2009 International Conference on Intelligent Human-Machine Systems and Cybernetics

Surface Reconstruction Based on Computer Stereo Vision Using Structured Light Projection

Lijun Li
College of Mechanical Engineering and Automation, HuaQiao University
Mold & Die Technology and Research Center
Quanzhou, Fujian, China
jlj@hqu.edu.cn

Yingjie Ke
College of Mechanical Engineering and Automation, HuaQiao University
Mold & Die Technology and Research Center
Quanzhou, Fujian, China
yingjieke@gmail.com

Kaiyong Jiang
College of Mechanical Engineering and Automation, HuaQiao University
Mold & Die Technology and Research Center
Quanzhou, Fujian, China
jiangky@hqu.edu.cn

Abstract—**Object reconstruction is one of the most important topics in computer vision due to its wide field of application. Reverse engineering has an increasing need for reconstruction of stereo parts. A system for reconstruction three dimensional (3-D) object's surface from its 2-D views using a coded structured light combining with the epipolar geometry that exists between the stereo image pair is presented. A set of multiple gray coded patterns and a set of sine wave patterns are combined to form a single "Gray Code Composite Pattern" and a single "Phase Shift Composite Pattern" respectively. These two structured light patterns are projected onto the target object through a DLP projector and the distorted strips patterns produced by its surface are captured by two CCD cameras. By demodulating the distorted patterns and adding epipolar constraints the match points between the image pair are found out and then the target object's 3-D point's coordinates on the surface are obtained. The system shows excellent linearity and the results shows good resolution.**

Keywords- computer stereo vision; reverse engineering; structured light; epipolar geometry; surface reconstruction

I. INTRODUCTION

The last few years has seen an increasingly use of computer vision in reverse engineering for rapid prototyping. Object construction is the one of the most important part in computer vision since it's widely used. Stereo vision is based on viewing the scene from two or more points of view and then finding correspondences between the pair images in order to obtain the 3-D position [1]. The obtaining of 3-D coordinates is possible if cameras have been calibrated beforehand. However, difficulties in finding the corresponding points between the pair of views arise, even when taking into account epipolar constraints. Coded structured light consists of one or two cameras and a device that projects light patterns onto the measuring surface. Nowadays the most commonly used devices are LCD and DLP video projectors. Such devices project an image with a certain structure so that a set of pixels are easily distinguishable by means of a local coding strategy. Therefore, when locating such coded points in the image grabbed by the cameras, the correspondence problem is solved with no requirement for geometrical constraints [2]. The projecting images are called patterns, as they are globally structured, and the images captured by cameras are called views.

In this paper, we devise a method to acquire object 3-D information with a high resolution based on stereo vision using a combination of gray-code and phase-shift structured light projection. The gray-code technique allows the unique description of 2n different coded strips of projection, but the measurement resolution is still low. Combined with the phase-shift approach, each pixel in every strip is coded by phase values: Because the phase is continuously distributed within its range of nonambiguity, a theoretically infinite height resolution can be obtained [3]. Now take the epipolar constraints into account, we can uniquely determine the corresponding point to a certain view point in left or right view.

The remainder of this paper is structured as follows. The next section describes the architecture of our system. Section III gives an overview of the algorithms used to determine unique code from the structured lighting. In section IV, we formulate the camera calibration method and the registration process between the pair of views. Section V presents a leaf blade surface reconstruction results. Finally, our conclusion and future work is suggested in Section VI.

II. THE SYSTEM

The goal of our system is to acquire 3-D point cloud of the target object surface and texture it from its 2-D views. Fig. 1 shows the steps for acquiring 3-D coordinates of the points on the target object surface. Our approach relies on a DLP projector that cast structured light patterns onto the target object and a pair of CCD cameras that capture the distorted strips images produced by it. By demodulating the distorted patterns and adding epipolar constraints the match points between the image pair are found out and then the stereo positions of these surface points are obtained by point reconstruction. The structured light patterns contain of a set of multiple gray coded patterns and a set of sine wave patterns which are respectively combined to form a "Gray Code Composite pattern" and a "Phase Shift Composite Pattern". We choose to perform the complete acquisition

978-0-7695-3752-8/09 $25.00 © 2009 IEEE
DOI 10.1109/IHMSC.2009.235

process of the views in a completely dark room. This reduces the intervention from light sources other than the projector light.

