

Three-dimensional absolute shape measurement by combining binary statistical pattern matching with phase-shifting methods

YATONG AN AND SONG ZHANG*

School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907, USA

*Corresponding author: szhang15@purdue.edu

Received 8 March 2017; revised 1 June 2017; accepted 1 June 2017; posted 2 June 2017 (Doc. ID 290301); published 26 June 2017

This paper presents a novel method that leverages the stereo geometric relationship between projector and camera for absolute phase unwrapping on a standard one-projector and one-camera structured light system. Specifically, we use only one additional binary random image and the epipolar geometric constraint to generate a coarse correspondence map between projector and camera images. The coarse correspondence map is further refined by using the wrapped phase as a constraint. We then use the refined correspondence map to determine a fringe order for absolute phase unwrapping. Experimental results demonstrated the success of our proposed method. © 2017 Optical Society of America

OCIS codes: (120.0120) Instrumentation, measurement, and metrology; (120.2650) Fringe analysis; (100.5070) Phase retrieval.

<https://doi.org/10.1364/AO.56.005418>

1. INTRODUCTION

Three-dimensional (3D) shape measurement can be used in many applications. For example, it can be used to do quality control in manufacturing, disease diagnoses in medical practices, as well as others [1,2].

In general, phase-based methods are much more accurate and more robust than intensity-based methods [3]. They can achieve higher spatial resolution and provide denser 3D points than intensity-based methods. The extensively employed phase-retrieval methods include those using Fourier transform [4–6] and phase-shifting-based fringe analysis algorithms [7]. The former methods can use sinusoidal grating [8], binary grating [9,10], or periodical lattice grid pattern [11–13] to carry the desired phase information. The latter primarily uses binary [14,15] or sinusoidal grating patterns [7] as the phase carrier. Typically, these phase-retrieval methods can provide a wrapped phase that ranges only from $-\pi$ to π with a modulus of 2π due to the use of an arctangent function. To obtain a continuous phase map for 3D reconstruction, a phase-unwrapping method is needed to remove those 2π discontinuities. In essence, phase unwrapping is to determine an integer number (or a fringe order) $k(x, y)$ of 2π 's for each pixel so that the problem of 2π discontinuities can be resolved.

Over the past many decades, many phase-unwrapping methods have been developed. In general, they can be classified into two categories: spatial and temporal phase unwrapping. Spatial phase-unwrapping methods identify 2π discontinuous locations on a wrapped-phase map and remove them by adding

or subtracting an integer number of 2π 's. They determine the fringe order $k(x, y)$ of a point by analyzing the phase values of its neighboring pixels on the wrapped-phase map. These methods typically yield only a relative phase because the unwrapped phase is relative to the starting point of the unwrapping process. The book edited by Ghiglia and Pritt [16] summarized many spatial phase-unwrapping methods, and Su and Chen [8] reviewed many robust quality-guided phase-unwrapping algorithms. Despite these developments, spatial phase unwrapping is fundamentally limited to measure “smooth” surfaces, i.e., they assume that there is no larger than π phase change introduced by object surface geometry between two successive pixels in at least one unwrapping path. Therefore, it is very challenging for spatial phase-unwrapping methods to be used if one wants to measure objects with abrupt depth changes, or to simultaneously measure absolute geometries of multiple isolated objects.

Temporal phase-unwrapping methods overcome the difficulties of spatial phase-unwrapping methods. Over the years, many temporal phase-unwrapping methods have been developed, including two- or multi-wavelength phase shifting [17–19], phase coding [20–22], and gray coding [23,24]. Unlike spatial phase-unwrapping methods where phase can be directly unwrapped from the wrapped-phase map, temporal phase-unwrapping methods require the acquisition of additional image(s). Therefore, temporal phase-unwrapping methods sacrifice measurement speed to resolve the fundamental limitation of spatial phase-unwrapping methods.

To alleviate the slowed measurement speed problem of temporal phase-unwrapping methods, researchers attempted to reduce the number of additional image acquisitions for absolute phase recovery. Li *et al.* [25] proposed a phase-unwrapping method that uses a single additional image that consists of 6 types of slits. These slits form a pseudo random sequence, and the fringe order is identified by the position of a sub-sequence of those slits from an entire sequence. Zhang [26] attempted to use a single-stair image where the stair changes coincide with those 2π discontinuities, such that the fringe order can be determined from the single-stair image. Zuo *et al.* [27] attempted to encode wrapped phase and base phase into four fringe patterns. All these attempts using one additional image for absolute phase unwrapping work well under certain conditions. However, they all require substantial subsequent image processing in the intensity domain, making it difficult for them to measure objects with rich texture.

In general, the state-of-the-art temporal phase-unwrapping methods ignore the inherent stereo geometric constraints [28] between projector and camera images for fringe-order determination. This paper proposes a novel method that leverages the stereo geometric relationship between projector and camera for absolute phase unwrapping. In brief, we project a binary random image onto an object using a projector and capture the random pattern by camera. Based on the epipolar geometric constraints and the additional binary random pattern, we generate a coarse correspondence map between projector and camera images. We then refine the coarse correspondence map using wrapped phase as a constraint. The refined correspondence map is then used to determine the fringe order $k(x, y)$ for phase unwrapping. Experimental results demonstrate the success of our proposed method.

Section 2 explains the principles of the proposed method. Section 3 presents experimental results to validate the proposed method. Lastly, Section 4 summarizes this paper.

2. PRINCIPLE

In this section, we will thoroughly explain the principles behind our proposed method. Specifically, we will briefly introduce the three-step phase-shifting algorithm and then elucidate the entire framework of our proposed method.

A. Three-Step Phase-Shifting Algorithm

Over the years, many fringe analysis methods have been developed, including phase-shifting-based fringe analysis methods and Fourier transform-based fringe analysis methods. Among all these fringe analysis methods, phase-shifting-based methods are extensively adopted due to their robustness and accuracies. The three-step phase-shifting algorithm is desirable for high-speed measurement applications because it uses the least number of fringe images for pixel-wise phase retrieval.

