3D surface reconstruction based on binocular vision using structured light

MA Zhi-feng(马志峰)[1,✉], HAN Fu-hai(韩福海)[1,2], WANG Teng-fei(王腾飞)[1]

(1. School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China;
2. Troop 63600 People's Liberation Army, Jiuquan 732750, Gansu, China)

Abstract: A 3D surface reconstruction method using a binocular stereo vision technology and a coded structured light, which combines a gray code with phase-shift has been studied. The accuracy of the 3D surface reconstruction mainly depends on the decoding of gray code views and phase-shift views. In order to find the boundary accurately, gray code patterns and their inverses are projected onto a human eye plaster model. The period dislocation between the gray code views and the phase-shift views in the course of decoding has been analyzed and a new method has been proposed to solve it. The splicing method is based on feature points. The result of the 3D surface reconstruction shows the accuracy and reliability of our method.

Key words: 3D surface reconstruction; structured light; gray code; phase-shift

CLC number: TP 311.1 **Document code**: A **Article ID**: 1004-0579(2016)03-0413-05

A coded structured light technology is one of the non-contact 3D surface reconstruction technologies[1]. This technology is based on projecting a light pattern and viewing the illuminated scene from one of several points of view. Since the pattern is coded, correspondences between image points and points of the pattern can be easily found. Decoded points can be calculated by a classical triangulation strategy and 3D information can thus be obtained [2].

Pattern projection techniques differ in the way how each point is identified in the pattern. Salvi J [3] et al. surveyed pattern projection techniques and classified them according to their codification strategy. Ming-June Tsai [4] et al. combined the gray code encoding and sub-pixel edge detection to obtain the positions of the fringe edges and centers with a sub-pixel accuracy. However, using the gray code can only uniquely label pixels in a small region. For a higher measuring resolution, Bergman[5] et al. integrated the gray code method and the phase-shift method, which projects a sine wave pattern onto the scene in succession. It exploits higher spatial resolution than the gray code as the phase is continuously distributed within its range of nonambiguity. Li Li-jun[6] et al. used Bergman's method in the binocular system, which includes two cameras and one projector, and eliminated the projector calibration process. In practice, however, this does not work well. Daniel Scharsterin[7] et al. solved this problem by projecting a combination of gray code patterns and their inverses.

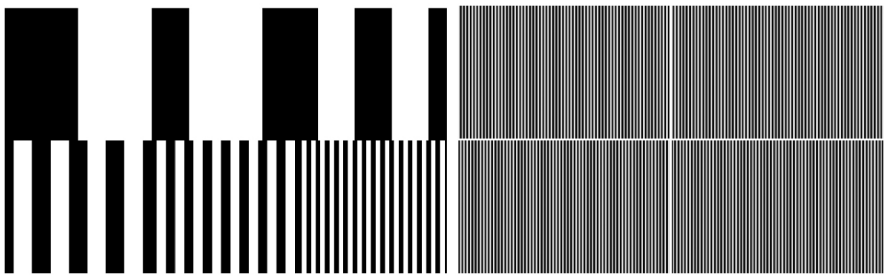
In this paper, we devised a method to acquire object 3D surface information with a high resolution based on binocular vision and projecting a combination of the gray code and phase-shift structured light.

1 Combination of gray code and phase-shift

Our method relies on a DLP projector that first projects Gray code structured light patterns onto a target object. The number of the gray code struc-

Received 2014-12-23
✉ Author for correspondence, lecturer, Ph. D.
E-mail: mazhifeng@bit.edu.cn
DOI: 10.15918/j.jbit1004-0579.201625.0315

tured light patterns is 6, and Fig. 1a shows all of them. Then we project 4 phase-sift structured light patterns which are shown in Fig. 1b onto the target object. At the same time, a pair of CCD cameras captures 10 distorted strips images produced by the projector. The gray code patterns are projected to divide the measuring surface of the plaster into 2^6 regions. For every given pixel (x, y) in same regions, we can get a same period value k after decoding first 6 images. The phase-shift patterns are projected to subdivide each one of the 2^6 regions. After decoding phase-shift views, a given pixel (x, y) in every region will be labeled by a relative phase value $\theta(x, y)$. Then pixels in the measuring surface can be labeled by absolute phase that is calculated by combining the period value k and the relative phase value θ. This process is known as phase unwrapping.



