

# High-speed absolute three-dimensional shape measurement using three binary dithered patterns

William Lohry and Song Zhang\*

Department of Mechanical Engineering, Iowa State University, Ames, IA, 50011, USA

[\\*song@iastate.edu](mailto:song@iastate.edu)

**Abstract:** This paper describes a method to reconstruct high-speed absolute three-dimensional (3D) geometry using only three encoded 1-bit binary dithered patterns. Because of the use of 1-bit binary patterns, high-speed 3D shape measurement could also be achieved. By matching the right camera image pixel to the left camera pixel in the object space rather than image space, robust correspondence can be established. Experiments demonstrate the robustness of the proposed algorithm and the potential to achieve high-speed 3D shape measurements.

© 2014 Optical Society of America

**OCIS codes:** (120.0120) Instrumentation, measurement and metrology; (110.6880) Three-dimensional image acquisition; (120.5050) Phase measurement.

---

## References and links

1. X. Mei, X. Sun, M. Zhou, H. Wang, X. Zhang, "On building an accurate stereo matching system on graphics hardware," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on* pp. 467–474 (2011).
2. J. Davis, R. Ramamoorthi, and S. Rusinkiewicz, "Spacetime stereo: A unifying framework for depth from triangulation," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* **2**, 359 (2003).
3. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision* **47**, 7–42 (2002).
4. A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Computer Vision–ACCV 2010* (Springer, 2011), pp. 25–38.
5. M. Bleyer, C. Rhemann, and C. Rother, "PatchMatch Stereo-Stereo Matching with Slanted Support Windows," in *BMVC* **11**, 1–11 (2011).
6. P. Heise, S. Klose, B. Jensen, and A. Knoll, "PM-Huber: PatchMatch with Huber Regularization for Stereo Matching," in *Computer Vision (ICCV), 2013 IEEE International Conference on* pp. 2360–2367 (2013).
7. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence* **24**, 603–619 (2002).
8. B. Scholkopf, S. Mika, C. J. Burges, P. Knirsch, K. Muller, G. Ratsch, and A. J. Smola, "Input space versus feature space in kernel-based methods," *IEEE Trans. Neural Networks* **10**, 1000–1017 (1999).
9. G.-Q. Wei, W. Brauer, and G. Hirzinger, "Intensity-and gradient-based stereo matching using hierarchical Gaussian basis functions," *IEEE Trans. Pattern Analysis and Machine Intelligence* **20**, 1143–1160 (1998).
10. X. Zhou and P. Boulanger, "New eye contact correction using radial basis function for wide baseline videoconferencing system," in *Advances in Multimedia Information Processing–PCM 2012* (Springer, 2012), pp. 68–79.
11. Y. Xu, D. Wang, T. Feng, and H.-Y. Shum, "Stereo computation using radial adaptive windows," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on* **3**, 595–598 (2002).
12. S. Mattoccia, S. Giardino, and A. Gambini, "Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering," in *Computer Vision–ACCV 2009* (Springer, 2010), pp. 371–380.

13. C. Reich, R. Ritter, and J. Thesing, "3-D shape measurement of complex objects by combining photogrammetry and fringe projection," *Optical Engineering* **39**, 224–231 (2000).
  14. D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* **1**, I–195 (2003).
  15. C. Bräuer-Burchardt, P. Kühmstedt, and G. Notni, *Combination of Sinusoidal and Single Binary Pattern Projection for Fast 3D Surface Reconstruction* (Springer, 2012).
  16. W. Lohry, V. Chen, and S. Zhang, "Absolute three-dimensional shape measurement using coded fringe patterns without phase unwrapping or projector calibration," *Opt. Express* **22**, 1287–1301 (2014).
  17. Z. Yang, Z. Xiong, Y. Zhang, J. Wang, and F. Wu, "Depth Acquisition from Density Modulated Binary Patterns," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on* pp. 25–32 (2013).
  18. P. Wissmann, R. Schmitt, and F. Forster, "Fast and accurate 3D scanning using coded phase shifting and high speed pattern projection," in *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011 International Conference on* pp. 108–115 (2011).
  19. B. Li, Y. Wang, J. Dai, W. Lohry, and S. Zhang, "Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques," *Opt. Laser Eng.* **54**, 236–246 (2014), (invited); doi:10.1016/j.optlaseng.2013.07.010.
  20. H. Hirschmuller and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," *IEEE Trans. Pattern Analysis and Machine Intelligence* **31**, 1582–1599 (2009).
  21. E. A. Nadaraya, "On estimating regression," *Theory of Probability & Its Applications* **9**, 141–142 (1964).
  22. H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* **2**, 807–814 (2005).
- 