Mathematically, three phase-shifted fringe patterns with equal phase shifts can be described as

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\varphi(x, y) - 2\pi/3], \quad (1)$$

$$I_2(x, y) = I'(x, y) + I''(x, y) \cos[\varphi(x, y)], \quad (2)$$

$$I_3(x, y) = I'(x, y) + I''(x, y) \cos[\varphi(x, y) + 2\pi/3], \quad (3)$$

where $I'(x, y)$ is the average intensity, which is also *texture* (a photograph of an object). $I''(x, y)$ is the intensity modulation, and $\varphi(x, y)$ is the phase that is used for 3D reconstruction. By solving Eqs. (1)–(3), we can obtain

$$I'(x, y) = \frac{I_1(x, y) + I_2(x, y) + I_3(x, y)}{3}, \quad (4)$$

$$\varphi(x, y) = \tan^{-1} \left\{ \frac{\sqrt{3}[I_1(x, y) - I_3(x, y)]}{2I_2(x, y) - I_1(x, y) - I_3(x, y)} \right\}. \quad (5)$$

Due to the use of an arctangent function, the phase obtained from Eq. (5) is a wrapped phase whose value ranges from $-\pi$ to π with a modulus of 2π . To remove those 2π discontinuities and obtain a continuous phase map, a phase-unwrapping algorithm is needed. Mathematically, the relationship between the wrapped phase φ and the unwrapped phase Φ can be described as

$$\Phi(x, y) = \varphi(x, y) + 2\pi \times k(x, y), \quad (6)$$

where $k(x, y)$ is often referred to as fringe order. The goal of phase unwrapping is to obtain an integer number $k(x, y)$ for each pixel, such that the unwrapped phase obtained from Eq. (6) is continuous without 2π discontinuities.

As discussed in Section 1, numerous spatial and temporal phase-unwrapping methods have been developed. The spatial phase unwrapping gives only relative phase and is limited to measure “smooth” surfaces; and the temporal phase unwrapping can be used to measure arbitrary surfaces and obtain absolute phase, but requires additional image(s). To our knowledge, none of the existing temporal phase-unwrapping methods fully utilizes the geometric constraints of the structured light system for fringe-order determination. In this paper, we propose a new absolute phase-unwrapping method that leverages the inherent stereo constraints between projector and camera images, which will be discussed next.

B. Proposed Method

1. Framework Overview

Figure 1 shows the overall framework of our proposed method. We use four patterns in total: three phase-shifted patterns and one random binary pattern. From three phase-shifted images, we compute the texture I' using Eq. (4) and the wrapped phase φ using Eq. (5). We then binarize the camera-captured random image by comparing its gray values with the gray values of the texture image I' pixel by pixel. Utilizing the binarized camera random image I^b and the original projector binary random image, we generate a correspondence map between projector and camera images through block matching based on stereo geometric constraints. The correspondence map created in this stage cannot achieve pixel-level accuracy due to the fundamental limitation of the correspondence calculation algorithm and the quality of these two input random images, and we call this level of correspondence as *coarse* correspondence. To generate a precise correspondence map, we refine the coarse correspondence map using the wrapped phase φ as a constraint to adjust the corresponding point for each pixel. Once a precise corresponding map is obtained, we can determine the precise corresponding projector pixel for each camera pixel and thus the fringe order $k(x, y)$ for each pixel. Finally, we unwrap the entire

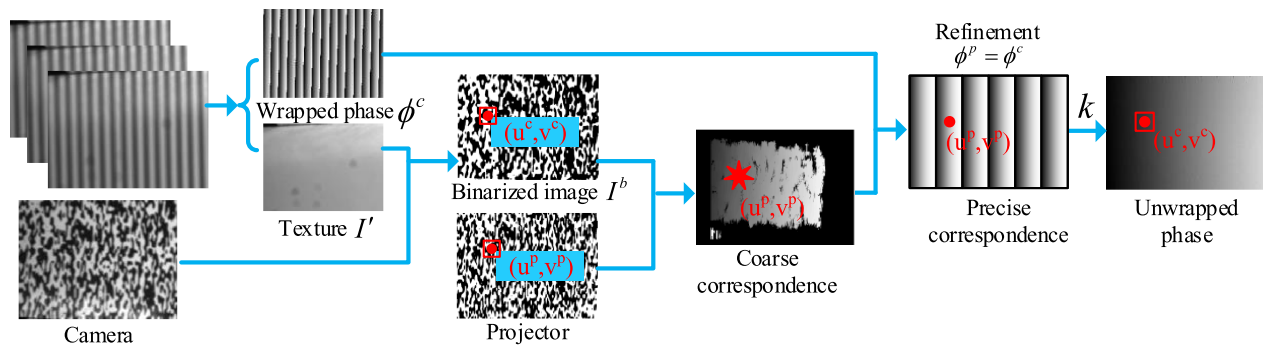


Fig. 1. Framework overview of the proposed phase-unwrapping method. We use three phase-shifted patterns and one binary random pattern. From three phase-shifted images, we generate wrapped-phase ϕ , and texture I' . Then we binarize the camera random image by comparing it with the texture image. Based on the binarized camera random image and the original projector random pattern, we generate a coarse correspondence. The wrapped-phase constraint is utilized to do refinement and get correspondence that is more precise. From the precise correspondence map, we determine a fringe order, and it can be used for the final phase unwrapping.

phase map for 3D reconstruction. The rest of this section details the entire framework we proposed.

2. Epipolar Geometry

To understand our proposed method, the epipolar geometry of the standard stereo-vision system has to be introduced first. This subsection explains the general concepts of epipolar geometry and how to use epipolar geometry as constraints for stereo correspondence establishment.

