A New Method of Three-Dimensional Measurement Based on Dual Cameras

Li Zhongyang, Cai Lei, Luo Peng, Wang Baitong[1] and Huang Xiuxiang[1]
[1]Faculty of Engineering and Design, Hunan Normal University, Changsha, Hunan, 410081, China
hxxpxj01@hotmail.com

Abstract - We proposed a novel method for a three-dimensional measurement based on the structured light with stereo vision system which mainly consists of two cameras and a projector. This new method could realize shape recovery without projector calibration which is usually a complicated and time-consuming procedure. Furthermore, the new method is also developed to realize the correspondence of the image points between the two cameras. We showed the measurement system and described the principle of the system briefly. Finally, a 3-D model figured out by the system is presented to verify its performance.

Index Terms – Three-dimensional measurement, dual cameras, structured light

I. Introduction

Three-dimensional shape recovery is one of the most important research topics in optical metrology due to its noncontact and noninvasive[1-3]. Some examples of its applications are Medical imaging and diagnosis, industrial inspection of manufactured parts, reverse engineering, object recognition and 3D map building[4]. However the conventional three-dimensional measurement systems with single camera and single projector have many drawbacks[5]. A) The projector should be calibrated as it is regarded as an inverse camera in a conventional 3-D measurement and the calibration procedures are always time-consuming and complicated. B) The projector calibration has a lower precision than camera calibration because the images for the projector calibration are usually figured out with the images captured by a camera. C). The field of view is limited in a conventional 3-D measurement system. Hence, the single camera and single projector 3-D measurement system may be impractical in applications of fast and high resolution.

One of most important procedure in the measurement is to figure out the corresponding point between the projected images and the captured images[6-7]. It is very easy to realize it in the single camera and single project system since the projector is regarded as an inverse camera. But it is not the same in the double camera system.

In this paper, we present a novel method to acquire 3-D shape with the double cameras and single projector system. We can get a more accurate model by the new system with an efficient calibration since it only needs to calibrate the two cameras. In addition, the correspondence could be realized by calculating the cross point between a line and a curve which is the central line of the deformed fringe.

II. Description of the setup

Fig. 1 shows the measurement system which contains a projector projecting fringe images onto the objects, two cameras capturing deformed fringe images in different view fields, a PC generating fringe images and controlling the two cameras and a turn table spinning the objects for the different areas measuring. Fig. 2 shows the photograph of the dual cameras measurement system.

At the beginning, the projector project 9 fringe images and the two cameras capture each deformed fringe image after the projection[8-9]. The captured images are stored in PC. After the 9 fringe images are captured, the object is spun an angle of 45° by the turn table. Then repeat the fringe projection and images capturing. After 7 times rotations with same angle, the object is measured in all the field of view and 72 images are captured by each camera.

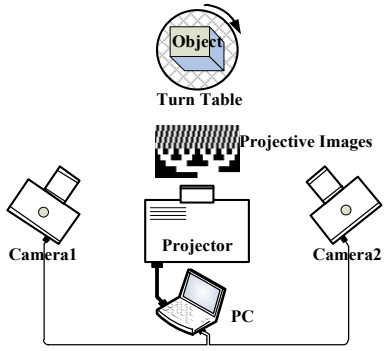

Fig.1 Structured light measurement system

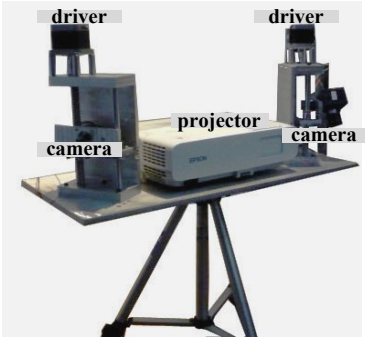

Fig.2 Photograph of the measurement

III. Three-dimensional measurement principle

Camera model has been widely studied these years. A camera model is often regarded as a pinhole model with intrinsic parameters, such as the parameters of focal length, principle point, optical axis, pixel skew effect and the pixel size and with extrinsic parameters such as the rotation matrix and the translation vector from the world coordinate system to the camera coordinate system [10-12].

