Flexible real-time natural 2D color and 3D shape measurement

Pan Ou,[1,2] Beiwen Li,[1] Yajun Wang,[1] and Song Zhang[1,*]

[1]Department of Mechanical Engineering, Iowa State University, Ames, IA 50011, USA
[2]School of Instrumentation Science and Opto-electronics Engineering, Beihang University, Beijing 100191, China
[*] song@iastate.edu

Abstract: The majority of existing real-time 3D shape measurement systems only generate non-nature texture (i.e., having illumination other than ambient lights) that induces shadow related issues. This paper presents a method that can simultaneously capture natural 2D color texture and 3D shape in real time. Specifically, we use an infrared fringe projection system to acquire 3D shapes, and a secondary color camera to simultaneously capture 2D color images of the object. Finally, we develop a flexible and simple calibration technique to determine the mapping between the 2D color image and the 3D geometry. Experimental results demonstrate the success of the proposed technique.

References and links

1. G. Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon. **3**(2), 128–160 (2011).
2. S. Zhang, "Recent progresses on real-time 3-D shape measurement using digital fringe projection techniques," Opt. Lasers Eng. **48**(2), 149–158 (2010).
3. S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, "Real-time 3D model acquisition," ACM Trans. Graph. **21**(3), 438–446 (2002).
4. L. Zhang, N. Snavely, B. Curless, and S. M. Seitz, "Spacetime faces: high-resolution capture for modeling and animation," ACM Trans. Graph. **23**(3), 548–558 (2004).
5. S. Zhang and P. S. Huang, "High-resolution real-time three-dimensional shape measurement," Opt. Eng. **45**(12), 123601 (2006).
6. R. Höfling and P. Aswendt, "Real time 3D shape recording by DLP-based all-digital surface encoding," in *Proc. SPIE*, **7210**, 72,100E1–8 (2009).
7. K. Liu, Y. Wang, D. L. Lau, Q. Hao, and L. G. Hassebrook, "Dual-frequency pattern scheme for high-speed 3-D shape measurement," Opt. Express **18**(5), 5229–5244 (2010).
8. M. Schaffer, M. Grosse, and R. Kowarschik, "High-speed pattern projection for three-dimensional shape measurement using laser speckles," Appl. Opt. **49**(18), 3622–3629 (2010).
9. 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**(21), 21628–21635 (2010).
10. S. Zhang, D. Van Der Weide, and J. Oliver, "Superfast phase-shifting method for 3-D shape measurement," Opt. Express **18**(9), 9684–9689 (2010).
11. J. Pan, P. S. Huang, and F.-P. Chiang, "Color phase-shifting technique for three-dimensional shape measurement," Opt. Eng. **45**(12), 013602 (2006).
12. L. Zhang, B. Curless, and S. M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," in *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, 24–36 (2002).
13. Z. Zhang, C. E. Towers, and D. P. Towers, "Phase and colour calculation in color fringe projection," J. Opt. A, Pure Appl. Opt. **9**(6), S81–S86 (2007).
14. Z. Zhang, C. E. Towers, and D. P. Towers, "Time efficient color fringe projection system for 3D shape and color using optimum 3-frequency selection," Opt. Express **14**(14), 6444–6455 (2006).
15. S. Zhang and S.-T. Yau, "Simultaneous three-dimensional geometry and color texture acquisition using single color camera," Opt. Eng. **47**(12), 123604 (2008).
16. X. Liu, X. Peng, H. Chen, D. He, and B. Z. Gao, "Strategy for automatic and complete three-dimensional optical digitization," Opt. Lett. **37**(15), 3126–3128 (2012).
17. Y. Wang, J. I. Laughner, I. R. Efimov, and S. Zhang, "3D absolute shape measurement of live rabbit hearts with a superfast two-frequency phase-shifting technique," Opt. Express **21**(5), 5822–5832 (2013).
18. S. Zhang and P. S. Huang, "Novel method for structured light system calibration," Opt. Eng. **45**(8), 083601 (2006).

19. Z. Y. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000).
20. A. Geiger, F. Moosmann, O. Car, and B. Schuster, "A toolbox for automatic calibration of range and camera sensors using a single shot," *In Proc. of International Conference on Robotics and Automation (ICRA),* 3936–3943 (2012).