For a standard stereo vision system that includes two cameras, each camera can be mathematically modeled as a well-established pinhole system [28]. If both cameras are calibrated under the same coordinate system, the projection from the world coordinates to each camera pixel coordinates can be established. After calibration, there exists the inherent geometric relationship between a point on one camera and all its possible corresponding pixels on the other camera; the method to establish the geometric constraints is often referred to as *epipolar geometry*.

The epipolar geometry mainly constrains the corresponding pixels of one camera image pixel to be a line on the other camera image. Figure 2 illustrates the epipolar geometric constraints. Here O^l and O^r , respectively, indicate the focal point of the left camera lens and the focal point of the right camera lens; the intersection points between line $\overline{O^l O^r}$ and the image planes E^l and E^r are called *epipoles* in the stereo system. For a pixel P^l on the left camera image, it can correspond to multiple points P_1, P_2, P_3 at different depths in a 3D space.

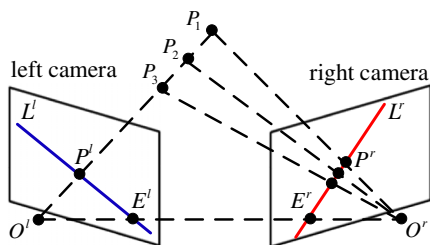


Fig. 2. Epipolar geometry constraint: one pixel on one image can correspond to only one line (called epipolar line) on the other image.

Though these points are different in a 3D space, they all fall on the same line on right camera image L^r , which is called the *epipolar line*. From similar geometric relationships, all pixels on line L^l on the left image can correspond only to points on line L^r on the right camera image. More details about epipolar geometry can be found in Ref. [28].

The advantage of using epipolar geometry is that it can: (1) improve computational efficiency when searching for correspondence between two images, and (2) reduce the chances of finding false corresponding points. For a pair of images captured from different perspectives, the corresponding pixels of one pixel on one camera image can lie only on the epipolar line of the other camera image. Thus, the correspondence searching is reduced to a simpler 1D problem instead of the original complex 2D problem.

To facilitate the correspondence searching, we usually rotate and translate the original images such that the corresponding epilines are on the same row; this process is called *image rectification*. Figure 3 illustrates the results after image rectification. All epipolar lines in both images become horizontal, and the corresponding pixels on one image can appear only on the same row of the other image. Image rectification can be done by using some well-established algorithms or toolboxes, such as those provided by OpenCV. Because correspondence can happen only on the same row (or column) on a pair of rectified images, searching for correspondence will be more efficient.

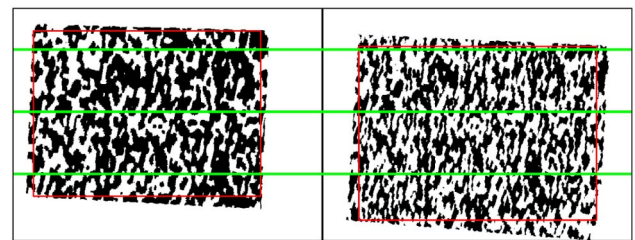


Fig. 3. Image rectification process aligns the epipolar lines on the same row to speed up the searching process. Green lines shown on the image are epipolar lines, including all possible corresponding points for those pixels.

Since the transformation matrices from the original image to rectified image coordinate systems are known, it is straightforward to determine the correspondence points on the original images by inverting the transformation process.

The disparity between two corresponding pixels on different images is related to the depth of an object point. Intuitively, the disparity increases as the distance from an object to the system decreases. As shown in Fig. 2, when an object point P moves closer to the system from P_1 to P_3 , length difference between $\overline{E^l P^l}$ and $\overline{E^r P^r}$ increases. In other words, the disparity between corresponding pixels P^l and P^r increases as the object point moves closer to the system. Those apparent corresponding pixel differences in a pair of images is often called a *disparity map*. Similarly, if we have a disparity map between a pair of images, we can infer the correspondence relationship between pixels on those two images.

3. Coarse Correspondence Generation

The epipolar geometry concepts were primarily developed in the computer vision field, and most open-source disparity map generation algorithms were developed for a standard stereo system: two cameras have the same sensor size, the same optics, and the same pixel size, and the true corresponding points on two camera images roughly follow the Lambertian model (i.e., an object has equal reflectance in different perspective angles). However, a standard structured light system consisting of a projector and a camera usually does not conform to those assumptions for the standard stereovision system.

1. *The sensor parameters are different.* The projector and the camera usually have different sensor sizes, different physical pixel sizes, and different system lenses.

2. *They have different image generation mechanisms.* Projector images are ideal computer-generated images that are not affected by the lens, the object, the environment settings, etc. In contrast, camera images are affected by all of these.

Those two major differences violate the basic assumptions of many existing disparity map generation algorithms developed for a standard stereovision system, and thus it is difficult to directly adopt any of them without modifications.

For most intensively employed disparity map generation algorithms developed in the field of stereovision, two input images are expected to have a similar field of view (FOV) and same resolution. Yet, the difference in sensor parameters can cause different FOVs and different resolutions between projector and camera images. To mimic a standard stereovision system, we crop and down sample (or up sample) either projector or camera images to match the resolution for a similar FOV. We first crop either projector or camera images to ensure that their FOVs are similar for a designed working distance; and then we down sample (or up sample) one of the images to match the number of pixels between two cropped images. This preprocessing step allows the use of many existing disparity map generation algorithms to generate reasonable quality results.

The Lambertian surface assumption of the standard stereovision system does not hold for a structured light system since the projector image is computer generated while the camera image is practically captured. One of the differences is that the camera image could be distorted, depending on the location

of the object, while the projector image is not. For example, part of the camera image could be blurred, as shown in Fig. 4(a), while the projector image is always perfectly focused, as shown in Fig. 4(c). Therefore, the correspondence searching is very challenging. To address that problem, we propose to use a random binary pattern and binarize the camera-captured image to preserve the reliable features. The binary image is suggested because of its robustness to variations of non-ideal situations where the camera image is generated.

