IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 20, 2021, accepted February 15, 2021, date of publication February 23, 2021, date of current version March 4, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3061415

Improving the Performance of 3D Shape Measurement of Moving Objects by Fringe Projection and Data Fusion

CHENGPU DUAN[1], JUN TONG[1], LEI LU[2], JIANGTAO XI[1], (Senior Member, IEEE), YANGUANG YU[1], (Senior Member, IEEE), AND QINGHUA GUO[1]

[1]School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia
[2]College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China

Corresponding author: Jun Tong (jtong@uow.edu.au)

The work of Lei Lu was supported by the Key Research Fund in Science and Technology of Henan Province, China, under Grant 202102210314.

ABSTRACT Fringe projection profilometry (FPP) has been widely studied for measurements of three-dimensional (3D) shapes. However, the measurement of moving objects is still a challenging problem associated with commonly used methods. Phase-shifting profilometry (PSP) suffers from motion-induced errors when applied to moving objects, while Fourier transform profilometry (FTP) is sensitive to background light and noise. This paper proposes a novel data fusion approach to improve the measurement of moving objects by alleviating noise-induced errors. Firstly, a set of coarse measurements on the shape of the object are obtained at different time instants by FTP. Then the parameters associated with the motion of the object over these time instants are retrieved from these coarse estimates by applying the iterative closest point (ICP) method. Finally, the parameters are utilized to adaptively fuse the coarse measurements based on their quality. Simulations and experiments verify that the proposed approach can significantly improve the measurement accuracy.

INDEX TERMS Fourier transform profilometry, dynamic object measurement, motion parameter estimation.

I. INTRODUCTION

As an effective technology for 3D shape measurement, fringe projection profilometry (FPP) has attracted extensive research over the past two decades. Among various approaches for FPP, Fourier transform profilometry (FTP) and phase-shifting profilometry (PSP) are of the most commonly used [1]–[8]. Both methods face challenges for measuring objects in motion [5]–[8]. PSP utilizes a sequence of fringe patterns, which does not suffer from the influence of ambient light. However, PSP expects objects to be static during the projection and acquisition of multiple fringe patterns and is thus prone to motion errors. In contrast to PSP, FTP employs a single fringe pattern and thus is suitable for fast measurement of objects in motion, but it suffers from the accuracy degradation due to such factors as background light and noise in the captured fringe patterns. These challenges have been addressed from various perspectives.

For PSP, it is possible to accelerate the measurements using fast cameras and projectors or projector defocusing techniques [9], [10] so that the motion artifact can be alleviated. Alternatively, one can use colored fringe patterns and devices with multiple channels [11] or extra cameras [12] to capture simultaneously multiple patterns required by PSP. These schemes, however, may significantly increase the hardware costs. Zhang and Yau [13] alleviate the motion-induced errors by using PSP scheme that requires fewer fringe patterns, while [14], [15] use composite phase-shifting patterns to implement the PSP. Though alleviating the errors due to motion, these schemes are more sensitive to noise than traditional PSP, and motion-induced errors can still exist if the object moves fast. Besides, PSP schemes that integrate the motion models into the phase retrieval process have also attracted significant interests. Lu, *et al.* [16] propose a modified PSP approach for objects

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

with 2D (planar) motions. The motion parameters (i.e., the rotation matrix and translation vector) are estimated using markers attached, then utilized to compensate for the phase shifts caused by motion. Lu, *et al.* [17] extend the method to 3D motion that involves translation in the height direction (i.e., the normal direction of the reference plane) and uses the least squares method to retrieve the phase. Lu, *et al.* [18] also apply the scale-invariant feature transform (SIFT) algorithm to retrieve the feature points first to estimate the motion parameters such that the human intervention in [16] can be removed and automated measurements can be achieved. In general, the methods of [16]–[18] are restricted to objects with 2D motion or special cases of 3D motion.

One class of methods for improving the measurements of 3D shapes of moving objects is to enhance the FTP. The π phase-shifting method [19] uses phase-shifting sub-patterns and two line-scan cameras to improve the measurement. With this method, each point is measured twice during movement, and the shape reconstruction is improved by alleviating the influence of background light. This approach is only suitable for the case where the object moves at a constant speed in a fixed direction. Guo and Huang [20] propose a modified FTP approach to improve the accuracy by projecting an extra flat image. In [21], 2D Hanning windowing is applied to 2D FTP to improve the separation of the height information from noise, which helps the measurements of objects in motion and also objects with discontinuous surfaces. Similar approaches based on windowed Fourier transform and Gaussian filtering are also studied in [22]–[24]. However, all these modified FTP approaches are still limited in their performance for moving objects due to their single-shot nature of projection.

This paper proposes a novel approach to enhancing the performance of 3D shape measurement of moving objects by fusing multiple height maps retrieved using standard FTP systems. We consider rigid objects moving freely in 3D space. Initial coarse measurements are obtained by FTP, and robust principal component analysis (RPCA) is employed to alleviate the noise-induced errors. The results are then utilized to estimate the 3D motion parameters by applying the iterative closest point (ICP) method. Finally, we fuse the multiple height maps to obtain the final result using weights adapted to the quality of the captured fringe patterns.

The paper is organized as follows. The system model and problem formulation are introduced in Section II. The proposed method is presented in Section III, and the simulation and experimental results are discussed in Section IV. Finally, conclusions are drawn in Section V.

II. SYSTEM MODEL AND PROBLEM STATEMENT

Consider an FTP system consisting of a projector projecting fringe patterns and a camera capturing the patterns deformed by the object, as illustrated in Fig. 1. The height information of the object can be retrieved from the phase difference between the reference fringe patterns and deformed fringe patterns.

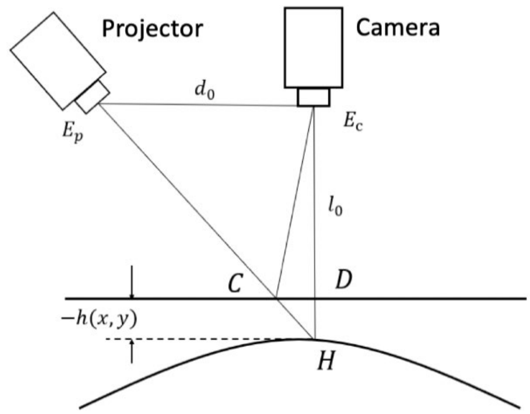

FIGURE 1. The FTP system.

FTP retrieves the object-induced phase difference from a deformed fringe pattern, which can be modeled as

$$d(x, y) = r(x, y) \sum_{n=-\infty}^{\infty} A_n \exp\left(i\left(2\pi n f_0 x + n\varphi(x, y)\right)\right), \tag{1}$$

and the reference fringe pattern, which can be modeled as

$$s(x, y) = r_0(x, y) \sum_{n=-\infty}^{\infty} A_n \exp\left(i\left(2\pi n f_0 x + n\varphi_0(x, y)\right)\right), \tag{2}$$

where $r(x, y)$ is the nonuniform reflectivity of the object surface, $r_0(x, y)$ is the reflectivity on the reference plane, A_n is a Fourier series coefficient, and f_0 is the fundamental frequency of the projected fringe pattern. The object-induced difference between the phases $\varphi(x, y)$ and $\varphi_0(x, y)$, which carries the height information of the object relative to the reference plane, can be obtained as:

$$\Phi^h(x, y) = \mathrm{Im}\{\log[\hat{d}(x, y) \times \hat{s}^*(x, y)]\}, \tag{3}$$

where $*$ denotes conjugate. In (3), $\hat{d}(x, y)$ and $\hat{s}(x, y)$ are the components of the patterns at the fundamental frequency, which are obtained by applying bandpass-filtering to $d(x, y)$ and $s(x, y)$ [25].

