Aperture Detection and Alignment Control Based on Structured Light Vision*

LI Yuan[12], WANG Qinglin [1], FAN Zhun [2], CHEN Hui [1], SUN Yong [1]

1 School of Automation, Beijing Institute of Technology, Beijing 100081, P.R. China
2 Department of Management Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
E-mail: liyuan@bit.edu.cn, wangql@bit.edu.cn, zhfa@man.dtu.dk, learningpub202@126.com, 20906093@bit.edu.cn

Abstract: Apertures play a critical role as the marks for workpiece positioning on assembly lines. The detection and alignment of an aperture on workpiece surface is a typical manipulation task for assembling robots. In this paper, a vision sensor based on laser structured light is presented to detect an aperture and align it to a robotic end-effector with an eye-in-hand vision system. The images with laser stripes are segmented via adaptive threshold and the stripes skeleton are extracted, then the cross stripes are detected by Hough transformation. The feature points of the aperture are extracted according to intensity distribution of pixels on the cross. A set of visual features and a partitioned visual control law are presented for aperture alignment based on structured light. Experiments have been conducted to test the effectiveness of the algorithms developed.

Key Words: Robot Vision, Structured Light, Features Extraction, Aperture Detection.

1 INTRODUCTION

The detection and location of circular aperture on workpieces is an important task for assembly robot, which is the foundation of automatic assembly [1,2]. For robotic dimensional measurement and objective positioning tasks, there exist several basic methods based on robot vision, such as stereo vision description, structured light scanning, and motion based vision measurement, etc. Among these vision measurement methods, structured light vision is regard as a promising method because of its simplicity, good performance in real-time and abundant information. Structured light vision has been successfully introduced to industrial fields such as welding robots, assembly robots and spraying robots, etc. A good illustration is vision control in seam tracking for welding robots based on laser scanning[3]. Structured light vision is not only used in welding robot for positioning weldment, also in dimensional measurement of workpieces and surface defects inspection [4~6].

Some vision systems were developed for apertures detection and position on workpiece surface. In [7], a conic laser was adopted to scan the surface of workpiece, and an endoscope above the circular aperture acquired the image for bore surfaces inspection. The aperture was reconstructed in three-dimension space via processing the sequence images. For an assembly robot, the workpiece alignment task includes two aspects: one is the apertures visual detection and position, the other is the reaching strategy for robotic alignment, i.e. the vision control law design. For the former task, structured light vision, especially laser stripe scanning, is a competitive solution; for the latter, a robust and precise vision control law is highly desirable.

At present, some exciting and fruitful research results have been achieved for robot visual servoing, e.g. the classical position based visual servoing (PBVS) and image based visual servoing (IBVS); they have the advantages and disadvantages in 2D and 3D space. PBVS performs well and is stable in Cartesian space but sensitive to calibration error, sometimes even out-of view may occur; IBVS has good trajectory in image space but the singular pose is a problem, which would lead to servoing failure [8]. To improve the performance of vision servoing, advanced approaches such as hybrid visual servoing were presented, among which 2.5D visual servoing is regarded as a representative approach [9]. Another advanced approach is partitioned visual servoing [10]. P. Corke presented partitioned control for plane alignment task, in which the translational DOF of z axis was decoupled to the two rotational DOFs. J. Pages [11] presented a visually positioning method for a plane-to-plane alignment task using four pointed laser emitters, by which three degrees of freedoms of a plane were visual controlled via image features. Furthermore, global stability and robustness to vision sensor error was demonstrated in [12].

In our opinions, the following factors should be combined in vision control law design: the vision task, sensing method, sensor configuration and visual feature selection. In this paper, a visual sensor based on structured light is presented to detect and align to the aperture on workpiece surface. Some technical issues, the vision sensor configuration, image features extraction and vision control law will be addressed.

The rest of the paper is organized as follows. The processing of structured light image of visual sensor is illustrated in Section II. In Section III, the aperture brim features extraction algorithm based on distribution of intensity is introduced. In Section IV, the control law is presented for aperture alignment based on the visual features. The experimental results are shown in Section V. Finally, Section VI concludes the paper and gives an expectation of the future work.