The binary random image is generated using band-limited $1/f$ noise where $\frac{1}{20 \text{ pixels}} \leq f \leq \frac{1}{5 \text{ pixels}}$. The range of f controls the granularity of the random pattern, and the selection of f is optimized for each system, depending upon settings of the hardware and configuration of the system.

Figure 4(c) shows the ideal random binary pattern, denoted by I^p , which is projected onto a flat, white board by the projector. Figure 4(a) shows the corresponding image captured by the camera, denoted by I^c . The non-ideal camera image shows the nonuniform brightness and nonuniform focus level across the entire image, while the projector image does not depict any of these nonuniform issues. We binarize the camera image by comparing the texture image, I' , generated from phase-shifted fringe images, with the camera-captured random image, I^c pixel by pixel:

$$I^b(u, v) = \begin{cases} 255, & I^c(u, v) > I'(u, v), \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Figure 4(d) shows the binarized camera image shown in Fig. 4(a). Clearly, the binarized camera visually appears closer to the projector-projected image than the original camera image appears. Therefore, intuitively, the binarization preprocess can increase the success of conventional stereo disparity map generation algorithms.

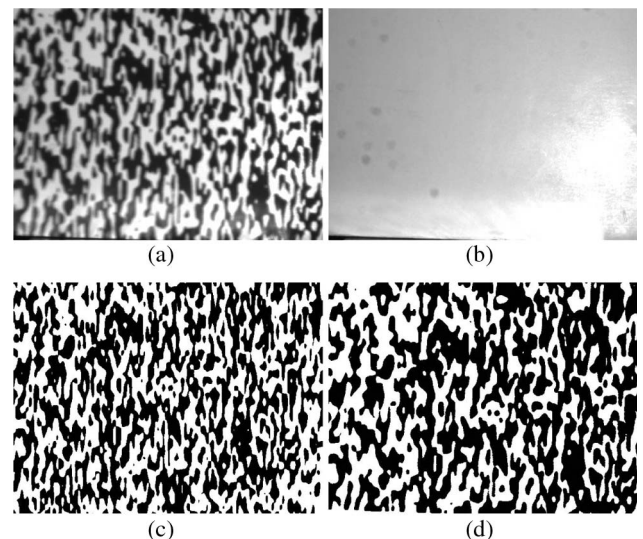


Fig. 4. Random pattern and preprocessing of the camera-captured image. (a) Random image captured by a camera I^c ; (b) texture I' computed from phase-shifted fringe images; (c) binary random projector pattern I^p ; (d) binarized camera image I^b from the camera image shown in (a).

In this research, we used the block-matching algorithm [29,30] to generate a disparity map between the projector image I^p and the binarized camera image I^b . Basically, the block matching divides one image into small rectangle blocks. For a block of pixels on one image, blocking matching calculates a cost function at each possible location in a search area (along the epipolar lines) on the other image. The location with the lowest cost is the best correspondence point.

Though block matching is commonly used because of its simplicity, the sizes of image blocks and search areas have a strong impact on its matching accuracy. A small block size generates more accurate matching, but produces redundancy and can be easily affected by noise. A large block size can tolerate more noise but results are less accurate since it is influenced by a lot more surrounding pixels. The optimal block size varies for different scenes and different parts of an object. Usually an optimal block size is determined by checking an overall disparity map quality, but detail parts always contain many artifacts, such as holes and missing boundary. In addition, block matching usually cannot achieve pixel-level accuracy due to noise in two input images and the use of a block of pixels for matching. Figure 5(a) shows an example of the disparity map in the rectified image coordinate system between Figs. 4(c) and 4(d) after applying the block-matching algorithm. On this disparity map, the gray value of each pixel represents the position difference between corresponding camera and projector pixels. Obviously, there are many holes, and the boundary is very rough. To address these problems, a refinement stage is required, which will be discussed next.

4. Refinement

To bring back the lost information (i.e., holes and some missing boundary) of a raw disparity map directly generated from the block-matching algorithm, we perform cubic spline interpolation along both row and column directions. We call this step *hole filling*. Taking a row of pixels as an example, we extract all the indices of pixels that have reliable disparity values, such as $\{(i_{k_1}, d_{k_1}), (i_{k_2}, d_{k_2}), \dots, (i_{k_n}, d_{k_n})\}$, where i_{k_j} is the index of the pixel position in a row, and d_{k_j} represents the disparity value of pixel i_{k_j} . We use points with reliable disparity values as “knots” to interpolate missing points between a pair of knots, (i_{k_j}, d_{k_j}) and $(i_{k_{j+1}}, d_{k_{j+1}})$, with cubic polynomials $d_k = q_k(i)$. It is well known that the cubic function $q_k(i)$ is in the form of $a_k + b_k i + c_k i^2 + d_k i^3$, or $d_k = a_k + b_k i + c_k i^2 + d_k i^3$, where a_k , b_k , c_k , and d_k are parameters of a cubic function. To obtain these four unknown coefficient parameters, we fit an initial cubic polynomial function q_1 based on the first four points on a row $\{(i_{k_1}, d_{k_1}), (i_{k_2}, d_{k_2}), (i_{k_3}, d_{k_3}), (i_{k_4}, d_{k_4})\}$, and determine those four unknowns a_1, b_1, c_1, d_1 by simultaneously solving four equations:

$$d_{k_j} = a_1 + b_1 i_{k_j} + c_1 i_{k_j}^2 + d_1 i_{k_j}^3, \quad (8)$$

where $j = 1, \dots, 4$. For each following point (i_{k_j}, d_{k_j}) , where $j = 5, \dots, n$, we also fit a cubic polynomial function and denote it as q_{j-3} . The parameters of these new cubic polynomials are determined by the following boundary conditions:

$$q_{j-3}(i_{k_{j-1}}) = d_{k_{j-1}}, \quad (9)$$

$$q_{j-3}(i_{k_j}) = d_{k_j}, \quad (10)$$

$$\left. \frac{d}{di}(q_{j-3}) \right|_{i=i_{k_{j-1}}} = \left. \frac{d}{di}(q_{j-4}) \right|_{i=i_{k_{j-1}}}, \quad (11)$$

$$\left. \frac{d^2}{di^2}(q_{j-3}) \right|_{i=i_{k_{j-1}}} = \left. \frac{d^2}{di^2}(q_{j-4}) \right|_{i=i_{k_{j-1}}}, \quad (12)$$

where $a_{j-3}, b_{j-3}, c_{j-3}, d_{j-3}$ are four unknowns. Once the cubic polynomial function is determined, the missing disparity point i_m can be interpolated based on the cubic function. For example, if i_m locates between two knots i_{k_t} and $i_{k_{t+1}}$, then the interpolated disparity value $d_m = q_{t-3}(i_m)$.

