

RESEARCH

Open Access

Binary defocusing technique based on complementary decoding with unconstrained dual projectors



Xuexing Li¹  and Wenhui Zhang^{1,2*}

Abstract

Binary defocusing technique can effectively break the limitation of hardware speed, which has been widely used in the real-time three-dimensional (3D) reconstruction. In addition, fusion technique can reduce captured images count for a 3D scene, which helps to improve real-time performance. Unfortunately, it is difficult for binary defocusing technique and fusion technique working simultaneously. To this end, our research established a novel system framework consisting of dual projectors and a camera, where the position and posture of the dual projectors are not strictly required. And, the dual projectors can adjust defocusing level independently. Based on this, this paper proposed a complementary decoding method with unconstrained dual projectors. The core idea is that low-resolution information is employed for high-resolution phase unwrapping. For this purpose, we developed the low-resolution depth extraction strategy based on periodic space-time coding patterns and the method from the low-resolution order to high-resolution order of fringe. Finally, experimental results demonstrated the performance of our proposed method, and the proposed method only requires three images for a 3D scene, as well as has strong robustness, expansibility, and implementation.

Keywords: Three-dimensional reconstruction, Binary defocusing technique, Fusion technique, Dual projection system

Introduction

Fringe projection profilometry (FPP) plays an important role in some academia and applied fields, such as product inspection, reverse engineering, and computer animation, because of its advantages of being high precision, full field, and nondestructive [1–5]. A series of fringe patterns generated by the computer are casted on an object by a projector. Then, a camera acquires the patterns distorted by the object surface geometry, as well as the three-dimensional (3D) profile of the object is obtained by the reconstruction algorithms [6]. With the development of the technical requirement, the real-time performance is increasingly important [7]. Generally, in

order to enhance the real-time capability, we inevitably require to increase hardware speed or decrease the number of captured images to reconstruct a 3D scene.

Recently, with the development of projector technique, the refresh rate of special projector can exceed 10,000 frames/s in 1-bit mode. The novel technique has benefited from the digital-mirror-device (DMD) technique developed by Texas Instruments (TI). For example, the TI D4100 projector can reach 3.255×10^4 frames/s in binary mode with 1024×768 resolution [4]. Based on this, researchers have proposed a binary defocusing technique (BDT), which breaks the limitation of the maximum refresh rate of the commercial projector. Around 2009, Lei et al. [8] developed the flexible 3D reconstruction in projector defocusing, which successfully made rate breakthroughs. Nevertheless, the quality is unsatisfactory due to the high-order harmonics influences. In

* Correspondence: megres.li@foxmail.com

¹School of Mechanical Technology, Wuxi Institute of Technology, Wuxi 214121, China

²Wuxi GNFR information technology Co, Ltd, Wuxi 214000, China

2010, researchers introduced pulse width modulate (PWM) techniques into BDT [9]. The method alleviates the high-order harmonics influences, but it also has two defects: (1) the width of fringe limits its improvement, and (2) the one-dimensional (1D) fringe of the PWM technique is not fit for the two-dimension of BDT. In 2012, Lohry et al. [10] proposed to approximate the triangular waveform by modifying 2×2 pixels so that the casted patterns are better for BDT. Lately, researchers introduced the dithering techniques into BDT in order to significantly improve the quality of the casted fringe patterns. Because the dithering techniques represent grayscale images with binary images for printing and processing, it results in a higher quality of fringe pattern than other previous 1D defocusing method. The relevant researches mainly include the Bayer method [11], the error-diffusion method [12] and the iteration method [13]. In conclusion, the speed of the projector has been greatly improved because of BDT, so the hardware speed limitation falls on the camera. For the camera, if its refresh rate is increased, the real-time performance can be enhanced to some extent. However, the refresh rate is the inherent property of camera. Hence, the more effective means is to reduce the number of captured images for a 3D scene reconstruction.

It usually has unsatisfactory resolution and precision, with only one captured frame decoded by spatial information. As to this problem, some researchers considered decoding with multiple images, for example, phase-shifting measurement profilometry (PMP). In PMP, especially, three-step phase-shifting (3-PS) algorithm can achieve rapidity since it captures the minimum number of phase-shifting (PS) patterns to reconstruct a 3D scene. Furthermore, the algorithm inevitably needs to retrieve wrapped phase. Generally, there are two types of methods with unwrapping phase: spatial phase unwrapping and temporal phase unwrapping. Spatial phase unwrapping does not require additional decoding images, but it fails when the reconstructed surfaces are abrupt, missing, or discrete. Hence, temporal phase unwrapping is still necessary, and the fusion strategies are introduced to reduce the number of captured images. Normally, the fusion strategies can be divided into the patterns fusion and the multi-device fusion. In the patterns fusion, there are two ways: the fusion based on single channel (i.e. gray) [14] and multichannel (i.e. color) [15]. The gray fusion is a superposition according to different spatial frequencies or angles, and the color fusion obtains irrelevance of color. It is noted that both have some defects, such as low-resolution, low-contrast or color influence. More importantly, they are not suitable for BDT, i.e., it is impossible for a single projector to have multiple defocusing levels at the same time. In the multi-device fusion, an additional device is introduced. Zhang

et al. [16] proposed a fusion framework which consists of a Kinect and a PMP system. The framework used low-resolution depth information of Kinect to guide wrapped phase retrieving. The method captures three images to reconstruct a 3D scene, but the system rate is limited due to the use of Kinect. We also proposed a reconstruction method for binary defocusing technique based on complementary decoding with dual projectors [17], which includes two strategies: the strict-alignment method and the loose-alignment method. These two strategies have some requirements for the position and posture of the dual projectors, which lead to a lack of flexibility and high algorithmic complexity.

In this paper, we present a novel binary defocusing technique based on complementary decoding with unconstrained dual projectors. A novel system framework with dual projectors and a camera is built. In our research, the position and posture of the introduced projectors are not strictly required. And, the common projection area just needs to cover the measured volume. The core idea is that low-resolution information is help for high-resolution phase unwrapping. The low-resolution information is mainly the low-resolution order of fringe in PMP, which is obtained by the low-resolution depth relation. The dual projectors have their own calibration relationship with the single camera, and there is the same calibration reference for the whole system. Therefore, the relation can be expressed as an equation between the low-resolution depth calculated by the binary PMP model and the low-resolution depth computed by our proposed low-resolution depth extraction strategy based on periodic space-time coding patterns (hereafter called LDES). In practice, through optimizing the relation, the low-resolution order is obtained. Moreover, we also propose the strategy from the low-resolution order to high-resolution order. Finally, the high-resolution phase unwrapping is easy to get from the high-resolution order. The experimental results are presented to prove the success of our proposed method, which only requires three captured images for a 3D scene and has strong robustness, flexibility, and implementation since the constraint of the position and posture of the introduced two projectors is avoided.

The organization of this paper is as follows. In Section 2, the principle of the proposed method is given in detail, which includes the theoretical system framework and the complementary decoding method with unconstrained dual projectors. In Section 3, the experiments are reported, which are used to evaluate the proposed method. Section 4 makes a conclusion for this paper.

Principle

In this section, we theoretically represent the BDT based on complementary decoding with unconstrained dual

projectors. A system framework with a camera and unconstrained dual projectors is established. Based on this, we show the core idea of this paper, i.e., low-resolution information is employed for high-resolution phase unwrapping. In addition, we explain in detail (1) the operator from wrapped phase to depth in low-resolution, (2) the proposed low-resolution depth extraction strategy based on periodic space-time coding patterns, and (3) the strategy from the low-resolution order to high-resolution order of fringe.