2 SENSOR IMAGE PROCESSING

2.1 Vision Sensor Configuration

The structure of visual sensor is shown in Figure .1. The vision sensor is configured as follow manners. Two laser emitters are employed to project perpendicular laser light planes to form a cross stripe feature on the workpiece, and the laser stripes will be distorted with the aperture brim. A camera with optical filter amounted on a robot end-effector captures the feature stripes. Robotic task is to align the camera to the aperture on the workpiece surface. In desirable position, the camera optical axis is vertical to the workpiece surface and aligns to the aperture. In camera frame C, x_c and y_c are parallel to the two laser planes Π_1 and Π_2, respectively. The feature points of the image are extracted for detecting the aperture, and the pose of the camera is controlled by visual features.
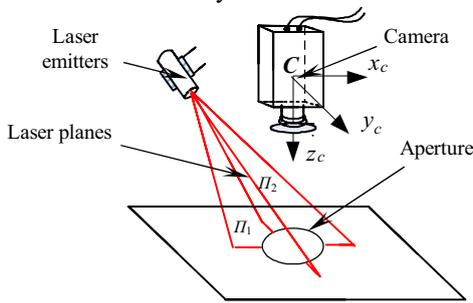
Fig. 1 Aperture detection sensor based on laser structured lights

The gray image captured by the camera is with size of 640×480 pixels. The image processing algorithms include image segmentation, binarization, skeleton extraction and detection of cross stripes.

2.2 Image Segmentation

The typical image of laser structured light is shown as Figure.2. When the perpendicular laser planes were projected on the surface of workpiece with an aperture, there is a cross feature which is composed of line l_1 and l_2. There are partly translated stripes (the broken points) near the aperture because of the variety of the surface height. The edge of the circle aperture is figured by the translation of the cross stripe, which would be extracted as image features.
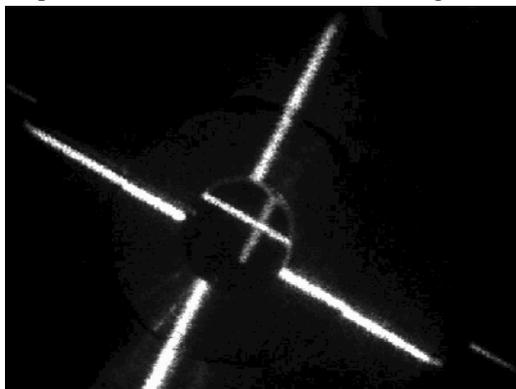
Fig. 2 Structured light image of the vision sensor

Because the stripe contrasts clearly with the dark background, it is proper that the cross is segmented from the background by intensity threshold. Figure.3(a)is an intensity histogram of the original image. It is shown that the intensities of background pixels converge on the area of low intensity; the pixels on the stripe converge on the area of very high intensity.
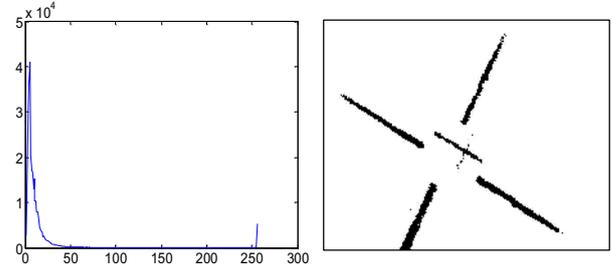
Fig. 3 (a) Intensity histogram of the original image;
(b) Image segmentation result with threshold

Since the pixels of stripe have the higher intensity level, the number of pixels belonged to every intensity level are counted from high level of intensity. When the accumulation of the pixels reaches specified threshold, the intensity level at that moment is regarded as the image segmentation threshold [3], shown as (1).

$$\begin{cases} Ft = \sum_{x=0}^{x_{max}} \sum_{y=0}^{y_{max}} P(x,y,i) \\ P(x,y,i) = \begin{cases} 1 & I(x,y) \geq i \\ 0 & else \end{cases} \\ Threshold = K; \quad if \quad Ft > P \end{cases} \quad (1)$$

Where $I(x, y)$ is the intensity of pixel with coordinates of (x, y); Ft is the counter of pixels which intensities levels lie in $[K, 255]$; P is the specified threshold of the number of the pixels with the high intensity level.