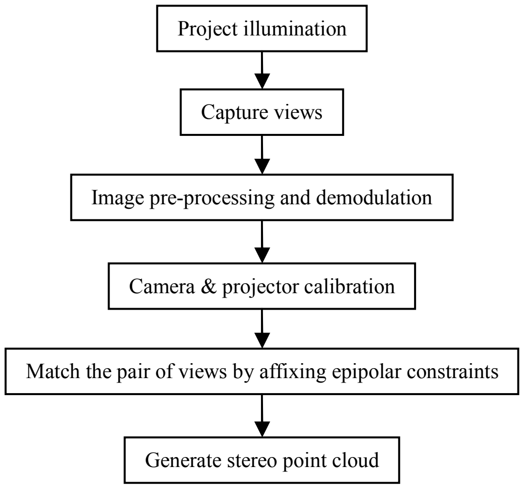

Figure 1. The steps of coordinates of the surface point acquisition.

III. STRUCTURED LIGHT

To uniquely label pixels in a small region, we project a set of structured light patterns onto the object surface, and then decode the set of projected intensities at each pixel to give it a unique label. In order to distinguish among all the projected stripes, a Gray Code Composite pattern is projected in order to label each region. A sine wave Phase Shift Composite Pattern is projected in succession in order to label pixels on a line strip in each region. At this section the design of these two patterns as well as the decoding of the Gray Code Composite Pattern and the demodulation of the Phase Shift Composite Pattern are described in detail.

A. *Design of Gray Code Composite Pattern*

Gray code is the code in which the numbers are represented as binary patterns and the consecutive numbers differ by only one bit position. Gray codes are well suited for such binary position encoding, since only one bit changes at a time, and thus small mislocalizations of 0-1 changes cannot result in large code changes [4].The first step in generating a Gray Code Composite Pattern is to create Gray code structured light patterns. Gray code patterns only contain black and white (on/off) pixel values. The division of 2n strips in the scene requires n Gray-code patterns. We choose TOSHIBA TDP-T420 to doing the projection job in consideration of its excellent performance. The influence of environmental light is too slight to taking in to account because this DLP projector has 4000 ANSI Lumens which can provide a high contrast in the projection. For our projector with 1024×768 pixels, it is sufficient to illuminate the scene with 6 vertical patterns, and the thinnest strips are 16 pixels width. Fig. 2 shows the Gray Code Composite Pattern.

B. *Dedign of the Phase Shift Composite Pattern*

Unfortunately, using Gray the Code Composite Pattern can only uniquely label pixels in a small region. So we project another sine wave pattern onto the scene in succession. Because the phase is continuously distributed within its range of nonambiguity, a theoretically infinite height resolution can be obtained. The sine wave pattern is projected 4 times by shifting it 1/4 of its half period so that by using a adequate linear combination of the 4 images the albedo of the object can be canceled and the phase of the pixels on a vertical line strip in each minimal region can be calculated through the illuminance values. Fig. 3a shows the sine wave pattern. Fig. 3b shows the Phase Shift Composite Pattern enlarged views.