Not only for each row, we also do cubic spline interpolation for each column. Therefore, those points inside holes are interpolated twice: along row and column directions, and we take the average of these two interpolation results. For boundary missing disparities, usually they are interpolated only in either row or column direction, and thus we directly take one direction interpolation result. Figure 5(b) shows the result after hole filling.

Meanwhile, as mentioned in Section B.3, a disparity map can provide only a coarse correspondence, but not a precise correspondence. To achieve higher correspondence accuracy, we use the wrapped-phase value as the constraint for refinement. This is because if the correspondence is precise, the corresponding pixels should have the same wrapped-phase values as from the camera and the projector. This constraint is used to adjust pixel correspondence between projector and camera images, and we call it *phase constraint* for the rest of this paper; the step of using phase constraint to refine the correspondence determination is called *refinement*.

Specifically, the process to refine a correspondence map based on the phase constraint works as follows. Suppose there is a pixel (u^c, v^c) on a camera image; the corresponding projector pixel is (u^p, v^p) , according to a coarse disparity map generated from block matching. For the coarse correspondence given by a disparity map, it is possible that $\varphi(u^c, v^c) \neq \varphi(u^p, v^p)$, and thus we should slightly adjust the value of u^p or v^p to make sure

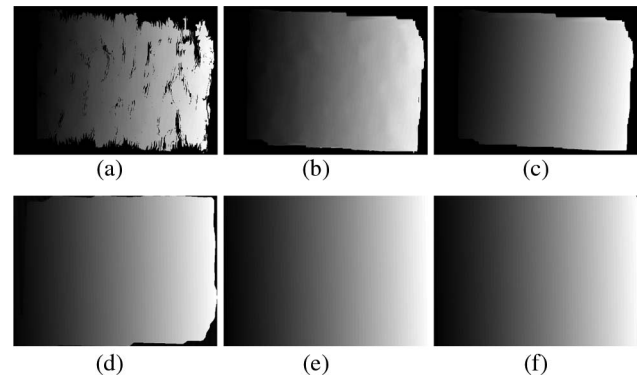


Fig. 5. Disparity map generation and refinements. (a) Raw disparity map directly generated from block matching; (b) disparity map after hole filling; (c) refined disparity map based on phase constraint; (d) unwrapped phase from (b); (e) unwrapped phase after extrapolation; (f) final unwrapped phase map. The difference between the first-row and second-row images is that (a)–(c) are in the rectified image coordinate system, and (d)–(f) are in the original camera image coordinate system.

that $\varphi(u^c, v^c) = \varphi(u^{p*}, v^{p*})$, where (u^{p*}, v^{p*}) is the refined corresponding projector pixel around (u^p, v^p) . By enforcing the phase constraint, we can make sure that (u^{p*}, v^{p*}) on the projector is the precisely corresponding pixel for the camera pixel (u^c, v^c) .

Figure 5(c) shows the disparity map after filling missing holes and refining the coarse corresponding map with phase constraint. It is clearly much smoother than the coarse disparity map shown in Fig. 5(b), as expected. However, it still shows missing areas on the boundaries, comparing the texture image shown in Fig. 4(b). The reason for this problem is that the boundary information on the original disparity map does not present, and the hole-filling step cannot fill them either. Figure 5(d) shows the unwrapped phase in the original camera image coordinate system based on the refined disparity map shown in Fig. 5(c). The missing boundaries on a disparity map can lead to the zigzag boundaries on the unwrapped-phase map. A straightforward way to solve this problem is to extrapolate the disparity map to fill those values. However, from our experience, extrapolation on a disparity map is quite challenging because disparity values have no clear property that we can utilize. In contrast, an unwrapped phase is guaranteed to be monotonous in either row or column direction. Thus, we propose to do extrapolation on the unwrapped phase to fill those still missing boundary points.

Specifically, we use a linear or low-order polynomial fitting to fill the missing boundary unwrapped-phase values. Take a row from the unwrapped phase map as an example; we extract all pixels that have an unwrapped phase, such as $\{(i_m, \Phi_m), (i_{m+1}, \Phi_{m+1}), \dots, (i_k, \Phi_k)\}$, where m is the starting point, and k is the ending point of a segment on which we have unwrapped-phase information. To fill the unwrapped phase values for boundary pixels i_1, i_2, \dots, i_{m-1} and $i_{k+1}, i_{k+2}, \dots, i_n$, we fit a linear or low-order polynomial function based on pixels from i_m to i_k and their unwrapped phase values $\Phi_m, \Phi_{m+1}, \dots, \Phi_k$. Based on the fitted polynomial function, we then predict the missing boundary unwrapped-phase values $\Phi_1, \Phi_2, \dots, \Phi_{m-1}$ and $\Phi_{k+1}, \Phi_{k+2}, \dots, \Phi_n$. Since the predicted unwrapped-phase values could be slightly different from the truth, we again refine the unwrapping result using the phase constraint that guarantees the same wrapped-phase values. Finally, we calculate fringe orders k based on those extrapolated unwrapped-phase values, and unwrap the entire phase. Figure 5(e) shows the results after extrapolation, and Fig. 5(f) shows the final unwrapped phase that can be used for 3D reconstruction.