Fig. 3 shows a typical diagram of system coordinate, where P is an arbitrary point with coordinates (X_w, Y_w, Z_w), (X_{c1}, Y_{c1}, Z_{c1}), (X_{c2}, Y_{c2}, Z_{c2}) in the world coordinate system and the two camera coordinate system respectively[13]. The coordinate of its projection in the two image planes is (U_{C1}, V_{C1}) and (U_{C2}, V_{C2}) respectively and the pixel coordinate in the two images is (U_1, V_1) and (U_2, V_2) respectively.

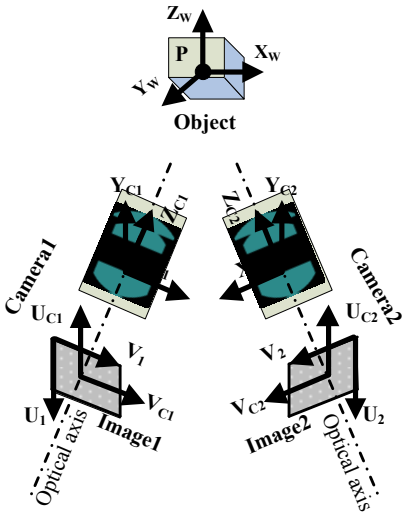

Fig. 3 The algorithm of the system

In camera1, the relationship between a point in the pixel coordinate and image plane coordinate can be described as follows just as coordinate translation:

$$\begin{pmatrix} U_1 \\ V_1 \end{pmatrix} = \begin{pmatrix} U_{C1} \\ V_{C1} \end{pmatrix} + \begin{pmatrix} U_{01} \\ V_{01} \end{pmatrix} \tag{1}$$

Where (U_{01}, V_{01}) is the origin of the pixel coordinate. The relationship between a point in the camera coordinate and image plane coordinate can be described as

$$\begin{cases} \dfrac{U_{C1}}{f_1} = \dfrac{X_{C1}}{Z_{C1}} \\ \dfrac{V_{C1}}{f_1} = \dfrac{Y_{C1}}{Z_{C1}} \end{cases} \tag{2}$$

Where f_1 is the focal length of Camera1. While the horizontal focal length and the vertical focal length are not equal. We suppose f_{1x} and f_{1y} is the horizontal and vertical focal length respectively. After operating, it is obtained

$$Z_{C1} \cdot \begin{pmatrix} U_1 \\ V_1 \\ 1 \end{pmatrix} = K \cdot \begin{pmatrix} X_{C1} \\ Y_{C1} \\ Z_{C1} \end{pmatrix} \tag{3}$$

Where $K = \begin{pmatrix} f_{1X} & d & U_{01} \\ 0 & f_{1Y} & V_{01} \\ 0 & 0 & 1 \end{pmatrix}$, is usually called intrinsic parameter.

The relationship between a point in the camera coordinate and world coordinate can be described as follows[14]:

$$\begin{pmatrix} X_{C1} \\ Y_{C1} \\ Z_{C1} \end{pmatrix} = R * \begin{pmatrix} X_W \\ Y_w \\ Z_W \end{pmatrix} + T \tag{4}$$

Where $R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$, $T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$ is called rotation matric and translation vector respectively. So, summarising, from equations (3) and (4) we obtain:

$$Z_{C1} \begin{pmatrix} U_1 \\ V_1 \\ 1 \end{pmatrix} = \begin{pmatrix} f_{1X} & 0 & U_{01} \\ 0 & f_{1Y} & V_{01} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{pmatrix} \begin{pmatrix} X_W \\ Y_w \\ Z_W \\ 1 \end{pmatrix} \tag{5}$$

That is:

$$Z_{C1} \begin{pmatrix} U_1 \\ V_1 \\ 1 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} & k_{14} \\ k_{21} & k_{22} & r_{23} & k_{24} \\ k_{31} & k_{32} & k_{33} & k_{34} \end{pmatrix} \begin{pmatrix} X_W \\ Y_w \\ Z_W \\ 1 \end{pmatrix} \tag{6}$$

Where $\begin{bmatrix} k_{ij} \end{bmatrix} = \begin{pmatrix} f_{1X} & 0 & U_{01} \\ 0 & f_{1Y} & V_{01} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{pmatrix}$

Obviously, the second camera has the same properties as the first one:

$$Z_{C2} \begin{pmatrix} U_2 \\ V_2 \\ 1 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{pmatrix} \begin{pmatrix} X_W \\ Y_w \\ Z_W \\ 1 \end{pmatrix} \tag{7}$$

