The following publication Budianto and D. P. K. Lun, "Robust Fringe Projection Profilometry via Sparse Representation," in IEEE Transactions on Image Processing, vol. 25, no. 4, pp. 1726-1739, April 2016, doi: 10.1109/TIP.2016.2530313.

Robust Fringe Projection Profilometry via Sparse Representation

Budianto, *Student Member, IEEE,* and Daniel P.K. Lun, *Senior Member, IEEE*

***Abstract*—In this paper, a robust fringe projection profilometry (FPP) algorithm using the sparse dictionary learning and sparse coding techniques is proposed. When reconstructing the three-dimensional (3D) model of objects, traditional FPP systems often fail to perform if the captured fringe images have a complex scene, such as having multiple and occluded objects. It introduces great difficulty to the phase unwrapping process of an FPP system that can result in serious distortion in the final reconstructed 3D model. For the proposed algorithm, it encodes the period order information, which is essential to phase unwrapping, into some texture patterns and embeds them to the projected fringe patterns. When the encoded fringe image is captured, a modified morphological component analysis (MCA) and a sparse classification procedure are performed to decode and identify the embedded period order information. It is then used to assist the phase unwrapping process to deal with the different artifacts in the fringe images. Experimental results show that the proposed algorithm can significantly improve the robustness of an FPP system. It performs equally well no matter the fringe images have a simple or complex scene, or are affected due to the ambient lighting of the working environment.**

***Index Terms*—Fringe projection profilometry, 3D model reconstruction, morphological component analysis (MCA), dictionary learning, sparse coding**

I. Introduction

Fringe projection profilometry (FPP) is a popular approach for reconstructing the three-dimensional (3D) model of objects. It has been generally applied to various 3D scanning systems including 3D scene reconstruction [1, 2], medical tomography [3], 3D face scanning [4], industrial quality control [5, 6], and 3D reconstruction of complex objects [7-12]. In a typical FPP setup (see Fig.1a), a projector is employed to project a spatially periodic fringe pattern onto an object. The deformed fringe pattern as shown on the object surface encodes the height information of the object (see Fig.1c). It is captured by a camera and the resulting fringe image is analyzed. The deformation of the fringe pattern can be mathematically modeled as the phase changes of a periodic signal. Thus by estimating the phase changes in the captured fringe image, the 3D model of the object can be inferred.

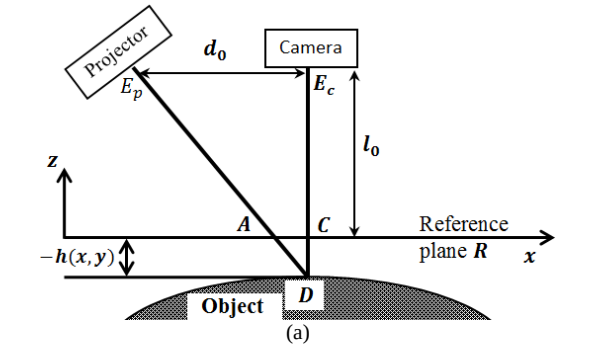
The work described in this paper was fully supported by the Hong Kong Research Grant Council (under research account: PolyU 5210/13E).

Budianto and D. P.-K. Lun* are with the Centre for Signal Processing, Department of Electronic and Information Engineering, the Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: enpklun@polyu.edu.hk)

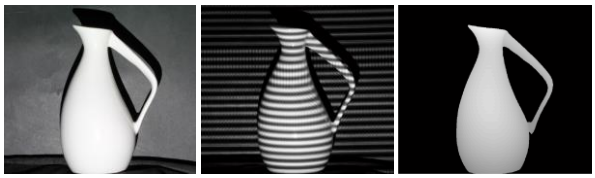
In the last decade, many methods have been developed for measuring the phase changes in the fringe images. For instance, in the phase shifting profilometry method (PSP), three fringe patterns with fixed phase shifts are projected onto the target object [13, 14]. The phase changes are measured using a minimum of three shifted fringe images. In the Fourier transform profilometry (FTP) method [3, 15], the phase is estimated after filtering out the unwanted frequencies of the fringe pattern in the Fourier domain. In the wavelet transform profilometry (WTP) [16-20] and the short time Fourier transform profilometry (STFTP) methods [21, 22], the phase is tracked by analyzing the ridge in the spatial-frequency domain. And in the PSP with FTP approach, a prefiltering stage is employed in conjunction to PSP methods [6, 23].

For most phase imaging systems, it is essential to obtain the true phase from the incomplete and noisy phase estimations so as to infer the physical measurement of interest. It is also the case for FPP. All the aforementioned FPP methods can only provide a modulo-2π estimation of the phase, which is the so-called wrapped phase. A phase unwrapping procedure is needed to recover the true phase from the wrapped phase estimation. Traditional phase unwrapping methods are often based on the Itoh analysis, which suggests to integrate the wrapped phase differences to obtain the true phase [24]. This approach assumes the smoothness of the true phase and the availability of all wrapped phase data. However such assumption is not valid in case of occlusion or incomplete observations. Severe distortion will thus result in the reconstructed 3D model. To solve the problem, an intuitive solution is to project additional fringe patterns which contain the period order information for phase unwrapping. By period order information, we refer to the integer number of 2π jumps in the phase angle that is hidden in the wrapped phase data. If the period order is known, the true phase information can be readily obtained. Some examples of the methods in this category include the gray-code plus phase shifting method [25], the phase code plus phase shifting method [26], the temporal phase unwrapping algorithm [27], and the gray code plus strip-edge shifting method [5]. However the requirement of additional fringe projections is a major disadvantage of these approaches. It does not only slow down the acquisition process of the fringe images, but also requires more memory and processing time to reconstruct the 3D model.

To avoid projecting additional fringe patterns, recently some approaches were proposed to embed the period order



(a)



(b) (c) (d)

Fig. 1. (a) FPP setup in crossed optical axes geometry [15]. (b) A testing object. (c) Fringe image of the object. (d) Reconstructed 3D model of the object based on the fringe image in (c).

information into the fringe patterns. In these approaches, the period order information is first encoded using, for instance, patterns of multiple wavelengths or multiple frequencies [28-30], multiple color [31], random patterns [32, 33], or structural markers [13, 20, 23, 34, 35], etc. They are then embedded into the fringe patterns and projected onto the object. Theoretically, these approaches allow correct phase unwrapping without additional fringe projections. But the approaches in [13, 20, 28-31, 34, 35] can only be used for fringe images of simple scene, e.g. only have a single simple object. For the methods in [23, 32, 33], the accuracy of the period order estimation is low hence additional procedure such as the voting strategy [33] is required to obtain the correct result.

In this paper, the problem of true phase estimation is addressed by exploiting the sparsity of different components of a fringe image when it is embedded with period order information. It is achieved through a dictionary learning approach. A dictionary consists of a set of words (vectors) which provides a sparse representation of a given signal through a procedure known as the sparse coding. The dictionary learning and sparse coding techniques have been used successfully for classification tasks such as object classification [36], face recognition [37-39], texture segmentation [39-41], etc. However, to our knowledge, the dictionary learning approach has not been used in FPP before. For the proposed approach, different code patterns that carry the period order information are first embedded into the fringe patterns and projected onto the target objects. The deformed fringe images are captured and the code patterns that carry the period order information are detected by performing a modified morphological component analysis (MCA). It differs from the traditional MCA [42-44] by considering also the blurring effect and noise in fringe images. A discriminative dictionary is then used to estimate the period order information from the detected code patterns. To reconstruct the 3D model of the object, a multi-level quality guided phase unwrapping algorithm [45] is adopted. It is assisted by the estimated period order information so that it can deal with the different artifacts in the fringe images. The proposed algorithm improves over the traditional methods in that, first, it does not require additional projections to embed period order information; second, it is robust due to the use of the modified MCA and the dictionary learning methods for detecting the period order information; and finally, the proposed algorithm can perform equally well for fringe images of simple and complex scene, such as those with multiple complex objects. They are all verified by our experimental results.

The rest of the paper is organized as follows. First the background of FPP, MCA, and dictionary learning are presented in Section II. The details of the proposed algorithm are described in Section III. Experimental results are reported in Section IV and finally, conclusions are drawn in Section V.

II. Background

A. Principle of FPP Method

In this section, we first outline the basic concept of the conventional FFP method [15]. The typical setup of the FPP method is shown in Fig.1a. In the figure, a projector at position E_p projects a fringe pattern and a camera located at position E_c captures the deformed fringe pattern as shown on the object surface. The projector and camera are placed at a distance of d_0 from each other and they are at a distance of l_0 from a reference plane *R*. The deformed fringe pattern captured by the camera at E_c can be modeled mathematically as a sinusoidal function as follows [15]:

$$g(x,y) = a(x,y) + b(x,y)\cos\left[\varphi(x,y)\right] + n(x,y) \qquad (1)$$

In (1), $g(x,y)$ represents the pixels, in x and y directions, of the fringe image captured by the camera. a is the bias caused by the surface illumination of the object; b is the local amplitude of the fringe; and n is the noise generated in the image capture process. $\varphi(x,y) = 2\pi f_0 x + \phi(x,y)$ is the phase angle in which f_0 is the carrier frequency, and $\phi(x,y)$ is the phase shift of the fringe in x and y directions. It has a relationship with the object height h as follows:

$$h(x,y) \approx -\frac{l_0}{2\pi f_0 d_0}\Delta\phi(x,y) \qquad (2)$$

In (2), $\Delta\phi(x,y) = \phi(x,y) - \phi_o(x,y)$, where ϕ_o is the ϕ when there is no object. It is assumed to be known in the initial calibration process. Hence if φ is known, ϕ and also $\Delta\phi$ can be determined. Then the object height profile h and in turn the 3D model of the object can be reconstructed. However, any trigonometric methods that directly retrieve the true phase map φ from the fringe image g will end up with only the wrapped version of φ, which is always in the range of $-\pi$ to π. It is caused by the fact that, as shown in (1), all φ separated by 2π will have the same g. Let us denote the wrapped phase map to

be $\hat{\varphi}$ such that,

$$\hat{\varphi} = \Pi(\varphi) = \mathrm{mod}(\varphi + \pi, 2\pi) - \pi \;\; with \;\; \hat{\varphi} \to (-\pi, \pi] \quad (3)$$

where $\mathrm{mod}(x, y)$ is the modulus function of x with the divisor y. A phase unwrapping process is needed for retrieving φ from $\hat{\varphi}$. Many of the existing unwrapping algorithms [46] are based on the Itoh analysis [24], which suggests to integrate the wrapped phase differences in order to calculate the desired true phase map φ. Denote $\varphi_y(x) \equiv \varphi(x, y)$ and $\hat{\varphi}_y(x) \equiv \hat{\varphi}(x, y)$. The phase unwrapping process based on the Itoh analysis can be described as follows:

$$\varphi_y(x) = \hat{\varphi}_y(0) + \sum_{x=1}^{m} \Delta_y(x) \quad (4)$$

where $\Delta_y(x) = \hat{\varphi}_y(x) - \hat{\varphi}_y(x-1)$ and $-\pi < \Delta_y(x) \le \pi$. However, (4) is valid only if $\Delta_y(x)$ is known for all *x*. If for a particular point *x'* such that $\Delta_y(x')$ is estimated with error or even missing, the unwrapped $\varphi_y(x)$ will have error for all $x \ge x'$. Such situation unfortunately is common in many FPP setups.