After the object-induced phase difference is retrieved by FTP and unwrapped into the absolute object phase value $\hat{\Phi}^h(x, y)$, the object height is computed as

$$\hat{h}(x, y) = \frac{l_0 \hat{\Phi}^h(x, y)}{\hat{\Phi}^h(x, y) - 2\pi f_0 d_0}, \tag{4}$$

where l_0 denotes the distance between the lens of the camera and the reference plane and d_0 is the distance between the lens of the projector and that of the camera, as shown in Fig.1.

Similarly to [16], now let us consider the case that the surface of a rigid object is probed by fringe patterns at K discrete time instants, during which the object changed its position due to its movement. Consider a point P on the object surface, whose coordinates (x_k^P, y_k^P, h_k^P) at time instant k can

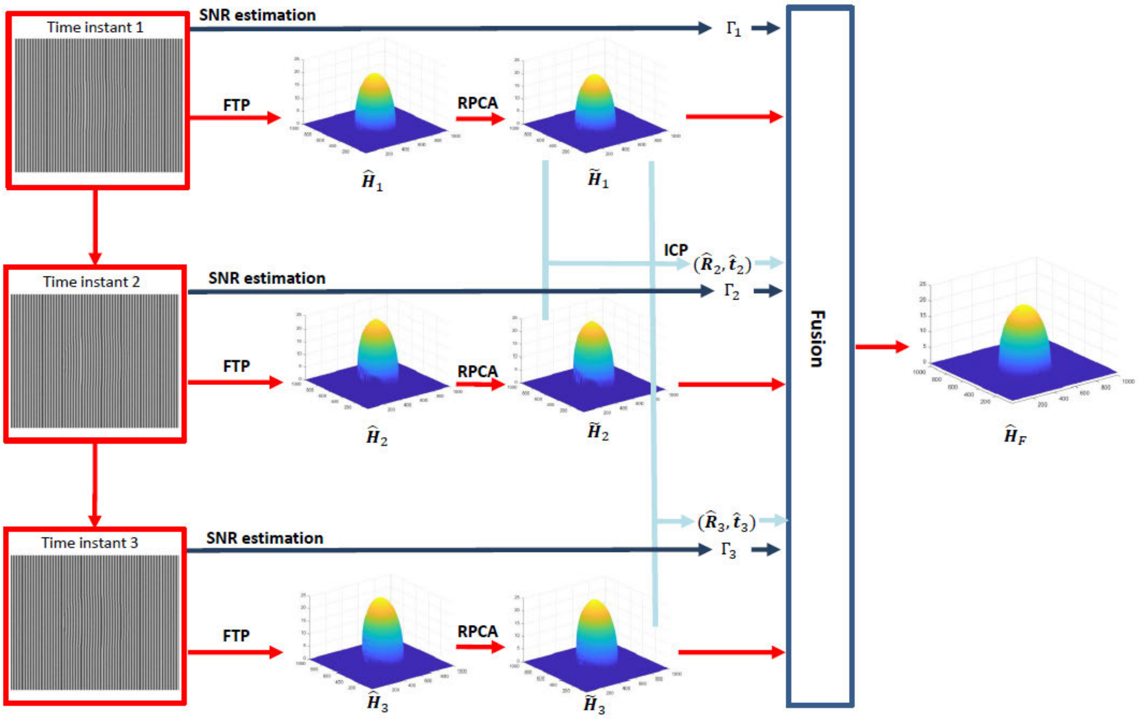

FIGURE 2. Workflow of the proposed approach when height maps acquired at $K = 3$ time instants are fused.

be described by

$$\begin{bmatrix} x_1^P \\ y_1^P \\ h_1^P \end{bmatrix} = \boldsymbol{R}_k \begin{bmatrix} x_k^P \\ y_k^P \\ h_k^P \end{bmatrix} + \boldsymbol{t}_k, \begin{bmatrix} x_k^P \\ y_k^P \\ h_k^P \end{bmatrix} = \bar{R}_k \begin{bmatrix} x_1^P \\ y_1^P \\ h_1^P \end{bmatrix} + \bar{\boldsymbol{t}}_k, \quad k = 1, 2, \cdots, K, \tag{5}$$

where $\boldsymbol{R}_k$ and $\boldsymbol{t}_k$ denote the rotation matrix and translation vector, respectively, and

$$\bar{\boldsymbol{R}}_k = \boldsymbol{R}_k^{-1}, \bar{t}_k = -\boldsymbol{R}_k^{-1}\boldsymbol{t}_k. \tag{6}$$

As mentioned above, FTP requires the projection of only one fringe pattern for each measurement and the performance is insensitive to the motion, and hence it can be applied to each time instant k to obtain a height map. However, the performance of FTP is sensitive to the strong noise from background light and the discontinuity of the object surface which can lead to spectrum aliasing. Various techniques such as [20]–[24] may be used to improve the performance of FTP. In general, the measurement accuracy of FTP is still limited by noise due to its single-shot nature. This paper aims to alleviate the noise-induced errors for FTP in measuring moving objects.

III. PROPOSED METHOD

In this section, we propose a novel method to improve the performance of the FTP approach for measuring moving objects by fusing multiple coarse measurements of the height map acquired at different time instants. The overall workflow is illustrated in Fig. 2, assuming that FTP measurements at $K = 3$ time instants are fused as an example. Meanwhile, the method can be easily extended to other values of K and to the cases where multiple measurements are performed simultaneously using multiple cameras. At each time instant k, one fringe pattern is projected and a height map $\hat{\boldsymbol{H}}_k$ is retrieved using FTP and then denoised by using robust principal component analysis (RPCA). The resulting height maps $\{\tilde{\boldsymbol{H}}_k\}$ are first used to produce estimates $(\hat{\boldsymbol{R}}_k, \hat{\boldsymbol{t}}_k)$ of the 3D motion parameters by applying the ICP algorithm, and then fused with adaptive weights computed from the signal-to-noise ratios (SNR) $\{\Gamma_k\}$ estimated for the captured fringe patterns to optimize the overall accuracy.

A. RETRIEVAL OF INITIAL HEIGHT MAPS WITH DENOISING

A standard FTP process is applied to retrieve an object height map at each time instant. Since a single pattern is used for acquiring each height map, there can be significant errors in the retrieved height maps due to noise, etc. Such errors can lead to poor results for further processing.