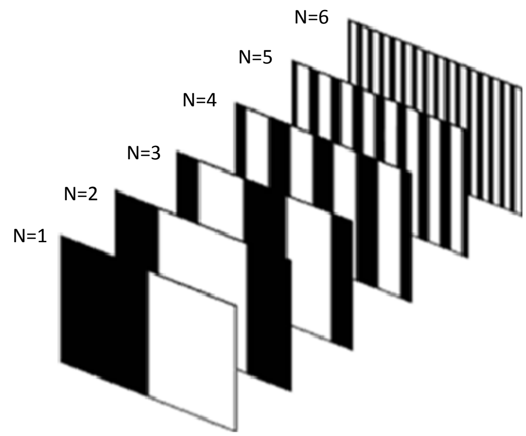

Figure 2. The Gray Code Composite Pattern

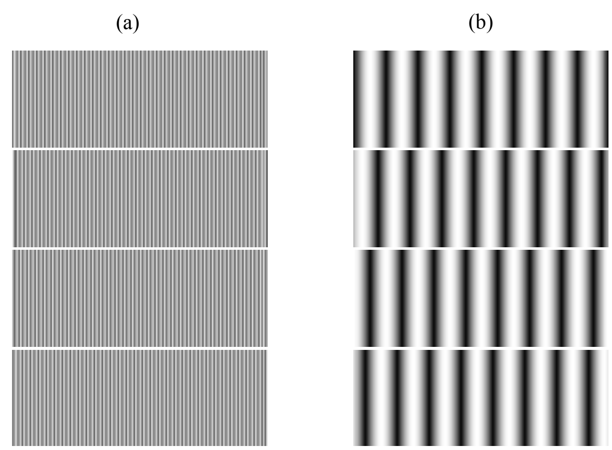

Figure 3. (a) Sine wave pattern and its three steps phase shift views
(b) The Phase Shift Composite Pattern enlarged views.

C. *Image pre-processing and composite patterns demodulation*

1) Noise removal

Before we start the decoding process, we assume that the images we are dealing with are noisy. Such noise is due to unfavorable lighting circumstances, object textures, camera CCD capturing noise and especially the low pixel density of the projector's DLP grid.

If the noise is not eliminated from the image, it will lead to significant error in the decoding process and directly influences the resolution of point cloud reconstruction. An adaptive smooth filter algorithm is adopted in our system to solve this problem, the main idea of the algorithm is to carry out iterative convolution operation of the image with a mean weighted template, and the weighted coefficients of the template is determined by the gradient function of the pixel [7].The advantage of this algorithm is as follows [8]:

- The iterative operation in adaptive smooth filter can sharp the edge of image, high location precision will be obtained when the edge is detected.
- With multiple iterative operations, the image finishes adaptive smoothing according to the block; the fuzzy edge can be avoided.

2) Binarising and Decoding with Gray code views

In practice, however, the bright strips seems more width than the black ones on every Gray coded pattern due to interreflections in the scene and "fogging" inside the projector [5] (adding a low-frequency average of intensities to the projected pattern), which causes confusion at the boundary between bright strips and black ones. So the Gray code views need to be binarized using thresholding method which is if the pixel value for each pixel in the demodulated pattern is greater than a threshold value then that pixel value is set equal to 1 else it is set equal to 0.

In order to get the codeword of each minimal strip on the Gray Code Composite Pattern view, the decoding of the Gray Code Composite Pattern is carried out. The conversion from Gray code word to decimal code word needs the help of binary code word. The canonical way to convert Gray code to binary code is to XOR the bits one at a time, starting with the two highest bits, using the newly calculated bit in the next XOR [9]. Convent these binary codes to decimal codes, and we have accomplished the decoding process.

3) Demodulation of the phase shift views

Given the images of the object illuminated with these patterns, how do we compute the phase and hence (u,v) coordinates at each pixel? Assuming a linear image formation process, we have the following image formation equation.

$$I_n(u,v) = B(u,v) + A(u,v)\cos\left(\emptyset - \frac{n\pi}{N}\right), n = 1\sim 4 \quad (1)$$

Where $B(u,v)$ is the average intensity of the nth projected pattern, $A(u,v)$ is the albedo corresponding to the object surface pixel (u,v), and $\emptyset$ is its phase.

According to equation (1) we have the following four patterns of sine wave phase shift image formation equations.

$$I_1(u,v) = B(u,v) + A(u,v)\cos\left(\emptyset - \frac{\pi}{4}\right) \quad (2)$$

$$I_2(u,v) = B(u,v) + A(u,v)\cos\left(\emptyset - \frac{\pi}{2}\right) \quad (3)$$

$$I_3(u,v) = B(u,v) + A(u,v)\cos\left(\emptyset - \frac{3\pi}{4}\right) \quad (4)$$

$$I_4(u,v) = B(u,v) + A(u,v)\cos(\emptyset - \pi) \quad (5)$$

Demodulation of the phase shift views is accomplished when the phase $\emptyset$ of a certain pixel (u,v) is calculated from the four equations above.