It is important to note the visualization difference between the first-row images and the second-row of images in Fig. 5. The perspective difference is a result of showing them in different coordinate systems where images shown in Figs. 5(a)–5(c) are in the rectified image coordinate system, and Figs. 5(d)–5(f) show images in the original camera image coordinate system.

3. EXPERIMENT

We built a structured light system to experimentally verify the capability of our proposed method. The system consists of one charge-coupled device (CCD) camera (Model: The Imaging Source DMK 23U618) and one digital-light processing (DLP) projector (Model: DELL M115HD). The camera resolution is

640 × 480, and it is attached with a 12-mm-focal-length lens (Model: Computar M1214-MP2). The projector's native resolution is 1280 × 800 with a 14.95-mm-fixed-focal-length lens. The baseline between the projector and camera is approximately 60 mm, and the distance between the measured objects and the structured light system is approximately 460 mm. We adopted the method proposed by Zhang and Huang [31] to calibrate our structured light system.

We first measured a single sphere. The images are pre-processed to address the problems associated with different resolutions, pixel sizes, and FOVs between projector and camera images. We cropped the projector image to be 640 × 480 pixels for this experiment, such that the camera and projector have similar FOV and the same resolution at the location of the sphere. On the cropped projector image area, we project three phase-shifted patterns and one binary random pattern. The fringe period of phase-shifted patterns is 18 pixels. Figure 6(a) shows one of the three fringe images. The random binary pattern we used is the same as the one shown in Fig. 4(c). Figure 6(b) shows the random pattern image captured by the camera. From three phase-shifted fringe images, we can obtain the texture and wrapped-phase map, which are shown in Figs. 6(c) and 6(d), respectively.

We generated a coarse correspondence map between the binarized camera random image and projector random pattern by applying the block-matching algorithm. As discussed in Section B.3, we binarize the random image by comparing it with the texture using Eq. (7). Figure 7(a) shows the binarized image. Figure 7(b) shows the original projector binary random pattern. Both the camera and projector binary images are rectified so that the correspondences are on the same row, such as pixels on the green lines in Figs. 7(a) and 7(b). Through block matching, we generated a coarse disparity map, which is shown in Fig. 7(c). On this disparity map, the gray value of each pixel

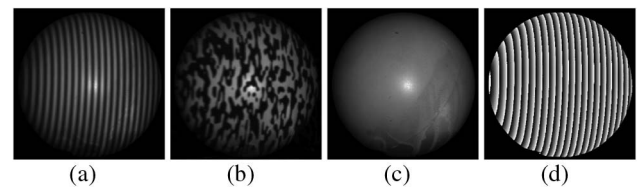


Fig. 6. Experiment on a sphere. (a) One of the three fringe images; (b) random pattern captured by the camera; (c) texture computed from fringe images; (d) wrapped phase from fringe images.

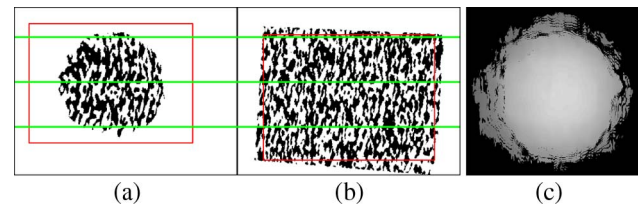


Fig. 7. Coarse correspondence generation. (a) Rectified binarization result of camera random image; (b) rectified projector binary random pattern; (c) rough disparity map generated by block matching.

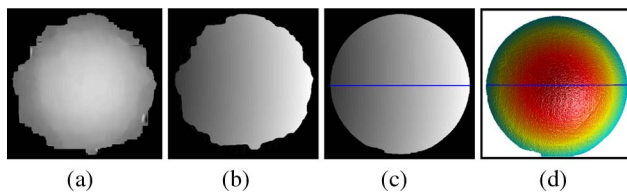


Fig. 8. Experiment results of a sphere. (a) Refined disparity map; (b) unwrapped phase using the refined disparity map and phase constraint; (c) boundary filling result using polynomial fitting and phase constraint; (d) 3D reconstruction result.

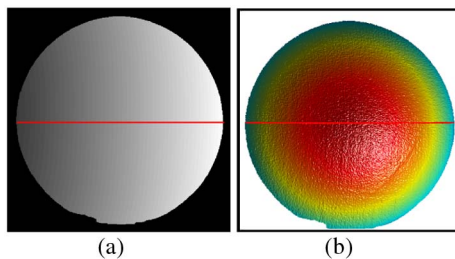


Fig. 9. Measurement result of the sphere using the gray-coding method. (a) Unwrapped phase map; (b) 3D reconstructed geometry reconstructed.

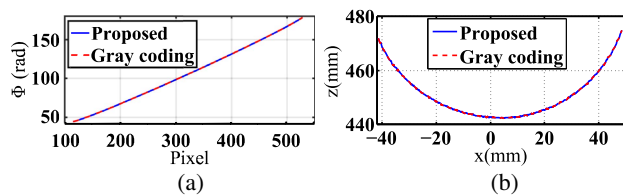


Fig. 10. Comparison between the results from the gray-coding method and our proposed method for the single sphere measurements. (a) Cross sections comparison on the unwrapped phase shown in Fig. 8(c), marked as the solid blue line, and Fig. 9(a), marked as the dashed red line; (b) cross sections comparison on the unwrapped phase shown in Fig. 8(d), marked as the solid blue line, and Fig. 9(d), marked as the dashed red line.

represents the position difference between corresponding camera and projector pixels.