Denote by matrix $\hat{\boldsymbol{H}}_k = \{\hat{h}_k(x, y)\}$ the height map obtained using FTP at time instant k. We can model $\hat{\boldsymbol{H}}_k$ as the summation of the true object shape matrix $\boldsymbol{H}_k$, an impulsive error matrix $\boldsymbol{S}_k$, and a random, unstructured error matrix $\boldsymbol{W}_k$, i.e.,

$$\hat{\boldsymbol{H}}_k = \boldsymbol{H}_k + \boldsymbol{S}_k + \boldsymbol{W}_k, \quad k = 1, 2, \cdots, K. \tag{7}$$

For many applications, the object to be measured have a continuous surface shape, and in this case, it is reasonable to assume that $\boldsymbol{H}_k$ is a low-rank matrix, i.e., it can be represented in a low-dimensional subspace using a small number of singular vectors. Meanwhile, the impulsive error $\boldsymbol{S}_k$ is sparse, i.e., only a small subset of its entries is non-zero with significant magnitudes. The above distinct features of $\boldsymbol{H}_k$ and $\boldsymbol{S}_k$ can be applied to recover them from $\hat{\boldsymbol{H}}_k$ with high accuracy. Intuitively, $\boldsymbol{S}_k$ can be estimated by thresholding the entries of $\hat{\boldsymbol{H}}_k$ while $\boldsymbol{H}_k$ can be retrieved by projecting $\hat{\boldsymbol{H}}_k$ onto the low-dimensional subspace spanned by its principal singular vectors. These processes, however, are subject to errors due to the mixing of $\boldsymbol{H}_k$ and $\boldsymbol{S}_k$ in $\hat{\boldsymbol{H}}_k$. In order to enhance accuracy, iterative algorithms can be applied to refine the estimates of $\boldsymbol{H}_k$ and $\boldsymbol{S}_k$. Various implementations of RPCA have been proposed in the past decade, e.g. [26], [27], which have been successfully applied to many areas such as computer vision.

The original RPCA can be formulated as a matrix decomposition problem, which incorporates rank and sparsity constraints. This problem is non-convex and NP-hard. In order to obtain tractable solutions, one can heuristically reformulate RPCA as the following principal component pursuit (PCP) problem [27]:

$$\min_{\boldsymbol{S}_k,\boldsymbol{H}_k} \gamma \|\boldsymbol{S}_k\|_{l_1} + \|\boldsymbol{H}_k\|_* \quad \text{s.t.} \left\|\boldsymbol{H}_k + \boldsymbol{S}_k - \hat{\boldsymbol{H}}_k\right\|^2 \leq \epsilon^2. \quad (8)$$

In the above, $\|\cdot\|_{l_1}$ is the l_1 norm, defined as the sum of the absolute values of the entries of a matrix, and $\|\cdot\|_*$ is the nuclear norm, defined as the sum of the singular values. The PCP reformulation is motivated by the fact that the l_1 norm and nuclear norm can act as surrogate functions of the sparsity level and rank, respectively. Therefore, minimizing these norms as in (8) can induce a sparse $\boldsymbol{S}_k$ and a low-rank $\boldsymbol{H}_k$, subject to a tolerance of residual errors. The rank, sparsity level, and level of residual errors of the solutions can be controlled by the parameters $\gamma > 0, \epsilon^2 > 0$. In particular, a smaller value of ϵ^2 requires the residual error to be smaller. The parameter γ controls the rank of the recovered $\boldsymbol{H}_k$. A smaller γ leads to a lower rank of $\boldsymbol{H}_k$, which may oversmooth the result and discard details of the shape information. A larger γ leads to a larger rank of $\boldsymbol{H}_k$, which tends to include more errors in the recovered $\boldsymbol{H}_k$. The choice of γ has been a topic intensively studied in the literature. However, there is no universally optimal solution available. In this paper, we resort to numerical studies by simulating common objects under different shape complexity and noise levels and identify a suitable range for γ.

Various low-complexity algorithms have been proposed to solve (8). We here apply the alternating direction method (ADM) of [28] for its simplicity, but other methods [26], [27] can be alternatively used. For completeness, the ADM method is sketched as follows. Firstly, a multiplier $\boldsymbol{Z}_k$ is introduced to construct a cost function from (8) as

$$J(\boldsymbol{S}_k, \boldsymbol{H}_k, \boldsymbol{Z}_k) = \gamma \|\boldsymbol{S}_k\|_{l_1} + \|\boldsymbol{H}_k\|_* - \langle \boldsymbol{Z}_k, \boldsymbol{S}_k + \boldsymbol{H}_k - \hat{\boldsymbol{H}}_k\rangle + \frac{\rho}{2}\left\|\boldsymbol{S}_k + \boldsymbol{H}_k - \hat{\boldsymbol{H}}_k\right\|^2, \quad (9)$$

where $\rho > 0$ is a penalty parameter for controlling the tolerance of the residual error, and $\langle\cdot\rangle$ represents the standard trace inner product. The optimization problem then becomes the minimization of $J(\boldsymbol{S}_k, \boldsymbol{H}_k, \boldsymbol{Z}_k)$. Let $P^t(\cdot)$ be a thresholding operator with threshold $t > 0$, that is, $P^t(\boldsymbol{X})$ returns a matrix obtained by thresholding all the entries of $\boldsymbol{X}$ into the range $[-t, t]$. Let n denotes the number of iteration of the ADM method, and $(\boldsymbol{S}_k^n, \boldsymbol{H}_k^n, \boldsymbol{Z}_k^n)$ the current solution. The solutions are then updated iteratively as follows:

- Generate $\boldsymbol{S}_k^{n+1}$:

$$\boldsymbol{S}_k^{n+1} = \left(\frac{1}{\rho}\boldsymbol{Z}_k^n - \boldsymbol{H}_k^n + \hat{\boldsymbol{H}}_k\right) - P^{\frac{\gamma}{\rho}}\left(\frac{1}{\rho}\boldsymbol{Z}_k^n - \boldsymbol{H}_k^n + \hat{\boldsymbol{H}}_k\right). \quad (10)$$

- Generate $\boldsymbol{H}_k^{n+1}$:

$$\boldsymbol{H}_k^{n+1} = \boldsymbol{U}^{n+1} \operatorname{diag}\left(\max\left\{\sigma_i^{n+1} - \frac{1}{\rho}, 0\right\}\right)\left(\boldsymbol{V}^{n+1}\right)^T, \quad (11)$$

  where $\boldsymbol{U}^{n+1}$, $\boldsymbol{V}^{n+1}$ and σ_i^{n+1} are, respectively, the left singular vector matrix, right singular vector matrix and singular values of $\hat{H}_k - \boldsymbol{S}_k^{n+1} + \frac{1}{\rho}\boldsymbol{Z}_k^n$.
- Update the multiplier:

$$\boldsymbol{Z}_k^{n+1} = \boldsymbol{Z}_k^n - \rho(\boldsymbol{S}_k^{n+1} + \boldsymbol{H}_k^{n+1} - \hat{\boldsymbol{H}}_k). \quad (12)$$

  The above iterative algorithm can be terminated based on a pre-set number of iterations or variation of the solutions. We refer the readers to [28] for detailed discussion about the properties and implementation of the ADM method.

B. MOTION PARAMETER ESTIMATION

After applying the FTP and RPCA steps in Section III-A, K height maps $\{\tilde{\boldsymbol{H}}_k\}$ are obtained, which may still suffer from errors. Fusing these height maps may improve the overall accuracy of the 3D reconstruction. In order to achieve this, we first estimate the rotation matrix $\boldsymbol{R}_k$ and translation vector $\boldsymbol{t}_k$ in (5) from $\{\tilde{\boldsymbol{H}}_k\}$. We propose to apply the ICP algorithm [29], [30], which can be used to register 3D measurement of a moving rigid object and obtain the associated motion parameters. Starting from initial coarse estimates of the motion parameters, the correspondence between the points in the different measurements is established using a minimum Euclidean distance criterion. The motion parameters are then refined based on the established correspondence using the least squares method. The above correspondence matching and motion parameter estimation are carried out in an iterative manner until a stopping criterion is met. In contrast to the methods using markers and SIFT [16]–[18], the ICP approach eliminates the needs of human intervention and feature points extraction for establishing the correspondence. We assume that the height of the reference plane is low and known. A threshold is then set to distinguish the object and the background and extract the point set corresponding to the object. In general, the object point set $\tilde{\boldsymbol{H}}_k$ each may contain a large set of points. Uniform sampling is used to extract a smaller subset of $\tilde{\boldsymbol{H}}_k$ before ICP is applied, such that the computational complexity can be reduced. Let $\boldsymbol{Q} = \{\boldsymbol{q}_n, n = 1, 2, \ldots, N\}$ and $\boldsymbol{G} = \{\boldsymbol{g}_m, m = 1, 2, \ldots, M\}$ be the sets of 3D points sampled from the retrieved object surfaces at instant 1 and instant k, respectively. Let

$$\boldsymbol{q}_n \triangleq \begin{bmatrix} x_n \\ y_n \\ h_n \end{bmatrix}, \quad \boldsymbol{g}_m \triangleq \begin{bmatrix} u_m \\ v_m \\ z_m \end{bmatrix}, \quad \begin{matrix} n = 1, 2, \ldots, N, \\ m = 1, 2 \ldots, M \end{matrix} \tag{13}$$