## 1. Introduction

The three-dimensional (3D) shape measurement field is rapidly growing and finds applications in a diverse set of areas. Recent progress for depth measurement using two cameras is of particular interest because of the lower cost of cameras and ease of calibration. The class of stereo capture techniques can be split into both passive [1] and active [2] methods.

Early techniques in stereo vision determined each pixel's disparity by computing a cost within a window as the pixel is shifted across the target image [3]. While this technique can work well for areas with high texture variations, it suffers when a template can have multiple matches across the image. Further, the technique implicitly assumes the disparity within the template window is constant, which is violated for areas with high geometry variation.

More recent techniques in stereo vision have tried to account for this non-constant disparity within a window [4–6]. A recent algorithm developed by Geiger et al. uses corresponding feature points between images to generate triangles that can be compared [4]. Because the cost is computed on the interpolated triangle, it is less prone to areas of large variance of disparity. In addition, there are other techniques account for the disparity variance by using slanted windows [5, 6].

Kernel-based techniques have been extensively used for image analysis and some have been employed for stereo matching [7–9]. Comaniciu used grayscale or color intensity information to segment images with a Nadaraya-Watson estimator [7]. A comparison between image space and feature space is provided in [8]. Wei introduced a disparity optimization technique using weighted Gaussian functions, which allows disparity estimation to occur in an image deformation space [9]. However, because it works in the image domain, it works best for regions with high-texture variation and applies a smoothness constraint to non-feature areas. Another algorithm uses radial basis function (RBF) as a post-processing step, allowing regions of inaccurate disparity to be filled in by neighboring pixels with accurate disparities, more accurate disparities [10]. However, the RBF method is not applied to the initial disparity computation itself. Other algorithms use a radially-weighted support window for computation [11, 12] but still require a high texture variation for accurate matching.

Despite these numerous techniques developed for stereo vision, relying solely on the object's inherent texture information presents particular challenges. These challenges could be some-

what alleviated by adding a projected texture and especially phase information. Stereo depth maps have been obtained using two cameras with absolute phase [13, 14], but doing so requires spatial or temporal unwrapping.

Phase-shifting and a coded pattern within four images (three phase-shifted fringes and an additional coded pattern) provided stereo depth information in [15]. To eliminate the additional pattern, techniques for an embedded coded-pattern in both three-step [16, 17] and four-step [18] phase-shifted patterns have been introduced. However, these techniques have some limitations. In [17], a low-pass filter is applied to the captured image itself, resulting in loss of high-frequency information in the geometry. In [16], the encoded pattern aids an existing stereo algorithm, and thus the phase is not used during matching itself. Finally, [18] requires more than the minimum three images for phase-shifting, which is not desirable for high-speed capture.

The aforementioned phase-based stereo-vision techniques obtain phase maps by projecting sinusoidal fringe images, which are often represented using 8-bit grayscale values, limiting the measurement speeds. In contrast, if only 1-bit binary patterns are needed, the measurement speed can be at least one order of magnitude faster [19].

This paper introduces a novel computational framework that only uses three phase-shifted binary patterns for absolute 3D shape measurement. A phase-based stereo-vision technique is used to acquire high-speed geometry even in the presence of geometric distortion. Rather than requiring phase unwrapping or projector calibration, an embedded pattern within three binary dithered phase-shifted fringes provides correspondence between the images. A potential match between two cameras is compared by undistorting the correspondence pattern using the captured phase, so that comparison occurs in 3D object space rather than image space. This undistortion accounts for geometric distortion caused by the parallax between cameras, and reduces false matches. In this research, we have verified the success of the proposed technique by developing a system capable of projecting structured patterns at 4,000 Hz, resulting in a potential 3D shape acquisition rate of 1,333 Hz. However, due to the speed limitation of the camera we used, the system can only achieve 50 Hz 3D shape measurement speed.