B. MCA based on Variable Splitting using SALSA

Given an observed signal Y, which is formed by the summation of two signals X_1 and X_2, i.e. $Y = X_1 + X_2$. Assume that X_1 and X_2 have different morphologies. The goal of MCA is to estimate X_1 and X_2 by solving the following optimization problem:

$$\arg\min_{w1,w2} \left\| Y - \Phi_1 w_1 - \Phi_2 w_2 \right\|_2^2 + \lambda_1 \left\| w_1 \right\|_1 + \lambda_2 \left\| w_2 \right\|_1 \quad (5)$$

where Φ_1 and Φ_2 are the transform matrices of X_1 and X_2 such that the transform coefficients w_1 and w_2 (i.e. $X_1 = \Phi_1 w_1$ and $X_2 = \Phi_2 w_2$) are sparse. In [47], the MCA is realized by using the rational-dilation wavelet transform (RADWT) [48] and the iterative thresholding based on the split augmented Lagrangian shrinkage algorithm (SALSA) developed in [49, 50]. More specifically, the problem in (5) can be formulated as,

$$f_1(v) = \left\| Y - \Phi v \right\|_2^2, \quad f_2(w) = \lambda_1 \left\| w_1 \right\|_1 + \lambda_2 \left\| w_2 \right\|_1 \quad (6)$$

with

$$\Phi = [\Phi_1 \quad \Phi_2], \; v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \; w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \; \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \quad (7)$$

Then X_1 and X_2 can be obtained by iteratively performing the following computation. Given the initial v^0 and δ^0,

1. $z^{k+1} = \mathcal{S}_\lambda(v^k + \delta^k) - \delta^k$;
2. $\delta^{k+1} = \dfrac{1}{(\mu + 2)} \Phi^T (Y - \Phi z^{k+1})$;
3. $v^{k+1} = \delta^{k+1} + z^{k+1}$

where $z = w - \delta$; z^{k+1} stands for the *z* after *k*+1 iterations; and $\mathcal{S}_\lambda(a) = \mathrm{sgn}(a) \max(|a| - \lambda, 0)$ is the standard soft thresholding operation for each element in the matrix with threshold λ. The above computation is repeated until convergence. Then $X_i = \Phi_i w_i^{converge}$, where $w_i^{converge}$ is the w_i when converged; and *i* = 1,2. It is shown in [47] that the above algorithm can achieve the global minimum when converged.

C. Dictionary Learning

For the traditional dictionary learning methods, a set of training signals $\dot{X} = [\dot{x}_1, \dot{x}_2, ..., \dot{x}_N]$ in $\mathbb{R}^{M \times N}$ is employed to learn an over-complete dictionary $D \in \mathbb{R}^{M \times L}$ (hence *L* > *M*) that can give a sparse representation $\gamma = [\gamma_1, ..., \gamma_N] \in \mathbb{R}^{L \times N}$ of $\dot{X}$. The learning process can be compactly written as a minimization problem as follows:

$$\arg\min_{D,\gamma} \left\| \dot{X} - D\gamma \right\|_2^2 \; s.t. \; \forall i \quad \left\| \gamma_i \right\|_0 \le T_0 \quad (8)$$

where T_0 is a threshold; and $\|\cdot\|_0$ is the pseudo l_0-norm, which can be implemented by counting the number of non-zero entries. One of the popular approaches to solve the above optimization problem is the K-SVD algorithm [51]. Instead of l_0-norm, we can also choose to use the l_1-norm since both the l_0-norm and l_1-norm promote sparsity. When l_1-norm is used, methods such as the basis pursuit [52] or lasso [53] can be used as in online dictionary learning (ODL) [54].

Given a dictionary *D*, the sparse representation $\gamma = [\gamma_1, ..., \gamma_N] \in \mathbb{R}^{L \times N}$ of a signal $X = [x_1, x_2, ..., x_N]$ in $\mathbb{R}^{M \times N}$ can be obtained through a process known as the sparse coding, which can be expressed as another minimization problem as follows:

$$\arg\min_{\gamma_i} \left\| x_i - D\gamma_i \right\|_2^2 \; s.t. \; \left\| \gamma_i \right\|_0 \le T_0 \quad \forall i \quad (9)$$

(9) can be solved by using the orthogonal matching pursuit algorithm (OMP) [55]. Since the sparse codes of a signal can also be considered as its features, recently there are many works trying to use the dictionary learning methods in signal classification applications [36-41, 56]. However, traditional dictionary learning methods are for signal reconstruction purposes (such as compressive sensing). The discriminative power of the dictionaries needs to be improved for using them in classification applications. To do so, researchers have developed several approaches to learn a classification-oriented dictionary in a supervised learning fashion by exploring the class label information. More specifically, denote $W \in \mathbb{R}^{m \times L}$ to be a model parameter which assigns the class label for a given sparse code, where m denotes the number of classes to be discriminated. Two approaches are commonly used for assigning the class labels, either by minimizing the class-specific residue [41] or by a linear classification [37, 38]. In this research, we employ the later approach. W can be jointly learned with the dictionary D based on a cost function as follows:

$$\arg\min_{D,W,\gamma} \left\| \dot{X} - D\gamma \right\|_2^2 + \lambda_2 f(W) + \lambda_3 f(D) \quad s.t.\ \forall i \quad \left\| \gamma_i \right\|_0 < T_0 \tag{10}$$

$f(D)$ denotes a function that forces the sub-dictionaries in D for different classes to be as incoherent as possible [37, 41]. $f(W)$ denotes a classifier function, e.g., a linear classifier [38], a label consistent term [37], a logistic function [36], Fisher discrimination criterion [57]. And $\{\lambda_i\}_{i=1,2,3}$ denote the regularization parameters. Comparing with (8), it can be seen that additional constraints are added in (10) for improving the discriminative power of the learned dictionary. In this work, we use a similar approach to learn a dictionary that can effectively detect the period order information of the fringe images and in turn facilitate the reconstruction of the 3D model of the object.

III. Proposed Algorithm

A. Encoding of the Period Order Information

Recall from Section II-*A* that our goal is to accurately evaluate the true phase map φ from the wrapped phase map $\hat{\varphi}$ computed from the fringe image. Once φ is obtained, the 3D model of the object at each scene point can be evaluated using (2). In fact, there is a direct relationship between φ and $\hat{\varphi}$ as follows [24]:

$$\varphi = \hat{\varphi} + K 2\pi \tag{11}$$

where K is also called the K-map. It gives the missing period order information (or the k-value) of $\hat{\varphi}$. If it is known, the phase unwrapping problem is solved. The key idea of the proposed method is to encode the K-map with some unique textural patterns. They are then embedded into the fringe patterns and projected onto the object. Thus the captured fringe image can be formulated as,

$$X = X_1 + X_2 . \tag{12}$$

where X_1 denotes the original sinusoidal fringe pattern (i.e. g in (1)); and X_2 denotes the code patterns that encode the k-value. It is defined by,

$$\begin{aligned} X_2 &= \Theta\{K(\varphi)\} \\ K &: \mathbb{R} \to \mathbb{Z}^+ \\ \varphi &\mapsto \left\lfloor \frac{\varphi + \pi}{2\pi} - \pi \right\rfloor \end{aligned} \tag{13}$$

where Θ is an encoding function for each k-value; and $\lfloor \cdot \rfloor$ is the floor function that gives the closest smaller integer number. Hence X_2 is unique for each k-value.

When designing the encoding function Θ, it is important to ensure that: 1) Θ can generate some code patterns X_2 which have different morphological structures from X_1, e.g., different frequencies, colors, etc.; and 2) X_2 should have a unique feature for each k-value. In this paper, we propose to construct the code patterns by a concatenation of image patches (or textons). An example of five unique textons and their corresponding code patterns are shown in Fig. 2. As can be seen in the figure, each texton has the size of 5x5 pixels and has a unique orientation and scale. It will give a unique response when applying to a 2D band-pass multiresolution transform filter, e.g. the Gabor filter. As shown in column 3 of Fig. 2, one unique code pattern is assigned to each k-value of the fringe. We shall discuss in the next few sections how the k-value can be efficiently decoded by using a dictionary learning method.

B. Overview of the Proposed Decoding Algorithm

As illustrated in Fig. 3, the proposed decoding algorithm consists of an offline stage performed during the system calibration and an online stage when the object's 3D model is actually reconstructed. At the offline stage, a large number of training fringe patterns, which have been encoded with different code patterns, are first applied to a modified MCA procedure so that the fringe patterns and codes patterns are separated. Since the k-values that the code patterns represent are known, a supervised label consistent K-SVD (LC-KSVD) training process [37] can be carried out for learning a discriminative dictionary *D* and a linear classifier *W* which will be used at the online stage.

At the online stage, the encoded fringe images of the testing object are fed to a modified MCA procedure to separate the fringe patterns and the code patterns. The wrapped phase map $\hat{\varphi}$ is then computed from the fringe patterns while the k-values of $\hat{\varphi}$ are determined by feeding the code patterns to the discriminative dictionary *D* and a linear classifier *W* learned at the offline stage. These k-values thus obtained are used to guide a multilevel phase unwrapping procedure to evaluate the true phase map φ and in turn reconstruct the 3D model of the object. Details of the algorithm will be explained in the following subsections.

C. The modified MCA

For a typical FPP setup, a fringe pattern needs to go through various optical devices before being captured by the camera. Hence the captured fringe images are often noisy and blurred,

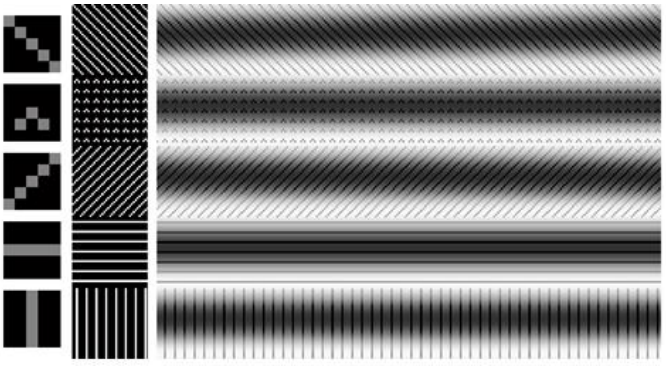


Fig. 2. 5×5 pixel binary textons (most left); the code pattern generated by the corresponding texton (second column); and the encoded fringe pattern (column three).

which can be formulated as,

$$Y = \mathcal{K}(X_1 + X_2) + \varepsilon \tag{14}$$

where Y is the captured image; $\mathcal{K}$ is the blurring kernel; ε is a Gaussian white noise of known variance. X_1 and X_2 are defined as in (12) and (13). Y can be interpreted as the convolution of the original encoded fringe pattern with the blurring kernel plus additive Gaussian white noise.