The ICP algorithm aims to obtain the motion parameters and establish the correspondence relationships of the model point set $\boldsymbol{Q}$ and the data point set $\boldsymbol{G}$. Let the Euclidean distance $d(\boldsymbol{q}_n, \boldsymbol{g}_m)$ between two points $\boldsymbol{q}_n$ and $\boldsymbol{g}_m$ be

$$d(\boldsymbol{q}_n, \boldsymbol{g}_m) = \sqrt{(x_n - u_m)^2 + (y_n - v_m)^2 + (h_n - z_m)^2}. \tag{14}$$

For each $\boldsymbol{q}_n$, the corresponding point in $\boldsymbol{G}$ is estimated as

$$\boldsymbol{g}_n^* = \operatorname{argmin}_{\boldsymbol{g}_m \in \boldsymbol{G}} d\left(\boldsymbol{q}_n, \boldsymbol{g}_m\right). \tag{15}$$

Assume that $\boldsymbol{G}' \triangleq \{\boldsymbol{g}_n^*\}$ is the set containing the estimated corresponding points of $\boldsymbol{Q}$. Then the rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$ are estimated by minimizing

$$E(\boldsymbol{R}, \boldsymbol{t}) = \sum\nolimits_{n=1}^{N} \left\| \boldsymbol{q}_n - \left(\boldsymbol{R}\boldsymbol{g}_n^* + \boldsymbol{t}\right) \right\|^2. \tag{16}$$

Let

$$\boldsymbol{q}' = \frac{1}{N} \sum\nolimits_{n=1}^{N} \boldsymbol{q}_n, \quad \boldsymbol{q}_{cn} = \boldsymbol{q}_n - \boldsymbol{q}', \ n = 1, 2, \cdots, N, \tag{17}$$

$$\boldsymbol{g}' = \frac{1}{N} \sum\nolimits_{n=1}^{N} \boldsymbol{g}_n^*, \quad \boldsymbol{g}_{cn} = \boldsymbol{g}_n^* - \boldsymbol{g}', \ n = 1, 2, \cdots, N. \tag{18}$$

The cross-correlation matrix is then defined as

$$\boldsymbol{C} = \sum\nolimits_{n=1}^{N} \boldsymbol{g}_{cn} \boldsymbol{q}_{cn}^T \tag{19}$$

with the singular value decomposition (SVD)

$$\boldsymbol{C} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{V}^T \tag{20}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the orthogonal left and right singular vector matrices, respectively. Then the rotation matrix and translation vector that minimize (16) can be found as

$$\hat{\boldsymbol{R}}' = \boldsymbol{V}\boldsymbol{U}^T, \tag{21}$$

$$\hat{\boldsymbol{t}}' = \boldsymbol{q}' - \hat{\boldsymbol{R}}'\boldsymbol{g}'. \tag{22}$$

The accuracy of $\hat{\boldsymbol{R}}'$ and $\hat{\boldsymbol{t}}'$ depends on the correspondence established using (15). Such correspondence may be poor initially but can be improved iteratively by updating the data points using the estimated motion parameters as $\boldsymbol{g}_m \leftarrow \hat{\boldsymbol{R}}' \boldsymbol{g}_m + \hat{\boldsymbol{t}}'$, $m = 1, 2, \ldots, M$. The operations described by (15)-(22) are then repeated based on the updated data point set $\boldsymbol{G} = \{\boldsymbol{g}_m\}$, yielding updated $\hat{\boldsymbol{R}}'$ and $\hat{\boldsymbol{t}}'$. Let $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{t}}$ denote the overall rotation matrix and translation vector and initialize them to $\hat{\boldsymbol{R}} = \boldsymbol{I}_{3\times3}$, $\hat{\boldsymbol{t}} = \boldsymbol{0}_{3\times1}$. They are then updated iteratively as $\hat{\boldsymbol{R}} \leftarrow \hat{\boldsymbol{R}}'\hat{\boldsymbol{R}}$, $\hat{\boldsymbol{t}} \leftarrow \hat{\boldsymbol{R}}'\hat{\boldsymbol{t}} + \hat{\boldsymbol{t}}'$. The above process will be repeated for a preset number of iterations or until the value of the objective function $E\left(\hat{\boldsymbol{R}}', \hat{\boldsymbol{t}}'\right)$ becomes lower than a threshold. By letting $\boldsymbol{Q}$ and $\boldsymbol{G}$ be data points from $\tilde{\boldsymbol{H}}_1$ and $\tilde{\boldsymbol{H}}_k$ and applying the ICP algorithm, we can obtain the estimates $(\hat{\boldsymbol{R}}_k, \hat{\boldsymbol{t}}_k)$ of the motion parameters $(\boldsymbol{R}_k, \boldsymbol{t}_k)$ defined in (5) for each time instant.

C. HEIGHT MAPS FUSION

The motion parameters $(\hat{\boldsymbol{R}}_k, \hat{\boldsymbol{t}}_k)$ for $k = 2, 3, \cdots, K$ can now be used for fusing the measurement results at different time instants. Let $\hat{r}_k^{ij}$ denote the (i, j)th entry of $\hat{\boldsymbol{R}}_k$ and $\hat{t}_k^n$ the nth entry of $\hat{\boldsymbol{t}}_k$. According to (5), we can generate K estimates of the height map at instant $k = 1$ from $(\hat{\boldsymbol{R}}_k, \hat{\boldsymbol{t}}_k)$ and $\tilde{\boldsymbol{H}}_k$ as

$$\tilde{h}_{1,k}^P \left(x_1^P, y_1^P\right) = \hat{r}_k^{31} x_k^P + \hat{r}_k^{32} y_k^P + \hat{r}_k^{33} \tilde{h}_k^P + \hat{t}_k^3, \quad k = 1, 2, \cdots, K. \tag{23}$$