Section 2 explains the principles of the proposed computational framework that includes binary pattern generation, phase computation, stereo coded image remapping, and phase-based stereo-matching. Section 3 shows experimental results verifying the performance of the proposed computational framework. Section 4 summarizes this paper.

## 2. Principle

In this research we propose a novel computational framework that could achieve both high-speed and high accuracy 3D shape measurement. The framework contains three major components:

1. Binary dithered pattern generation method that generates three phase-shifted binary dithered sinusoid patterns embedded with coded information in the quality map;
2. Phase remapping algorithm that remaps the code from one camera to match the perspective of the second camera; and,
3. Stereo matching algorithm that establish precise correspondence by aggregating the cost from the remapped coded pattern.

The rest of this section will detail each of these major components.

## 2.1. Binary dithered pattern generation

This paper used three-step patterns modified in two ways: first, by embedding a coded pattern for correspondence to those three fringe patterns; second, by applying the random dithering technique to generate the desired binary patterns. The embedded pattern reduces errors in stereo matching, while the random dithering allows high-speed capture.

The intensities for three-step phase shifting patterns are written as:

$$I_k(x,y) = I'(x,y) + I''(x,y) \cos \left[ \phi(x,y) - \frac{2\pi k}{3} \right]. \quad (1)$$

Here  $I'(x,y)$  is the average intensity,  $I''(x,y)$  the intensity modulation, and  $\phi(x,y)$  the projected phase.

Similar to the method discussed in our previous paper [16], to improve correspondence using stereo matching, a coded pattern can be embedded within the projected fringe pattern. In these experiments, band-limited  $1/f$  noise where  $\frac{1}{20\text{pixels}} < f < \frac{1}{5\text{pixels}}$  with intensity  $0.5 < I_p(x,y) < 1$ . This pattern is used to modulate the amplitude, so that the modified three-step phase shifting pattern is:

$$I_k(x,y) = I'(x,y) + I_p(x,y)I''(x,y) \cos \left[ \phi(x,y) - \frac{2\pi k}{3} \right]. \quad (2)$$

The encoded pattern can be recovered by computing the intensity modulation

$$\gamma(x,y) = \frac{I''(x,y)}{I'(x,y)} = \frac{\sqrt{3(I_1 - I_3)^2 + (2I_2 - I_1 - I_3)^2}}{I_1 + I_2 + I_3} \quad (3)$$

From Eq. (1), the coded pattern was embedded with fringe images. The coded pattern computed from Eq. (3) is normalized by the mean image intensity. Therefore, it is less sensitive to the differences in lighting or exposure between each of the stereo cameras. This helps improve the accuracy of the matching step.

The images given in Eq. (2) use an 8-bit grayscale representation, limiting its speed of projection and thus 3D shape measurement speeds. To remedy this problem, this research converts the 8-bit grayscale images to 1-bit binary patterns by using the random dithering algorithm. Briefly, the random dithering technique is to compare the intensity of the 8-bit image with uniform random noise, and for each pixel if the grayscale value of the fringe pattern is larger than the random noise, the pixel value is set to be 1, otherwise 0, as described here:

$$I_k^d(x,y) = \begin{cases} 1 & \text{if } I_k(x,y) > \text{Unif}(0,1) \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

Here  $\text{Unif}(0,1)$  represents uniform random noise between 0 and 1 and  $I_k^d(x,y)$  the binary value in the dithered pattern at pixel  $(x,y)$ .

Figure 1 shows an example of the phase-shifted fringe patterns. Figure 1(a) shows the coded pattern embedded into the quality map of the phase-shifted sinusoidal patterns illustrated in Fig. 1(b). The resultant 8-bit grayscale fringe pattern with embedded coded pattern is shown in Fig. 1(c). After applying the random dithering technique, the resulting binary pattern is shown in Fig. 1(d). Since the random dithering process produces artifacts in the form of high-frequency noise, the projector is defocused during capture. This preserves the underlying pattern while still enabling a fast capture rate.