Given a predefined transform matrices Φ_1 and Φ_2, the problem in (14) can be formulated as an optimization problem similar to (5) as follows:

$$\arg\min_{\alpha_1,\alpha_2} \left\| Y - \mathcal{K}(\Phi_1\alpha_1 + \Phi_2\alpha_2) \right\|_2^2 + \lambda_1 \left\|\alpha_1\right\|_1 + \lambda_2 \left\|\alpha_2\right\|_1 \tag{15}$$

where α_i are the sparse representations of $\{X_i\}_{i=1,2}$ and λ_i are the regularization parameters. (15) can still be solved by using the MCA with a slight modification. First, we employ

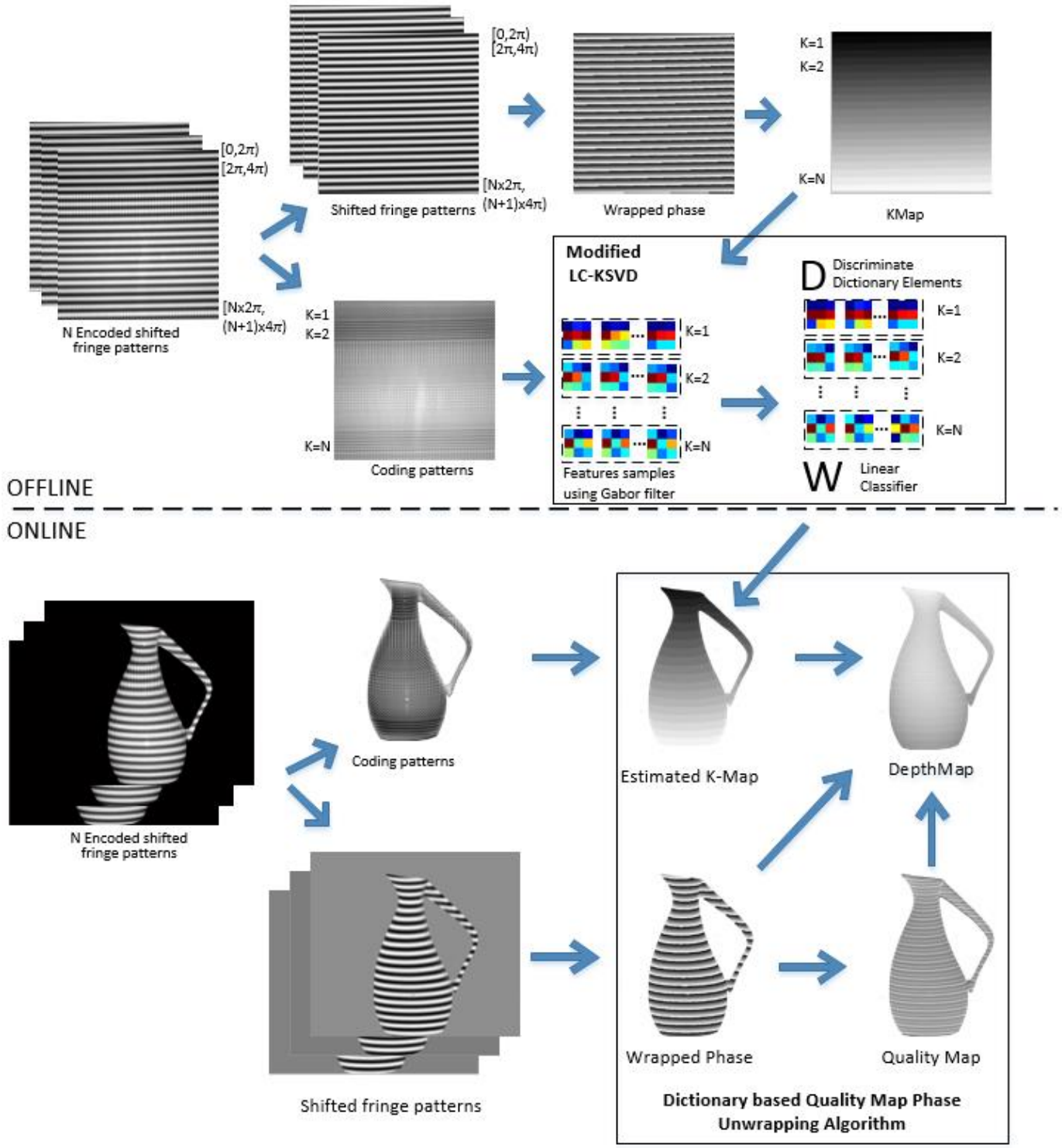


Fig. 3. Flowchart of the proposed FPP method for 3D model reconstruction

similar to [47, 58] the tuned-Q wavelet transform (TQWT) for the implementation of Φ_1 and Φ_2. They fulfill the mutual incoherence requirement and can efficiently capture the structures of interest (i.e. having sparse α_i). For performing the MCA, we first utilize the splitting variable approach. Let,

$$f_1(v) = \|Y - \mathcal{K}\Phi v\|_2^2, \quad f_2(\alpha) = \lambda_1 \|\alpha_1\|_1 + \lambda_2 \|\alpha_2\|_1 \tag{16}$$

with

$$\Phi = [\Phi_1 \quad \Phi_2], \; v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \; \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \tag{17}$$

(15) can be written as the following iterative optimization procedure: Given the initial v^0 and δ^0,

$$\begin{aligned} \alpha^{k+1} &= \arg\min_{\alpha} \lambda_1 \|\alpha_1\|_1 + \lambda_2 \|\alpha_2\|_1 + \frac{\mu}{2}\left\|\alpha - v^k - \delta^k\right\|_2^2 \\ v^{k+1} &= \arg\min_{v} \|Y - \mathcal{K}\Phi v\|_2^2 + \frac{\mu}{2}\left\|\alpha^{k+1} - v - \delta^k\right\|_2^2 \\ \delta^{k+1} &= \delta^k - (\alpha^{k+1} - v^{k+1}) \end{aligned} \tag{18}$$

Since the tight-framed TQWT is used for the implementation of Φ_1 and Φ_2; and $\mathcal{K}$ can be assumed to be a circular convolution operator, $\mathcal{K}$, Φ_1 and Φ_2 can be factorized as,

$$\mathcal{K} = U^T H U, \quad \Phi_i = U^T C_i U \quad \forall i \in 1,2 \tag{19}$$

where U represents the discrete Fourier transform (DFT), $U^T = U^{-1}$ is its inverse; H and C_i are some diagonal matrices. As shown in Appendix A, (18) can be simplified to the following iterative algorithm such that the fringe pattern and the embedded code pattern (i.e. X_1 and X_2 as defined in (12) and (13)) can be obtained from the observed blurred and noisy fringe image Y:

Algorithm I: Modified MCA

Given the initial v^0 and δ^0,

1. $z^{k+1} = \mathcal{S}_{\lambda/\mu}(v^k + \delta^k) - \delta^k$
2. $v^{k+1} = \frac{1}{\mu}\left(I - U^T F U\right)\left(\Phi^T \mathcal{K}^T Y + \mu z^{k+1}\right)$
3. $\delta^{k+1} = v^{k+1} - z^{k+1}$

For the algorithm, we assume that Φ are in a tight frame, i.e. $\Phi\Phi^T = 2I$. The definition of most parameters is the same as in Section II-*B*, except a new parameter *F* is introduced which is defined as follows:

$$F = C^T H^* \left(\mu + 2|H|^2\right)^{-1} HC \tag{20}$$

where H^* is the complex conjugate of H; $|H|^2$ is the squared absolute values of the entries of the diagonal matrix H; and $C = [C_1 \quad C_2]$. In (20), the expression $H^*(\mu + 2|H|^2)^{-1}H$ is actually a Wiener filter in the frequency domain. It helps to enhance the blurred and noisy fringe image to facilitate the decomposition.

One of the important features of **Algorithm I** is its fast computation speed. In the algorithm, the term $\Phi^T \mathcal{K}^T Y$ can be pre-computed before the iteration; and F can be computed efficiently since it is performed in the Fourier domain. The total computation cost is only $O(rN \log_2 N)$, where N is the total number of pixels of the fringe image and r is the redundancy factor which is set to 2. Fig. 4 shows the results of using **Algorithm I** to decompose an encoded fringe image with various regularization parameter λ. As shown in the figure, the fringe pattern X_1 and the code patterns X_2 can be clearly extracted from the encoded fringe image when λ_1 is set to ~0.3 ($\lambda_1 + \lambda_2 = 1$). The parameter λ_1 and λ_2 are set at the offline stage empirically.

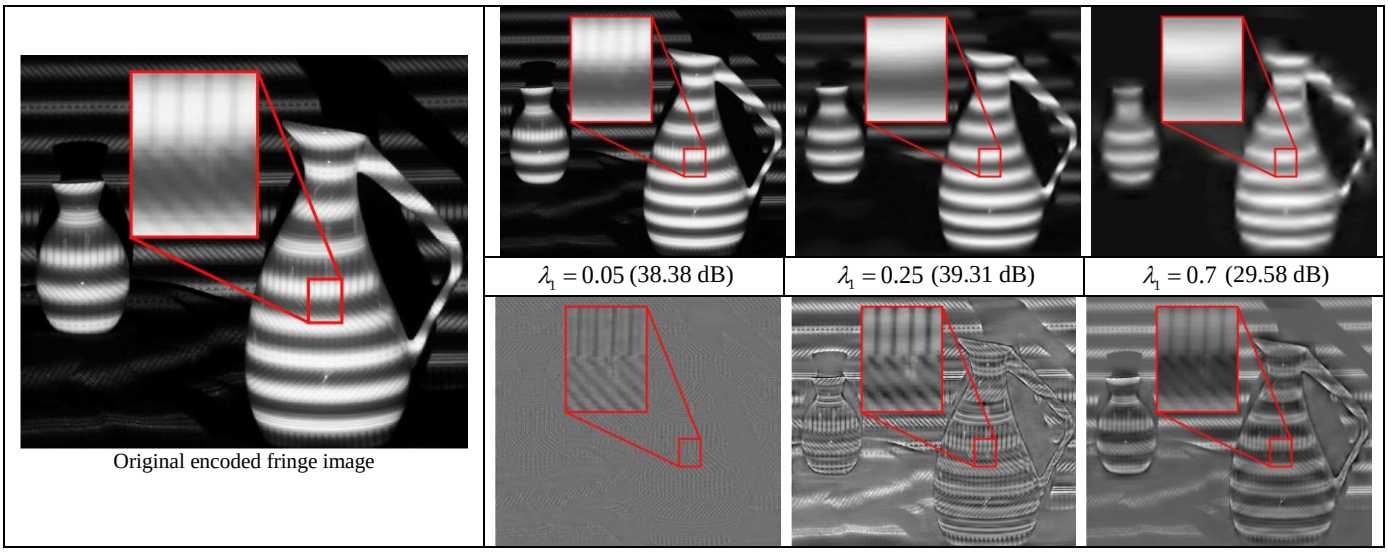


Fig. 4. Decomposition of the encoded fringe image using the modified MCA with different regularization parameter λ: The input encoded fringe image (the most left column); the decomposed fringe pattern X_1 (top row); the decomposed code patterns X_2 (bottom row); with different regularization parameter λ.