In the above, (x_k^P, y_k^P) denote the coordinates of the point P at the time instant k, and $\tilde{h}_k^P$ the corresponding height in map $\tilde{\boldsymbol{H}}_k$. Linear interpolation is used to obtain $\tilde{h}_{1,k}^P$ when (x_k^P, y_k^P) do not fall on the center of a pixel. In the case of $k = 1$ for the first instant, $\hat{\boldsymbol{R}}_1 = \boldsymbol{I}_{3\times3}$, $\hat{\boldsymbol{t}}_1 = \boldsymbol{0}_{3\times1}$. Finally, we fuse the multiple height maps in (23) to further alleviate noise-induced errors and improve accuracy. Note that due to the variations of the signal-to-noise-ratio (SNR), reflectivity, and orientation of the object with respect to the fringe patterns, the measured heights for the same point at different time instants have different accuracies. A weighted combination of the multiple height maps is employed with weights $\{w_k\}$:

$$\hat{h}_F \left(x_1^P, y_1^P\right) = \sum\nolimits_{k=1}^{K} w_k \tilde{h}_{1,k}^P \left(x_{1,k}^P, y_{1,k}^P\right), \tag{24}$$

where we use the coordinates at time instant $k = 1$ to present the final height map $\hat{\boldsymbol{H}}_F = \{\hat{h}_F \left(x_1^P, y_1^P\right)\}$. The optimal choice of $\{w_k\}$ is an open issue. In this work, we consider a heuristic approach similar to the maximum ratio combining (MRC) [31] that chooses $\{w_k\}$ according to estimates of the SNR in the fringe images $d(x, y)$ captured by the camera, which can be also modeled as:

$$d(x, y) = a + b\cos(\varphi(x, y)) + n(x, y) \tag{25}$$

where a is the average intensity, b is the intensity modulation, $\varphi(x, y)$ is the reference phase, and $n(x, y)$ is the noise. Assume that the captured deformed fringe patterns contain the zero-mean noises $n(x, y)$ with variance σ^2, which are independent of the fringe signals. Then a and b can be estimated by averaging the captured images and curve fitting, respectively. We can verify that

$$\frac{1}{\hat{Q}\hat{T}} \sum\nolimits_{x} d^2(x, y) \approx a^2 + \frac{b^2}{2} + \sigma^2, \tag{26}$$

where $\hat{T}$ is the period of the fringe and $\hat{Q}$ is the number of fringes. Therefore, we can obtain an estimate of the SNR of

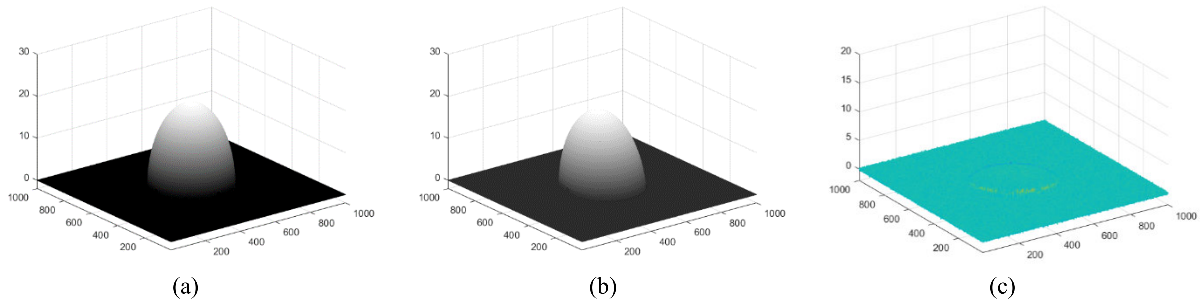

FIGURE 3. Verification of the ICP algorithm at SNR = 30dB. (a): Object before movement. (b): Object after movement. (c): Difference between the original object in (a) and its prediction obtained from (b) using the estimated motion parameters.

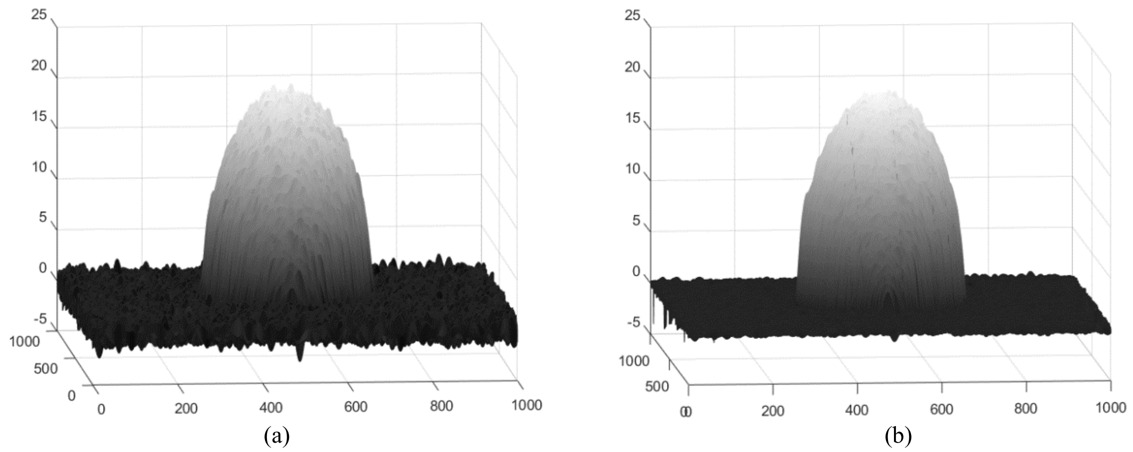

FIGURE 4. Verification of proposed approach for SNR = 25dB. (a): Results of FTP for a 3D dynamic object. (b): Proposed approach.

the captured image as:

$$\Gamma_k \triangleq \frac{\frac{1}{\hat{Q}\hat{T}}\sum_x \left(a + b\cos\left(\varphi\left(x, y\right)\right)\right)^2}{\frac{1}{\hat{Q}\hat{T}}\sum_x n^2\left(x, y\right)} \approx \frac{a^2 + \frac{b^2}{2}}{\frac{1}{\hat{Q}\hat{T}}\sum_x d^2\left(x, y\right) - \left(a^2 + \frac{b^2}{2}\right)}. \quad (27)$$

Finally, we set the combining weights in (28) as

$$w_k = \frac{\sqrt{\Gamma}_k}{\sum_{n=1}^{K}\sqrt{\Gamma}_k}. \quad (28)$$

IV. SIMULATION AND EXPERIMENTAL RESULTS

This section presents the simulation and experimental results to demonstrate the effectiveness of our proposed approach. In order to quantitatively assess an estimated height map $\hat{h}\left(x, y\right)$, the NMSE (normalised mean squared error) is computed:

$$\mathrm{NMSE} = \frac{\sum_{y=1}^{Y}\sum_{x=1}^{X}\left(\hat{h}\left(x, y\right) - h(x, y)\right)^2}{\sum_{y=1}^{Y}\sum_{x=1}^{X}\left(h\left(x, y\right)\right)^2}, \quad (29)$$

where $h(x, y)$ is the true height map, X and Y are the numbers of pixels along the x-axis and y-axis, respectively.

We first examine the feasibility of applying the ICP algorithm to estimate the motion parameters. Consider an FPP system with $l_0 = 4000$mm and $d_0 = 600$mm. A spatially symmetric hemi-ellipsoidal object (with radius 200mm and height 20mm) is simulated as shown in Fig.3(a) and then transformed by a known rotation matrix and translation vector into that in Fig.3(b). We treat the height map as a two-dimensional signal with average power P_h. Noise with average power P_n is added to the height distribution and the signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} = 10\log\frac{P_h}{P_n}$. The motion parameters are estimated by using the ICP algorithm from the two point clouds from Fig.3(a) and Fig.3(b). They can then be utilized to transform the height map in Fig.3(b) to predict that in Fig. 3(a). The prediction error refers to the difference between the height map in Fig. 3(a) and its prediction obtained from Fig.3(b) using the estimated motion parameters, which is shown in Fig.3(c) for SNR = 30 dB. The NMSE of the prediction is found to be 0.0032, 0.0108 and 0.0214 respectively under the SNR of 30 dB, 25dB and 20 dB. These results demonstrate the feasibility of using the ICP algorithm to estimate the motion parameters from noisy height maps. From Fig. 3(c), there can be a slight difference between the original map and the predicted map using the motion parameters estimated. This

TABLE 1. Rotation angles and translation distance.