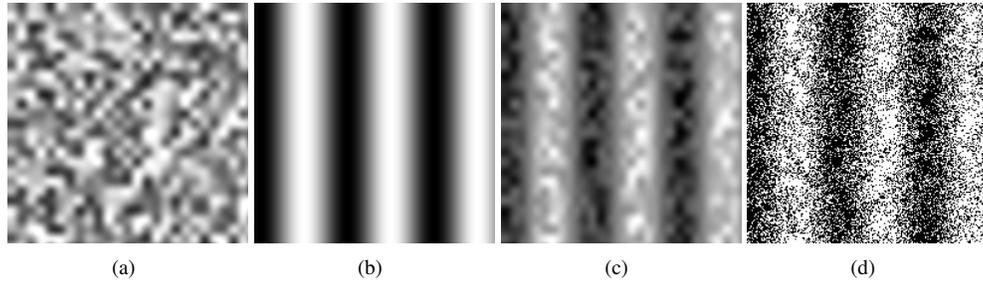


Fig. 1. Example of the 8-bit grayscale patterns and the resultant dithered patterns. (a) Random pattern; (b) 8-bit sinusoidal fringe pattern; (c) Modified sinusoidal fringe pattern by combining (a) and (b) using Eq. (2); (d) Final 1-bit dithered pattern after applying the random dithering technique.

## 2.2. Phase remapping

Common windowed stereo metrics such as sum of absolute differences (SAD) and normalized cross-correlation (NCC) compare the intensity values of a template window with that of a target [20]. Although fast, they rely on the assumption that the disparity within a window is constant. However, often the surface will have non-constant disparity values and thus the values of the pixel at  $(i, j)$  in the left window will not correspond to the pixel at  $(i, j)$  in the right window.

To account for this in-window discrepancy, we introduce the idea of phase remapping. The purpose of phase remapping is to map the coded pattern captured from one camera to match the perspective of the other camera, reducing error caused by geometric distortion.

Several properties of the captured images allow us to do this. First, the phase is monotonically increasing, with the exception of  $2\pi$  jumps. Second, the camera images are calibrated and rectified to satisfy the epipolar constraint and guarantee that the  $y$ -value of the remapped coded pattern will be the same in both the left and right images. Finally, the coded pattern at a certain phase value is the same in both the left and right images. That is, if the point  $(x, y)$  in the left image corresponds with point  $(x - d, y)$  in the right image, then  $\phi_L(x, y) = \phi_R(x - d, y)$  and  $I_L(x, y) = I_R(x - d, y)$ .

Using these three properties, the right image at a given disparity  $d$  can be remapped to match the perspective of the left image. If the disparity is correct, then the left and remapped right coded patterns will have pixel-to-pixel correspondence and thus have a low matching cost. Because the effects of geometric distortion have been eliminated with the phase remapping, its matching cost will be lower than a standard windowed cost.

The typical approach treats the image intensity as a function of spatial position in the image. However, because phase is monotonically increasing, the image intensity can instead be considered a function of the phase. In doing so, remapping can be achieved using the common kernel smoothing introduced by Nadaraya and Watson [21]. The kernel smoother is used to estimate a weighted average. In this paper, we propose to use a weight based on phase similarity. This weight allows the kernel smoother to search for the target point in the image that matches the phase of the source and provide an interpolated coded pattern at that point.

Let  $K_\epsilon(\phi_L, \phi_R)$  represent a kernel for a given disparity  $d$ . The left phase  $\phi_L(x, y)$  is used as a reference to predict the intensity  $\hat{I}_R(x - d, y)$  of the coded pattern of the right image around a

window of size  $2N + 1$ :

$$\hat{I}_R(x-d, y) = \frac{\sum_{i=-N}^N w_i(x, y, d) I_R(x-d+i, y)}{\sum_{i=-N}^N w_i(x, y, d)}, \quad (5)$$

$$w_i(x, y, d) = \gamma' K_\varepsilon(\phi_L(x, y), \phi_R(x-d+i, y)). \quad (6)$$

Here  $w_i$  is the weight given by the estimator,  $\hat{I}_R(x-d, y)$  the estimated intensity of the right coded pattern,  $I_R(x-d, y)$  the actual intensity of the right coded pattern.  $N$  should be chosen to be roughly one-quarter the period of the sinusoid in the captured image; in these experiments  $N = 7$  was used. To reduce the effects of different camera sensors, a Laplacian of Gaussian (LoG) filter can be applied to both the left and right camera images after remapping. The gamma-dependent weight  $\gamma'$  given in Eq. 6 is given by:

$$\gamma'(x, y) = \begin{cases} t & \text{if } \gamma(x, y) \geq t \\ \gamma(x, y) & \text{otherwise} \end{cases}. \quad (7)$$

Since the correspondence pattern is encoded within the quality map,  $t$  should be chosen so that  $\gamma'$  provides approximately equal weights to pixels with a high quality as those with a low quality imposed by the correspondence pattern.