(a) Gray code patterns (b) Phase-shift patterns

Fig. 1 Coded structured light

2 Phase unwrapping

We can get the period value of a given pixel (x, y) from decoding the gray code views, and get the relative phase from demodulating the phase-shift views.

2.1 Thresholding and decoding gray code views

After projecting the gray code patterns onto the target object, we project their inverses. Different from the Daniel Scharstein's method, we set the pixel value to 1 or 0 according to the pixel value difference between the Gray code patterns and theirs inverses.

After thresholding the gray code views precisely, it is easy to decode the gray code views. Then conversion of these binary codes to decimal codes is carried out, and the period value k of the measuring surface can be acquired.

2.2 Decoding phase-shift views

The periodic intensity pattern is projected four times by shifting 1/4 of the period, each time, as Fig. 1b shown. Assuming a linear image formation process, we have the following 4 image formation equations as

$$I_1(x,y)=I'(x,y)+I''(x,y)\cos\left[q(x,y)\right] \tag{1}$$

$$I_2(x,y)=I'(x,y)+I''(x,y)\cos\left[q(x,y)+p/2\right] \tag{2}$$

$$I_3(x,y)=I'(x,y)+I''(x,y)\cos\left[\theta(x,y)+\pi\right] \tag{3}$$

$$I_4(x,y)=I'(x,y)+I''(x,y)\cos\left[\theta(x,y)+3\pi/2\right] \tag{4}$$

where $I_i(x,y)$ is the grey levels of pixel (x,y) from images corresponding to the four phase-shift patterns, $I'(x,y)$ is the average intensity of the patterns, $I''(x,y)$ is the albedo corresponding to the object surface pixel (x,y), and $\theta(x,y)$ is its phase.

$$\theta(x,y)=\arctan\frac{I_{10}(x,y)-I_8(x,y)}{I_7(x,y)-I_9(x,y)},\ -\pi\leqslant\theta\leqslant\pi \tag{5}$$

2.3 Absolute phase

Through the two procedures above, integrate period value k and phase $\theta(x,y)$, and absolute phase $\Psi(x,y)$ is acquired as

$$\Psi(x,y)=2k\pi+\theta(x,y) \tag{6}$$

Then we need a period rectification because of the period dislocation.

2.4 Period rectification

In ideal conditions, the increase of a pixel

(x, y) period value k must be consistent with the mutation of period value $\theta(x, y)$ from π to $-\pi$, because the gray code and the phase-shift have the same period. Unfortunately, we often find that there is a big dislocation in the periodic boundary. For example, the phase of a given pixel $P(x_c, y_c, z_c)$ is π, and the phase of its neighborhood pixel $P(x+1, y)$ is $-\pi$, however, the difference of their period does not equal 1. Similarly, when the difference of their period equals 1, but the phase does not change from π to $-\pi$. This phenomenon is called a period dislocation. Since we project the gray code patterns and their inverses, the periodic boundary of the gray code views is more reliable than the periodic boundary of the phase-shift views. According to this, we can adjust the periodic boundary in the phase-shift views to be consistent with the period boundary in the Gray code views.

A pixel $S(x, y)$ and its neighboring pixel $P(x, y)$ are located on both sides of the periodic boundary in gray code views. If their phases meet one of the following three conditions, $\theta_S(x, y) < \theta_P(x, y)$, $\theta_S(x, y) < 0$, $\theta_P(x, y) > 0$, it means that a period dislocation has happened. Then, looking for the mutation point of its phase value, the phase values of the points between it and the boundary are obtained by the straight line fitted to this point and opposite points.

3 Image splicing

Image splicing is the 3D points cloud splicing. We need calculate 3D points cloud of each surface. Then we splice all the 3D points cloud together.