The image is segmented by the adaptive threshold K, and the binary image is shown as Figure 3 (b).

2.3 Cross Stripe Extraction

The cross line features of the stripe can be extracted from the binary image segmented by adaptive threshold. Some algorithms of lines detection such as Hough transformation (H-T) are robust and reliable for the task [8], but there is a problem that it takes much time for H-T to deal with lots of candidate pixels. A solution is that only edge points of strips are used for line detection, so that the number of points to be transformed and the cost of computation are decreased. However, the width of stripe usually is more than 10 pixels, so it is not precise enough to detect the cross lines by the edge points. Here the skeleton of stripe is extracted as proper points of central lines. The binary image is scanned to check the connectivity of each foreground points with neighborhood pixels. The edge points of the stripe are erased gradually. Then the stripe points are eroded one by one until the skeleton is left, as shown in Figure 4. This process can also be done via a thinning algorithm, by which

the both sides of the stripe edge were detected, thus the central lines can be obtained by averaging two edge points.
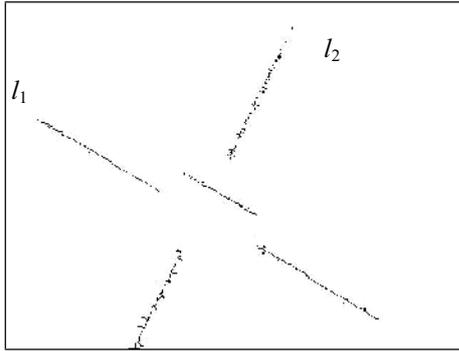
Fig. 4 Skeleton extraction of stripes

The skeleton of the stripe is an approximation of the cross stripe central lines, and the computing complexity of H-T is greatly decreased by using processed candidate points. The main lines l_1, l_2 of the cross tripe can be gained by H-T. The coordinates of the points on the skeleton are mapped into the parameters matrix $A(\theta, r)$ of the line function, as shown in (2), and the (θ, r) with the two maximums at main peaks in A is the parameters of the functions of lines l_1 and l_2 .

$$
\begin{cases}
A(\theta,r) = \sum_{k=1}^{M} \sum_{\theta=0}^{\pi} B(\theta,r) \\
B(\theta,r) = \begin{cases} 1 & r = x\cos\theta + y\sin\theta \\ 0 & others \end{cases}
\end{cases}
\tag{2}
$$

Where M is the number of points on the skeleton figure; x, y are the pixels coordinates on the skeleton of the stripe; and $\theta \in [0, \pi]$.
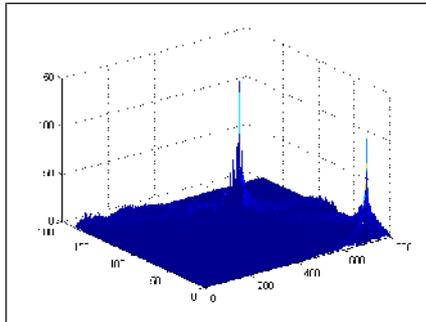
Fig. 5 Parameter matrix of H-T for detect cross lines

The cross lines are principal line components in image space, so there are two main peaks in the parameters matrix of the line functions, as shown in Figure 5. Since the slope of two lines is great different, the two main peaks of parameter can be distinguished and extracted respectively.

3 APERTURE DETECTION

In structured light image, there are four feature points formed by the translation of the stripes at the aperture brim. In this section, the feature points would be extracted according to the intensity distribution of the pixels arrayed along the central lines l_1 and l_2, which were detected in the above section.