D. Dictionary Learning

In this section, we present the dictionary learning procedure and how the learned dictionary and the linear classifier are used in the k-value estimation.

1) Training Set for Dictionary Learning

To construct a reliable and robust dictionary, a large number of high quality image patches is needed for the training process. For the proposed algorithm, the image patches are selected during the offline stage by first projecting the encoded fringe pattern on a flat surface and capturing the resulting fringe image following the standard FPP approach. Since the object is a flat surface, there will not be any phase jump hence all k-values can be easily determined from the fringe image. They will be used as the ground truth for training the dictionary. As mentioned above, the encoded fringe image will go through the modified MCA such that the fringe pattern and code patterns are separated. For each region R^k in the separated code pattern (where k is the k-value of that region), a set of code pattern patches and their Gabor features can be obtained. Let,

$$P^k = \{p_i^k\}_{i=1...N^k} \text{ and } \zeta^k = \{\zeta_i^k \mid \zeta_i^k = G(p_i^k)\}_{i=1...N^k} \quad (21)$$

where P^k and ζ^k are a patch set and a patch feature set, respectively. N^k is the number of training patches in P^k. In our experiments, we randomly select 256 patches ($N^k = 256$) for each k. Suppose we have K k-values, i.e. $k = 1,...K$, the total number of training patches is $N^k K$.

In (21), $G(\cdot)$ is a function that extracts the features of a given patch. For the proposed system, the Gabor transform is used to extract the features of the patches. The Gabor features are obtained by convolving a patch p with a Gabor kernel of J scales, i.e. $j = 1,..., J$; and θ orientations, i.e. $\theta_o = 1,..., \theta$. It results in a complex vector which can be written as, $\rho_{j,\theta}^k = p^k * G_{j,\theta}$, where '*' denotes the convolution operator. The final feature set ζ^k is formed by taking the mean of the complex vectors $\rho_{j,\theta}^k$ of different scales and orientation defined as follows:

$$\zeta^k = \left[\overline{|\rho_{1,1}^k|}, ..., \overline{|\rho_{1,\theta}^k|}, ..., \overline{|\rho_{J,1}^k|}, ..., \overline{|\rho_{J,\theta}^k|} \right] \quad (22)$$

where $\overline{|\rho|}$ is the mean of the magnitude of ρ. From our experiments, we observe that good results can be obtained by setting $J = 3$ and $\theta = 6$.

2) Discriminative Dictionary Learning and Linear Classifier

To learn a discriminative dictionary D, we employ the label consistent K-SVD algorithm version 1 (LC-KSVD1) [37] which minimizes the following objective function,

$$\arg\min_{D,A,\gamma} \|Z - D\gamma\|_2^2 + \|B - A\gamma\|_2^2 \ s.t. \ \forall i \ \|\gamma_i\|_0 \le T \quad (23)$$

where $Z = [\zeta^1, \zeta^2, ..., \zeta^K] \in \mathbb{R}^{\theta J \times N^k K}$ is the training feature set. Each column of Z is a patch feature $\zeta_i^k = G(p_i^k)$ with $p_i^k \in X_2$. As it is indicated in [37], (23) will learn a dictionary $D \in \mathbb{R}^{\theta J \times L}$ which gives the sparse codes $\gamma \in \mathbb{R}^{L \times N^k K}$ of Z. Besides, the second term of (23) enhances the discriminability of γ by minimizing the difference between the linear transformation of γ and a discriminative block diagonal binary matrix $B \in \mathbb{R}^{N^k K \times N^k K}$ defined as follows:

$$B = \begin{bmatrix} 1_{N^k \times N^k} & 0 & 0 & 0 \\ 0 & 1_{N^k \times N^k} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1_{N^k \times N^k} \end{bmatrix} \quad (24)$$

where $1_{N^k \times N^k}$ is a square matrix of size $N^k \times N^k$ with all elements equal to 1. It forces the code pattern patches of the same period order *k* to have very similar sparse representations. It results in good classification performance even using a simple linear classifier as will be discussed below. The matrix D and A are initialized using the discrete cosine transform (DCT) basis and updated using the approach in [37].

When a discriminative sparse code of a patch is obtained, we need a classifier to determine the k-value that the code represents. As mentioned above, it is sufficient to use a linear classifier which can be obtained by solving the following minimization problem:

$$\arg\min_{W} \|H - W\gamma\|_2^2 + \lambda_W \|W\|^2 \quad (25)$$

where W is the required classifier and γ is obtained from (23). λ_W is a constant to control the contribution of the corresponding term; and

$$H = \begin{bmatrix} 1_{1 \times N^k} & 0 & 0 & 0 \\ 0 & 1_{1 \times N^k} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1_{1 \times N^k} \end{bmatrix} \quad (26)$$

(25) has a close form solution as follows:

$$W = H\gamma^T(\gamma\gamma^T + \lambda_W I)^{-1} \quad (27)$$

E. Applying to Phase Unwrapping

Both the dictionary D and the linear classifier W will be used at the online stage to determine the k-values of the code patterns. At the online stage, the input encoded fringe image will go through a modified MCA procedure so that the fringe pattern and code patterns are separated as indicated in Fig. 3. Patches are extracted from the code patterns and their Gabor features are obtained using the same approach as in (22). Given a patch feature vector $\tilde{\zeta}$, the following sparse coding method can be used to obtain its sparse code $\tilde{\gamma}$:

$$\arg\min_{\tilde{\gamma}} \|\tilde{\zeta} - D\tilde{\gamma}\|_2^2 \ s.t. \ \|\tilde{\gamma}\|_0 \le T_0 \quad (28)$$

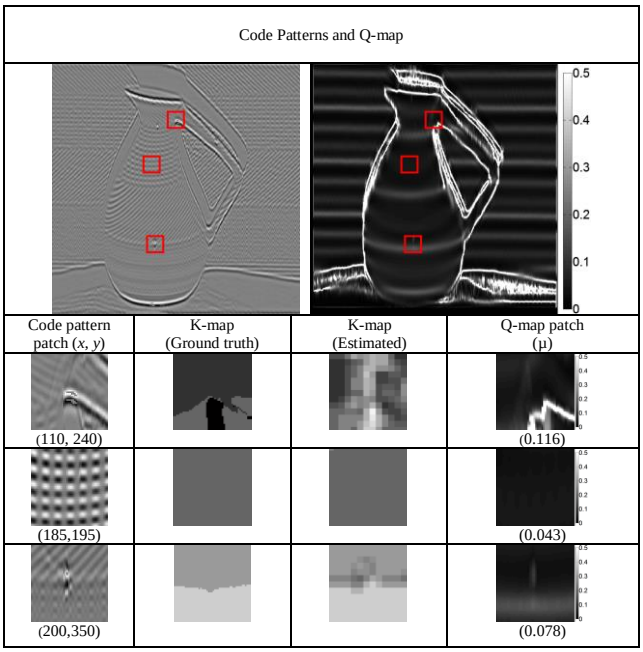


Fig. 5. Code patterns and Q-map at various locations

(28) can be solved by using the OMP [55]. The k-value encoded by the code patch can be obtained by,

$$\hat{k} = \max_i \left(W\tilde{\gamma} \right) \tag{29}$$

where $\max_i(.)$ returns the index i of the coefficient in the vector $W\tilde{\gamma}$ of which the value is the maximum.

The above sparse k-value classification algorithm can be applied to all code patches to obtain their k-value; however the process can be time consuming. We propose to apply it together with the traditional multilevel quality guided phase unwrapping algorithm [45] to achieve a higher computational efficiency. The proposed algorithm first randomly chooses a "good" point in the wrapped phase map $\hat{\varphi}$. To determine if the point is "good", we make use of the approach given in [45, 46] as below,

$$Q(i,j) = \max\left\{ \max\left\{ \left|\Delta^x_{i,j}\right|, \left|\Delta^x_{i-1,j}\right| \right\}, \max\left\{ \left|\Delta^y_{i,j}\right|, \left|\Delta^y_{i,j-1}\right| \right\} \right\} \tag{30}$$

where $\Delta^x_{i,j} = \Pi(\hat{\varphi}(i+1,j) - \hat{\varphi}(i,j))$ and $\Delta^y_{i,j} = \Pi(\hat{\varphi}(i,j+1) - \hat{\varphi}(i,j))$ are the wrapped phase differences in the horizontal and vertical directions; $\Pi(\cdot)$ is the phase wrapping operator defined in (3). In (30), $Q \in [0, 2\pi)$ is called the quality map (or Q-map), which gives the prediction of the quality of each unwrapped phase data. Note that we consider the quality is good if the Q value is low, and vice versa.

Traditionally, the Q-map is used to guide the phase unwrapping process. However, we found that for a particular position where the Q-map indicates the quality is low (e.g. the Q value has a high mean), the result of k-value classification at that position also will not be good. An example is shown in Fig. 5 which depicts the code patterns and Q-map obtained in an FPP experiment of an object. Three code pattern patches and Q-map patches are extracted. The sparse k-value classification algorithm as mentioned above is applied to each code pattern patch and the resulting K-maps are shown. For the Q-map patch centered at position (x, y), the mean $\mu = \overline{Q(x,y)}$ of the patches is also evaluated. In the second row, the code pattern patch and Q-map patch are near the boundary of the object. Hence some of the Q values (fourth column) are rather high which means that they are not suitable for phase unwrapping. It can be seen that the corresponding Q-map patch has a high mean value. One can also see that the K-map estimated in the corresponding location (third column) is quite far away from the ground truth (second column). The same can be seen in the results in the final row. The code pattern patch and Q-map patch are at the location where distortion is found due to the specular reflection to the object. It can be seen that the corresponding Q-map patch has a relatively higher mean value. And the K-map estimated in the corresponding location (third column) is also somewhat different from the ground truth (second column). On the contrary, the code pattern patch and Q-map patch in the third row are at the smooth region of the object. The mean of the Q-map patch is low and the K-map estimated is very close to the ground truth. The above observation is expected since for the code pattern patches located at the positions where there are abrupt changes in the fringe pattern, both the MCA and the classification algorithm which heavily rely on the second order statistics in the optimization process will have difficulty to obtain statistically stationary data. Hence the estimation is prone to error.