3.1 3D points cloud calculating

A binocular stereo vision measurement is based on the parallax. A triangular restricted relationship is created by the image plane of the two cameras and the object. Then, three-dimensional surface information can be obtained.

In Fig. 2, two cameras are exactly the same. Baseline distance B is the distance between two connections from the optical center of the camera, f is the focal length of the camera.

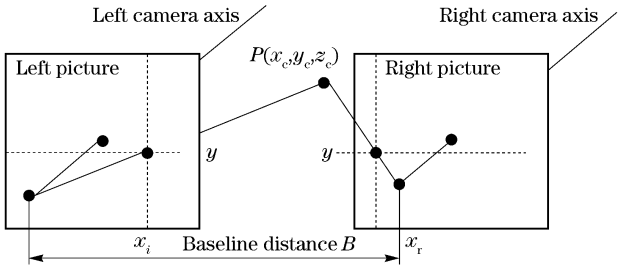


Fig. 2 Binocular stereo measurement principle

$P(x_c, y_c, z_c)$ is one 3D point on the surface of the object, p_l and p_r are the image points of the two cameras, and their physical coordinates are $p_l = (x_l, y_l)$, $p_r = (x_r, y_r)$.

We know that $y_l = y_r = y$ by rectification.

$$\begin{cases} x_l = fx_c / z_c \\ x_r = f(x_c - B) / z_c \\ y = fy_c / z_c \end{cases} \tag{7}$$

Parallax $D = x_l - x_r$. The 3D point of P in the left camera coordinate system is

$$\begin{cases} x_c = Bx_l / D \\ y_c = By / D \\ z_c = Bf / D \end{cases} \tag{8}$$

We need to filter out noise due to the error.

If $z_c \neq 0$, $z_{c-1} = 0$ and $z_{c+1} = 0$, then $x_c = 0$, $y_c = 0$, $z_c = 0$;

If $z_c \neq 0$, $|z_c - z_{c-1}| > 0.5$ and $|z_c - z_{c+1}| > 0.5$, then $x_c = 0$, $y_c = 0$, $z_c = 0$ (9)

3.2 3D points cloud splicing

We use the splicing method based on markers and manually select feature points.

3.2.1 Recognition of feature identifying points

The premise based on the feature identifying points splicing method is to obtain accurate three-dimensional data of feature identifying points. Since the number of markers is at least three, it can not be a multi-point collinear case. Fig. 3 shows a plaster model with markers affixed.

Because the image is converted to a gray image, so the gray value of each of its points is between 0 and 255. We artificially select a threshold and binarization. Fig. 4 shows the results of Fig. 3

after binarization. We repeat it until an appropriate value selection is found.

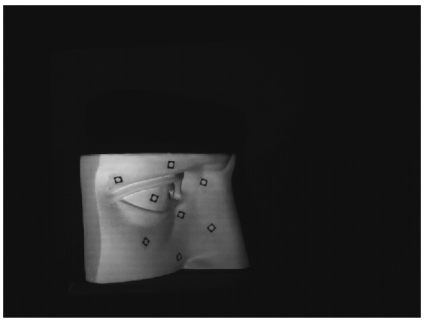


Fig. 3 Plaster model with markers affixed

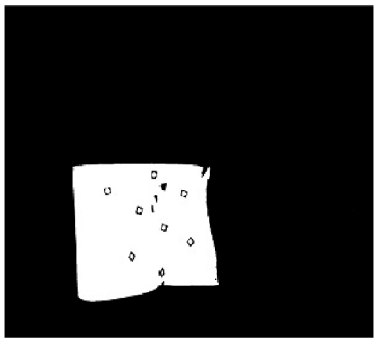


Fig. 4 Image after binarization

3.2.2 Multi-angle splicing

Measurement data points having a part of the overlapping region is obtained at two different viewing angles Ⅰ, Ⅱ. In Fig. 5, a set of points of measurement in region Ⅰ is P, a set of points of measurement in region Ⅱ is Q, coordinate $Oxyz$ is the world coordinate, $O_1x_1y_1z_1$ is the camera coordinate of the left camera at position A, $O_2x_2y_2z_2$ is the camera coordinate of the left camera at position B. The camera coordinate of the right camera is omitted.

