it is interesting to know how their behaviors will be in quadrilateral grids.

As shown in Fig. 9(c) and (d), the HFLI and CTI methods get obvious shape errors. Although they are already improved from each lane, they are intentionally designed to solve the integration issues in rectangle grids. The errors indicate their performance is limited in quadrilateral grids. Numerical differentiation formats for non-uniform grids may improve the reconstruction accuracy of HFLI, but it is complicated and inconvenient in practical use. The HFLI and CTI methods are consequently not recommended in such situations.

In sharp contrast to the two methods above, good behaviors of RBFI method are observed when the data grids are in quadrilaterals (Fig. 9(e)) or even irregular shapes (triangular grids) (Fig. 10).

3.2. Speed with different data sizes

Apart from accuracy, the algorithm speed is usually in the list of concerned aspects. Actually, the computation time of different methods changes differently when the data size changes. In order to see the whole picture clearly, these three methods run for a comparison in various sizes of complete gradient data from a small dataset with 32 (pixel) × 32 (pixel) × 2 (direction) to a huge dataset with 2048 (pixel) × 2048 (pixel) × 2 (direction). The result is shown in Fig. 11, in which each running time result is the average value of 30 repeats in MATLAB® with Intel® Core™ i7-4600M CPU @2.90 GHz.

Obviously, the HFLI method runs relatively fast when the data size is small (*e.g.* $64 \times 64 \times 2$) or even not that slow with a medium size (*e.g.* $256 \times 256 \times 2$) dataset, but its computation time extremely increases along with the data getting to a large size (*e.g.* $1024 \times 1024 \times 2$).

Benefited from the fast transform algorithm, the CTI method makes very nice speed performance and it is the fastest method among these three methods in our comparison within all the data sizes.

A tiny size data (*e.g.* $32 \times 32 \times 2$) can be directly handled by RBFI method without stitching, but it runs relatively slow. To integrate larger size data with RBFI may need to split the dataset into subsets to integrate and then stitch them together. If the geometry is the same in each subset and the data are complete with no missing samples, a large intermediate matrix in computation will be same for each subset and there is no need to perform recalculations. This trick can save much computation time. With the enlarging of data size, the running time of RBFI method increases slower than that of HFLI method does. As a result, RBFI method becomes faster and faster than HFLI method while dealing with a large size gradient data (*e.g.* $2048 \times 2048 \times 2$). However, if the geometry is not the same for each subset or the data are incomplete, RBFI method slows down extremely and could be the slowest one.

4. Discussion

Based on our comparison of these three integration methods respect to accuracy and speed as shown in Fig. 12, several observations are addressed as follows.

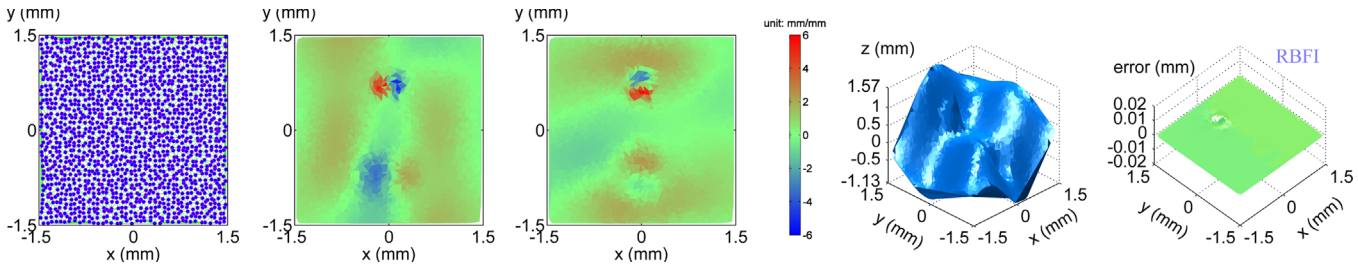

Fig. 10. The RBFI method can deal with gradient data in irregular grids.