3.1 Intensity Mapping

The intensity distribution often discontinues at the distorted parts of the stripes, so detecting the distribution breaking of intensity is proper for features extraction. The intensities of the pixels along the cross lines and its neighborhood are accumulated as (3).

$$
\begin{cases}
I_s(u,v) = \sum_{i=-\Delta}^{\Delta} \sum_{j=-\Delta}^{\Delta} I(u+i, v+j) \\
r = u\cos\theta + v\sin\theta \\
u \in [\Delta, Width - \Delta] \\
v \in [\Delta, Height - \Delta]
\end{cases}
\tag{3}
$$

Where θ and r are the parameters of the lines functions; Δ is the neighborhood of accumulation operation; *Width* and *Height* are the size of the image. Intensity-mapping curve may be disturbed by some noise on the stripes, so it is necessary to filter the curves to weaken the noise. The median filter and smoothing filter are adopted to reduce the pulse disturbance and measure noise respectively.
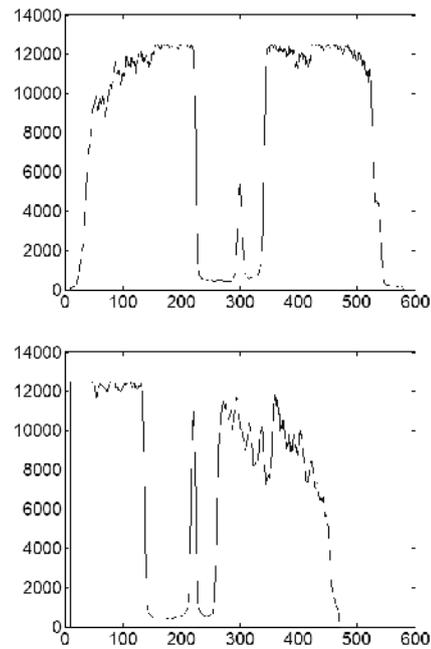
Fig. 6 Intensity mapping curve along line l_1 and l_2

The intensity mapping curves of two cross lines after filtering are shown in Figure 6. There is low intensity area in the aperture because of broken stripes, and there is intensity pulse disturbance caused by distorted stripes in the aperture area, which should be rejected.

3.2 Features Extraction

There are several intensity breaking points in the mapping curve, which are caused by aperture brim, the distorted stripe in the aperture area, and the imaging noise on the workpiece surface. Firstly, the difference of intensity mapping curves is calculated, and each peaks of difference value would be the candidate features, shown as Figure 7.

3775
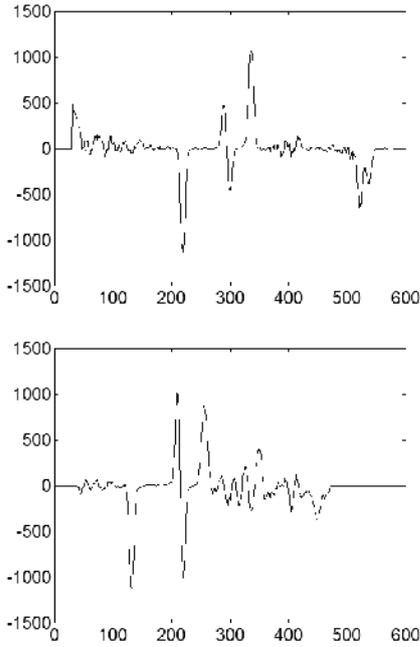
Fig. 7 Difference curves of intensity mapping of line l_1 and l_2

A threshold is set so that some lower peaks caused by noise are rejected. Because the values of peaks caused by aperture brim and those in the aperture are similar, it would incur confusion to judge the validity of features by the value of peaks. Since the pair of peaks caused by disturbance is of narrow range, and the range of true pair of peaks is related to the size of aperture, the rule of judging validity is presented as following manner. The pair of positive and minus peaks of false breaking points is rejected by comparing the distances between those candidate points. In this way, the proper pair of peaks with maximum distance is extracted, and the four feature points are located by the breaking points on the intensity curve respectively.