Consequently, the proposed algorithm first chooses a point (x, y) in the wrapped phase map $\hat{\varphi}$ such that the mean value of the Q-map patch at the same position, i.e. $\overline{Q(x,y)}$, is low. The sparse k-value classification algorithm is then applied to the code patterns to obtain the k-value at this position. Then starting from this position of the wrapped phase map, the traditional multilevel scanline phase unwrapping algorithm is used to unwrap the phase data. The whole procedure can be summarized as follows:

Algorithm II

Inputs: The wrapped phase map $\hat{\varphi}$, the quality map Q
Output: The unwrapped phase map φ

1. **Initialize** *minQThreshold* with a small value
2. **Repeat**
   (a) Randomly select a point (x,y) where $\overline{Q(x,y)} < minQThreshold$
   (b) If the true phase and k-value of a neighbor of (x,y) has been obtained, directly compute $\varphi(x,y)$ and its k-value based on those of its neighbor.
   (c) Otherwise, estimate the k-value using the proposed sparse k-value classification algorithm (i.e. (28) and (29)). Then based on the estimated k-value, compute $\varphi(x,y)$.
   (d) Starting from (x,y), scan all other pixels of which the Q value also smaller than *minQThreshold*. Repeat step (a), (b) and (c) for all these pixels.
   (e) If no more pixel whose Q value has a mean smaller than *minQThreshold*, increase the value of *minQThreshold* by a fixed amount.
3. **Until** no more pixel to be unwrapped

Similar to the traditional multi-level scanline phase unwrapping algorithm [45], **Algorithm II** is very simple that

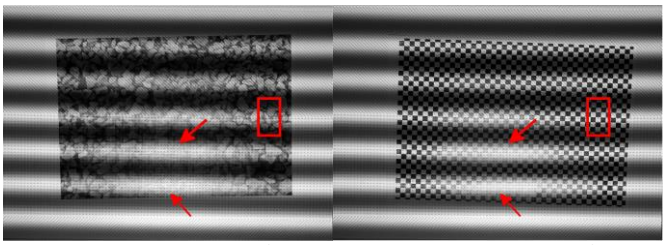


Fig. 6. The captured encoded fringe patterns with a stone texture (left) and a checkered texture (right). The highlights regions due to the global illumination are pointed out by arrows.

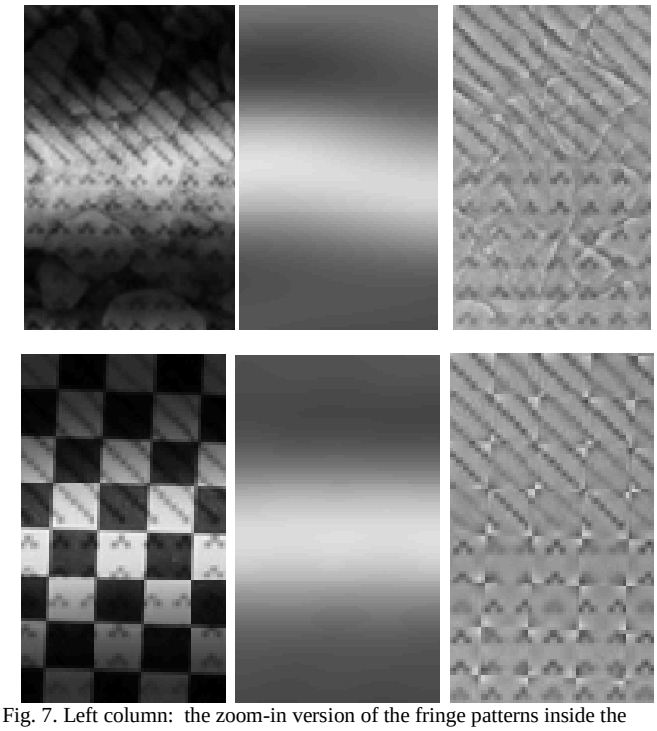


Fig. 7. Left column: the zoom-in version of the fringe patterns inside the rectangular boxes in Fig. 6. Middle and right columns: the result after separation using the modified MCA.

can be implemented with real time performance. In addition, it allows true phase estimation even when there are multiple disconnected regions in the wrapped phase map, which is common for fringe images of complex scene. In such situation, traditional algorithms such as [45] will fail since in these regions, their wrapped phase data cannot find any neighbors with known true phase or k-value hence unwrapping cannot be carried out. And since most traditional algorithms do not have the remedy such as Step 2c of **Algorithm II**, errors cannot be avoided. For **Algorithm II**, the k-value and hence the true phase in these regions can be obtained based on the sparse k-value classification algorithm. It allows the algorithm to be applicable for fringe images of complex scene. Some examples will be given in Section IV to illustrate the performance of the proposed algorithm.

IV. Experiments And Results

To verify the proposed algorithm, a series of FPP experiments were conducted using different real objects. To implement these experiments, an FPP hardware setup that contains a DLP projector and a digital SLR camera was adopted. The projector has a 2000:1 contrast ratio with light output of 3300 ANSI lumens and the camera has a 22.2 x 14.8mm CMOS sensor and a 17-50mm lens. Both devices were connected to a computer with a 3.4GHz CPU and 16GB RAM for image processing. They were placed at a distance of 700mm-1200mm from the object. The program code was written in MATLAB running on a personal computer at 3.4 GHz. The resolution of the testing fringe images is 728×640 pixels.

A. Quantitative Evaluation

The first experiment was to reconstruct the 3D model of a shiny flat board. Since the ground truth of the flat board can be easily measured, it allows us to have an objective comparison of the accuracy of different methods. For this experiment, the size of the board is 500mm×400mm. Different methods were tested including the conventional phase shifting profilometry (PSP) with the Goldstein phase unwrapping algorithm (PSP+Goldstein) [13, 14, 59, 60], the PSP method with speckle-embedded fringe patterns (PSP-Speckle) [33], and the proposed algorithm. All methods use three phase shifted fringe patterns to implement the PSP. While PSP+Goldstein is the conventional approach, PSP-Speckle was recently proposed and indeed similar to the proposal algorithm in the sense that it also embeds code patterns into the fringe patterns to carry the period order information. However, it does not use the MCA method to separate the code patterns and fringe patterns; and it also does not use the dictionary method to decode the code patterns. To understand the importance of the modified MCA in the proposed algorithm, we try to replace it by using the traditional method as in PSP-Speckle, i.e. simply summing up all frames to remove the fringe pattern and obtain the code

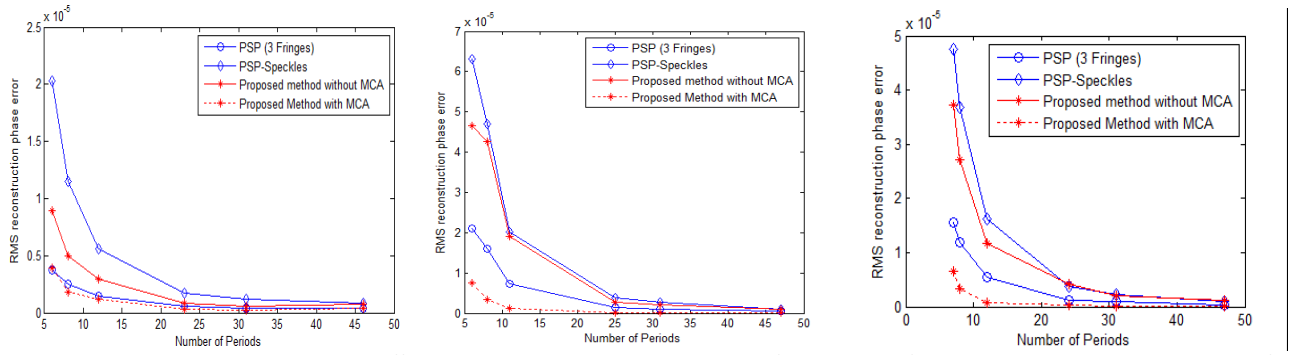
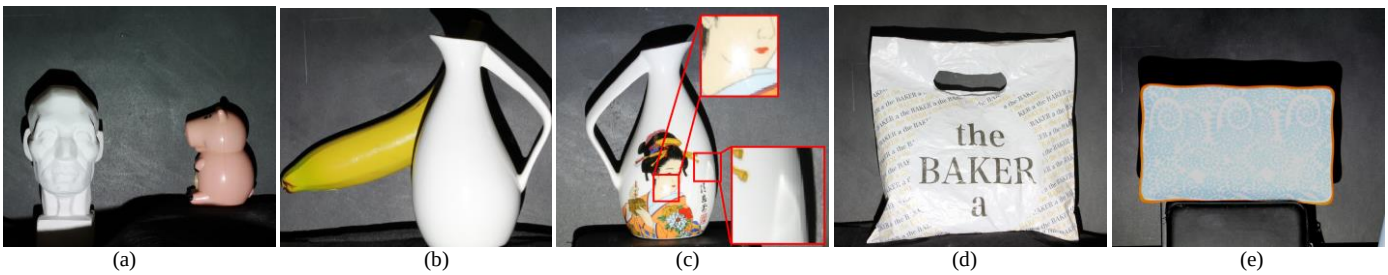


Fig. 8. Reconstruction error in RMS when using different methods with respect to the number of periods in the fringe pattern. The reconstruction errors of the flat board with shiny surface (left), checkered pattern (middle) and stone pattern (right).



(a) (b) (c) (d) (e)

Fig. 9. Objects that form complex scenes. (a) A head sculpture and a plastic toy; (b) a plastic banana occluded by a jar; (c) a jar with texture and highlights; (d) a plastic bag with regular textual pattern; and (e) a ceramic plate with complex texture cover the whole object.

patterns. The method is then compared with the proposed algorithm and other conventional approaches. In this experiment, the board was scanned for 50 times using those approaches. The total number of reconstructed points is 678,665 by ignoring a few pixels at the boundary which contain some artifacts. The root means square (RMS) of the reconstruction phase error is obtained by averaging the whole surface. The comparison result is shown in Fig. 8 (left). As it is seen in the figure, the PSP-Speckle method shows low accuracy particularly when the number of periods is small. It is because the embedded speckles introduce distortion to the original fringe patterns that affect the accuracy. Meanwhile the accuracy of the proposed algorithm without the modified MCA only shows a slightly better performance than the PSP-Speckle method but is worse than the conventional PSP. It shows that the embedded code patterns in the proposed algorithm also introduce error to the fringe patterns similar to the PSP-Speckle method. It can affect the final result even using the learned dictionary and linear classifier in the proposed method. However once the modified MCA is applied, the phase error of the proposed algorithm is similar or even slightly better for all period numbers as illustrated in Fig. 8 (left). It shows that the embedded coded patterns in the proposed algorithm can be separated successfully using the modified MCA.