1. Introduction

Three-dimensional (3D) shape measurement is crucial for many areas ranging from industrial practices to scientific study [1]. Over the past decades, a number of techniques have been developed including stereo vision, structured light, and time of flight. The structured-light technique has emerged as one of the mainstreams because of its simplicity and speed [2].

Lately, real-time 3D shape measurement is increasingly available [3–10]. However, the majority of the existing real-time 3D shape measurement systems use a visible light, though successful, they have limitations in applications such as homeland security and biometrics, where the visible light might not be desirable. To our knowledge, there is little study on high-resolution, real-time 3D shape measurement using an infrared fringe projection technique. Moreover, most of the aforementioned real-time 3D shape measurement can simultaneously provide a black-and-white (b/w) texture by analyzing the structured patterns that are also used for 3D reconstruction. However, this texture generation approach is not natural, meaning that the texture is not captured without the directional projection light. This usually introduces shadows on the acquired texture due to the measured geometric shapes. To acquire a natural texture image, the projection light must be turned off such that only the ambient light is on when the texture image is captured. This can be done by capturing an additional image without projection light on, but sacrificing the measurement speed, and often drastically.

Simultaneously acquiring 2D color texture is more difficult, though viable, by using a color camera for both 3D geometry and 2D color texture capture [11–15], or by adding an additional color camera only for color texture capture [5, 16], and establish the mapping between the color camera and the b/w camera. The former usually sacrifices 3D measurement quality due to the inherent color induced problems (e.g., color coupling), and the latter typically requires a complicated hardware setup [5], or a sophisticated calibration routine [16] to create the mapping between the color camera and the b/w camera.

To our knowledge, there is no system that can simultaneously capture natural 2D color texture and high-resolution 3D geometry in real time. This paper bridges the gap in the real-time 3D shape measurement field. Specifically, we use a near infrared camera/projector pair to perform 3D shape measurement, and a secondary color camera to simultaneously capture 2D color images solely illuminated by ambient lights. Since these two light regimes do not overlap, simultaneous 3D shape and natural 2D texture capture can be realized without sacrificing speed; and because the color camera only captures visible light without being interfered by the infrared light for 3D shape measurement, nature 2D color images can be obtained. Moreover, to increase the flexibility of such a dual-camera system, we also develop a simple calibration method that quickly determines the mapping from the color camera to the 3D geometry. Experimental results will be presented to verify the success of the proposed technique for real-time 3D shape measurement applications.

Section 2 explains the phase-shifting technique. Section 3 presents the proposed mapping method. Section 4 shows experimental results, and finally Section 5 summarizes the paper.

2. Absolute coordinates recovery from two-frequency phase-shifting technique

Over the years, numerous phase-shifting algorithms have been developed including three step, four step, double three step, and least squares. We use a five-step phase-shifting algorithm with equal phase shifts instead of a three-step phase-shifting algorithm to reduce the noise of the system. Five phase-shifted images can be described as,

$$I_n(x, y) = I'(x, y) + I''(x, y)\cos[\phi(x, y) + 2\pi n / 5]. \tag{1}$$

Where $I'(x,y)$ is the average intensity, $I''(x,y)$ the intensity modulation, and $\phi(x,y)$ the phase to be solved for, and $n = 1,2,\ldots,5$. Using a least square method, we obtain the phase

$$\phi(x,y) = \tan^{-1}\left[\frac{\sum_{n=1}^{5} I_n(x,y)\sin(2\pi n/5)}{\sum_{n=1}^{5} I_n(x,y)\cos(2\pi n/5)}\right]. \tag{2}$$

Equation (2) provides the phase ranging $[-\pi,\pi)$ with 2π discontinuities. The 2π discontinuities can be removed by adopting a spatial or temporal phase unwrapping algorithm. In this research, we adopted a two-frequency phase-shifting technique using the binary defocusing technique described in [17]. Briefly, a single fringe of the lower frequency patterns covers the whole range of measurement, and thus there is no phase unwrapping is necessary. This phase obtained from the low frequency fringe patterns is used to unwrap the high-frequency phase point by point. Once the high-frequency phase map is unwrapped, (x,y,z) coordinates for each point can be recovered once the system is calibrated. In this research, we adopted a calibration method similar to that discussed in [18].