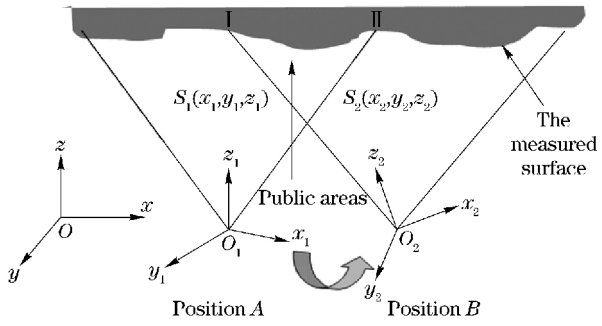


Fig. 5 Schematic data splicing

If we extract a subset $\{m_i \mid m_i \in P, i=1,2,\cdots,N\}$ containing N data points in the set of data points P, there is a subset $\{m'_i \mid m'_i \in Q, i=1,2,\cdots,N\}$ containing N data points in the set of data points Q corresponding to each of its data points. Matching process is solving the coordinate conversion $\boldsymbol{R}$ and $\boldsymbol{T}$ of the three-dimensional relationship between the data points in two different coordinate systems. $\boldsymbol{R}$ is the rotation matrix and $\boldsymbol{T}$ is the translation matrix.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} + \boldsymbol{T} \tag{10}$$

Generally, $\boldsymbol{R}$ and $\boldsymbol{T}$ can be obtained by minimizing the objective function as

$$J = \sum_{i=1}^{N} \| m'_i - (\boldsymbol{R}m_i + \boldsymbol{T}) \|^2.$$

The following describes the four elements method.

①Solving the centroid of the set of two feature points $\{m_i \mid m_i \in P, i=1,2,\cdots,N\}$ and $\{m'_i \mid m'_i \in Q, i=1,2,\cdots,N\}$,

$$\begin{cases} u = \dfrac{1}{N}\sum_{i=1}^{N} m_i \\ u' = \dfrac{1}{N}\sum_{i=1}^{N} m'_i \end{cases} \tag{11}$$

②Making the feature point set $\{m_i \mid m_i \in P, i=1,2,\cdots,N\}$ and $\{m'_i \mid m'_i \in Q, i=1,2,\cdots,N\}$ a translation with respect to the centroid, respectively,

$$\begin{cases} u = m_i - u \\ u' = m'_i - u' \end{cases} \tag{12}$$

③ Calculating the correlation matrix $\boldsymbol{K}$ by the set of points $\{P_i\}$ and $\{P'_i\}$ after moved,

$$\boldsymbol{K} = \frac{1}{N}\sum_{i=1}^{N} P_i (P'_i)^{\mathrm{T}} \tag{13}$$

④Constructing a four-dimensional symmetric matrix $\boldsymbol{K}_\alpha$ by each element $K_{ij}(i,j=1,2,3)$ in correlation matrix $\boldsymbol{K}$,

$$\boldsymbol{K}_\alpha = \begin{bmatrix} K_{11}+K_{22}+K_{33} & K_{32}-K_{23} & K_{13}-K_{31} & K_{21}-K_{12} \\ K_{32}-K_{23} & K_{11}-K_{22}-K_{33} & K_{12}+K_{21} & K_{31}+K_{13} \\ K_{13}-K_{31} & K_{12}+K_{21} & -K_{11}+K_{22}-K_{33} & K_{23}+K_{32} \\ K_{21}-K_{12} & K_{31}+K_{13} & K_{23}+K_{32} & -K_{11}-K_{22}+K_{33} \end{bmatrix} \tag{14}$$