The kernel function is typically a RBF depending only on the radius  $r = \|x - x_i\|$  and in this paper is represented by the minimum angle between both phases. The RBF used for these experiments is an inverse quadratic function:

$$K_\varepsilon(\phi_L, \phi_R) = \frac{1}{1 + (\varepsilon r)^2}, \quad (8)$$

plotted visually in Fig. 2. In these experiments,  $\varepsilon = 10/\text{rad}$  was used to remap the image. Once the image has been remapped for a given disparity, the matching cost for a pixel  $(x, y)$  at disparity  $d$  can be computed using the sum of absolute differences (SAD):

$$SAD(x, y, d) = \sum_{i=-N}^N \sum_{j=-N}^N |I_L(x+i, y+j) - \hat{I}_R(x-d+i, y+j)|. \quad (9)$$

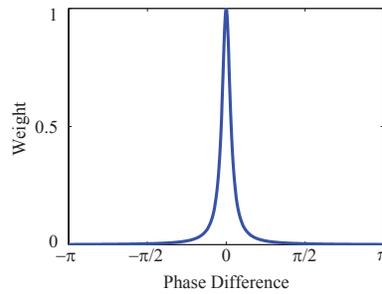


Fig. 2. Radial basis function (RBF) used to remap coded image to a reference phase with  $\varepsilon = 10/\text{rad}$ .

Because SAD computation is linearly separable, it is equivalent to compute the costs for each pixel only once for each disparity and apply a box filter of size  $2N + 1$  to the result [3].

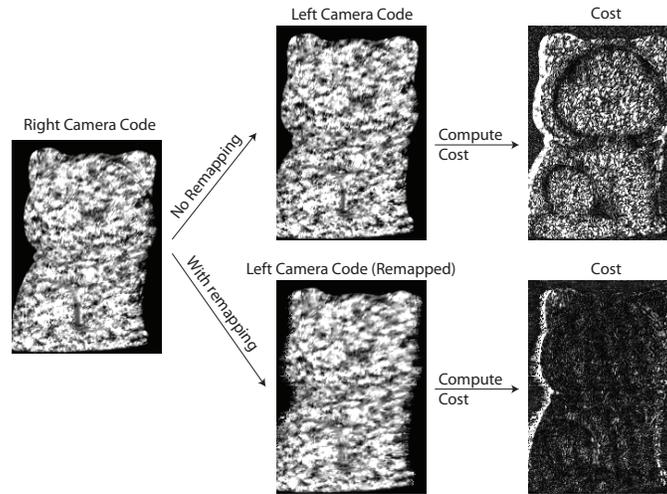


Fig. 3. The encoded pattern from the right camera is compared to the left camera either without remapping (top row) or with remapping (bottom row). Using the traditional sliding window approach causes “rings” of low cost in regions of large disparity variation. By applying remapping (bottom row), the cost is minimized throughout the entire region, reducing the possibility of a false match. A video of the remapping technique applied to multiple disparity levels is shown in [Media 1](#).

### 2.3. Stereo matching

To determine the disparity for a given method, the simplest method is to use a winner-takes-all (WTA) strategy, assigning the disparity with the lowest cost over all tested disparities. Because WTA does not take into account global information, i.e. pixels outside of the immediate  $N \times N$  window, it is possible that the disparity with the lowest cost is not the correct disparity.

To reduce this error, Hirschmuller introduced a semi-global method (SGM) that optimizes disparity values over scanlines in 16 different directions [22]. The optimization assigns a cost associated with switching:

$$L'_r(\mathbf{p}, \mathbf{d}) = \mathbf{C}(\mathbf{p}, \mathbf{d}) + \min \left[ L'_r(\mathbf{p} - \mathbf{r}, \mathbf{d}), \right. \quad (10) \\ \left. L'_r(\mathbf{p} - \mathbf{r}, \mathbf{d} - \mathbf{1}) + \mathbf{P}_1, \right. \\ \left. L'_r(\mathbf{p} - \mathbf{r}, \mathbf{d} + \mathbf{1}) + \mathbf{P}_1, \right. \\ \left. \min_i L'_r(\mathbf{p} - \mathbf{r}, \mathbf{i}) + \mathbf{P}_2 \right]$$