Time instant	(α, β, θ)	$\boldsymbol{t}$(mm)
1	$0°, 0°, 0°$	(0,0,0)
2	$-0.32°, -0.48°, 0.32°$	(3,3,2)
3	$-0.64°, -0.64°, 0.64°$	(5,6,4)

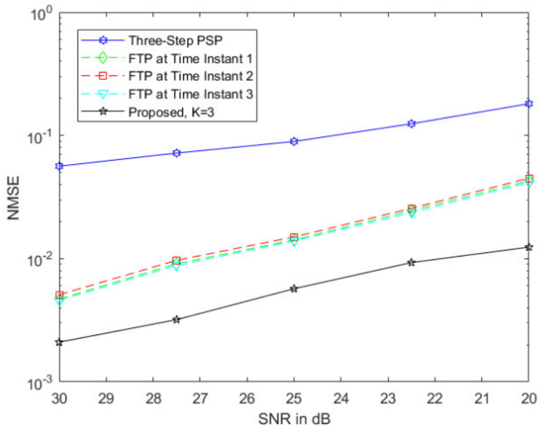

FIGURE 5. Comparison of the proposed approach with three-step PSP and FTP.

difference is partly caused by the modeling errors due to imperfect motion parameters estimation and linear interpolation. Note also that for spatially symmetric objects the same motion of the object may be described using different models. In this case, the motion parameters found by the ICP algorithm are not unique. In the noisy cases, ICP aims to find the motion parameters that minimize the prediction errors described above.

The second simulation aims to verify the effectiveness of the proposed method. The workflow as demonstrated in Fig.2 is considered. The rotation matrix is given by

$$\boldsymbol{R}(\alpha, \beta, \theta) = \boldsymbol{R}_z(\theta)\, \boldsymbol{R}_y(\beta)\, \boldsymbol{R}_x(\alpha) \tag{30}$$

where

$$\boldsymbol{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{31}$$

$$\boldsymbol{R}_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \tag{32}$$

$$\boldsymbol{R}_z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{33}$$

and α, β and θ are the rotation angles along the x, y, and z axis, respectively. The movement parameters at the three time instants are listed in Table 1.

The standard FTP is first tested and its reconstruction result at SNR = 25 dB is shown in Fig.4(a), achieving an NMSE of 0.0245 for the height measurement. In all simulations, we set $\gamma = 0.1$ for the RPCA algorithm. The proposed method fuses the three height maps as shown in Fig. 4(b), achieving an NMSE of 0.0065. The NMSE comparisons under different SNRs are shown in Fig. 5. It is confirmed that the proposed approach can improve the performance compared to traditional FTP and PSP schemes. Note that the PSP scheme

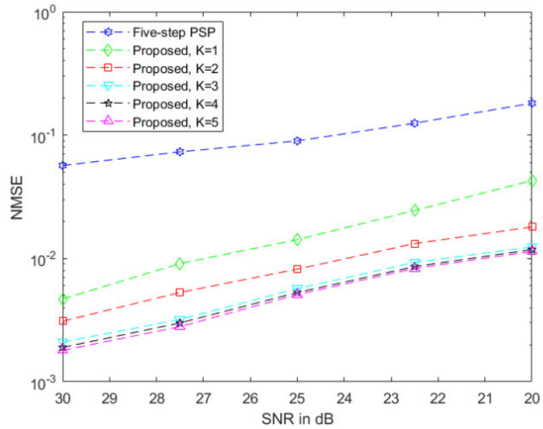

FIGURE 6. Performance of the proposed approach with different numbers of height maps fused.

FIGURE 7. The gourd shaped object.

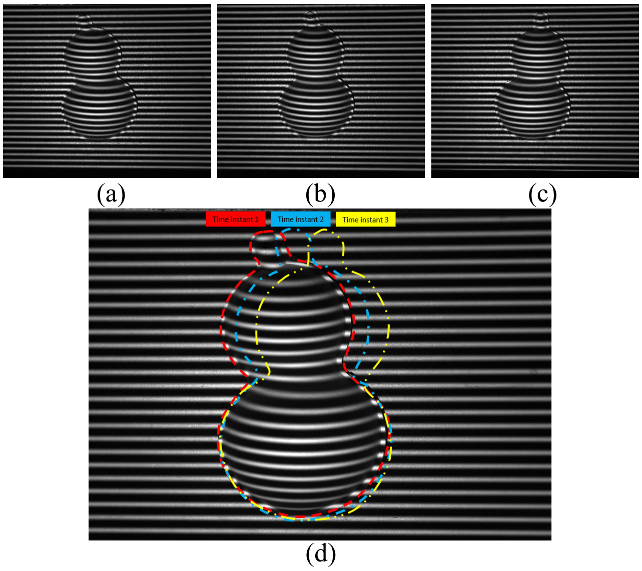

FIGURE 8. Dynamic 3D gourd at three time instants and movement trajectory.

suffers from the motion artifacts. In Fig. 6, we also simulate the fusion of different numbers of height maps. The overall NMSE performance of the proposed approach depends on the measurement error of each height map, the motion parameter error from the ICP algorithm and linear interpolation error from height fusion. When K is small, the overall performance is dominated by the original measurement error of each height map. When K is large, the performance is dominated by the modeling errors which cannot be mitigated by increasing K. Overall, fusing more height maps can achieve better performance but the improvement of NMSE becomes minor when K is large. In order to balance the computation time and reconstruction accuracy, we set $K = 3$ in the experiment.

In the experiment, a gourd shaped object shown in Fig. 7 is captured at $K = 3$ time instants, with its instantaneous shape illustrated in Fig. 8(a), 8(b) and 8(c). The movement of the object at the three time instants is shown in Fig. 8(d). Fig. 9(a) shows the performance of a three-step PSP scheme utilizing fringe patterns captured at the three time instants. Clearly, due to the motion artifact, such a PSP scheme performs poorly. The proposed approach applies FTP to obtain three height maps of the dynamic object, and the coarse height map acquired at time instant 1 is shown in Fig.9(b). The PRCA algorithm is then applied to filter the coarse height map to produce a refined height map, as shown in Fig.9(c). The experimental results presented are based on the empirical choice of $\gamma = 0.1$ determined from numerical simulations. The ICP algorithm works on the error-suppressed height maps to estimate the rotation matrix and translation vector. With the uniform sampling rate of 1/500 and around 2000 points from each point cloud, ICP takes about 20 seconds when implemented using MATLAB on an Acer Aspire V 15 laptop with a 2.8 GHz CPU and 16G memory. Finally, the three height maps are fused using the motion parameters and SNR-based weights, as shown in Fig.9(d).

In order to evaluate the performance improvement of the proposed technique over the traditional FTP, we also calculate the NMSE result for the experimental results presented above. As the true shape of the gourd is not known, the nine-step PSP result in Fig. 10(a) (when the gourd is kept static) is assumed as the true shape. The NMSE with respect to Fig. 9 of the cases displayed above are obtained in Table 2.

TABLE 2. Comparison of the NMSE results of the gourd shaped object.