Supposing that P_1 is an arbitrary point with the pixel coordinate (U_1, V_1) on image1. P and P_2 is the corresponding point on the object with the world coordinate (X_w, Y_w, Z_w) and corresponding point on the image2 with the pixel coordinate (U_2, V_2). Obviously, (U_1, V_1) is known and (U_2, V_2) is unknown and (X_w, Y_w, Z_w) is the coordinate we want to figure out. The two matrix [k_{ij}] and [t_{ij}] can be calculated by the intrinsic matrix K, rotation matrix R and translation vector T of the two cameras which can be obtained through camera

calibration. So there are 7 unknown numbers in equation 6 and 7 as X_w, Y_w, Z_w, U_2, V_2, Z_{C1}, Z_{C2}. Here we suppose V_2 is a known number and try to figure out U_2. Finally we get:

$$V_2 = mU_2 + n \tag{8}$$

$$\text{where } m = \frac{C_3D_2 - C_2D_3}{C_3D_1 - C_1D_3}, \quad n = \frac{C_2}{C_3} = \frac{C_3D_2 - C_2D_3}{C_3D_1 - C_1D_3} \cdot \frac{C_1}{C_3}$$

$$C = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} * \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}^{-1} \cdot \begin{pmatrix} U_1 \\ V_1 \\ 1 \end{pmatrix}$$

$$D = \begin{pmatrix} t_{14} \\ t_{24} \\ t_{34} \end{pmatrix} - \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} * \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}^{-1} \cdot \begin{pmatrix} r_{14} \\ r_{24} \\ r_{34} \end{pmatrix}$$

Where we can find the coordinate of the pixel in image2 is in a line $V_2 = mU_2 + n$. In practice, the coordinate is also in a curve which is on the object and projected by the projector as shown in Fig. 4. We can locate the specific curve since the coded structured light can label each fringe. For example, if we get an arbitrary pixel point on the image1, then we can figure out the number of the fringe where the point located by the coded structured light labeling. We can locate the specific fringe on the image2 according to the number figured out. So the pixel coordinate of the point on image2 can be figured out by calculating the cross point between the line and the curve.

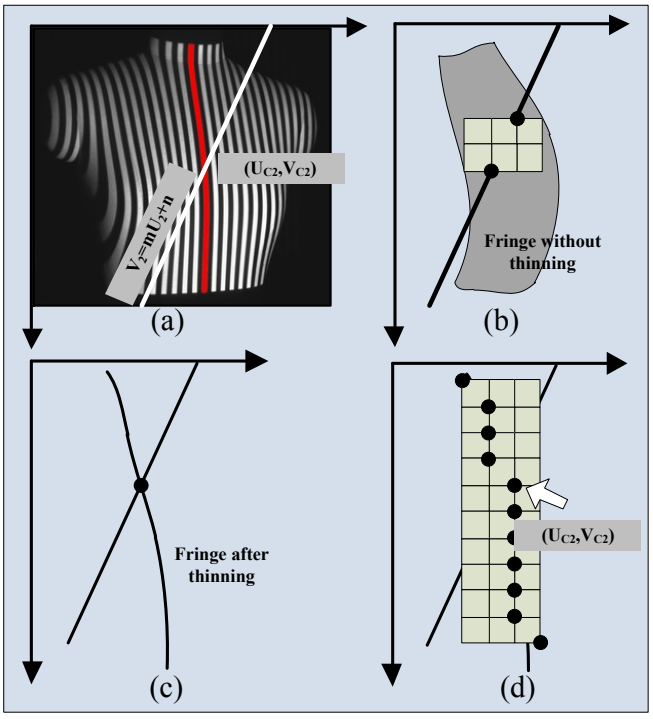

Fig.4 Corresponding Point Calculation: (a) Algorithm (b) Ambiguous Point without Thinning (c) Fringe Thinning (d) Calculation

While the curve is not an ideal curve since it has certain width in practice as shown in Fig.4 (b), hence, there may be one or more cross point can be found. The best solution is stripe thinning which can generate a curve with one pixel width shown in Fig. 4(c). In this case, the number of cross point is not more than one. Since the pixel coordinate is discontinuous, there may be no cross point could be found on the image2. Here we use the closest pixel point to replace the ideal cross point and only a sub-pixel error would be introduced shown in Fig.4(d).