3. Mapping from 2D color texture to 3D geometry

For a pinhole camera model, the geometric relationship between the world coordinates and the image coordinates is essentially a projection from a 3D space to a 2D plane. In homogenous coordinate system, the relationship between the world coordinates and the image coordinates can be mathematically represented as [19],

$$s\mathbf{I} = \mathbf{A}[\mathbf{R},\mathbf{t}]\mathbf{X}^w = \mathbf{H}\mathbf{X}^w. \tag{3}$$

Where $\mathbf{I} = [u,v,1]^T$ is the homogeneous coordinates of a point on the camera image plane, $\mathbf{X}^w = [x^w, y^w, z^w, 1]^T$ is the corresponding homogeneous world coordinates for that point, s is a scale factor, $\mathbf{A}$ is the 3×3 camera intrinsic matrix, $\mathbf{R}$ is the 3×3 rotation matrix, $\mathbf{t}$ is a 3×1 translation vector, and $\mathbf{H} = \mathbf{A}[\mathbf{R},\mathbf{t}]$ is called the camera matrix. It should be noted that Eq. (3) describes a linear camera model; the nonlinear effects can be accounted for by adopting a nonlinear camera model.

The camera matrix $\mathbf{H}$ is a 3×4 matrix. Hence, the relationship between the world coordinates and the image coordinates can be represented as

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H}\mathbf{X}^w = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix}\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix}. \tag{4}$$

Since the camera matrix $\mathbf{H}$ is given up to a certain scale factor, s, there are 11 unknowns that have to be determined for this camera model. For a given set of 3D points and their corresponding (u, v) coordinates, Eq. (4) can be reformulated as:

$$\mathbf{Q}\mathbf{h} = 0, \tag{5}$$

where $\mathbf{h}$ is a linear composition of elements of the matrix $\mathbf{H}$:

$$\mathbf{h} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{21} & h_{22} & h_{23} & h_{24} & h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix}, \tag{6}$$

and $\mathbf{Q}$ is a matrix:

$$\mathbf{Q} = \begin{bmatrix} x_1^w & y_1^w & z_1^w & 1 & 0 & 0 & 0 & 0 & -u_1 x_1^w & -u_1 y_1^w & -u_1 z_1^w & -u_1 \\ 0 & 0 & 0 & 0 & x_1^w & y_1^w & z_1^w & 1 & -v_1 x_1^w & -v_1 y_1^w & -v_1 y_1^w & -v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ x_N^w & y_N^w & z_N^w & 1 & 0 & 0 & 0 & 0 & -u_N x_N^w & -u_N y_N^w & -u_N z_N^w & -u_N \\ 0 & 0 & 0 & 0 & x_N^w & y_N^w & z_N^w & 1 & -v_N x_N^w & -v_N y_N^w & -v_N z_N^w & -v_N \end{bmatrix}. \quad (7)$$

Here, the coordinates of the k-th point are denoted as $X_k^w = [x_k^w, y_k^w, z_k^w, 1]^T$, $I_k = [u_k, v_k, 1]^T$. Solving Eq. (5) can be done by the singular value decomposition $\mathbf{Q} = \mathbf{S V D}^T$ method. It is simply a column of $\mathbf{D}$ which corresponds to an entry of $\mathbf{V}$ with a smallest value.

The mapping from 3D coordinates obtained from the infrared system and the 2D color camera is essentially to determine matrix $\mathbf{H}^c$. In this research, we developed a simple method to determine this matrix by leveraging the 3D information we already have from the 3D shape measurement system. To calibrate this mapping matrix, we use the same checkerboard we used to calibrate the infrared 3D shape measurement. The checkerboard was placed in six different orientations. At each orientation, the checkerboard was measured by the infrared 3D shape measurement system to obtain its 3D shape, and was imaged by the color camera to capture color texture. The checkerboard corners were then extracted from both infrared image and the 2D color image using the algorithm described in [20]. From the corresponding corner pairs, the (u, v) coordinates on the color camera correlates to the 3D world coordinates (x^w, y^w, z^w) obtained by 3D shape measurement system. From which matrix $\mathbf{H}^c$ were solve using the method discussed above.
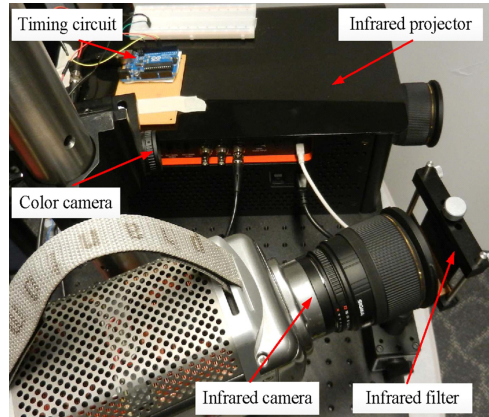
4. Experimental results

Fig. 1. Photograph of measurement system.