Theoretical system framework

Figure 1 shows the schematic diagram of system framework. The framework mainly includes a multiband camera and dual projectors. The dual projectors cast two different wavebands patterns in defocusing or focusing, respectively. In our research, we select Light-Crafter 3000 evaluation module (see red dotted circle in Fig. 1) of Texas Instruments (hereafter called EVM) as the projector since it has good programming and synchronization. The two EVMs are placed side by side, and cast blue and red patterns respectively. Next, the patterns are fused in the measured volume (see green dotted circle in Fig. 1). A color camera substitutes for the multiband camera, because of its three channels: red, green, and blue. The color camera captures the fusion images distorted by the object surface geometry. Then, the captured fusion images are separated and 3D profile of the object is reconstructed. Here, the position and posture of the two EVMs are not strictly required, but it

is noted that the common projection area should cover the measurement area.

Complementary decoding method with unconstrained dual EVMs

The proposed novel method only obtains three captured images for a 3D scene reconstruction regardless of the position and posture between the two EVMs. The core idea is that low-resolution information is employed for high-resolution phase unwrapping. Moreover, there are two important points to notice: (1) the common projection area should cover the measurement area; (2) the dual EVMs should have their own calibration relationship with the single camera, but there is the same calibration reference for the whole system. Let us suppose, the points $P(x_{low}, y_{low})$ are the low-resolution pixels, and there exists a following relation in theory.

$$D_{LDES}(x_{low}, y_{low}) = D_{PMP}(x_{low}, y_{low}) = Cal[n(x_{low}, y_{low})], \tag{1}$$

where $Cal[\cdot]$ denotes a binary PMP model, which calculates the depth $D_{PMP}(x_{low}, y_{low})$; $n(x_{low}, y_{low})$ represents the low-resolution order of fringe. The depth $D_{LDES}(x_{low}, y_{low})$ is computed by a low-resolution depth extraction strategy. In theory, because of the same calibration reference, the relation $D_{LDES}(x_{low}, y_{low}) = D_{PMP}(x_{low}, y_{low})$ is true in Eq. (1). As we all know, the purpose of the phase unwrapping is to solve the unknown ambiguity order of fringe. Therefore, in practice, Eq. (1) needs

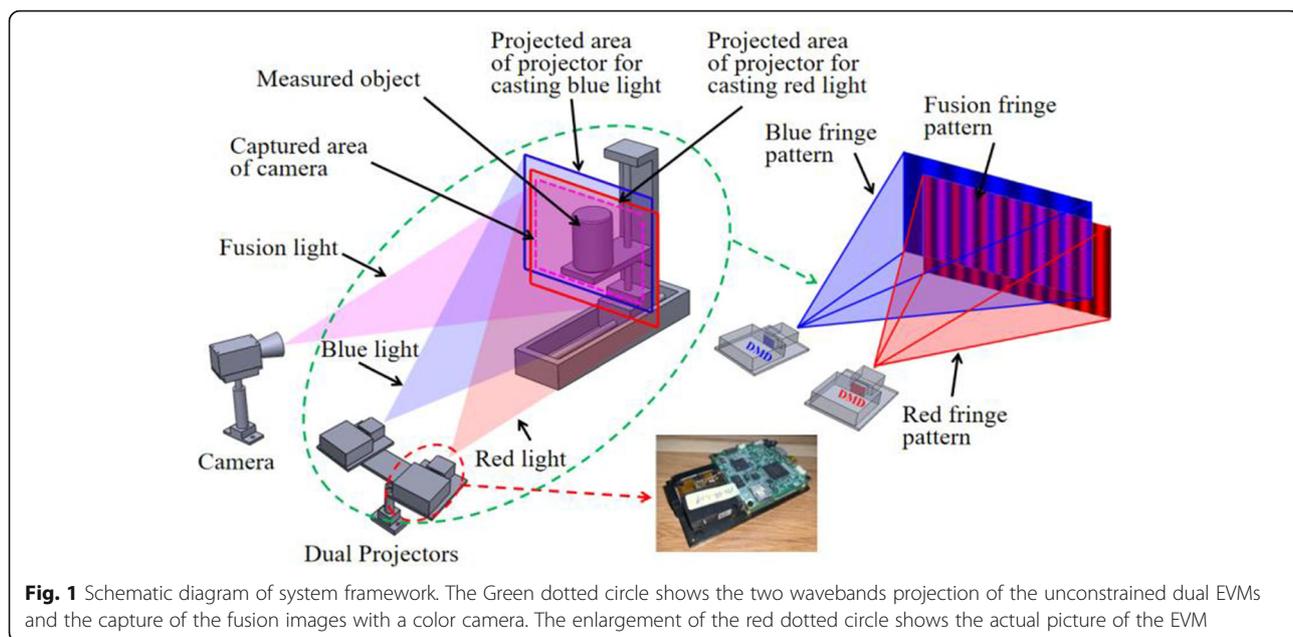


Fig. 1 Schematic diagram of system framework. The Green dotted circle shows the two wavebands projection of the unconstrained dual EVMs and the capture of the fusion images with a color camera. The enlargement of the red dotted circle shows the actual picture of the EVM

to be transformed into the optimization objective function, which can be express as:

$$\min_{n(x_{low}, y_{low})=0,1,\dots,N-1} \|D_{LDES}(x_{low}, y_{low}) - Cal[n(x_{low}, y_{low})]\|_F \quad (2)$$

The formula above means to achieve the optimal $n(x_{low}, y_{low})$ so that $D_{PMP}(x_{low}, y_{low})$ and $D_{LDES}(x_{low}, y_{low})$ are as close as possible. Then, for high-resolution phase unwrapping, the optimal $n(x_{low}, y_{low})$ needs to be extended to $n(x_{high}, y_{high})$.

Section 2.B.1 illustrates the operator $Cal[\cdot]$, which can be regard as the binary PMP model. Section 2.B.2 proposes the low-resolution depth extraction strategy based on periodic space-time coding patterns, which is used to solve $D_{LDES}(x_{low}, y_{low})$. Section 2.B.3 mainly explains the strategy from $n(x_{low}, y_{low})$ to $n(x_{high}, y_{high})$, which is used for high-resolution phase unwrapping.

Illustration of operator $Cal[\cdot]$

The operator $Cal[\cdot]$ is that how the wrapped phase $\phi(x_{low}, y_{low})$ converts to the depth $D_{PMP}(x_{low}, y_{low})$, whose procedure can be described as a five-step process [see Fig. 2 and in order to simplify representation, the (x_{low}, y_{low}) is omitted in the figure]. The five steps are illustrated in detail below.

Step 1: The binary PMP model based on 3-PS algorithm is introduced since it requires the minimum number of phase-shifting patterns [13]. For the algorithm with a phase shift of $2\pi/3$, the fringe patterns can be mathematically expressed as,

$$\begin{cases} I_1 = a + b \cdot \cos[\phi(x_{low}, y_{low}) - 2\pi/3] \\ I_2 = a + b \cdot \cos[\phi(x_{low}, y_{low})] \\ I_3 = a + b \cdot \cos[\phi(x_{low}, y_{low}) + 2\pi/3] \end{cases}, \quad (3)$$

where I is the intensity, a is the average intensity, b is

the intensity modulation and $\phi(x_{low}, y_{low})$ is the wrapped phase. As shown in Fig. 2, I_1, I_2 and I_3 are the 3-PS patterns, which are projected into the measured volume.

Step 2: The camera captures the reflected images and the wrapped phases are calculated through the following equation:

$$\phi(x_{low}, y_{low}) = \tan^{-1} \frac{\sqrt{3}[I_1 - I_3]}{2I_2 - I_1 - I_3}. \quad (4)$$

In Fig. 2, ϕ denotes the wrapped phase map of the captured images.