4   APERTURE ALIGNMENT CONTROL

An aperture on a plane has five DOFs in a camera frame, which includes three translational DOFs and two rotational DOFs. The partitioned visual control is presented based on the mapping relationship between image space and 3D Cartesian space. The translational DOFs are controlled via the coordinates of the feature points in image space, and the rotational DOFs are controlled via the parameters of the two stripes in image space.

4.1 Structured Light Imaging

The imaging property of a camera is described with the pinhole model. The intrinsic model of a camera can be written as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix} = M_{in} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix} \quad (4)$$

where
(u, v) are the coordinates of a point in an image,

(a_x, a_y) are the magnification coefficients from the normalized focus imaging plane coordinates to the image coordinates,

(u_0, v_0) denote the image coordinates of the camera's principal point,

M_{in} is the intrinsic parameters matrix, and

(x_c, y_c, z_c) are the coordinates of a point in the camera frame. The two equations of laser structured light planes Π_1 and Π_2 in the camera frame are

$$a_1 x_c + b_1 y_c + c_1 z_c + 1 = 0 \quad (5)$$

and

$$a_2 x_c + b_2 y_c + c_2 z_c + 1 = 0 . \quad (6)$$

Where (a_1, b_1, c_1) and (a_2, b_2, c_2) are the coefficients of the two light plane equation.

According to triangular measurement principle, the field depth z_c can be calculated as

$$z_c = -\left(\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \right)^{-1} , \quad (7)$$

which is the constraint condition for the visual measurement based on structured light.

The plane equation of the surface in the camera frame is

$$A x_c + B y_c + C z_c + 1 = 0 . \quad (8)$$

Where (A, B, C) are the coefficients of the object surface. With the pinhole model of the camera and the constraint condition of the structured light, the equations of the two stripes on the normalized focus imaging plane are

$$(A - a_1)x + (B - b_1)y + C - c_1 = 0 . \quad (9)$$

and

$$(A - a_2)x + (B - b_2)y + C - c_2 = 0 . \quad (10)$$

According to the configuration of the sensor, we have $a_1 = 0$ and $b_2 = 0$. In this case, the stripe Π_1 is parallel to x_c-axis in the camera frame, and the stripe l_1 is horizontal in image space in the desirable poison. Similarly, Π_2 is parallel to y_c-axis and upright in image space when the aperture plane is aligned.

4.2 Partitioned Visual Control Law

The goal of robotic plane alignment task is to position a camera aligned to a planar object at a distance. IBVS defines the task error function in image feature space; therefore the robot velocities are regulated using visual feedback. The block diagram of IBVS is shown as Figure .8.
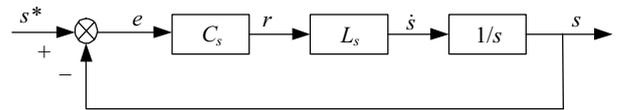
Fig. 8 Image based visual servo

where s is visual features; $s*$ is its value in desirable position; e is defined task function; C_s is the control law; r is the camera pose. The vision task function is defined by visual features $e = s - s^*$.

Imposing an exponential decrease of the task function

$$\dot{e} = -\lambda e , \quad (11)$$

3776

the following control law can be synthesized

$$v_c = -\lambda \hat{L}^+ e \qquad (12)$$

with λ is a positive gain, and $\hat{L}^+$ the estimate of the pseudo-inverse matrix of L_s, v_c is the camera velocity. With this control law, the closed-loop equation of the system in Figure 8 is

$$\dot{e} = -\lambda L_s \hat{L}_s^+ e = -\lambda M(e)e \,. \qquad (13)$$

The key issue of control law design for IBVS is the estimation of interaction matrix. The visual features set s is defined as a function of the pose parameters P of the object plane equation and the camera pose vector. Therefore the interaction matrix can be calculated as

$$L_s = \frac{\partial s}{\partial r} = \frac{\partial s}{\partial P} \frac{\partial P}{\partial r} \,. \qquad (14)$$

Select a set of visual features s as

$$s = (b_1, k_1, k_2, u_i, v_i) \,. \qquad (15)$$

where u_i, v_i are the coordinates of the aperture center in image space; k_1 is the slop of laser stripe l_1, and k_2 is the slop reciprocal of laser stripe l_2; b_1 is the intercept of l_1.