Figure 1 shows a photograph of the system we developed. It is composed of an infrared digital-light processing (DLP) projector (LightCommander, Logic PD, Inc.), a high-speed infrared CMOS camera (Phantom V9.1, Vision Research), and a color CCD camera (DFK 21BU04, The Imaging Source). The wavelength of infrared LED used in this projector is 850 nm. An infrared filter is placed in front of the high-speed infrared CMOS camera to block visible lights. Two cameras are triggered by an external circuit board that is synchronized with the projector. The infrared CMOS camera was set to capture images with a resolution of 576×576, and the color CCD camera was set with a resolution of 640×480.

We verified the proposed technique by measuring a dynamically deformable human facial expression. Since two-frequency phase-shifting techniques were adopted, the 3D shape measurement technique discussed above requires the capture of a sequence of fringe patterns

to reconstruct one 3D frame. Due to the low intensity of the infrared projector, the high-speed projector was set to project binary patterns at 200 Hz, and the infrared camera was precisely synchronized with the projector to capture 2D fringe patterns at 200 Hz. The 3D capturing rate is 20 Hz since it requires ten phase-shifted fringe patterns (five high-frequency patterns and five low-frequency patterns) to reconstruct one 3D frame. The color camera was hence to capture 2D color images at 20 Hz to precisely align with the 3D capture rate. The color camera and the infrared camera were precisely synchronized through the external trigger timing circuit.

Figure 2 illustrates the two frequency phase-shifting algorithms we adopted to obtain 3D shape. Figure 2(a) shows one of the five high-frequency phase-shifted fringe patterns on human face with an infrared projector Figs. 2(b) shows one of the five low-frequency phase-shifted fringe patterns. Figure 2(c) shows the wrapped phase using the high-frequency fringe patterns and the five-step phase-shifting algorithm, which possesses 2π phase discontinuities. The low-frequency wrapped phase, illustrated Fig. 2(d), is continuous without 2π discontinuities since a single fringe covers the whole measurement range. By referring to the phase map obtained from the low frequency patterns, the high-frequency wrapped phase can be unwrapped point by point using a temporal phase unwrapping algorithm. Figure 2(e) shows the unwrapped phase map.
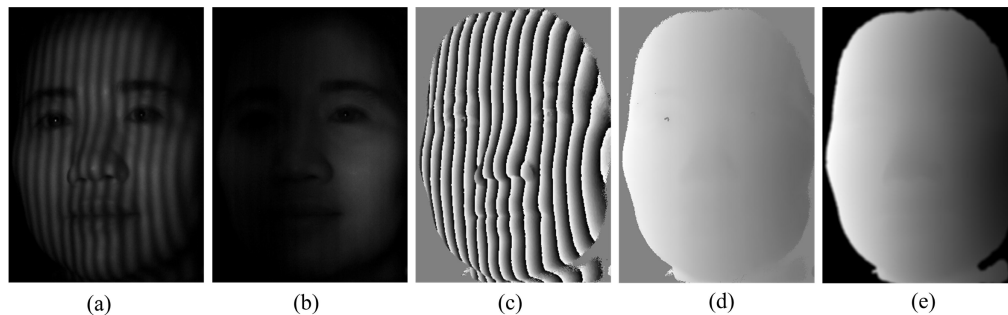

| (a) | (b) | (c) | (d) | (e) |

Fig. 2. Absolute phase retrieval using two frequency phase shifting technique. (a) High-frequency fringe pattern; (b) Low-frequency fringe pattern; (c) Wrapped phase map using high-frequency patterns; (d) Wrapped phase map using low-frequency patterns; (e) Unwrapped phase map for high-frequency fringe patterns.