Step 3: The Eq. (4) provides the phase ranging $[-\pi, +\pi]$ with 2π discontinuities. That is to say, we need to retrieve $\phi(x_{low}, y_{low})$, i.e. obtain the true value of $n(x_{low}, y_{low})$:

$$\Phi(x_{low}, y_{low}) = \phi(x_{low}, y_{low}) + 2\pi n(x_{low}, y_{low}), \quad (5)$$

where $\Phi(x_{low}, y_{low})$ is the unwrapped phase. In Fig. 2, Φ denotes the unwrapped phase map of the captured images.

Step 4: The calibration model needs to introduce the relative unwrapped phase from the reference plane, i.e., the relation can be expressed as

$$\Delta\Phi(x_{low}, y_{low}) = \Phi(x_{low}, y_{low}) - \Phi_0(x_{low}, y_{low}), \quad (6)$$

where $\Phi_0(x_{low}, y_{low})$ is the unwrapped phase map of the reference plane.

Step 5: The nonlinear phase-to-height mapping [18] is introduced to obtain the depth $D_{PMP}(x_{low}, y_{low})$, and the relation can be described as,

$$D_{PMP}(x_{low}, y_{low}) = \frac{1}{\alpha + \beta/\Delta\Phi(x_{low}, y_{low}) + \gamma/\Delta\Phi(x_{low}, y_{low})^2}. \quad (7)$$

where α, β and γ represent the calibration parameters,

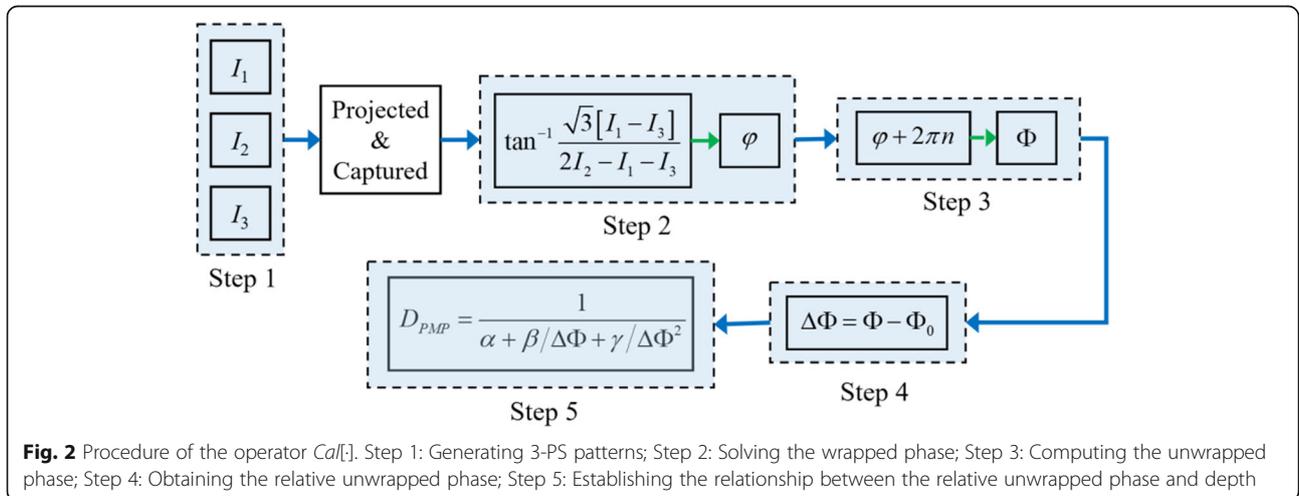


Fig. 2 Procedure of the operator $Cal[\cdot]$. Step 1: Generating 3-PS patterns; Step 2: Solving the wrapped phase; Step 3: Computing the unwrapped phase; Step 4: Obtaining the relative unwrapped phase; Step 5: Establishing the relationship between the relative unwrapped phase and depth

which are obtained by moving a flat board over known distance.

In conclusion, the procedure from step 1 to 5 represents the operator $Cal[\cdot]$. The Eq. (7) establishes the relationship between $\Delta\Phi(x_{low}, y_{low})$ and $D_{PMP}(x_{low}, y_{low})$. Moreover, based on the Eq. (5) and (6), $\Delta\Phi(x_{low}, y_{low})$ is related to the $n(x_{low}, y_{low})$ and $\phi(x_{low}, y_{low})$. $\phi(x_{low}, y_{low})$ can uniquely solve through Eq. (4). Therefore, the operator $Cal[\cdot]$ can also be described the relationship between $D_{PMP}(x_{low}, y_{low})$ and $n(x_{low}, y_{low})$, i.e., $D_{PMP}(x_{low}, y_{low}) = Cal[n(x_{low}, y_{low})]$.

Solution of depth $D_{LDES}(x_{low}, y_{low})$

For the depth $D_{LDES}(x_{low}, y_{low})$, we proposed a low-resolution depth extraction strategy based on periodic space-time coding patterns (hereafter called LDES). In the description of $Cal[\cdot]$ above, the 3-PS algorithm are introduced, i.e. one of both EVMs casts three PS patterns. Therefore, in order to balance the number of projected patterns of the dual EVMs, the other EVM also projects three patterns. That is, for the LDES, three patterns should be designed. As shown in Fig. 3, LDES has three parts: Encoding, Decoding and Calibration. Encoding is to generate three LDES patterns, which are projected into the measured volume; Decoding is to solve the decoded information from the captured LDES image; Calibration is to establish the mapping relationship from the decoded information to depth.

Encoding Figure 4 denotes a group of single period structures, which are the basis for designing the LDES patterns. In Fig. 4, the sub-figures (a), (b) and (c) show a synchronization structure and two coding structures. For the synchronization structure, the fringe edges are

used to determine the position of low-resolution points. And, according to the shift from black to white or from white to black, the code words can be represented in two states: [0, 1] or [1, 0]. Similarly, for the coding structure, there can be four states: [0, 0], [0, 1], [1, 0] or [1], where [0, 0] and [1] indicate that the both adjacent pixels are black and white, respectively. For example, in Fig. 4, the red dotted line crosses the boundary from white to black, the code words can be express as [1, 0]. Meanwhile, the red dotted line also crosses the coding structure A and B, the code words are [0, 1] and [0, 1]. Combining these three code words, the full word words can be written as [1, 0, 0, 1, 0, 1], which corresponds to the encoded value (37) from binary to decimal. Which corresponds to the encoded value: 37. Overall, a synchronization structure has two stages and each structure has four stages. So, a group of single period structures exists 32 (i.e., $32 = 2 \times 4 \times 4$) encoded states. The encoding table is shown in Table 1, which can also be used for decoding.

Generally, if a single period covers the full measurement area, the resolution will be low. Hence, the projected LDES patterns utilize the periodicity to horizontally tile the single period structures. However, the points of the same encoded value in different periods may interfere with each other. To this end, [19] represented a criterion, which determines the appropriate number of the periods, and it is as follows:

$$\begin{cases} \Delta \cdot n_t < N_x \cdot d_h/d \\ \Delta \cdot p \cdot n_t < N_x \end{cases}, \tag{8}$$

where Δ represents a stripe consisting of the number of pixels, n_t is the number of encoded values in a single