Finally we get the coordinate (u2,v2)of the pixel point on image2. So there are only 5 unknown numbers in the six equations. Therefore, only one of the two coordinate(u2,v2) of the pixel point on image2 has to be used to figure the point in the world coordinate. It can be deduces from equations(6) and(7):

$$\begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} = \begin{pmatrix} k_{11} - U_1k_{31} & k_{12} - U_1k_{32} & k_{13} - U_1k_{33} \\ k_{21} - V_1k_{31} & k_{22} - V_1k_{32} & k_{23} - V_1k_{33} \\ t_{11} - U_2t_{31} & t_{12} - U_2t_{32} & t_{13} - U_2t_{33} \end{pmatrix}^{-1} \cdot \begin{pmatrix} U_1k_{34} - k_{14} \\ V_1k_{34} - k_{24} \\ U_2t_{34} - t_{14} \end{pmatrix} \tag{9}$$

IV. System calibration and light coding

Zhang's method[15] is used in our system calibration, and the Matlab toolbox for camera calibration provided by Bouguet[16] is utilized to obtain the intrinsic parameter K and extrinsic parameters R and T. Fig. 5 shows the calibration images for camera1.

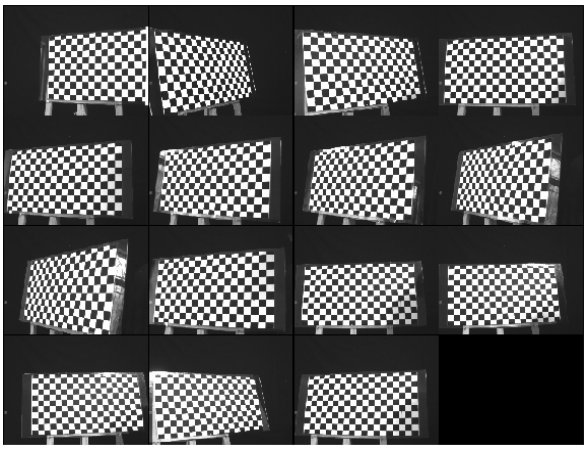

Fig. 5 Calibration images

We use the combination coding of Gray and phase shifting for the fringe coding[17]. Gray coding is a binary coding with strong robustness than the conventional binary coding. Phase shifting coding utilize the sine distribution of the grey scale of the projected light to label the different fringes in pixel scale. So, the combination coding with Gray and phase shifting can obtained pixel level resolution and lower noises [18]. Fig. 6 shows parts of the deformed images in which L-a-1 — L-a-8 is captured by camera1 and R-a-1 — R-a-8 is captured by camera2. These images are captured in different directions as shown in Fig. 6.

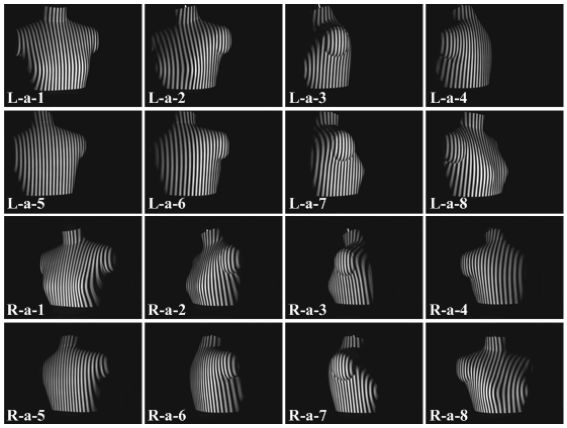

Fig. 6 Deformed images.

V. Result

As shown in Fig.1, the object can be spun by the turn table. At the experiment of measuring the dress form model, the two cameras capture images from 8 different directions shown in Fig.7(a) and each camera captures 9 images in a direction. A point cloud can be figured out by the 18 images captured in a direction. Finally we got 8 point clouds and the dress form model is modelled through the point cloud registration. Fig.7(b) shows the original photo of a dress form model. Fig.7(c) and (d) shows the modelling result and Fig. 7(e) shows the previous result built through the single camera and single projector system and we see the rough surface of the previous model because of the lower accuracy of the projector calibration.

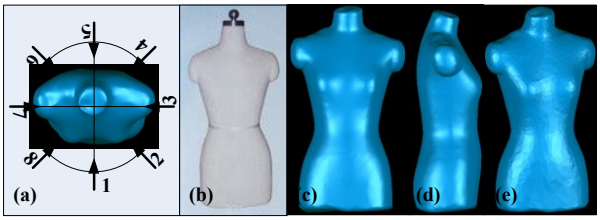

Fig.7 Measurement result

Table1 shows the comparison between the model result based on dual cameras and model result based on single camera. The proportional error of the model based on dual cameras and single projector is less than 0.32% while the proportional error from the single camera and single projector is less than 0.78%.