The image Jacobian matrix is estimated with the value at the desirable position, thus a constant control law is used for image based visual servoing.

$$L_s^* = diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \qquad (16)$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ are the constant gain.

5   EXPERIMENTAL RESULTS

Aperture detection and alignment experiments using a structured light vision were conducted with a plane workpiece with an aperture on its surface. The vision sensor and robot end-effector are shown in the Figure 9.
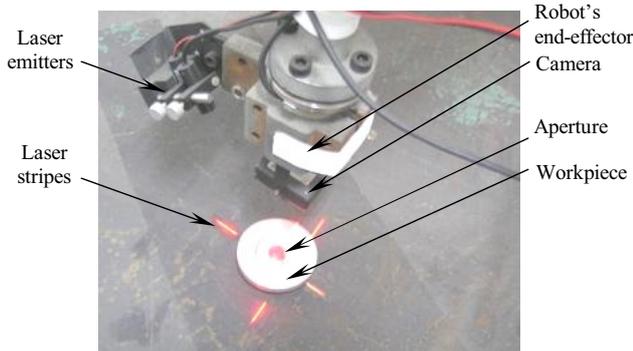
Fig. 9   Vision sensor for aperture detection and robot end-effector

The image processing results and feature points extraction for the aperture detection are shown in the Figure 10.

(a1)   (b1)

(a2)   (b2)
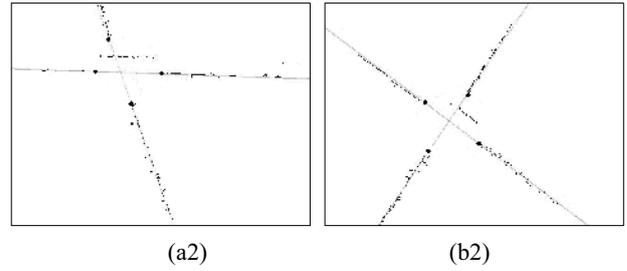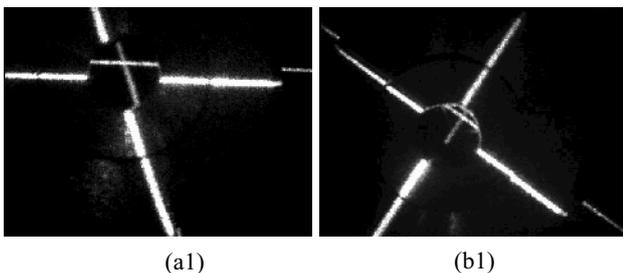
Fig. 10   Aperture detection results. (a1) and (b1) original images; (a2) and (b2) are features extraction of (a1) and (b1)

It can be seen that the feature points were extracted and the apertures were detected accurately and reliably with the disturbance of distorted stripe in the aperture area. It takes about 55 milliseconds to detect the aperture on a PC with 2.8GHz CPU clock rate, and image processing algorithm fulfills real time application.

The selected visual features show convergence by using the proposed visual control law, as shown in Figure .11 (a) and (b). The position and orientation of the object plane in camera frame is represented by the normal vector N and the distance D, shown in Figure 11(c) and (d). Both trends of robotic task in Cartesian space and image space show good performances in convergence and dynamic trajectory.