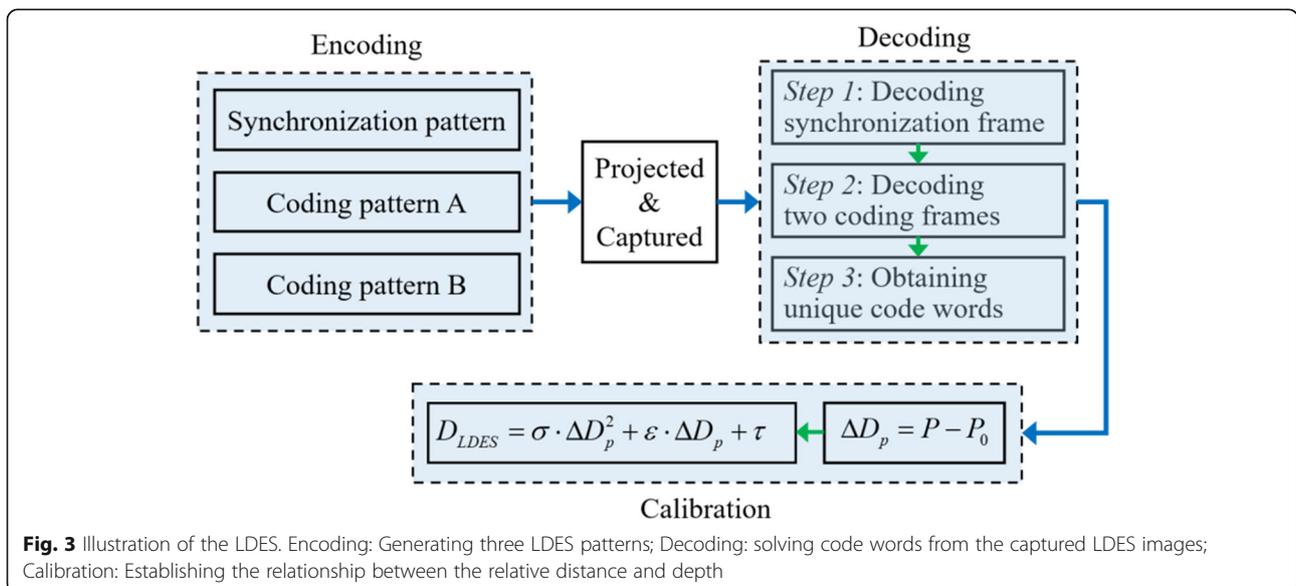
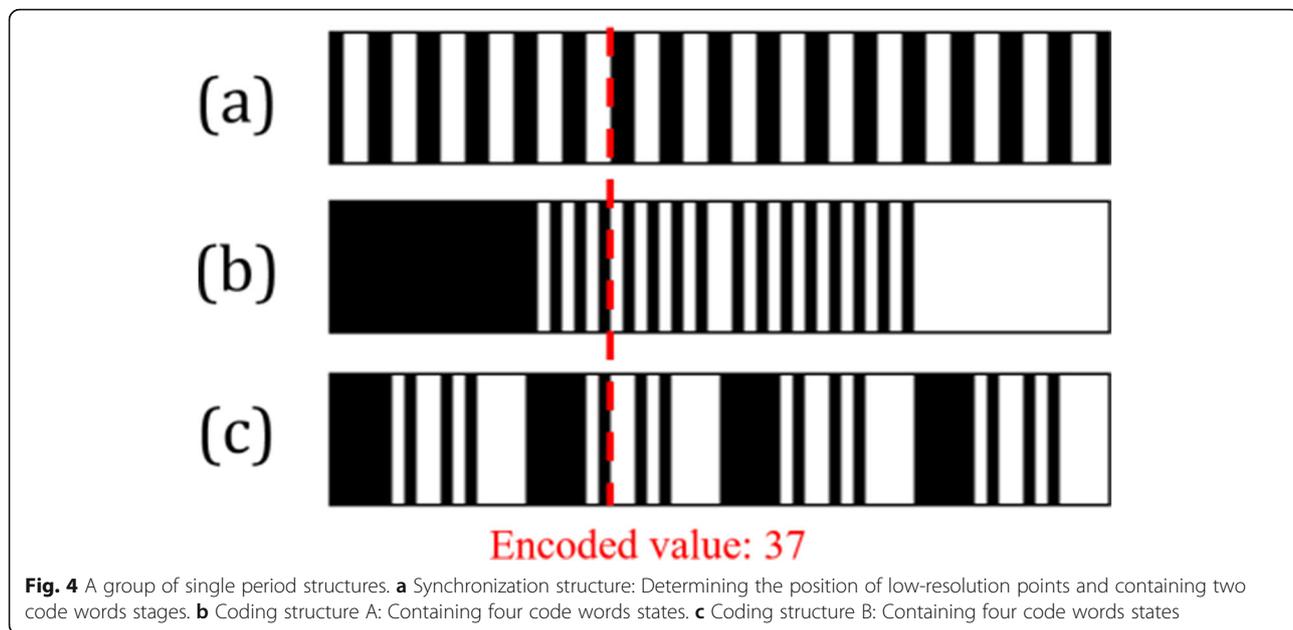


Fig. 3 Illustration of the LDES. Encoding: Generating three LDES patterns; Decoding: solving code words from the captured LDES images; Calibration: Establishing the relationship between the relative distance and depth



period, N_x is the camera horizontal resolution, d_h and d express the measurement depth and the working distance, p is the number of the periods. When the number of periods p meets Eq. (8), there are no confusions in the decoding process. i.e., the maximum value of p is crucial. In our experiment, the system parameters are approximately described as: $d_h \approx 150$ mm, $d \approx 800$ mm, $N_x = 640$ pixels, and $n_t = 32$. Based on Eq. (8), the parameter p should satisfy $p < 5.33$, and we set $p = 4$.

Decoding Three designed LDES patterns are projected into the measured volume. Then, the camera captures the reflected images. Next, the computer needs to decode them, and the procedure mainly includes following:

Step 1: For the captured synchronization image, the gradient is calculated, and according to the gradient, the edges are determined. The pixels on the edges are the low-resolution points $P(x_{low}, y_{low})$. If the gradient is

Table 1 Encoding table of 32 encoded stages for a group of single period structure

Syn.	Coding A	Coding B	Encoded value	Syn.	Coding A	Coding B	Encoded value
[0, 1]	[0, 0]	[0, 0]	16	[0, 1]	[1, 0]	[0, 0]	24
[1, 0]	[0, 0]	[0, 0]	32	[1, 0]	[1, 0]	[0, 0]	40
[0, 1]	[0, 0]	[0, 1]	17	[0, 1]	[1, 0]	[0, 1]	25
[1, 0]	[0, 0]	[0, 1]	33	[1, 0]	[1, 0]	[0, 1]	41
[0, 1]	[0, 0]	[1, 0]	18	[0, 1]	[1, 0]	[1, 0]	26
[1, 0]	[0, 0]	[1, 0]	34	[1, 0]	[1, 0]	[1, 0]	42
[0, 1]	[0, 0]	[1, 1]	19	[0, 1]	[1, 0]	[1, 1]	27
[1, 0]	[0, 0]	[1, 1]	35	[1, 0]	[1, 0]	[1, 1]	43
[0, 1]	[0, 1]	[0, 0]	20	[0, 1]	[1, 1]	[0, 0]	28
[1, 0]	[0, 1]	[0, 0]	36	[1, 0]	[1, 1]	[0, 0]	44
[0, 1]	[0, 1]	[0, 1]	21	[0, 1]	[1, 1]	[0, 1]	29
[1, 0]	[0, 1]	[0, 1]	37	[1, 0]	[1, 1]	[0, 1]	45
[0, 1]	[0, 1]	[1, 0]	22	[0, 1]	[1, 1]	[1, 0]	30
[1, 0]	[0, 1]	[1, 0]	38	[1, 0]	[1, 1]	[1, 0]	46
[0, 1]	[0, 1]	[1, 1]	23	[0, 1]	[1, 1]	[1, 1]	31
[1, 0]	[0, 1]	[1, 1]	39	[1, 0]	[1, 1]	[1, 1]	47

greater than zero, the point $P(x_{low}, y_{low})$ is on positive edge, i.e. the code words is [0, 1]. On the contrary, If the gradient is less than zero, the point $P(x_{low}, y_{low})$ is on negative edge, i.e. the code words is [1, 0].