Table 1 Proportional Error (unit: mm)

Results	Based on dual-camera			Based on single-camera		
Distance from the Bottom	200	400	800	200	400	800
Model Perimeter	684.5	834.6	1142.8	688.6	842.1	1134.0
Real Perimeter	683.3	836.1	1140.2	683.3	836.1	1140.2
Error	1.2	-1.5	2.6	5.3	6.0	-6.2
Proportional Error	0.32%	0.30%	0.32%	0.78%	0.72%	0.54%

VI. Conclusion

Using the dual cameras and single projector method to realize the shape recovery can avoid calibrating the projector which is complicated and time-consuming. We can get a more accurate measurement with the novel method since the camera calibration has a high accuracy than the projector calibration because projector calibration always relies on the camera calibration. Finally, we figured out a dress form model with the new measurement method and the proportional error is under 0.32% compared to the proportional error of 0.78% which is based on the single camera and single projector system.

References

[1] H. Xiuxiang, T. Qing, X. Yimin, L. ZeYuan and C. Lei, "Calibration method of 3D body measurement system using stripe grating projection," *Advanced Materials Research* 503, pp. 1265-1269, 2012

[2] H. Holt Olaf and S. Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects," *In Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference* on, vol. 2, pp. 359-366, 2001.

[3] T. Ming-June, and C. Hung, "Development of a high-precision surface metrology system using structured light projection." *Measurement* 38, no. 3, 236-247, 2005.

[4] R. Szymon, O. Hall-Holt, and M. Levoy, "Real-time 3D model acquisition." *In ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 438-446. ACM, 2002.

[5] S. Min, B. Dmitry Goldgof, and W. Kevin Bowyer. "Comparison of edge detector performance through use in an object recognition task." *Computer Vision and Image Understanding* vol. 84, no. 1, 160-178, 2001

[6] B. Kevin, C. Kranenburg and S. Dougherty. "Edge detector evaluation using empirical ROC curves," *Computer Vision and Image Understanding* vol. 84, no. 1, pp. 77-103. 2001.

[7] H. Yi Chih, "Decoding structured light patterns for three-dimensional imaging systems," *Pattern Recognition* 34, no. 2, pp. 343-349, 2001.

[8] J. Batlle, E. Mouaddib and J. Salvi, King, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey," *Pattern Recognition*, vol. 31, no. 7, pp. 963-982, 1998.

[9] J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition* vol. 37, no. 4, pp. 827-849,2004.

[10] S. Zhang and Y. Shing-tung, "Three-dimensional shape measurement using a structured light system with dual cameras," *Optical Engineering*, vol. 47, no. 1, pp. 01360401 – 01360412, 2008.

[11] S. Changming. "A fast stereo matching method," *Digital Image Computing: Techniques and Applications.* pp. 95-100, 1997.

[12] J. Davis, R. Ramamoorthi and S. Rusinkiewicz, "Spacetime stereo: a unifying framework for depth from triangulation," *Pattern Analysis and Machine Intelligence, IEEE Transactions* vol. 27, no. 2, pp. 296-302, 2005.

[13] T. Mitsuo and K. Mutoh. "Fourier transform profilometry for the automatic measurement of 3-D object shape," *Applied optics* vol. 22, no. 24, pp. 3977-3982, 1984

[14] C. Frank, G. M. Brown and S. Mumin, "Overview of three-dimensional shape measurement using optical methods," *Optical Engineering* vol. 39, no. 1 pp. 10-22, 2000.

[15] Z. Zhengyou, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions* on 22, no. 11 pp. 1330-1334. 2000.

[16] http://www.vision.caltech.edu/bouguetj/

[17] R. Carsten, R. Ritter and J. Thesing, "White light heterodyne principle for 3D-measurement," *Lasers and Optics in Manufacturing III*, pp. 236-244, 1997.

[18] R. Garcia and A. Zakhor, "Consistent stereo-assisted absolute phase unwrapping methods for structured light systems," *Selected Topics in Signal Processing, IEEE Journal of 6*, no. 5, pp. 411-424, 2012.