(a) With the development of these integration methods on their own lanes, the state-of-art methods (HFLI, CTI, and RBFI) from different lanes have been selected out. Although their ways to link the slope-shape relations are different, the shape reconstruction of each method is identically based on estimation with given gradient data in a least squares sense. All of them are basically able to handle various sizes of gradient data with noise, high-order components, and incomplete slope measurements, which meets general requirements in usual applications.

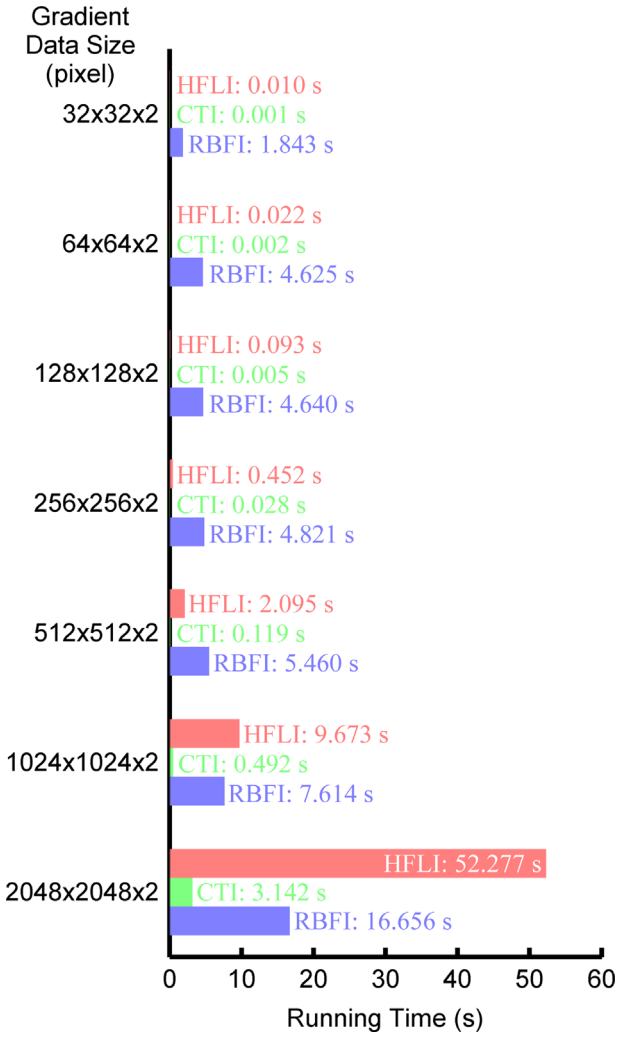

Fig. 11. The speed of different methods varies differently with increasing the gradient data size.

(b) The CTI method is not suggested to deal with tiny size dataset, because the ignored high-frequency terms under such low samplings are very likely not negligible.

(c) All the methods are invented to integrate the gradient field which should theoretically be conservative. However, a vector field from practical measurement may contain rotational components. A proper way to diminish the influence of the field with non-zero curl needs further investigations.

(d) The RBFI method can deal with the gradient data in irregular grids. The HFLI and CTI methods are not recommended if gradient data are not in rectangular grids.

(e) Benefited from the fast transform algorithms (*e.g.* FFT), the CTI method is fastest among these three methods at various data sizes. Comparing to the RBFI method, the HFLI method is faster at small and medium data size, but slower when handling large size gradient data, if the dataset is complete with the same grid geometry for each subset in the RBFI method. However, if the dataset is incomplete or the grid geometry is different for many subsets, the RBFI method can be extremely time consuming.

(f) The use of the HFLI and CTI methods are pretty easy and straightforward with no adjustable parameters. In the RBFI method, an improper effective range (specifically, the scaling factor) may affect the reconstruction result.

5. Conclusions

Two-dimensional integration methods for shape reconstruction from gradient data are reviewed in groups through the comparison. Three types of integration methods are mainly compared to show their different performance in accuracy and speed. Their advantages and drawbacks are presented in specific cases, *e.g.* higher order components, speed with various sizes, *etc.* It is of significance to properly choose a suitable integration method for a particular application.