An example of the proposed remapping is shown in Fig. 3 and the associated [Media 1](#). While the code for the right camera and left camera were both recovered, a simple subtraction for a given disparity causes a distinctive “ring” where the cost is low (top right). This can be avoided by remapping the left camera code (bottom) to match the phase of the right camera code. The result is a low cost for the entire region. When testing over a wide range of disparities, remapping avoids this problem of ringing as shown in Fig. 4. A video of the costs using the proposed method is shown in [Media 2](#) while that of SAD is shown in [Media 3](#).

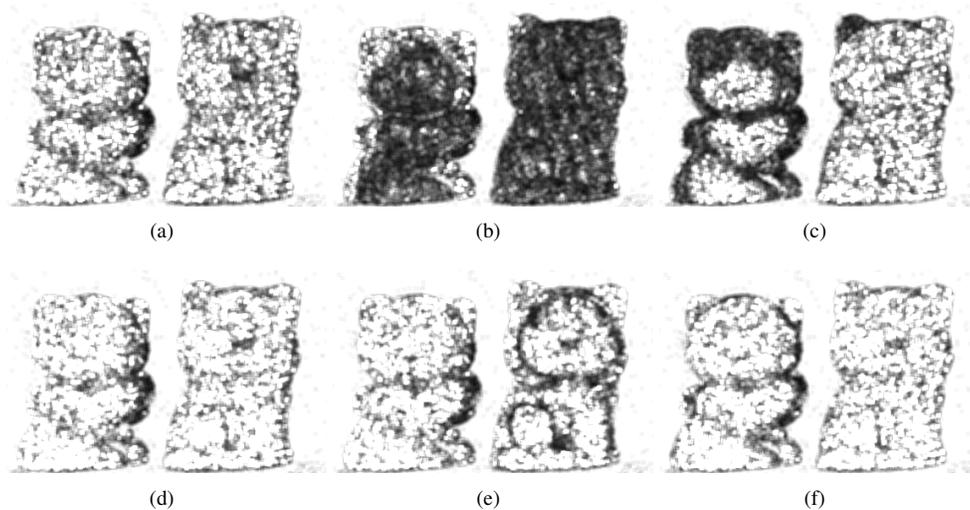


Fig. 4. Comparing results of using phase remapping and coded texture images. [Media 2](#) shows the animation of the costs for all disparities using the proposed method, and [Media 3](#) shows the animation of the costs for all disparities using SAD. (a)-(c) Costs for disparities of 35, 45, and 55 pixels using phase remapping; (d)-(f) Corresponding results of (a)-(c) by directly using the coded images.

### 3. Experimental results

Experiments were conducted to verify the success of the proposed technique. The system consisted of a digital-light-processing (DLP) Lightcrafter 4500 projector and two USB 3.0 Point-Grey Flea3 cameras (FL3-U3 13Y3M-C). Each camera has a 8 mm lens (Computar M0814-MP2) and the cameras remained untouched after calibration. The cameras were affixed approximately 47 mm apart on a small optics board and calibrated to satisfy the epipolar constraint using OpenCV stereo calibration. It is important to note that, throughout all the experiments, the projector remained uncalibrated.

To verify that the proposed method can properly measure multiple separate object, we first simultaneously measured two small statues that were placed approximately 300 mm from the cameras. In this research, only three phase-shifted fringe patterns are used and no spatial or temporal phase unwrapping is employed. It is well known that a single camera and single projector system cannot successfully measure such a system unless a temporal phase unwrapping algorithm is adopted.

Figure 5(a) shows the photograph of the system. Figure 5(b) and 5(c), respectively shows the recovered encoded pattern from the left camera and the right camera. Figure 5(d) shows the wrapped phase from the left camera, and Fig. 5(e) shows the wrapped phase from the right camera. These two phase maps are used to remap the pattern and compute the depth map shown in Fig. 5(f). Speckle noise shown in the original depth map can be alleviated by dynamic programming and median filtering. Figure 5(g) and 5(h) respectively shows the result after applying the dynamic programming and subsequently median filtering. This experiment clearly demonstrated that the proposed method can measure disjointed objects, achieving absolute 3D shape measurements. It should be noted that for this experiment, we select the camera image resolution of  $640 \times 512$ , and exposure time of 2 ms.