Step 2: For the two captured coding images, they are changed into binary images. In order to enhance robustness, a local neighborhood of the point $P(x_{low}, y_{low})$ is considered for decoding. The length of the local neighborhood is set to five, i.e. there is a binary list $L_p = [p_{-2}, p_{-1}, p_0, p_1, p_2]$, where p_0 represents the value (0 or 1) of the point $P(x_{low}, y_{low})$. Similar, p_{-2}, p_{-1}, p_1 and p_2 correspond to the values of the points $P(x_{low-2}, y_{low-2})$, $P(x_{low-1}, y_{low-1})$, $P(x_{low+1}, y_{low+1})$ and $P(x_{low+2}, y_{low+2})$, respectively. The decoding rules can be described as,

$$\begin{cases} \text{grad}(p_{-2}, p_{-1}, p_0, p_1, p_2) < 0 \\ \text{grad}(p_{-2}, p_{-1}, p_0, p_1, p_2) > 0 \\ p_{-2} + p_{-1} + p_0 + p_1 + p_2 = 0 \\ p_{-2} + p_{-1} + p_0 + p_1 + p_2 = 5 \end{cases}, \quad (9)$$

For the rules above, (a) if the gradient of the list is less than zero, the code words of the point $P(x_{low}, y_{low})$ is [1, 0]; (b) if the gradient of the list is greater than zero, the code words of the point $P(x_{low}, y_{low})$ is [0, 1]; (c) if the values from p_{-2} to p_2 are all zero, the code words of the point $P(x_{low}, y_{low})$ is [0, 0]; (d) if the values from p_{-2} to p_2 are all one, the code words of the point $P(x_{low}, y_{low})$ is [1].

Step 3: Combining the code words from Step 1 and Step 2, and retrieving in Table 1, the unique encoded values of the point $P(x_{low}, y_{low})$ can be obtained.

Calibration After decoding, we build a mapping relationship between the relative distance $\Delta D_p(x_{low}, y_{low})$ and the depth $D_{LDES}(x_{low}, y_{low})$, where $\Delta D_p(x_{low}, y_{low})$ stands for the relative distance between two points with the same encoded word. The relationship can be represented as a quadratic polynomial:

$$D_{LDES}(x_{low}, y_{low}) = \sigma \cdot \Delta D_p^2(x_{low}, y_{low}) + \varepsilon \cdot \Delta D_p(x_{low}, y_{low}) + \tau, \quad (10)$$

where σ , ε and τ are the calibration parameters. Here, it is important to note that Eq. (7) and (10) should be calibrated at the same time and adopt the same calibration reference. When a flat board is moved over known distance, the dual EVMs respectively project three PS patterns and three LDES patterns on the flat board surface at one time. Then, the fusion patterns are captured and separated into two kinds of calibration images. Finally, based on the obtained calibration images, the Eqs. (7) and (10) are respectively established.

Strategy from $n(x_{low}, y_{low})$ to $n(x_{high}, y_{high})$

Optimizing Eq. (2) can obtain the low-resolution $n(x_{low}, y_{low})$. In order to retrieve the wrapped phase $\phi(x_{high}, y_{high})$, $n(x_{low}, y_{low})$ needs to be extended to $n(x_{high}, y_{high})$, where $n(x_{high}, y_{high})$ represents the high-resolution order of fringe. Normally, there are two different phase discontinuities in wrapped phase map. (a) Because of the characterization of \tan^{-1} , the wrapped phase ranges from $-\pi$ to $+\pi$ and contains 2π discontinuity (the first discontinuity for short). (b) Due to the isolated objects or the abrupt changes, the wrapped phase contains uncertain discontinuity (the second discontinuity for short). Fortunately, the first discontinuity is easily handled by the spatial phase unwrapping method. And, except for the two types of discontinuities above, the phase of the other regions is smooth and continuous. Therefore, we proposed a strategy only for the second discontinuity. The strategy is described that the wrapped phase map are segmented into several connected regions, where the low-resolution orders $n(x_{low}, y_{low})$ extend to the high-resolution orders $n(x_{high}, y_{high})$.

Intervals segmentation The main idea of the intervals segmentation is that the connected intervals are segmented row-by-row to form a flag map. Here, the criterion of the connected intervals in each row is expressed as

$$\begin{cases} \max\{d(\phi, \phi_{l1}), d(\phi, \phi_{l2})\} < TH \\ \max\{d(\phi, \phi_{r1}), d(\phi, \phi_{r2})\} < TH \end{cases}, \quad (11)$$

where $d(\phi_i, \phi_j) = \min\{|\phi_i - \phi_j|, |\phi_i - \phi_j + 2\pi|, |\phi_i - \phi_j - 2\pi|\}$; TH is a predefined threshold value, which is determined by the measured objects. In practice, for convenient recording, we build a flag map. Firstly, the flag of the leftmost pixel in each row is identified as zero. Secondly, Scan from left to right, and meanwhile verify the criterion [see Eq. (11)] for each pixel. If the current scanned pixel is an invalid pixel (e.g., in the non-object region) or violates the criterion, add one to the flag of the current scanned pixel. Otherwise, the flag of the current scanned pixel is invariable. Figure 5 shows an example of a fan object. Figure 5(a) and (b) are the wrapped phase map and its corresponding flag map, where there are three sections: A_1-A_2 , B_1-B_2 and C_1-C_2 . Figure 5(c) denotes the flag curves of the three sections. As can be seen from the three curves, when the flag curve is sloping, the corresponding wrapped phase is in invalid interval or contains the second discontinuity; when the flag curve is flat, the corresponding wrapped phase only contains the first discontinuity.