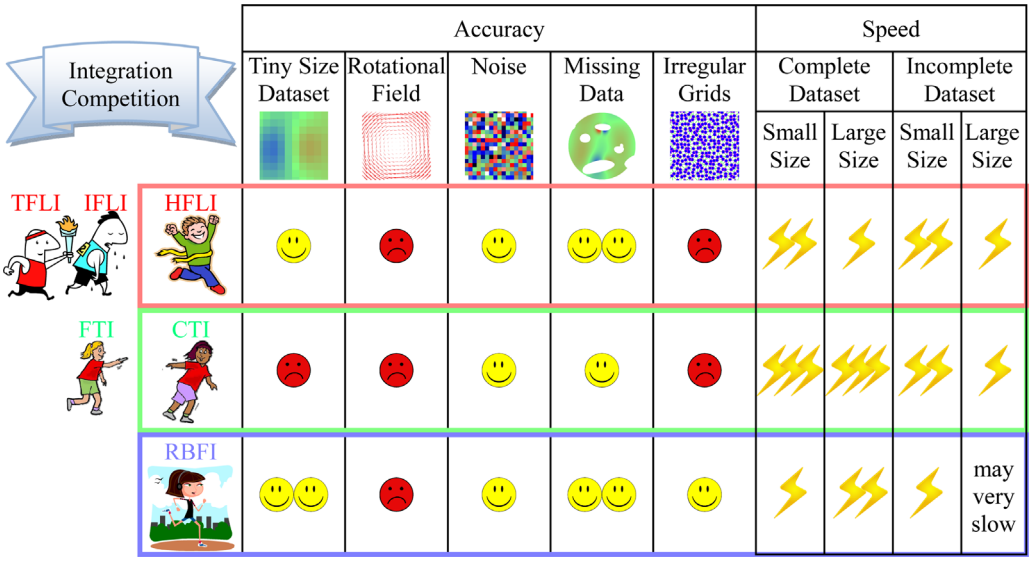

Fig. 12. The merits of drawbacks of the three integration methods are revealed obviously through comparisons.

Acknowledgements

The authors would like to thank Prof. Yuankun Liu in Sichuan University, China, Prof. Peng Su and Dr. Tianquan Su in University of Arizona for the helpful discussions.