(a)   (b)

(c)   (d)
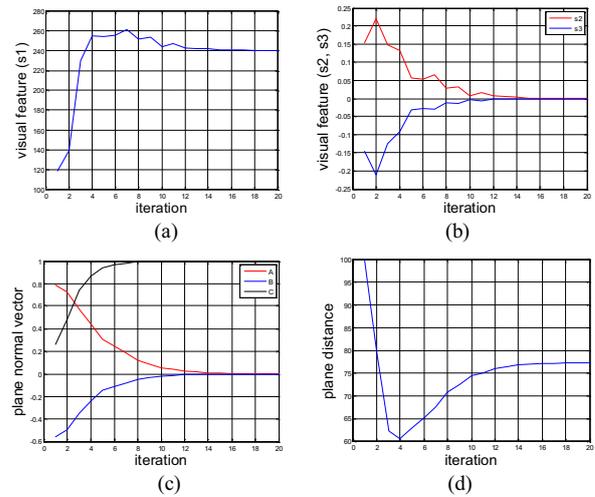
Fig. 11   Experimental results of workpiece alignment control

6   CONCLUSIONS

In this paper, a visual sensor with a cross laser emitters is designed for aperture detection and visual alignment. Image processing and features extraction algorithms are presented, and a partitioned control law is developed based on the visual features. Experimental results verified that the algorithms are effective and reliable, and it has good performance in real time. Because the cross lines can be detected by H-T robustly, the process of feature extraction is not sensitive to segmentation result, and the disturbances in other area of the image is excluded. Thanks to the selection of visual features, the control law was partitioned in visual aperture alignment.

The future work of this paper will be:

1. In the case of the apertures are located nearby the brim of workpiece, the cross stripes would be heavily distorted by the brim. The algorithm of extraction should be developed to eliminate the false breaking points on the stripes.

2. In features extraction, the stripe curvature based detection is another candidate method because the stripes may be continuous at the aperture brim.

REFERENCES

[1] G. Biegelbauer, M. Vincze, H. Nohmayer, Sensor based robotics for fully automated inspection of bores at low volume high variant parts. Proceedings of the 2004 IEEE International Conference on Robotics and Automation, 5:,. 4852-4857.

[2] R. J. Li, Y. C. Li, H. O. Sun, The application of image processing in the automatic inspection of the assembly holes of crossbeam, Journal of Changchun University of Technology (Natural Science Edition), 2006, 27(1): 43-45.

[3] D. Xu, Z. M. Jiang, L. K. Wang, *et al*, Features Extraction for Structured Light Image of Welding Seam with Arc and Splash Disturbance. The eighth International Conference on Control, Automation, Robotics, and Vision, Kunming, China, December, 2004,1559-1563.

[4] S. Amin-Nejad, J.S. Smith, J. Lucas. A visual servoing system for edge trimming of fabric embroideries by laser. Mechatronics, 2003, 13(6): 533-551.

[5] F. Chen, G. M. Brown, M. Song. Overview of three-dimensional shape measurement using optical methods. Optical Engineering. 2000. 39(1): 10-22.

[6] Y. Li, Q. L. Wang, Y. F. Li, *et al*, On-line visual measurement and inspection of welding bead using structured light, 2008 IEEE International Instrumentation & Measurement Technology Conference, British Columbia, Canada, 2038-2043.

[7] G. Biegelbauer, M. Vincze, Fast androbust bore automation detection in range image data for industrial automation. The 2nd International Symposium on 3D Data Processing, Visualization, and Transmission, Thessaloniki, Greece, 2004, 76-83.

[8] S. Huntchinson, G. Hager, P. I. Corke, A tutorial on visual servo control, IEEE Trans. on Robotics and Automation, 1996, 12(5): 651-670.

[9] E. Malis, F. Chaumette, S. Boudet, 2-1/2-D visual servoing, IEEE Trans. on Robotics and Automation, 1999, 15(2): 234-246.

[10] P. I. Corke, S. Huntchinson. A new partitioned approach to image-based visual servo control. IEEE Trans. on Robotics and Automation, 2001, 17(4): 507-515.

[11] J. C. Pagès, F. Chaumette, J. Salvi. Plane-to-plane positioning from image-based visual servoing and structured light, in 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004, 1004-1009.

[12] J. C. Pagès, C. Collewet, F. Chaumette, et al. Optimizing plane-to-plane positioning tasks by image-based visual servoing and structured sight, IEEE Trans. on Robotics, 2006, 22(5): 1000-1010.