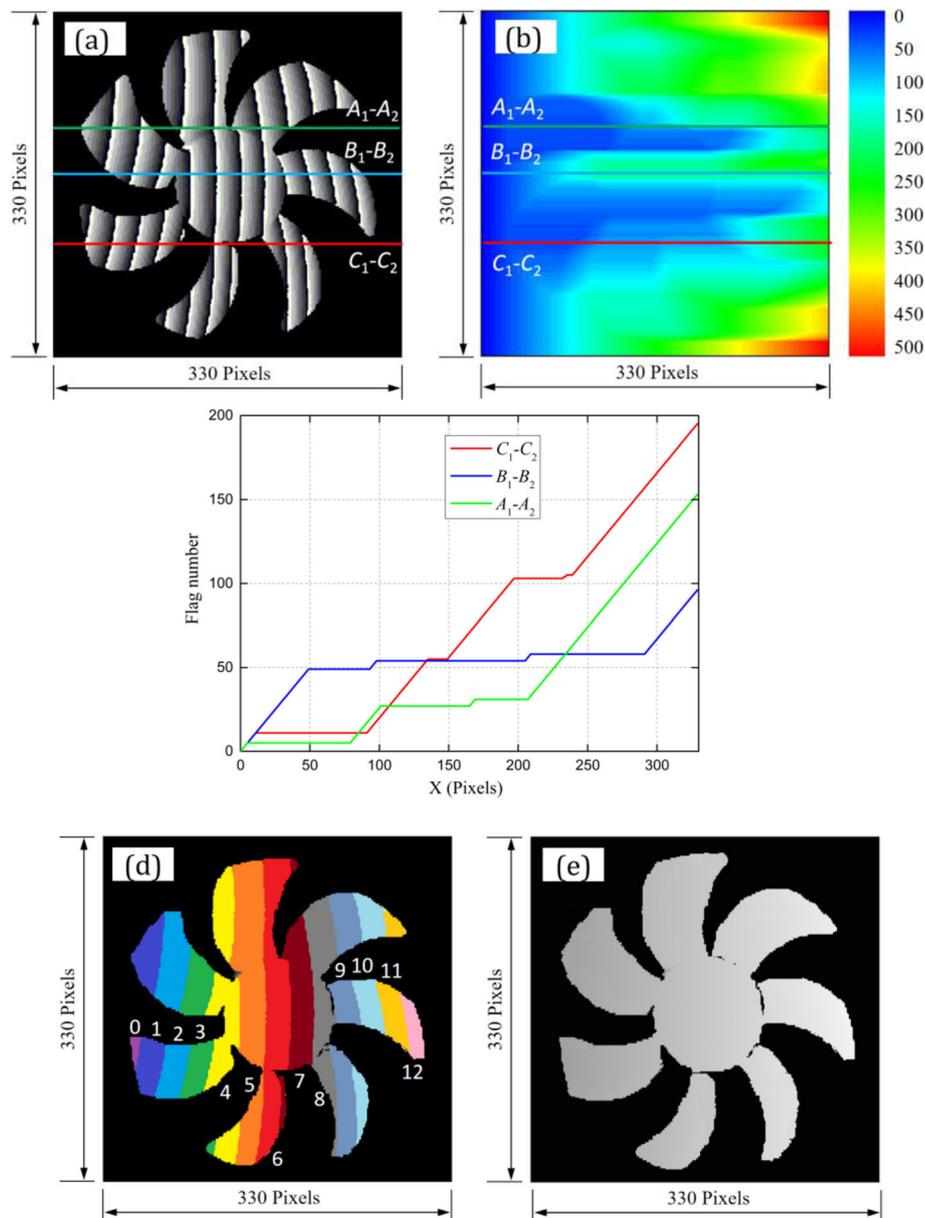


Fig. 5 Example of a fan object. **a** Wrapped phase map. **b** Corresponding flag map. **c** Flag curves of the three sections. **d** Pseudo-color map of the high-resolution orders (Numbers denote orders). **e** Unwrapped phase map

Order extension Through Sec. 2.B.1 and Sec. 2.B.2, the low-resolution orders $n(x_{low}, y_{low})$ of points $P(x_{low}, y_{low})$ are identified. Because the wrapped phase regions corresponding to the flat intervals in flag map only contains the first phase discontinuity, the low-resolution orders $n(x_{low}, y_{low})$ can be extended to the high-resolution orders $n(x_{high}, y_{high})$ with the spatial phase unwrapping method. Moreover, it is important to point out that because there may be multiple $n(x_{low}, y_{low})$ in a connected interval, the high-resolution orders $n(x_{high}, y_{high})$ of a pixel may be

multiple values as well, which cause the confusions in order extension. Hence, the identified high-resolution orders $n(x_{high}, y_{high})$ repeat most often, which is used to calculate the unwrapped phase $\Phi(x_{high}, y_{high})$. Moreover, if the low-resolution order $n(x_{low}, y_{low})$ is not found in a connected interval, the situation indicates that the interval has very few pixels, and the region should be ignored. Figure 5(d) shows the pseudo-color picture of the high-resolution orders map, and Fig. 5(e) denotes its unwrapped phase map.

Experiments

System setup

The experimental system setup mainly consists of dual EVMs and a high-speed color camera (AVT Pike F-032C). The camera attaches a 16 mm focal length ten Mega-pixel lens (VOC 16mm10MP), and its resolution is 640×480 with a maximum frame rate of 208 frames/s. The EVM has a resolution of 608×684 with a projection distance of 364–2169 mm. The dual EVMs respectively cast the red (waveband: about 600–750 nm) and the blue (waveband: about 370–550 nm) patterns. Each EVM is controlled through a different IP address (e.g., one is 192.168.1.129; the other is 192.168.1.119). In the calibration, the two waveband patterns are projected onto a white planar surface attached to a precision DAHENG motorized linear translation stage. For hardware synchronization, a dual channel signal generator (MingHe JDS 6600) is introduced.

Experimental procedure

Step 1: system configuration

The binary PS patterns and the designed LDES patterns are loaded into the dual EVMs respectively. The EVM for projecting the binary PS patterns is properly adjusted in defocusing, and the other EVM for casting

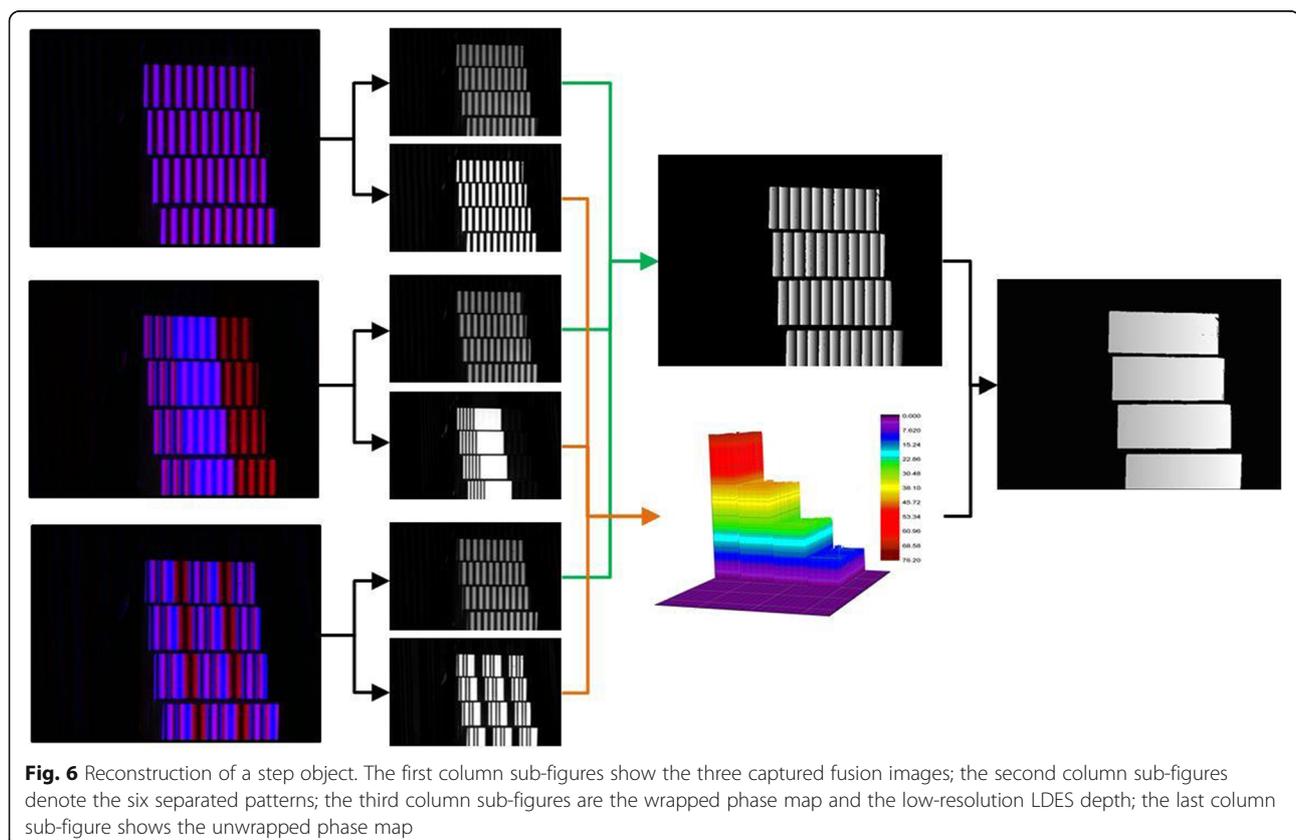
the LDES patterns projects in focusing. In addition, the output frequency of the dual channel signal generator is configured at 200 Hz, whose CH1 channel triggers the dual EVMs and CH2 channel does the camera.