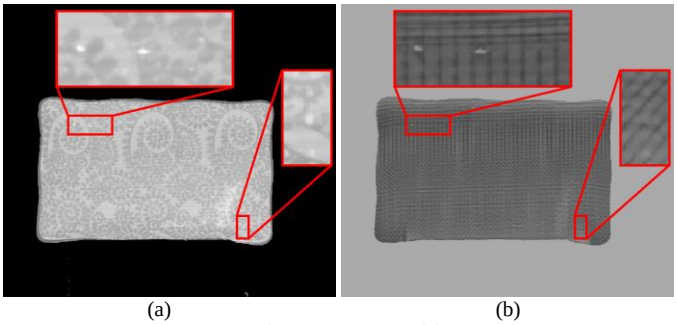
To further evaluate the robustness of the proposed modified MCA method, a checkered pattern paper and a stone pattern paper are attached to the flat board as shown in Fig. 6 and Fig. 7 left column). Besides, some highlight regions as indicated in Fig. 6 are introduced to the fringe patterns by adding another indirect light source. They introduce further difficulty in identifying the code patterns in the fringe images. Fig. 7 (middle and right column) shows the results of using the modified MCA. It is seen that the code patterns and the fringe patterns are successfully separated. We then repeated the experiments mentioned in the last paragraph but replacing the shiny flat board with the checkered and stone pattern boards. Then the results of using different PSP methods are shown in Fig. 8 (middle and right). They show that the proposed method with the modified MCA outperforms all competing approaches.

In this experiment, we have shown that the proposed algorithm can give the best performance when working on fringe images containing simple or complex texture patterns. We shall show in the next experiment that when working on complex scenes, the proposed method can also tackle the ambiguity problems while the other methods fail to deliver.

B. Qualitative Evaluation

For qualitative evaluation, objects that form a complex scene as depicted in Fig. 9 are considered. In these scenes, highlight regions, sudden intensity jump, occlusion, or bias due to the object's texture can be found. To be specific, the first scene consists of a head sculpture and a color plastic toy. The second scene consists of a plastic banana occluded by a jar. The third scene consists of a jar with texture and highlights. The fourth consists of a plastic bag with regular textural pattern. And the last scene consists of a ceramic plate with complex texture covering the whole object.

Similar to the first experiment, we compare the proposed algorithm with the conventional method (PSP+Goldstein) [13, 14, 59, 60] and PSP-Speckle [33]. The result of the comparison is illustrated in Fig. 11. In the figure, the ground truth is generated by scanning the scenes using 30 fringe pattern images with the period order determined manually. As shown in the figure, the resulting 3D models generated by the PSP+Goldstein method are erroneous for scene 1 and 2. When using the PSP+Goldstein method, it is traditionally assumed that a reference point (usually at the center of the image) is known so that a complete depth map can be grown from this reference point using the Goldstein method. Such assumption is valid only when there is only one object located at that reference point. For scene 1, both objects are not located at the center. And since there is no period order information provided in the fringe images, there is no way to estimate the absolute depth of the objects. For scene 2, it is noted that the plastic banana is located at the center of the image hence its absolute depth can be obtained from the reference. However, due to the shadow of the jar, the two objects are disconnected as can be seen in the second image in Fig. 9. Thus the depth estimation of the jar cannot be conducted through the reference at the plastic banana, and leads to the erroneous result. In practice, additional



(a) (b)

Fig. 10. (a) The speckle map of the object in the fifth scene; and (b) the code pattern of the same object resulted from the proposed modified MCA.

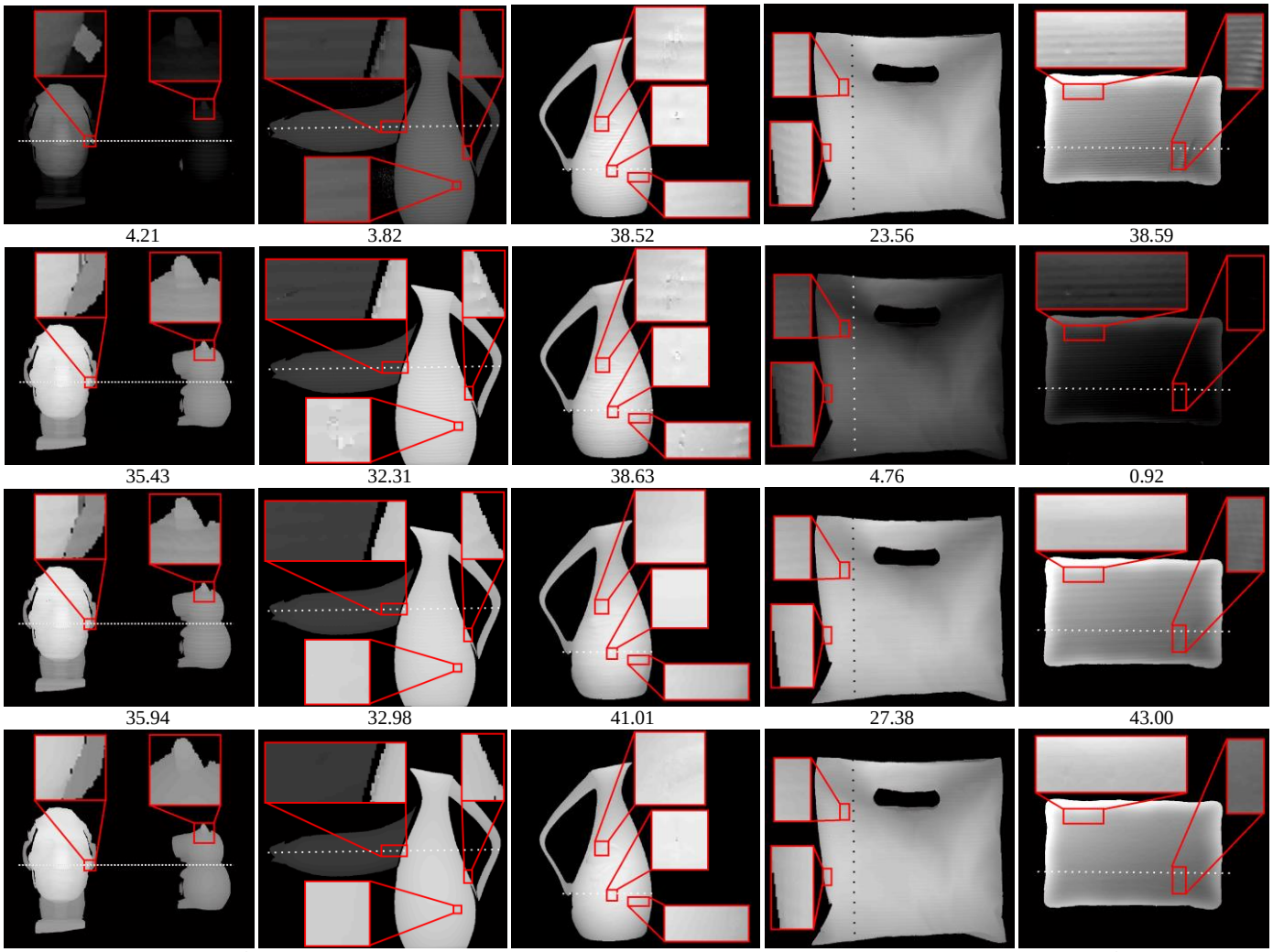


Fig. 11. 3D reconstruction of various complex scenes with multiple objects, occlusion, highlight regions, and textures. (First row) Results of the conventional PSP+Goldstein method; (second row) results of the PSP-Speckle method; (third row) results of the proposed method; and (fourth row) the ground truth. The numbers under the images of the first 3 rows are the SNR as compared with the ground truth.

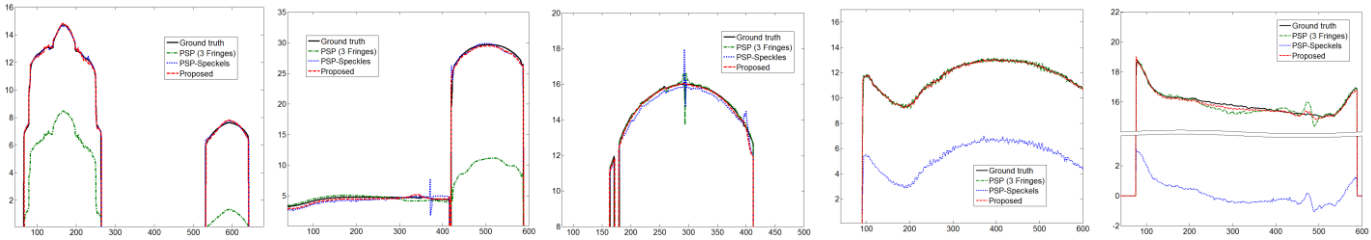


Fig. 12. A magnitude plot along the dash lines in Fig. 11.

hardware setup is required to give the reference for each disconnected region (such as using some tiny markers as in [13]). Without those reference points, unsatisfactory results will be obtained as in the first two images of the first row in Fig. 11.

Unlike the conventional PSP methods, both the PSP-Speckle method and the proposed algorithm provide the period order information and are able to obtain the true phase correctly as shown in Fig. 11 (the second and the third row). However, the PSP-Speckle method can give erroneous results when the scene is complex. As illustrated in Fig. 11 (the third to the fifth scene), distortion of different extents is found in the reconstructed 3D models using the PSP-Speckle method. Since the PSP-Speckle method does not have measure to remove the speckles from the fringe images, they can appear as noises in the reconstructed 3D model. Such noises may not only affect the smoothness of reconstructed 3D model surface, sometimes they can lead to erroneous reconstruction particularly in the regions where the fringe images have other artifacts, such as highlights (see the results of the third scene in Fig. 11 and its vertical gradient in Fig. 13).