⑤ Calculating unit feature vector $\boldsymbol{q}$ corre-

sponding with the largest eigenvalue of $\boldsymbol{K}_a$,

$$\boldsymbol{q}=(q_0,q_1,q_2,q_3)^{\mathrm{T}} \tag{15}$$

⑥Calculating rotation matrix $\boldsymbol{R}$,

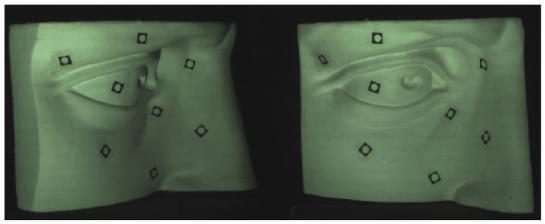
$$\boldsymbol{R}=\begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2+q_2^2-q_1^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2+q_3^2-q_1^2-q_2^2 \end{bmatrix} \tag{16}$$

⑦Calculating translation matrix $\boldsymbol{T}$,

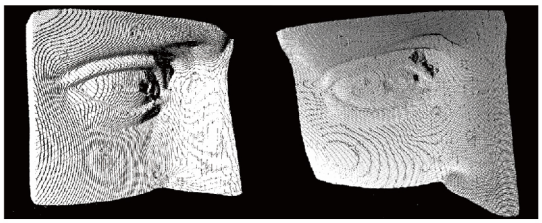
$$\boldsymbol{T}=u'-\boldsymbol{R}u \tag{17}$$

3.3 Experimental results

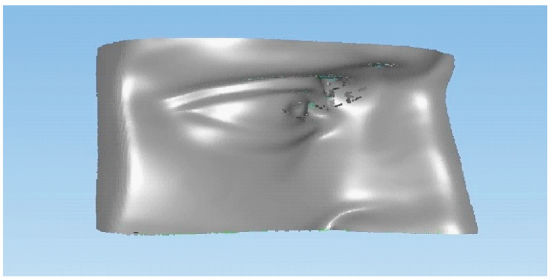
Using the method described in the previous section, we reconstructed the surface of a human eye plaster model (Fig. 6a). Fig. 6b shows the point cloud of the human eye plaster model. Fig. 6c shows the reconstruction result of 3D points cloud splicing.



(a) Human eye plaster model



(b) 3D points cloud model



(c) Reconstruction result of 3D points cloud splicing

Fig. 6 A group of human eye plaster model image and its results of 3D point cloud

4 Conclusion

In this paper, we have proposed a novel method to reconstruct an object surface using a binocular vision technology and a coded structured light combining the gray code with the Phase-shift. We analyzed the period dislocation when period value k and phase θ are integrated. The proposed method has been tested on a real object. Good results for the reconstruction surface of the plaster model indicate the accuracy and reliability of our method.

References:

[1] Sang Xinzhu. A method on 3-D shape measurement and its trend in development [J]. Measurement Transaction, 2001, 16(2): 36－37. (in Chinese)

[2] Salvi J, Pagès J, Batlle J. Pattern codification strategies in structured light systems [J]. Pattern Recognition, 2004, 37(4):827－849.

[3] Pagès J, Salvi J, García R, et al. Overview of coded light projection techniques for automatic 3D profiling [C]// IEEE International Conference on Robotic and Automation, Taipei, Taiwan, China, 2003.

[4] Tsai Ming june, Hung Chuancheng. Development of a high-precision surface metrology system using structured light projection [J]. Measurement, 2005, 38(3):236－247.

[5] Bergmann D, New approach for automatic surface reconstruction with coded light [C]// Proceedings of Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, SPIE, San Diego, CA, United States, 1995.

[6] Li Lijun, Ke Yingjie, Jing Kaiyong. Surface reconstruction based on computer stereo vision using structured light projection [C]// International Conference on Intelligent Human-Machine System and Cybernetics, Hangzhou, China, 2009.

[7] Daniel Scharstein, Richard Szeliski. Hight-Accuracy stereo depth maps using structured light [C]// IEEE Computer Conference on Computer Vision and Pattern Recognition (CVPR), Madison, WI, 2003.

(**Edited by** Cai Jianying)