From the unwrapped phase map, 3D shape can be recovered using the calibrated system parameters. This research used a similar structured light system calibration method discussed in [17]. Figure 3(a) and the associated video (Media 1) show the 3D reconstructed result. One may notice that the reconstructed 3D geometry is not very smooth (some bumps on the face). This might because the infrared light penetrates different depth of skin on different part of face, or the defocused sinusoidal patterns are not perfectly sinusoidal. From these fringe images, the b/w texture can be retrieved, as shown in Fig. 3(b), the b/w texture can be mapped on to the 3D geometry. Figure 3(c) and the associated video (Media 2) show the 3D results with texture mapping. The b/w is not natural due to the direct illumination of the projector (shadows on the side of nose are obvious). As a comparison, the color texture captured by the color camera (shown in Fig. 3(d)) does not have the same shadow problem. Figure 3(e) shows the mapped color texture pixel by pixel on the infrared camera. The color texture can then be mapped onto 3D geometry. Figure 3(f) and the associated video (Media 3) show the result in 3D space. This experiment clearly demonstrated that the proposed technique can successfully and simultaneously capture 3D geometry and natural 2D color texture in real time. It is important to notice that the mapping is usually not perfectly aligned with pixels due to the discrete effects of the cameras. Linear interpolation was adopted in this research to map color texture. It is also important to note that the mapping we used in this research is linear without considering the nonlinear distortion of the camera lens, but no obvious artifacts were created even with linear model.
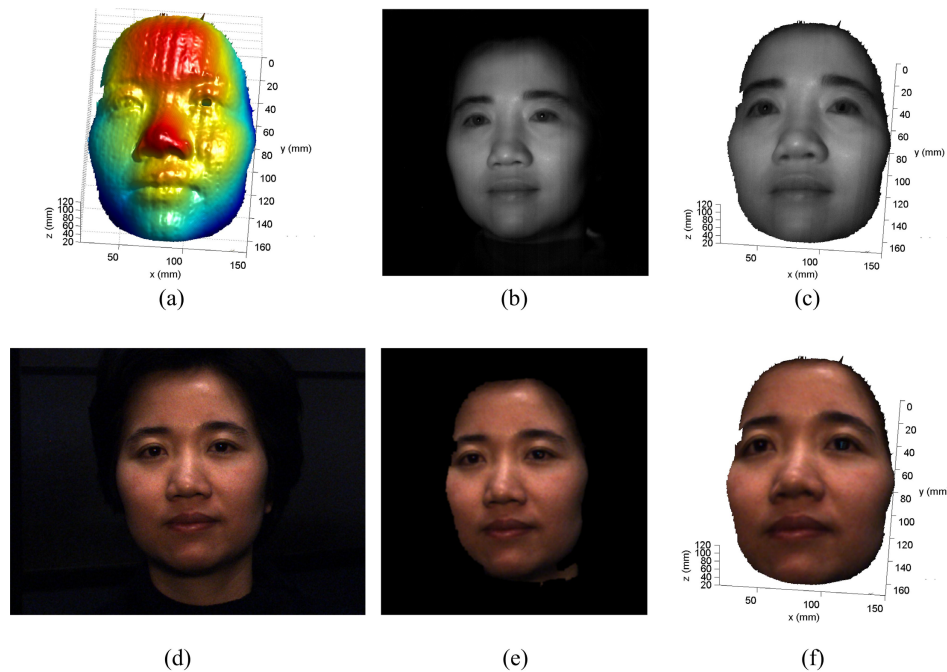
Fig. 3. Experimental results with texture mapping. (a) 3D reconstructed face using the infrared fringe patterns (Media 1); (b) Average b/w texture images obtained from the phase-shifted fringe patterns; (c) 3D results with b/w texture mapping (Media 2); (d) 2D color texture image captured by the color camera; (e) 2D color image mapped to the infrared camera; (f) 3D results with color texture mapping (Media 3).

5. Conclusion

This paper has presented a technology that can simultaneously capture natural 2D color texture and the 3D geometry in real time without using visible projection light. The major contributions are: (1) presented the first high-resolution, real-time 3D shape measurement system that combines a near infrared camera/projector pair to perform 3D shape measurement, and a secondary color camera to simultaneously capture 2D color images of the object solely illuminated by ambient visible lights; (2) developed a simple calibration method that could quickly determine the mapping from the color camera to the 3D captured geometry. A hardware system was developed that has verified the success of the proposed method. This technology could be significant to many applications including entertainment, biometrics, and cosmetics.

Acknowledgments