Traditional three-step PSP	Traditional FTP	RPCA filtered F TP	Proposed approach
0.6146	0.0135	0.0115	0.0054

Because the movement of the object leads to the loss of correspondence, the NMSE of three-step PSP is 0.6146, which indicates the failure of reconstruction. The NMSE of a coarse height map obtained by FTP is 0.0135. When the RPCA algorithm is employed, the NMSE is reduced to 0.0115. After fusing the multiple height maps, the NMSE result is reduced further to 0.0054. Fig.11 shows the reconstructed results for the cross sections at x = 800 as indicated in Fig.10(b), where the corresponding absolute errors can be observed. It is seen that, the proposed method is able to improve the performance of FTP for 3D dynamic objects.

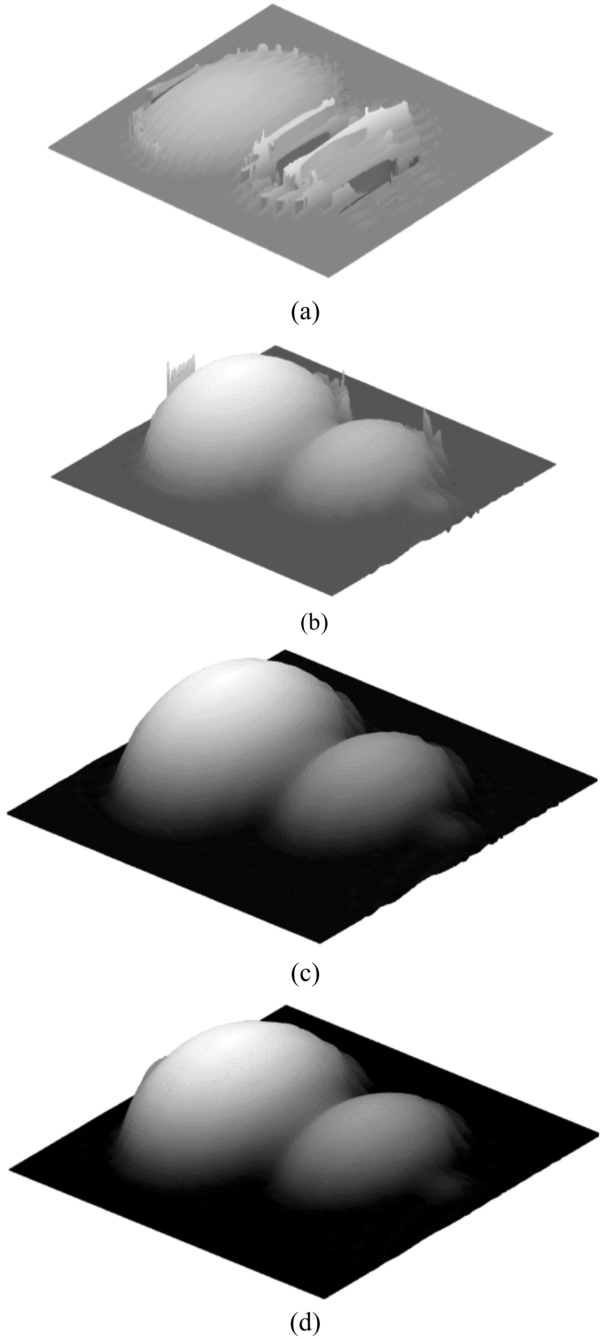

FIGURE 9. Comparison of the traditional FTP method and the proposed method when combined with FTP. (a): Results of three-step PSP using fringe patterns captured at the three different time instants in Fig. 7. (b): Reconstruction results from FTP. (c): Results obtained by refining (b) by using RPCA. (d): Results after fusing FTP results.

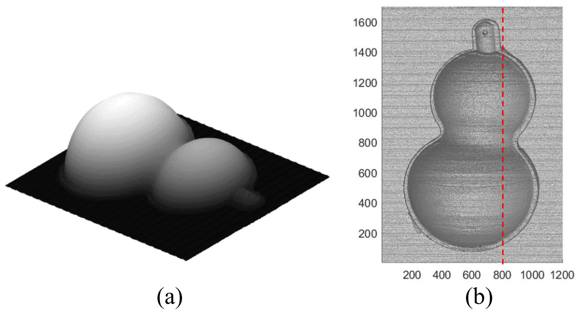

FIGURE 10. The result of nine-step PSP when the gourd shaped object is kept static.

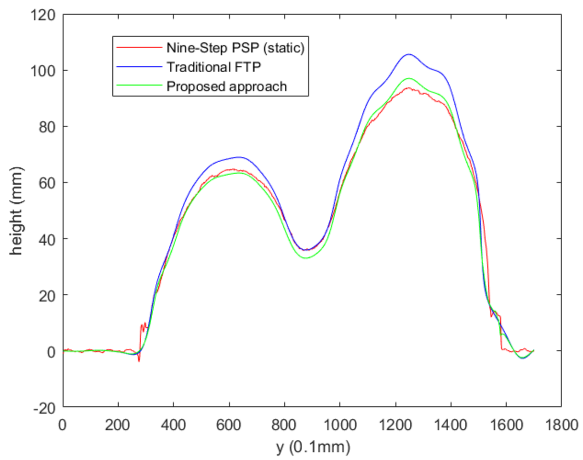

FIGURE 11. The result of the cross section at x = 800 as indicated in Fig. 10(b).

V. CONCLUSION

In summary, a novel method has been proposed to enhance the performance for reconstructing 3D dynamic objects. Initial height maps of the object are retrieved by using FTP at each time instant and then filtered using RPCA. The ICP algorithm is then applied to establish the correspondence relationship between 3D point clouds at different time instants. The multiple height maps are finally adaptively fused to improve the overall measurement accuracy. The effectiveness of the proposed method has been demonstrated by simulation and experimental studies. Future works may include the optimization of the fringe projection and phase retrieval for measuring complex dynamic objects with surface discontinuity, the improvement of the running efficiency of the ICP algorithm to perform real-time measurements, and the combination of the proposed method with other FPP techniques.

REFERENCES

[1] X. Su and Q. Zhang, "Dynamic 3D shape measurement: A review," *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 191–204, Feb. 2010.

[2] S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 149–158, Feb. 2010.

[3] S. S. Gorthi and P. Rastogi, "Fringe projection techniques: Whither we are?" *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 133–140, Feb. 2010.

[4] C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, and Q. Chen, "Phase shifting algorithms for fringe projection profilometry: A review," *Opt. Lasers Eng.*, vol. 109, pp. 23–59, Oct. 2018.

[5] P. S. Huang, "High-resolution, real-time three-dimensional shape measurement," *Opt. Eng.*, vol. 45, no. 12, Dec. 2006, Art. no. 123601.

[6] S. Yoneyama, Y. Morimoto, M. Fujigaki, and Y. Ikeda, "Three-dimensional surface profile measurement of a moving object by a spatial-offset phase stepping method," *Opt. Eng.*, vol. 42, no. 1, pp. 137–142, 2003.

[7] 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, Apr. 2015.

[8] X. Su, W. Chen, Q. Zhang, and Y. Chao, "Dynamic 3-D shape measurement method based on FTP," *Opt. Lasers Eng.*, vol. 36, no. 1, pp. 49–64, Jul. 2001.

[9] S. Zhang, D. Van Der Weide, and J. Oliver, "Superfast phase-shifting method for 3-D shape measurement," *Opt. Exp.*, vol. 18, no. 9, pp. 9684–9689, 2010.

[10] Y. Wang, S. Zhang, and J. H. Oliver, "3D shape measurement technique for multiple rapidly moving objects," *Opt. Exp.*, vol. 19, no. 9, pp. 8539–8545, 2011.