References

[1] Idir M, Mercere P, Modi MH, Dovillaire G, Levecq X, Bucourt S, et al. X-ray active mirror coupled with a Hartmann wavefront sensor. Nucl Instrum Methods Phys Res, Sect A 2010;616:162–71.
[2] Platt BC, Shack R. History and principles of Shack-Hartmann wavefront sensing. J Refractive Surg 2001;17:S573–7.
[3] Idir M, Kaznatcheev K, Dovillaire G, Legrand J, Rungsawang R. A 2 D high accuracy slope measuring system based on a Stitching Shack Hartmann Optical Head. Opt Express 2014;22:2770–81.
[4] Bon P, Maucort G, Wattellier B, Monneret S. Quadriwave lateral shearing interferometry for quantitative phase microscopy of living cells. Opt Express 2009;17:13080–94.
[5] Chanteloup JC. Multiple-wave lateral shearing interferometry for wave-front sensing. Appl Opt 2005;44:1559–71.
[6] Shen W, Chang M-W, Wan D-S. Zernike polynomial fitting of lateral shearing interferometry. Opt Eng 1997;36:905–13.
[7] Frankot RT, Chellappa R. A method for enforcing integrability in shape from shading algorithms. IEEE Trans Pattern Anal Mach Intell 1988;10:439–51.
[8] Agrawal A, Raskar R, Chellappa R. What is the range of surface reconstructions from a gradient field? In: Leonardis A, Bischof H, Pinz A, editors. Computer vision—ECCV. Berlin Heidelberg: Springer; 2006. p. 578–91.
[9] Ikeuchi K, Horn BKP. Numerical shape from shading and occluding boundaries. Artif Intell 1981;17:141–84.
[10] Knauer MC, Kaminski J, Häusler G. Phase measuring deflectometry: a new approach to measure specular free-form surfaces. (Optical Metrology in Production Engineering). Strasbourg, France: SPIE; 2004; 366–76.
[11] Bothe T, Li W, von Kopylow C, Jüptner WPO. High-resolution 3D shape measurement on specular surfaces by fringe reflection. (Optical Metrology in Production Engineering). Strasbourg, France: SPIE; 2004; 411–22.
[12] Huang L, Seng Ng C, Krishna Asundi A. Fast full-field out-of-plane deformation measurement using fringe reflectometry. Opt Lasers Eng 2012;50:529–33.
[13] Huang L, Ng CS, Asundi AK. Dynamic three-dimensional sensing for specular surface with monoscopic fringe reflectometry. Opt Express 2011;19:12809–14.
[14] Tang Y, Su X, Wu F, Liu Y. A novel phase measuring deflectometry for aspheric mirror test. Opt Express 2009;17:19778–84.
[15] Tang Y, Su X, Liu Y, Jing H. 3D shape measurement of the aspheric mirror by advanced phase measuring deflectometry. Opt Express 2008;16:15090–6.
[16] 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.
[17] Su P, Wang Y, Burge JH, Kaznatcheev K, Idir M. Non-null full field X-ray mirror metrology using SCOTS: a reflection deflectometry approach. Opt Express 2012;20:12393–406.
[18] 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.
[19] Southwell WH. Wave-front estimation from wave-front slope measurements. J Opt Soc Am 1980;70:998–1006.
[20] Hudgin RH. Optimal wave-front estimation. J Opt Soc Am 1977;67:378–82.
[21] Hudgin RH. Wave-front reconstruction for compensated imaging. J Opt Soc Am 1977;67:375–8.
[22] 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.
[23] Li W, Bothe T, von Kopylow C, Jüptner WPO. Evaluation methods for gradient measurement techniques. (Optical Metrology in Production Engineering). Strasbourg, France: SPIE; 2004; 300–11.
[24] Talmi A, Ribak EN. Wavefront reconstruction from its gradients. J Opt Soc Am A 2006;23:288–97.
[25] Ettl S, Kaminski J, Knauer MC, Häusler G. Shape reconstruction from gradient data. Appl Opt 2008;47:2091–7.
[26] Bon P, Monneret S, Wattellier B. Noniterative boundary-artifact-free wave-front reconstruction from its derivatives. Appl Opt 2012;51:5698–704.
[27] Huang L, Asundi A. Improvement of least-squares integration method with iterative compensations in fringe reflectometry. Appl Opt 2012;51:7459–65.
[28] Huang L, Asundi A. In: Gorecki C, Asundi AK, Osten W, editors. Improvement of least squares integration method with iterative compensation for shape reconstruction from gradient. Brussels, Belgium: SPIE; 2012. p. 84300S–9.
[29] 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.
[30] Noll RJ. Phase estimates from slope-type wave-front sensors. J Opt Soc Am A 1978;68:139–40.
[31] 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.
[32] Dai G-m. Modal wave-front reconstruction with Zernike polynomials and Karhunen? Loève functions J Opt Soc Am A 1996;13:1218–25.
[33] de Visser CC, Verhaegen M. Wavefront reconstruction in adaptive optics systems using nonlinear multivariate splines. J Opt Soc Am A 2013;30:82–95.
[34] Roddier F, Roddier C. Wavefront reconstruction using iterative Fourier transforms. Appl Opt 1991;30:1325–7.
[35] 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.
[36] 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.
[37] Tribolet J, Crochiere RE. Frequency domain coding of speech. IEEE Trans Signal Process 1979;27:512–30.
[38] Feig E, Winograd S. Fast algorithms for the discrete cosine transform. IEEE Trans Signal Process 1992;40:2174–93.
[39] Gerchberg RW. Super-resolution through error energy reduction. Optica Acta: Int J Opt 1974;21:709–20.
[40] Huang L, Asundi AK. Framework for gradient integration by combining radial basis functions method and least-squares method. Appl Opt 2013;52:6016–21.
[41] 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.
[42] L. Huang, M. Idir, C. Zuo, K. Kaznatcheev, L. Zhou, A. Asundi, Shape reconstruction from gradient data in arbitrarily-shaped pupils by iterative discrete cosine transforms in Southwell configuration, a manuscript.