To further test the system's capability of high-speed measurements, we carried out additional

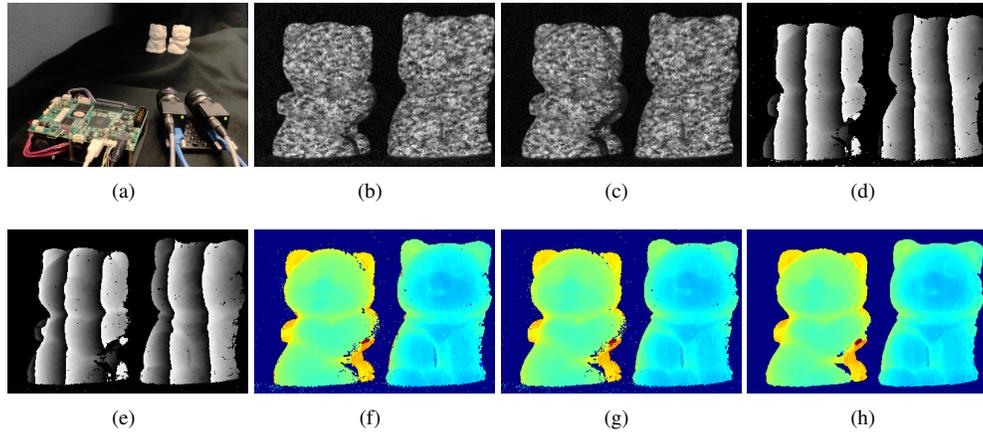


Fig. 5. Experimental results of measuring two separate statues. (a) System setup; (b) Coded pattern from left camera; (c) Coded pattern from right camera; (d) Phase map from left camera; (e) Phase map from right camera; (f) Remapped cost aggregation; (g) Result of (f) after applying dynamic programming; (h) Result of (g) after applying median filtering.

experiments. In the following two experiments, the resolution was chosen as  $1280 \times 1024$ , and the speed of the cameras as 150 Hz, which is the maximum speed that the Pointgrey cameras provide. The PointGrey cameras were synchronized with the DLP LightCrafter 4500 that refreshes the binary patterns at 150 Hz. Fringe images were loaded on to a DLP Lightcrafter 4500 projector, which was only slightly defocused at the object plane. While the projector was capable of projecting binary patterns at a 4KHz rate, the cameras limited the system speed, resulting in an acquisition speed of 150 Hz. Because the algorithm only requires 3 images for 3D shape measurement, the achieved 3D shape measurement speed is 50 Hz.

Figure 6 and the associated [Media 4](#) show the result of measuring moving fingers. It is clear that all fingers are properly captured, demonstrating that the algorithm can even reconstruct the disparity map from the moving images. It should be noted that for certain frames, there are some artifacts (i.e. vertical stripes) on the measurement results created from the motion. Although 50 Hz is quite high speed, it is still not sufficient when the fingers are moving rapidly.

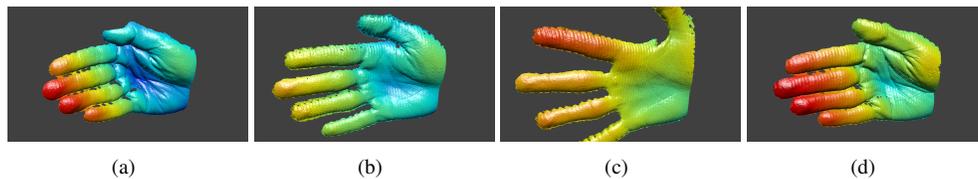


Fig. 6. Typical 3D frames of capturing a moving hand at 50 Hz (refer to [Media 4](#) for the complete video sequence). (a) Time (0 s); (b) Time (0.25 s); (c) Time (0.50 s); and (d) Time (0.75 s).

We measured a dynamic human face to demonstrate the capability of measuring more complex 3D shapes. Figure 7 and the associate [Media 5](#) show the results. Once again, the proposed algorithm can generate good quality results, demonstrating the success of the proposed algorithm.