IV. CALIBRATION AND REGISTRATION

A. Camera and Projector Calibration

In the camera calibration process one determines the intrinsic parameters of the camera i.e. the internal geometric and optical characteristics of the camera and/or the 3-D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters). The camera calibration provides a way to determine the position of a ray in a 3-D space that the object point must lie on, given the computer image coordinates. The same calibration process is applied for determining the intrinsic and extrinsic parameters of the projector.

We use the Zhang calibration method [6] as implemented by OpenCV API cvCalibrateCamera2() for intrinsic and extrinsic calibration of the camera. As proposed by OpenCV, a chessboard pattern is used for calibration, as Fig. 4 shows. This calibration model consists of focal length with respect to pixel widths and heights, the principal point and a radial and a tangential lens distortion modeled by two parameters each.

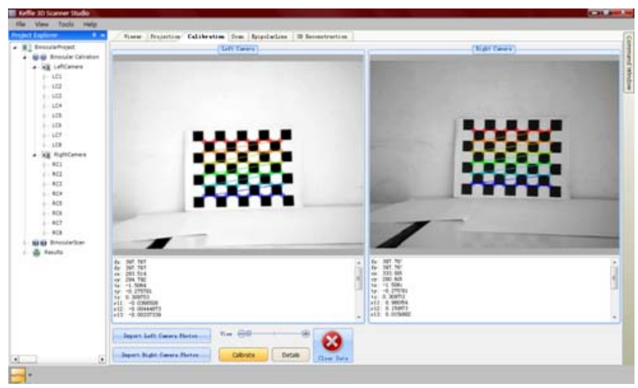

Figure 4. A snap of camera calibration process

We calibrate the projector as an inverse camera [9]. This means that instead of taking pictures of a chessboard with known geometry and detecting the corners inside the images, a chessboard pattern with known geometry is projected to different orientations and positions of a plane and the projections are measured with the calibrated camera. We project chessboard patterns and use the same corner detection algorithm to find the corners inside the images of the projections. In order to measure the 3-D position of the projected pattern with the calibrated camera, we attach a printed chessboard pattern to the projection plane and use extrinsic calibration to determine its position. The resulting points are fed into the OpenCV camera calibration routine.

B. Registration

Given an image point *x* in the left view, how does this constrain the position of the corresponding point x' in the right image? Epipolar constraints can figure this out. The epipolar geometry exists in a binocular system. Each epipolar line is a projection of the ray connecting the object point with the other camera position onto the current image plane. All epipolar lines pass their epipole. In this binocular system, the left epipole is the image of the projection center of the right camera and vice versa. The epipolar constraints are defined as the corresponding points can only lie on the epipolar line in the second image. Adding these constraints reduces 2-D searching of corresponding points to 1-D.

The epipolar line equations can be calculated from the two camera calibration matrices we obtained from calibration process. If an image point x in the left view is given, the corresponding epipolar line equation is determined by decomposing the two camera calibration matrices. As the corresponding point x' has the same codeword and phase in the Gray Code Composite Pattern and the Phase Shift Composite Pattern with point x, point x' definitely is the intersecting point of the certain phase line in the certain codeword strip and the epipolar line in the right view. The registration result after doing the image rectification is shown in Fig. 5.

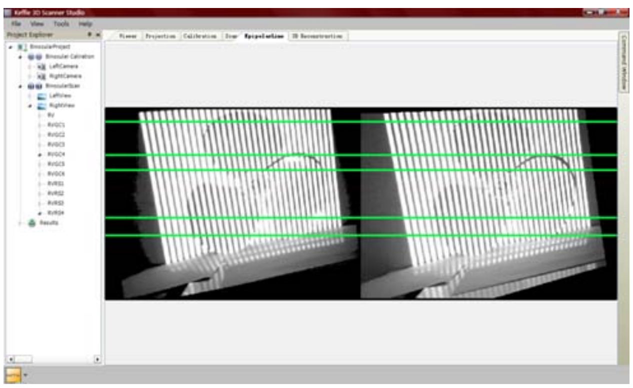

Figure 5. Registration result is shown above. The horizontal epipolar lines are drawn on the left and right views accurately.