Step 2: system calibration

The flat surface moves from 0 to 150 mm in intervals of 5 mm. At each position, the two waveband patterns are projected and reflected. Meanwhile, the fusion patterns are captured and separated into the PS images and the LDES images. Then, according to the relative phase $\Delta\Phi$ and the plate moving distances, the model parameters α , β and γ in Eq. (7) are solved. Similarly, according to the relative distance ΔD_p and the plate moving distances, the calibration parameters σ , ε and τ in Eq. (10) are obtained.

Step 3: practical measurement

The camera captures the three fusion patterns, which are distorted by the object surface geometry. Next, the computer separates them into the PS images and the LDES images. Then, the PS images are used to calculate the low-resolution wrapped phase $\phi(x_{low}, y_{low})$, and meanwhile $Cal[n(x_{low}, y_{low})]$ is established.

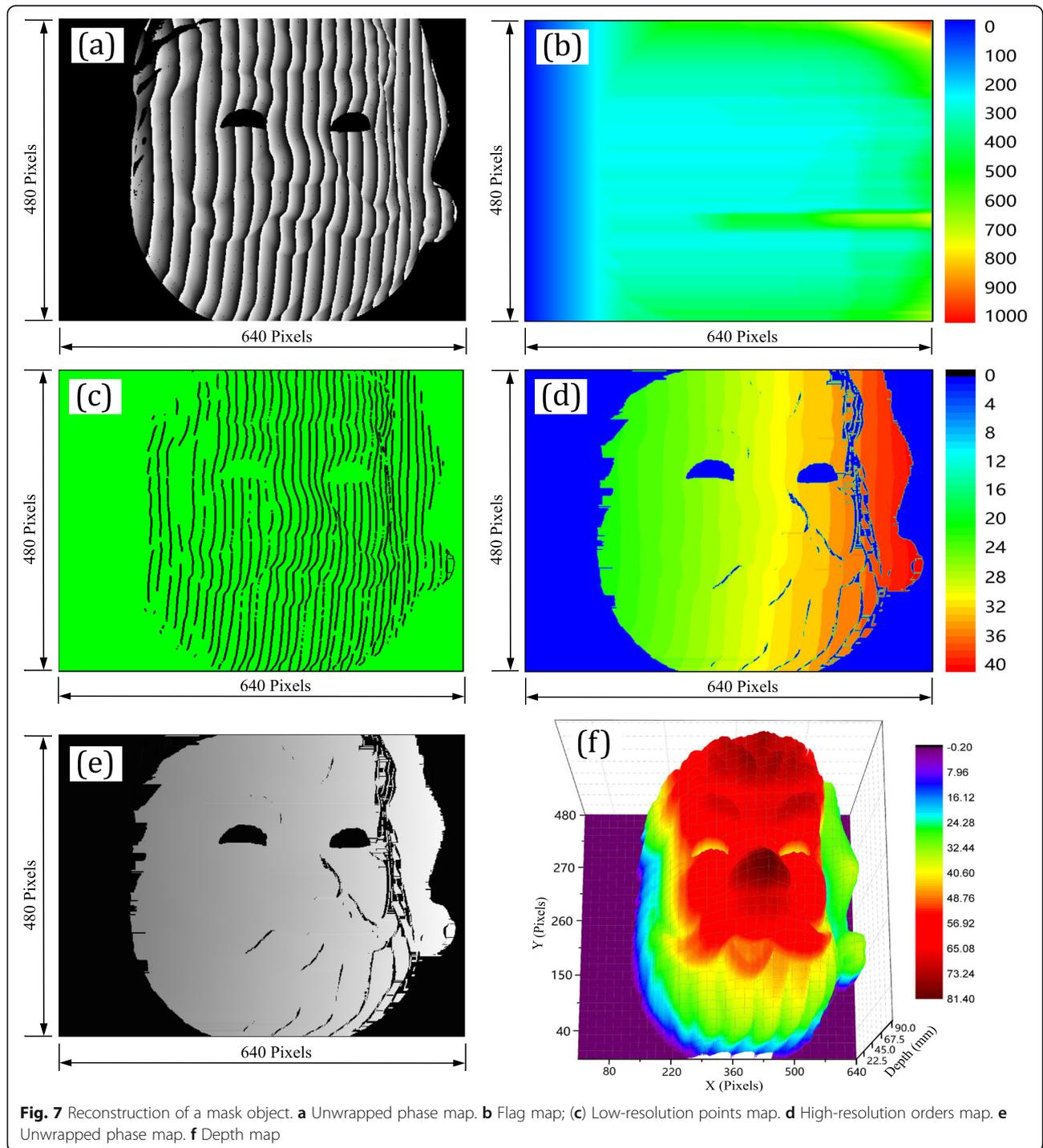


According to the decoding step and Eq. (10), the LDES images are utilized to compute the low-resolution depth $D_{LDES}(x_{low}, y_{low})$. Then, through optimizing Eq. (2) which consists of the obtained $D_{LDES}(x_{low}, y_{low})$ and $Cal[n(x_{low}, y_{low})]$, the low-resolution order $n(x_{low}, y_{low})$ is solved. Finally, based on the strategy from $n(x_{low}, y_{low})$ to $n(x_{high}, y_{high})$ in Sec. 2.B.3, the high-resolution order $n(x_{high}, y_{high})$ is obtained. Combining the PMP model,

the unwrapped phase map and the high-resolution depth is calculated.

Reconstruction of static and dynamic objects

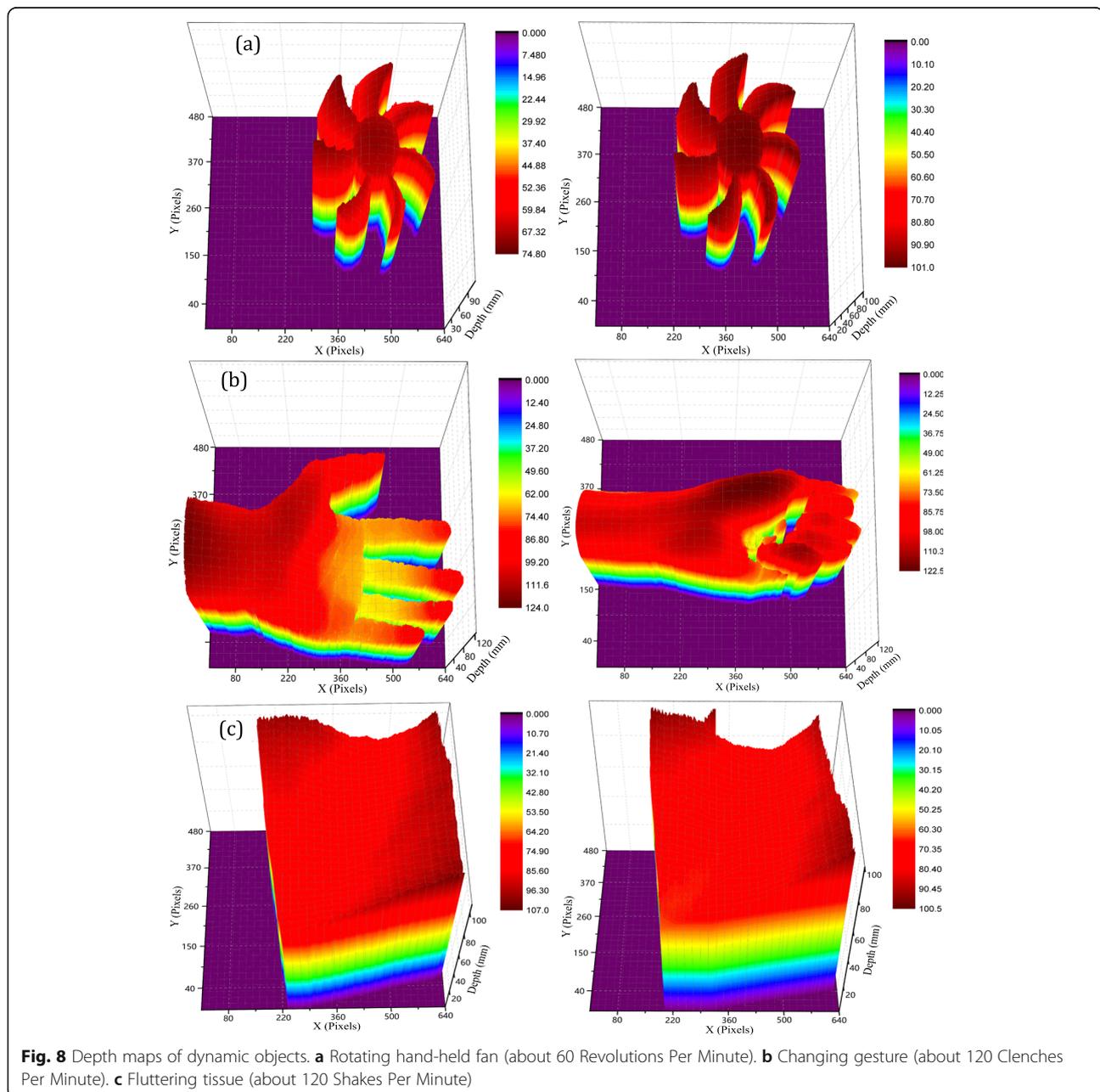
A step and a mask as the static objects are reconstructed. Figure 6 illustrates the procedure from the fusion images to the 3D profile of the step object. The first column sub-figures show the three captured fusion