Besides, incorrect period order information is generated when the object itself contains large textural pattern (such as the fourth and the fifth scene). They are confused with the speckles and lead to the erroneous k-value estimation as shown in Fig. 11 (second row, fourth and fifth columns) and Fig. 12

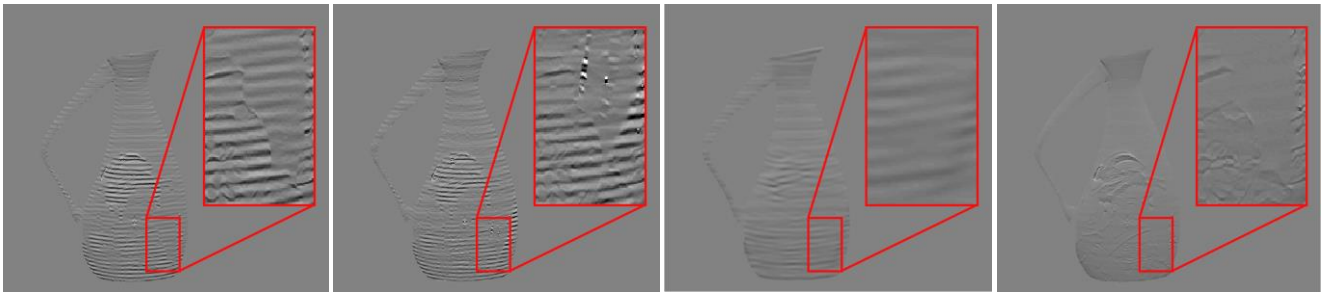
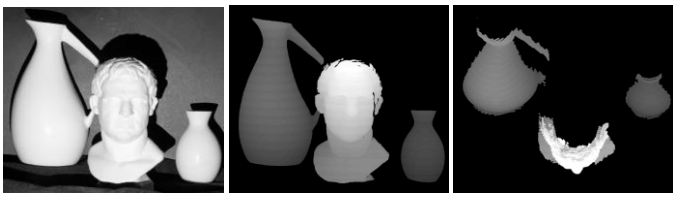


Fig. 13. The vertical gradient of Fig. 11 (column 3); (first column) PSP+Goldstein method; (second column) the PSP-Speckle method; (third column) the proposed method; (fourth column) the ground truth.

(blue dash dot line). To further illustrate this, we show in Fig. 10a the speckle map of the fifth scene after the fringe pattern is removed. It can be seen that the texture remained in the speckle map lets the speckles (i.e. the bright dots) be imperceptible. Most period order estimation procedures will fail in this situation. The same problem does not exist for the proposed algorithm since, by using the modified MCA, the code pattern can be effectively separated from the textural pattern of the object, as they have different morphological structures. The performance can be seen in Fig. 10b. Together with the dictionary learning method which significantly enhance the classification power, the proposed algorithm (the third row) can accurately recover the 3D model of objects and the results are close to the ground truth (as can be seen in the images in the third row of Fig. 11). More experimental results are shown in Fig. 14. They show the robustness of the proposed algorithm.

V. Conclusions

In this paper, we have presented a robust algorithm for fringe projection profilometry (FPP). The proposed approach solves the ambiguity problem by embedding codes patterns which carry the period order information to the fringe patterns. Unlike the traditional approaches, the proposed algorithm does not require special hardware setup or projecting additional fringe patterns. The key ingredients of the proposed algorithm include a modified morphological component analysis (MCA) for separating the code patterns and the fringe patterns; and a sparse dictionary learning and classification procedure for determining the k-values from the extracted code patterns. They are integrated to a multilevel quality guide phase unwrapping procedure to allow the phase unwrapping to be carried out in fringe images of complex scenes. Experimental results have demonstrated the superiority of the proposed algorithm over the traditional approaches in terms of robustness and accuracy. The proposed algorithm only requires approximately 300ms when running on a normal personal computer in the Matlab environment. We believe the computational efficiency can be further increased when working in industrial-grade computing systems.



(a) (b) (c)

Fig. 14. (a) A scene of complex objects. (b) Result of 3D reconstruction using the proposed algorithm. (c) The top view of the 3D reconstruction result.

Acknowledgement

The authors would like to thank Prof. Ivan Selesnick for making the TQWT codes available online, the associate editor, and two anonymous reviewers for providing us with valuable comments and insightful suggestions which have improved this manuscript.

Appendix A: Proof of Algorithm I

Consider the minimization problem as stated in (15),

$$\arg\min_{\alpha_1,\alpha_2} \left\|Y - \mathcal{K}(\Phi_1\alpha_1 + \Phi_2\alpha_2)\right\|_2^2 + \lambda_1\left\|\alpha_1\right\|_1 + \lambda_2\left\|\alpha_2\right\|_1$$

Applying the variable splitting method, the above problem can be reformulated as,

$$\arg\min_{v,\,\alpha} f_1(v) + f_2(\alpha) \quad s.t.\ v = \alpha$$

where

$$f_1(v) = \left\|Y - \mathcal{K}\Phi v\right\|_2^2, \quad f_2(\alpha) = \lambda_1\left\|\alpha_1\right\|_1 + \lambda_2\left\|\alpha_2\right\|_1$$

$$\Phi = [\Phi_1 \quad \Phi_2],\ v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},\ \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},\ \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

Using the augmented Lagrangian method (ALM), the solution of the above problem can be written as an iterative algorithm as follows:

Initialize $k \leftarrow 0$, $\mu > 0$, δ^0, and v^0

Repeat

$$\alpha^{k+1} \leftarrow \arg\min_{\alpha} \lambda_1\left\|\alpha_1\right\|_1 + \lambda_2\left\|\alpha_2\right\|_1 + \mu/2\left\|\alpha - v^k - \delta^k\right\|_2^2$$

$$v^{k+1} \leftarrow \arg\min_{v} \left\|Y - \mathcal{K}\Phi v\right\|_2^2 + \mu/2\left\|\alpha^{k+1} - v - \delta^k\right\|_2^2$$

$$\delta^{k+1} \leftarrow \delta^k - (\alpha^{k+1} - v^{k+1})$$

$$k \leftarrow k+1$$

Until meet the stopping criteria

In the above algorithm, line 3 can be solved using the soft-thresholding while line 4 is a least squares problem which

can be solved in its explicit form. By substituting $z = \alpha - \delta$, the above algorithm can be rewritten as,

Initialize $k \leftarrow 0$, $\mu > 0$, δ^0, and v^0
Repeat
$z^{k+1} = S_{\lambda/\mu}(v^k + \delta^k) - \delta^{k+1}$
$v^{k+1} = (\Phi^T \mathcal{K}^T \mathcal{K} \Phi + \mu I)^{-1} (\Phi^T \mathcal{K}^T Y + \mu z^{k+1})$
$\delta^{k+1} \leftarrow v^{k+1} - z^{k+1}$
$k \leftarrow k+1$
Until meet the stopping criteria

where $S_x(y)$ is the soft-thresholding of y with a threshold x. Now let us focus on the first term in line 4 of the above algorithm. Since we adopt a tunable Q-factor wavelet transform (TQWT) which has a tight frame ($\Phi_i \Phi_i^T = I$) for both Φ_1, and Φ_2, we have:

$$\Phi \Phi^T = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix} \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} = \Phi_1 \Phi_1^T + \Phi_2 \Phi_2^T = 2I$$

Thus applying the Sherman-Morrison-Woodbury matrix inversion lemma to the first term in line 4, we have

$$(\Phi^T \mathcal{K}^T \mathcal{K} \Phi + \mu I)^{-1} = \frac{1}{\mu} \left(I - \Phi^T \mathcal{K}^T (\mu I + 2 \mathcal{K} \mathcal{K}^T)^{-1} \mathcal{K} \Phi \right)$$

Also, since the TQWT is adopted, Φ_1 and Φ_2 can be implemented efficiently in the Fourier domain. Assume that $\mathcal{K}$ can be approximated as a circular convolution operator. Then $\mathcal{K}$, Φ_1, and Φ_2 can be factorized as,

$$\mathcal{K} = U^T H U, \Phi_i = U^T C_i U, \Phi_i^T = U^T C_i^T U, \forall i \in 1,2$$

where U represents the discrete Fourier transform (DFT), $U^T = U^{-1}$ is its inverse; H and C_i are some diagonal matrices. Therefore the above terms can be written as,

$$\begin{aligned} &\Phi^T \mathcal{K}^T (\mu I + p \mathcal{K} \mathcal{K}^T)^{-1} \mathcal{K} \Phi \\ &\quad = U^T C_i^T H^* U (\mu U^T U + 2 U^T H H^* U)^{-1} U^T H C_i U \\ &\quad = U^T \underbrace{C_i^T H^* \left(\mu + 2 |H|^2 \right)^{-1} H C_i}_{F} U \end{aligned}$$

where H^* is the complex conjugate of H; $|H|^2$ is the squared absolute values of the entries of the diagonal matrix H; and $C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$. By substituting the above term to the algorithm, **Algorithm I** is obtained as follows:

Initialize $k \leftarrow 0$, $\mu > 0$, δ^0, and v^0
Repeat
$z^{k+1} = \mathcal{S}_{\lambda/\mu}(v^k + \delta^k) - \delta^k$
$v^{k+1} = \frac{1}{\mu} \left(I - U^T F U \right) \left(\Phi^T \mathcal{K}^T Y + \mu z^{k+1} \right)$
$\delta^{k+1} = v^{k+1} - z^{k+1}$
$k \leftarrow k+1$
Until meet the stopping criteria. (Q.E.D.)

REFERENCES

[1] R. R. Garcia, and A. Zakhor, "Consistent Stereo-Assisted Absolute Phase Unwrapping Methods for Structured Light Systems," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 5, pp. 411-424, 2012.

[2] F. Sadlo, T. Weyrich, R. Peikert, and M. Gross, "A practical structured light acquisition system for point-based geometry and texture," in Eurographics/IEEE VGTC Symp. Proc. Point-Based Graphics, 2005, pp. 89-145.

[3] B. Shi, B. Zhang, F. Liu, J. Luo, and J. Bai, "360° Fourier Transform Profilometry in Surface Reconstruction for Fluorescence Molecular Tomography," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 3, pp. 681-689, 2013.

[4] S. Zhang, "High-resolution, High-speed 3-D Dynamically Deformable Shape Measurement Using Digital Fringe Projection Techniques," *Advances in Measurement Systems*, M. K. Sharma, ed., pp. 29-50: InTech, 2010.

[5] Z. Song, R. Chung, and X.-T. Zhang, "An Accurate and Robust Strip-Edge-Based Structured Light Means for Shiny Surface Micromeasurement in 3-D," *IEEE Trans. Ind. Electron.*, vol. 60, no. 3, pp. 1023-1032, 2013.

[6] T.-W. Hui, and G. K.-H. Pang, "3-D Measurement of Solder Paste Using Two-Step Phase Shift Profilometry," *IEEE Trans. Electron. Packag. Manuf.*, vol. 31, no. 4, pp. 306-315, 2008.

[7] Y. Wang, K. Liu, Q. Hao, D. L. Lau, and L. G. Hassebrook, "Period Coded Phase Shifting Strategy for Real-time 3-D Structured Light Illumination," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3001-3013, 2011.

[8] Y. Wang, K. Liu, Q. Hao, X. Wang, D. L. Lau, and L. G. Hassebrook, "Robust Active Stereo Vision Using Kullback-Leibler Divergence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 548-563, 2012.