[11] Z. Zhang, C. E. Towers, and D. P. Towers, "Time efficient color fringe projection system for 3D shape and color using optimum 3-frequency selection," *Opt. Exp.*, vol. 14, no. 14, p. 6444, 2006.

[12] T. Weise, B. Leibe, and L. V. Gool, "Fast 3d scanning with automatic motion compensation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[13] S. Zhang, "High-speed three-dimensional shape measurement system using a modified two-plus-one phase-shifting algorithm," *Opt. Eng.*, vol. 46, no. 11, Nov. 2007, Art. no. 113603.

[14] 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. Exp.*, vol. 18, no. 5, pp. 5229–5244, 2010.

[15] C. Zuo, Q. Chen, G. Gu, S. Feng, F. Feng, R. Li, and G. Shen, "High-speed three-dimensional shape measurement for dynamic scenes using bi-frequency tripolar pulse-width-modulation fringe projection," *Opt. Lasers Eng.*, vol. 51, no. 8, pp. 953–960, Aug. 2013.

[16] L. Lu, J. T. Xi, Y. G. Yu, and Q. Guo, "New approach to improve the accuracy of 3-D shape measurement of moving object using phase shifting profilometry," *Opt. Exp.*, vol. 21, no. 25, pp. 30610–30622, Dec. 2013.

[17] L. Lu, J. Xi, Y. Yu, and Q. Guo, "Improving the accuracy performance of phase-shifting profilometry for the measurement of objects in motion," *Opt. Lett.*, vol. 39, no. 23, pp. 6715–6718, 2014.

[18] L. Lu, Y. Ding, Y. Luan, Y. Yin, Q. Liu, and J. Xi, "Automated approach for the surface profile measurement of moving objects based on PSP," *Opt. Exp.*, vol. 25, no. 25, p. 32120, Dec. 2017.

[19] E. Hu and Y. He, "Surface profile measurement of moving objects by using an improved π phase-shifting Fourier transform profilometry," *Opt. Lasers Eng.*, vol. 47, no. 1, pp. 57–61, Jan. 2009.

[20] H. Guo and P. Huang, "3-D shape measurement by use of a modified Fourier transform method," *Proc. SPIE*, vol. 7066, Aug. 2008, Art. no. 70660E.

[21] X. Su, "Two-dimensional Fourier transform profilometry for the automatic measurement of three-dimensional object shapes," *Opt. Eng.*, vol. 34, no. 11, p. 3297, Nov. 1995.

[22] W. Chen, X. Su, Y. Cao, Q. Zhang, and L. Xiang, "Method for eliminating zero spectrum in Fourier transform profilometry," *Opt. Lasers Eng.*, vol. 43, no. 11, pp. 1267–1276, Nov. 2005.

[23] Q. Kemao, "Windowed Fourier transform for fringe pattern analysis," *Appl. Opt.*, vol. 43, no. 13, pp. 2695–2702, 2004.

[24] Q. Kemao, "Two-dimensional windowed Fourier transform for fringe pattern analysis: Principles, applications and implementations," *Opt. Lasers Eng.*, vol. 45, no. 2, pp. 304–317, Feb. 2007.

[25] M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," *J. Opt. Soc. Amer.*, vol. 72, pp. 156–160, Jan. 1982.

[26] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, Jul. 2018.

[27] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, p. 11, 2011.

[28] X. Yuan and J. Yang, "Sparse and low-rank matrix decomposition via alternating direction methods," *Pacific J. Optim.*, vol. 9, no. 1, pp. 167–180, 2013.
[29] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. DIM*, vol. 1, May/Jun. 2001, pp. 145–152.
[30] J. Minguez, L. Montesan, and F. Lamiraux, "Metric-based iterative closest point scan matching for sensor displacement estimation," *IEEE Trans. Robot.*, vol. 22, no. 5, pp. 1047–1054, Oct. 2006.
[31] V. A. Aalo and J. Zhang, "Performance analysis of maximal ratio combining in the presence of multiple equal-power cochannel interferers in a Nakagami fading channel," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 497–503, Mar. 2001.

CHENGPU DUAN received the B.E. degree (Hons.) from the University of Wollongong, Australia, and Zhengzhou University, China, in 2016. He is currently pursuing the Ph.D. degree with the University of Wollongong. His research interest includes 3-D shape measurement.

JUN TONG received the B.E. and M.E. degrees from the University of Electronic Science and Technology of China (UESTC) and the Ph.D. degree in electronic engineering from the City University of Hong Kong. He is currently a Senior Lecturer with the School of Electrical, Computer, and Telecommunications Engineering, University of Wollongong, Australia. His research interest includes signal processing and its applications to communication systems.

LEI LU received the B.E. degree from the Henan University of Science and Technology, China, in 2007, the M.E. degree from Zhengzhou University, China, in 2011, and the Ph.D. degree from the University of Wollongong, Australia, in 2015. From 2015 to 2016, he was working as a Postdoctoral Fellow with the Singapore University of Technology and Design. He is currently an Associate Professor with the Henan University of Technology, China. His research interests include optical 3-D image and 3D printing.

JIANGTAO XI (Senior Member, IEEE) received the B.E. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 1982, the M.E. degree in electrical engineering from Tsinghua University, Beijing, in 1985, and the Ph.D. degree in electrical engineering from the University of Wollongong, Wollongong, Australia, in 1996. From 1995 to 1996, he was a Postdoctoral Fellow with the Communications Research Laboratory, McMaster University, Hamilton, ON, Canada, and from 1996 to 1998, he was a Member of Technical Staff with Bell Laboratories, Lucent Technologies Inc., NJ, USA. From 2000 to 2002, he was the Chief Technical Officer of the TCL IT Group Company, China. In 2003, he rejoined as a Senior Lecturer with the University of Wollongong, where he is currently a Full Professor and the Head of the School of Electrical, Computer, and Telecommunications Engineering. His current research interests include signal processing and its applications in various areas, such as instrumentation and measurement, and communications.

YANGUANG YU (Senior Member, IEEE) received the B.E. degree from the Huazhong University of Science and Technology, China, in 1986, and the Ph.D. degree from the Harbin Institute of Technology, China, in 2000. She was with the College of Information Engineering, Zhengzhou University, China, on various appointments, including a Lecturer, from 1986 to 1999; an Associate Professor, from 2000 to 2004; and a Professor, from 2005 to 2007. From 2001 to 2002, she was a Postdoctoral Fellow with the Opto-Electronics Information Science and Technology Laboratory, Tianjin University, China. She also had a number of visiting appointments, including a Visiting Fellow at the Optoelectronics Group, Department of Electronics, University of Pavia, Italy, from 2002 to 2003; a Principal Visiting Fellow with the University of Wollongong, Australia, from 2004 to 2005; a Visiting Associate Professor and Professor with the Engineering School, ENSEEIHT, Toulouse, France, in 2004 and 2006. She joined the University of Wollongong, in 2007, where she is currently an Associate Professor with the School of Electrical, Computer, and Telecommunications Engineering. Her research interests include semiconductor lasers with optical feedback and their applications in sensing and instrumentations, secure chaotic communications, and signal processing and its applications to 3-D profile measurement and telecommunication systems.

QINGHUA GUO received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xi'an, China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, Hong Kong, in 2008. He is currently an Associate Professor with the School of Electrical, Computer, and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and also an Adjunct Associate Professor with the School of Engineering, The University of Western Australia, Perth, WA, Australia. He was a recipient of the Discovery Early Career Researcher Award from the Australian Research Council. His research interests include signal processing and telecommunications.