images. The second column sub-figures denote the six separated patterns, which include three PS images and three LDES images. The third column sub-figures are the wrapped phase map and the low-resolution LDES depth. The last column sub-figure shows the unwrapped phase map. Similarly, we also reconstructed a mask, as shown in Fig. 7, which includes the wrapped phase map, the flag map, the low-resolution points map, the high-resolution orders map, the unwrapped phase map, and the depth map. In summary, from the unwrapped phase maps and the depth maps of the two static objects, we can clearly see that the proposed method has a good

effectiveness and robustness when only three images are captured.

For reconstruction of dynamic objects, the slow rotating fan, changing gestures, and fluttering tissue were introduced. Both moments are selected to obtain the depth maps, as shown in Fig. 8. In the sub-figures about gestures, there are some small voids in the reconstructed surface since there may be not the low-resolution orders $n(x_{low}, y_{low})$ in these regions. Fortunately, these voids do not cause error transmission of phase unwrapping. For other two dynamic objects (fan and tissue), the phase unwrapping and depth all have a nice performance.



Again, the effectiveness of our proposed method is verified, and it has a certain ability of dynamic reconstruction.

Conclusion

This paper has presented a novel BDT based on complementary decoding with unconstrained dual projectors. A system framework with dual projectors and a camera is built, where the position and posture of the introduced projectors are not strictly required. The core idea is that low-resolution information is employed for high-resolution phase unwrapping. For this purpose, we explained in detail the operator $Cal[\cdot]$ from wrapped phase to depth in low-resolution, the proposed LDES and the strategy from the low-resolution order to high-resolution order of fringe. Finally, the experimental results prove that our proposed method only requires three captured images for a 3D scene and has strong robustness, flexibility, and implementation without considering the constraint of the position and posture of the introduced two projectors.

Acknowledgements

Authors would like to thank Xing Yanfeng in SUES, whose lab supported this research.

Authors' contributions

LXX drafted the work or substantially revised it. In addition, LXX configured experiments and wrote the codes. ZWH calculated the data, wrote the manuscripts and plotted the figures.

Authors' information

1. LXX received the B.Eng. degree from Donghua University, Shanghai, China, in 2011, and the Ph.D. degree in Signal and Information from Shanghai University, Shanghai, China in 2018. He has published more than 10 refereed journal and conference articles. His main research interests include machine vision, image point cloud deep learning.
2. ZWH received the B.S. degree in textile engineering from Hebei University of Science and Technology, Shijiazhuang, China, in 2013, and the M.S. degree in Textile Material and Textiles Design from Donghua University, Shanghai, China in 2016. Her research interests include machine vision, non-contact body measurement, 3D cutting.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 26 January 2021 Accepted: 28 June 2021

Published online: 10 July 2021

References

1. Chen, F., Brown, G.M., Song, M.M.: Overview of three dimensional shape measurement using optical methods. *Opt. Eng.* **39**(1), 10–22 (2000)
2. Blais, F.: Review of 20 years of range sensor development. *Electron. Imaging.* **5013**(1), 228–240 (2003)
3. Gorthi, S.S., Rastogi, P.: Fringe projection techniques: whither we are? *Opt. Lasers Eng.* **48**(2), 133–140 (2010)

4. Zhang, Z.H.: Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques. *Opt. Lasers Eng.* **50**(8), 1097–1106 (2012)
5. Huang, L., Chua, P.S., Asundi, A.: Least-squares calibration method for fringe projection profilometry considering camera lens distortion. *Appl. Opt.* **49**(9), 1539–1548 (2010)
6. Li, X.X., Zhang, Z.J., Yang, C.: Reconstruction method for fringe projection profilometry based on light beams. *Appl. Opt.* **55**, 9895(34), (2016)
7. Karpinsky, N., Zhang, S.: High-resolution, real-time 3D imaging with fringe analysis. *Real-Time Image Process.* **7**(1), 55–66 (2010)
8. Lei, S., Zhang, S.: Flexible 3-D shape measurement using projector defocusing. *Opt. Lett.* **34**(20), 2080–2082 (2009)
9. Ayubi, G.A., Ayubi, J.A., Di Martino, J.M.: Pulse-width modulation in defocused three-dimensional fringe projection. *Opt. Lett.* **35**(21), 3682–3684 (2010)
10. Lohry, W., Zhang, S.: 3D shape measurement with 2D area modulated binary patterns. *Opt. Laser Eng.* **50**(7), 917–921 (2012)
11. Wang, Y.J., Zhang, S.: Three-dimensional shape measurement with binary dithered patterns. *Appl. Opt.* **51**(27), 6631–6636 (2012)
12. Li, B.W., Wang, Y.J., Dai, J.F.: Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Opt. Laser Eng.* **54**(5), 236–246 (2014)
13. Li, X.X., Zhang, Z.J.: High-quality fringe pattern generation based on binary pattern optimization with projector defocusing. *J. Opt. Technol.* **84**(1), 22–28 (2017)
14. Novak, M., Millerd, J., Brock, N.: Analysis of a micropolarizer array-based simultaneous phase-shifting interferometer. *Appl. Opt.* **44**(32), 6861 (2005)
15. Pan, J., Huang, P.S., Chiang, F.P.: Color phase-shifting technique for three dimensional shape measurement. *Opt. Eng.* **45**(1), 013602 (2006)
16. Zhang, Y., Xiong, Z., Wu, F.: Fusion of time-of-flight and phase shifting for high-resolution and low-latency depth sensing, 2015. IEEE international conference on multimedia and expo (ICME). IEEE (2015)
17. Li, X.X., Zhang, W.H.: Reconstruction method for binary defocusing technique based on complementary decoding with dual projectors. *IEEE Access.* **9**, 29200–29206 (2021). <https://doi.org/10.1109/ACCESS.2021.3059659>
18. Li, W., Su, X., Liu, Z.: Large-scale three-dimensional object measurement: a practical coordinate mapping and image data-patching method. *Appl. Opt.* **40**(20), 3326–3333 (2001)
19. Xu, L., Zeng, D., Zhang, Z.J.: Alternate time-space coding for structured light system. *Opt. Eng.* **47**(12), 127201 (2008)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)