[9] M. Gupta, Q. Yin, and S. K. Nayar, "Structured Light in Sunlight," in Proc. IEEE Int. Conf. Comput. Vision (ICCV), 2013.

[10] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. Narasimhan, "A Practical Approach to 3D Scanning in the Presence of Interreflections, Subsurface Scattering and Defocus," *Int. J. Comput. Vision*, vol. 102, no. 1-3, pp. 33-55, 2013.

[11] T. Chen, H. P. A. Lensch, C. Fuchs, H.-p. Seidel, and M. Informatik, "Polarization and PhaseShifting for 3D Scanning of Translucent Objects," in Proc. IEEE Conf. Comput. Vision and Pattern Recognition (CVPR), 2007, pp. 1-8.

[12] C. Pengyu, X. Zhiwei, Z. Yueyi, Z. Shenghui, and W. Feng, "Dense single-shot 3D scanning via stereoscopic fringe analysis," in Proc. IEEE Int. Conf. on Image Process. (ICIP) 2013, pp. 3627-3631.

[13] S. Zhang, and S.-T. Yau, "High-resolution, real-time 3D absolute coordinate measurement based on a phase-shifting method," *Opt. Express*, vol. 14, no. 7, pp. 2644-2649, 2006.

[14] P. Huang, "A Fast Three-Step Phase Shifting Algorithm for Real-Time Three-Dimensional Shape Measurement," *Frontiers in Optics 2007/Laser Science XXIII/Organic Materials and Devices for Displays and Energy Conversion*, 2007.

[15] M. Takeda, and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," *Appl. Opt.*, vol. 22, no. 24, pp. 3977-3982, 1983.

[16] A. Dursun, S. Özder, and F. N. Ecevit, "Continuous wavelet transform analysis of projected fringe patterns," *Measurement Science and Technology*, vol. 15, no. 9, pp. 1768, 2004.

[17] T.-C. Hsung, D. Pak-Kong Lun, and W. W. L. Ng, "Efficient fringe image enhancement based on dual-tree complex wavelet transform," *Appl. Opt.*, vol. 50, no. 21, pp. 3973-3986, 2011.

[18] T.-C. Hsung, and D. P.-K. Lun, "On optical phase shift profilometry based on dual tree complex wavelet transform," in Proc. IEEE Int. Conf. on Image Process. (ICIP) 2010, pp. 337-340.

[19] B. Budianto, and D. P. K. Lun, "Inpainting For Fringe Projection Profilometry Based on Geometrically Guided Iterative Regularization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5531-5542, 2015.

[20] B. Budianto, P. K. D. Lun, and T.-C. Hsung, "Marker encoded fringe projection profilometry for efficient 3D model acquisition," *Appl. Opt.*, vol. 53, no. 31, pp. 7442-7453, 2014.

[21] Q. Kemao, "Windowed Fourier Transform for Fringe Pattern Analysis," *Appl. Opt.*, vol. 43, no. 13, pp. 2695-2702, 2004.

[22] Q. Kemao, H. Wang, and W. Gao, "Windowed Fourier transform for fringe pattern analysis: theoretical analyses," *Appl. Opt.*, vol. 47, no. 29, pp. 5408-5419, 2008.
[23] P. Cong, Z. Xiong, Y. Zhang, S. Zhao, and F. Wu, "Accurate Dynamic 3D Sensing With Fourier-Assisted Phase Shifting," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 3, pp. 396-408, 2015.
[24] K. Itoh, "Analysis of the phase unwrapping algorithm," *Appl. Opt.*, vol. 21, no. 14, pp. 2470, 1982.
[25] J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition*, vol. 37, no. 4, pp. 827-849, 2004.
[26] Y. Wang, and S. Zhang, "Novel phase-coding method for absolute phase retrieval," *Opt. Lett.*, vol. 37, no. 11, pp. 2067-2069, 2012.
[27] J. M. Huntley, and H. Saldner, "Temporal phase-unwrapping algorithm for automated interferogram analysis," *Appl. Opt.*, vol. 32, no. 17, pp. 3047-52, 1993.
[28] J. Gass, A. Dakoff, and M. K. Kim, "Phase imaging without 2π ambiguity by multiwavelength digital holography," *Opt. Lett.*, vol. 28, no. 13, pp. 1141-1143, 2003.
[29] W.-H. Su, and H. Liu, "Calibration-based two-frequency projected fringe profilometry: a robust, accurate, and single-shot measurement for objects with large depth discontinuities," *Opt. Express*, vol. 14, no. 20, pp. 9178-9187, 2006.
[30] K. Liu, Y. Wang, D. L. Lau, Q. Hao, and L. G. Hassebrook, "Dual-frequency pattern scheme for high-speed 3-D shape measurement," *Opt. Express*, vol. 18, no. 5, pp. 5229-5244, 2010.
[31] Y. Wang, S. Yang, and X. Gou, "Modified Fourier transform method for 3D profile measurement without phase unwrapping," *Opt. Lett.*, vol. 35, no. 5, pp. 790-792, 2010.
[32] Y. Zhang, Z. Xiong, Z. Yang, and F. Wu, "Real-Time Scalable Depth Sensing With Hybrid Structured Light Illumination," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 97-109, 2014.
[33] Y. Zhang, Z. Xiong, and F. Wu, "Unambiguous 3D measurement from speckle-embedded fringe," *Appl. Opt.*, vol. 52, no. 32, pp. 7797-7805, 2013.
[34] H. Cui, W. Liao, N. Dai, and X. Cheng, "A flexible phase-shifting method with absolute phase marker retrieval," *Measurement*, vol. 45, no. 1, pp. 101-108, 2012.
[35] S. Gai, and F. Da, "A novel phase-shifting method based on strip marker," *Optics and Lasers in Engineering*, vol. 48, no. 2, pp. 205-211, 2010.
[36] J. Mairal, F. Bach, and J. Ponce, "Task-Driven Dictionary Learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 791-804, 2012.
[37] Z. Jiang, Z. Lin, and L. S. Davis, "Label Consistent K-SVD: Learning a Discriminative Dictionary for Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2651-2664, 2013.
[38] Q. Zhang, and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in Proc. IEEE Conf. Comput. Vision and Pattern Recognition (CVPR), 2010, pp. 2691-2698.
[39] S. Yubao, L. Qingshan, T. Jinhui, and T. Dacheng, "Learning Discriminative Dictionary for Group Sparse Representation," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3816-3828, 2014.
[40] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Discriminative learned dictionaries for local image analysis," in Proc. IEEE Conf. Comput. Vision and Pattern Recognition (CVPR), 2008, pp. 1-8.
[41] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in Proc. IEEE Conf. Comput. Vision and Pattern Recognition (CVPR), 2010, pp. 3501-3508.
[42] J. Bobin, J. L. Starck, J. M. Fadili, Y. Moudden, and D. L. Donoho, "Morphological Component Analysis: An Adaptive Thresholding Strategy," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2675-2681, 2007.
[43] J. Bobin, J. L. Starck, J. Fadili, and Y. Moudden, "Sparsity and Morphological Diversity in Blind Source Separation," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2662-2674, 2007.
[44] L.-W. Kang, C.-W. Lin, and F. Yu-Hsiang, "Automatic Single-Image-Based Rain Streaks Removal via Image Decomposition," *IEEE Trans. Image Process.* , vol. 21, no. 4, pp. 1742-1755, 2012.
[45] S. Zhang, X. Li, and S.-T. Yau, "Multilevel quality-guided phase unwrapping algorithm for real-time three-dimensional shape reconstruction.," *Appl. Opt.*, vol. 46, no. 1, pp. 50-7, 2007.
[46] D. C. G. a. M. D. Pritt, *Two-dimensional phase unwrapping : theory, algorithms, and software*: John Wiley & Sons, 1998.
[47] I. W. Selesnick, "Resonance-based signal decomposition: A new sparsity-enabled signal analysis method," *Signal Processing*, vol. 91, no. 12, pp. 2793-2809, 2011.
[48] İ. Bayram, and I. W. Selesnick, "A Dual-Tree Rational-Dilation Complex Wavelet Transform," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6251-6256, 2011.
[49] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast Image Recovery Using Variable Splitting and Constrained Optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345-2356, 2010.
[50] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681-695, 2011.
[51] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311-4322, 2006.
[52] S. Chen, D. Donoho, and M. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM J. on Sci. Comput.*, vol. 20, no. 1, pp. 33-61, 1998.
[53] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, no. 3, pp. 273-282, 2011.
[54] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online Dictionary Learning for Sparse Coding," in Proc. Int. Conf. Machine Learning (ICML), New York, NY, USA, 2009, pp. 689-696.
[55] S. G. Mallat, and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397-3415, 1993.
[56] D.-S. Pham, and S. Venkatesh, "Joint learning and dictionary construction for pattern recognition," in Proc. IEEE Conf. Comput. Vision and Pattern Recognition (CVPR), 2008, pp. 1-8.
[57] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse Representation Based Fisher Discrimination Dictionary Learning for Image Classification," *Int. J. Comput. Vision*, vol. 109, no. 3, pp. 209-232, 2014.
[58] I. W. Selesnick, "Wavelet Transform With Tunable Q-Factor," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3560-3575, 2011.
[59] D. Malacara, *Optical Shop Testing (Wiley Series in Pure and Applied Optics)*: Wiley-Interscience, 2007.
[60] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, vol. 23, no. 4, pp. 713-720, 1988.

Budianto (S'08) received his B.Eng. degree from Sepuluh Nopember Institute of Technology, Indonesia and M.Eng. in Information and Communication Technology from the University of Western Australia, Australia in 2003 and 2009 respectively. He is currently pursuing Ph.D. degree at the Department of Electronic and Information Engineering, the Hong Kong Polytechnic University. His research interests include depth sensing and 3-D imaging.

Daniel, P.K. Lun received his B.Sc.(Hons.) degree from the University of Essex, U.K., and Ph.D degree from the Hong Kong Polytechnic University in 1988 and 1991, respectively. He is now an Associate Professor and Associate Head of the Department of Electronic and Information Engineering of the Hong Kong Polytechnic University. Dr Lun is active in research activities. He has published more than 125 international journals and conference papers. His research interest includes computational imaging, speech and image enhancement,

wavelet theories and applications. He was the Chairman of the IEEE Hong Kong Chapter of Signal Processing in 1999-00. He was the Finance Chair of 2003 IEEE International Conference on Acoustics, Speech and Signal Processing, and 2010 International Conference on Image Processing (ICIP2010). He was also the General Co-Chair and Technical Co-Chair of a number of international conferences. Dr Lun is a member of the DSP and VSPC Technical Committee of IEEE Circuits and Systems Society. He is a Chartered Engineer, a corporate member of IET, HKIE and a senior member of IEEE.