V. RECONSTRUCTION OF A LEAF BLADE SURFACE

Stereo surface reconstruction presuppositions knowing of the viewing geometry (extrinsic parameters) and the intrinsic parameters, and then reconstructs after finding correspondences exploiting epipolar geometry. In our approach we use the point reconstruction method. The stereo point X coordinates on the object surface is calculated by the equations of its two image points x and x' and the stereo calibration matrix. Fig. 6 shows the point cloud of our leaf blade surface reconstruction result.

VI. CONCLUSION AND FUTURE WORK

The concept of stereo point reconstruction based on a combination of Gray code and phase-shift light projection is introduced. Demodulation is carried out to the captured composite pattern to obtain Gray codeword of each minimal strip and phase of each point in any one of them. By adding the epipolar constraints the match points between the pair of views taken by two cameras are found out fast and accurately. The stereo points' coordinates is obtained from these corresponding image points using least squares method. This system has the advantage of speeding up the reconstruction process as well as showing excellent linearity. The reconstruction of this leaf blade surface also shows good resolution. Our future work is aim to texture the reconstructed surface and to generate the NC code program for controlling a specialized machine to finish the machining process and to accomplish a CAD/CAM system.

ACKNOWLEDGMENT

This paper is supported by State science and technology support projects of China (2006BAF01A23), key science and technology research project from Fujian province of China (2008H0085).

REFERENCES

[1] http://en.wikipedia.org/wiki/Stereo_vision.

[2] Jordi Pag`es and Joaquim Salvi, "Coded light projection techniques for 3-D reconstruction," EDP Sciences, vol. 4, May, 2005, doi: http://dx.doi.org/10.1051/bib-j3ea:2005801.

[3] Giovanna Sansoni, Matteo Carocci, and Roberto Rodella , "Three-dimensional vision based on a combination of gray-code and phase-shift light projection: analysis and compensation of the systematic errors," Applied optics, Vol. 38, pp. 6565-6573, Nov 1999.

[4] K. Sato and S. Inokuchi, "Three-dimensional surface measurement by space encoding range imaging," J. Robotic Systems, vol. 2, pp. 27-39 ,1985.

[5] Daniel Scharstein and Richard Szeliski, "High-Accuracy Stereo Depth Maps Using Structured Light," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03), 2003.

[6] Z. Zhang, "A Flexible New Technique for Camera Calibration," Microsoft Technical Report, http://research.microsoft.com/˜zhang, December, 1998

[7] HanJunwei, "Research and Application of An Edge Extraction Algorithm," Northwestern polytechnical University, PP.19-22, 2001.

[8] Xiaoji Liao and Kaiyong Jiang, "Upper Automatic Sewing Technology Based On Image," Proceedings of the IEEE International Conference on Information and Automation, PP. 524-529, 2008.

[9] http://www.dspguru.com/comp.dsp/tricks/alg/grayconv.htm.

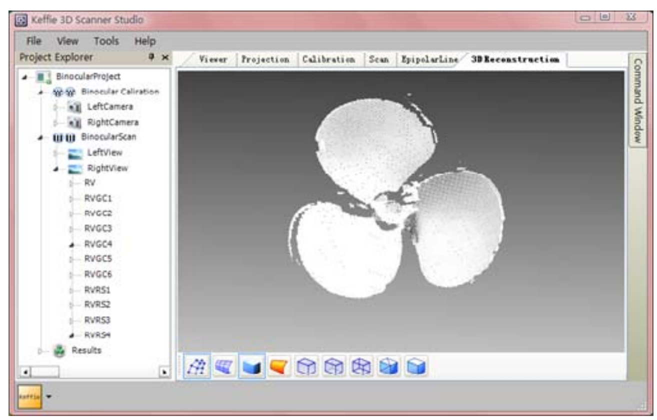

Figure 6. It shows the point cloud of our leaf blade surface reconstruction result.