In our previous paper we demonstrated that high-quality measurement could be obtained

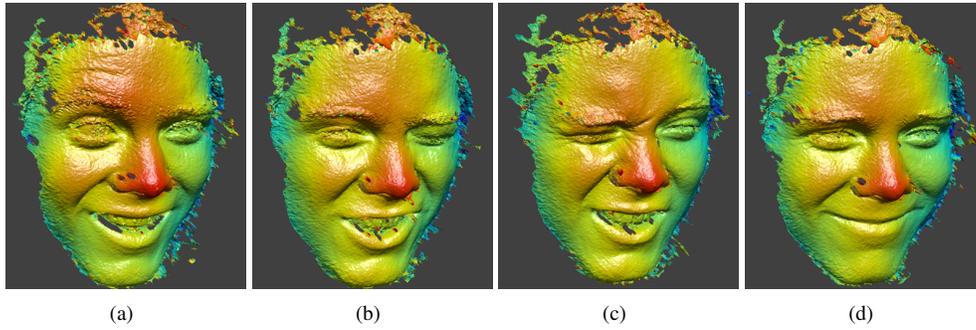


Fig. 7. Typical 3D frames of capturing a moving face at 50 Hz (refer to [Media 5](#) for the complete video sequence).

even in the presence of low-quality phase when using a two-camera approach. To test the phase quality, we projected these patterns onto a flat white surface placed 300 mm from the left camera and compared the computed geometry with a ground truth measurement. The ground truth phase was obtained by using 9-step binary fringes with a pitch of 18, which was then blurred using a Gaussian filter of size  $15 \times 15$  pixels and standard deviation  $15/3$  pixels. The two cameras were affixed 96 mm apart and calibrated to satisfy the epipolar constraint. The phase-to-height conversion was determined by moving the white plane 10 mm closer to the camera and recapturing the phase using a nine step phase-shifting algorithm.

Figure 8 shows the comparison results. These data demonstrated that the dithered fringes with an embedded pattern have an RMS error of 0.062 radians. Using the dithered phase directly with the reference plane method results in a depth error of 0.11 mm. However, when using the method proposed in this paper, the depth error is 0.049 mm. This suggests that the proposed method can acquire geometry with higher accuracy than directly using the phase captured by a single camera.

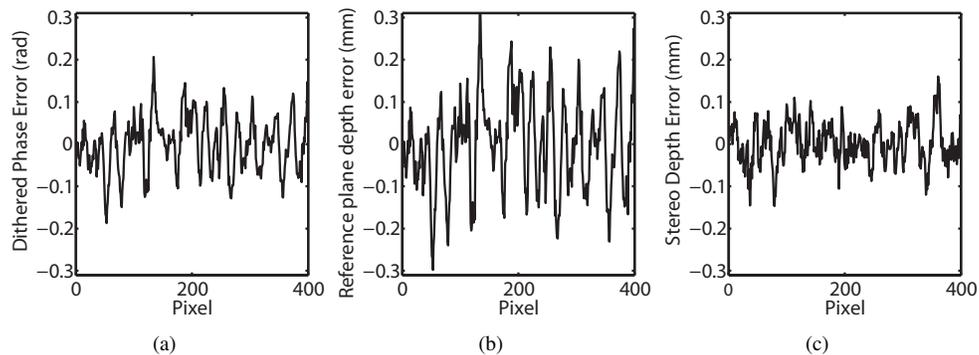


Fig. 8. Comparison of errors using a reference plane and using the proposed method. (a) phase error from the dithered pattern; (b) depth error using the dithered pattern and a reference plane; (c) depth error using the proposed technique.

#### **4. Conclusion**

This paper presented a framework for capturing stereo geometry using the minimum three 1-bit binary patterns without spatial or temporal phase unwrapping. The encoded pattern within the fringe images provide correspondence, and because the cost function is performed within the feature space rather than the image space, it is able to correctly identify the disparity for regions with geometric distortions or low texture variations. We have successfully developed a system that achieved 50 Hz 3D shape measurement speed, albeit much higher speed (e.g., 1,333 Hz) 3D shape measurement could be achieved should a higher speed camera is used.

#### **Acknowledgment**

The authors of this paper would like to thank Tyler Bell for the assistance to serve a models to test our system.

This study was sponsored by the National Science Foundation (NSF) under grant numbers: CMMI-1150711 and CMMI-1300376. The views expressed in this chapter are those of the authors and not necessarily those of the NSF.