The coarse disparity map shown in Fig. 7(c) is not precise and has artifacts (e.g., holes, missing boundary). We filled holes on the disparity map using the cubic spline interpolation along both row and column directions. Figure 8(a) shows the disparity map after hole filling. Since the coarse disparity map cannot

achieve pixel-level accuracy, the phase constraint was used to refine the correspondence map. Figure 8(b) shows the unwrapping result after applying the phase constraint to the coarse disparity map after hole filling. The unwrapped-phase map shown in Fig. 8(b) is accurate, but some boundary areas are still missing, as expected. We fitted a 2-order polynomial on each row and column to extrapolate those missing boundary areas, and again refined the extrapolated unwrapped-phase map using the phase constraint. Figure 8(c) shows the final unwrapped phase map after refinement. Figure 8(d) shows the final 3D reconstruction result based on the unwrapping phase. Clearly, the reconstructed geometry is continuous and smooth, suggesting our proposed method works well for a single object.

We further experimentally verified the phase quality by comparison with a traditional temporal phase-unwrapping method. Instead of using one random pattern, we use seven gray-coded binary patterns to determine fringe order k and temporally unwrap the wrapped phase [32]. Figure 9(a) shows the unwrapped-phase map for the sphere, and Fig. 9(b) shows the 3D reconstruction result.

We took a cross section of the unwrapped phase and 3D geometry and compared them with the results obtained from our proposed method. Figures 10(a) and 10(b) show the results. Clearly, our results perfectly overlap with the results obtained from the gray-coding method, confirming that our unwrapping result is correct and accurate. The gray-coding method is a well-known temporal phase-unwrapping method, and the recovered phase is absolute. Thus, our approach also generates an absolute phase map.

Since our method can obtain absolute phase, it can be used to simultaneously measure multiple isolated objects. To verify this capability, we measured two isolated objects, as shown in Fig. 11(a). These two objects were at a similar position as the sphere, and we did the same preprocessing of cropping to make sure the projector and camera had the same resolution and similar FOVs. Figure 11(b) shows the wrapped phase, and Fig. 11(c) shows the random pattern captured by the camera. We used block matching to generate a disparity map, and then applied the same hole filling, refinement, unwrapped-phase extrapolation, and refinement procedures as the sphere experiment. The final unwrapping phase is shown in Fig. 11(d), and Fig. 11(e) shows the 3D reconstruction result using our proposed method.

Once again, we compared our proposed method with the gray-coding method. The unwrapped phase through gray coding is shown in Fig. 12(a), and the reconstructed 3D geometry is shown in Fig. 12(b). They appear the same as the results we obtained from our proposed method. We then took cross

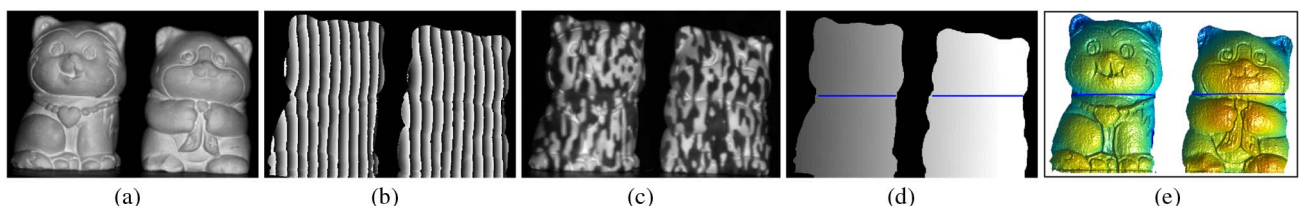


Fig. 11. Experiment result on two isolated objects. (a) Texture of the two isolated objects to be measured; (b) wrapped phase; (c) random pattern captured by the camera; (d) unwrapped-phase map using our proposed method; (e) 3D reconstruction result.

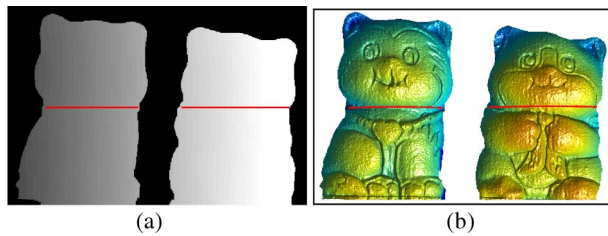


Fig. 12. Measurement result of two isolated objects using the gray-coding method. (a) Unwrapped-phase map for the sphere using the gray-coding method; (b) 3D geometry reconstructed using the gray-coding method.

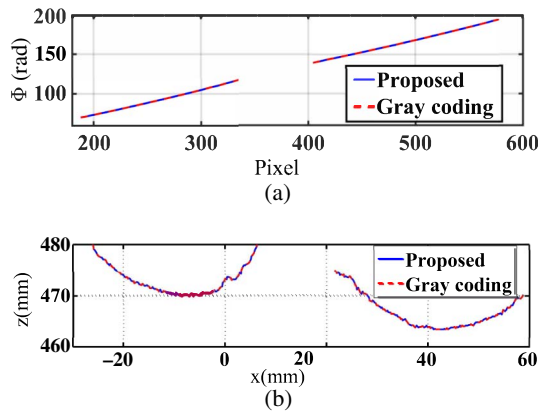


Fig. 13. Comparison between the results from the gray-coding method and our proposed method for the two isolated object measurements. (a) Cross section comparisons on the unwrapped phase shown in Fig. 11(d), marked as solid blue line, and Fig. 12(a), marked as dashed red line; (b) cross section comparisons on the unwrapped phase shown in Fig. 11(e), marked as solid blue line, and Fig. 12(b), marked as dashed red line.

sections of the unwrapped phase and 3D geometry and compared the results of our proposed method with the gray-coding results. Figures 13(a) and 13(b) show the results. Once again, our results perfectly overlap with the results from gray coding. All these experiments demonstrated the success of our proposed method for absolute phase unwrapping. Compared with gray coding, we use only one additional pattern instead of seven, thus our method is more applicable for high-speed measurement conditions.

4. SUMMARY

This paper has presented a new absolute phase-unwrapping method that utilizes the inherent stereo geometric relationship between projector and camera. The proposed method requires only one additional binary random image. Based on the epipolar geometric constraint and one additional binary random image, we generate a coarse correspondence map between projector and camera images. The coarse correspondence map is further refined by using the wrapped-phase constraint. We then use the refined correspondence map to determine the fringe order for absolute phase unwrapping. Experiments demonstrated

the success of our proposed method through measuring both a single object and multiple isolated objects. Since only one additional pattern is required to generate an absolute phase map, it has advantages for high-speed measurement applications.

Funding. Directorate for Engineering (ENG) (CMMI-1531048).

Acknowledgment. The views expressed in this paper are those of the authors and not necessarily those of the NSF.

REFERENCES

1. S. Zhang, *High-Speed 3D Imaging with Digital Fringe Projection Techniques*, 1st ed. (Taylor & Francis/CRC Press, 2016).
2. B. Li, Y. An, D. Cappelleri, J. Xu, and S. Zhang, "High-accuracy, high-speed 3D structured light imaging techniques and potential applications to intelligent robotics," *Int. J. Intell. Robot. Appl.* **1**, 86–103 (2017).
3. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," *Opt. Lasers Eng.* **48**, 149–158 (2010).
4. M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," *Appl. Opt.* **22**, 3977–3982 (1983).
5. K. Qian, "Windowed Fourier transform for fringe pattern analysis," *Appl. Opt.* **43**, 2695–2702 (2004).
6. B. Li, Y. An, and S. Zhang, "Single-shot absolute 3D shape measurement with Fourier transform profilometry," *Appl. Opt.* **55**, 5219–5225 (2016).
7. D. Malacara, *Optical Shop Testing* (Wiley, 2007).
8. X. Su and W. Chen, "Reliability-guided phase unwrapping algorithm: a review," *Opt. Lasers Eng.* **42**, 245–261 (2004).
9. Y. Gong and S. Zhang, "Ultrafast 3-D shape measurement with an off-the-shelf DLP projector," *Opt. Express* **18**, 19743–19754 (2010).
10. W. Lohry and S. Zhang, "Fourier transform profilometry using a binary area modulation technique," *Opt. Eng.* **51**, 113602 (2012).
11. M. Hÿtch, E. Snoeck, and R. Kilaas, "Quantitative measurement of displacement and strain fields from HREM micrographs," *Ultramicroscopy* **74**, 131–146 (1998).
12. M. Takeda and J. Suzuki, "Crystallographic heterodyne phase detection for highly sensitive lattice-distortion measurements," *J. Opt. Soc. Am. A* **13**, 1495–1500 (1996).
13. Z. Liu, X. Huang, H. Xie, X. Lou, and H. Du, "The artificial periodic lattice phase analysis method applied to deformation evaluation of TiNi shape memory alloy in micro scale," *Meas. Sci. Technol.* **22**, 125702 (2011).
14. S. Lei and S. Zhang, "Flexible 3-D shape measurement using projector defocusing," *Opt. Lett.* **34**, 3080–3082 (2009).
15. T. Xian and X. Su, "Area modulation grating for sinusoidal structure illumination on phase-measuring profilometry," *Appl. Opt.* **40**, 1201–1206 (2001).
16. D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software* (Wiley, 1998).
17. Y. Y. Cheng and J. C. Wyant, "Two-wavelength phase shifting interferometry," *Appl. Opt.* **23**, 4539–4543 (1984).
18. Y. Y. Cheng and J. C. Wyant, "Multiple-wavelength phase-shifting interferometry," *Appl. Opt.* **24**, 804–807 (1985).
19. C. E. Towers, D. P. Towers, P. David, and J. D. Jones, "Optimum frequency selection in multifrequency interferometry," *Opt. Lett.* **28**, 887–889 (2003).
20. Y. Wang and S. Zhang, "Novel phase-coding method for absolute phase retrieval," *Opt. Lett.* **37**, 2067–2069 (2012).
21. D. Zheng and F. Da, "Phase coding method for absolute phase retrieval with a large number of codewords," *Opt. Express* **20**, 24139–24150 (2012).
22. C. Zhou, T. Liu, S. Si, J. Xu, Y. Liu, and Z. Lei, "An improved stair phase encoding method for absolute phase retrieval," *Opt. Lasers Eng.* **66**, 269–278 (2015).

23. G. Sansoni, M. Carocci, and R. Rodella, "Three-dimensional vision based on a combination of gray-code and phase-shift light projection: analysis and compensation of the systematic errors," *Appl. Opt.* **38**, 6565–6573 (1999).
24. Q. Zhang, X. Su, L. Xiang, and X. Sun, "3-D shape measurement based on complementary Gray-code light," *Opt. Lasers Eng.* **50**, 574–579 (2012).
25. Y. Li, C. Zhao, Y. Qian, H. Wang, and H. Jin, "High-speed and dense three-dimensional surface acquisition using defocused binary patterns for spatially isolated objects," *Opt. Express* **18**, 21628–21635 (2010).
26. S. Zhang, "Composite phase-shifting algorithm for absolute phase measurement," *Opt. Lasers Eng.* **50**, 1538–1541 (2012).
27. C. Zuo, Q. Chen, G. Gu, S. Feng, and F. Feng, "High-speed three-dimensional profilometry for multiple objects with complex shapes," *Opt. Express* **20**, 19493–19510 (2012).
28. R. Hartley and A. Zisserman, eds., *Multiple View Geometry in Computer Vision*, 2nd ed. (Cambridge University, 2004).
29. A. Kaehler and G. Bradski, eds., *Learning OpenCV*, 2nd ed. (O'Reilly Media, 2014).
30. C. Je and H. M. Park, "Optimized hierarchical block matching for fast and accurate image registration," *Signal Process.* **28**, 779–791 (2013).
31. S. Zhang and P. S. Huang, "Novel method for structured light system calibration," *Opt. Eng.* **45**, 083601 (2006).
32. S. Zhang, "Flexible 3D shape measurement using projector defocusing: extended measurement range," *Opt. Lett.* **35**, 934–936 (2010).