DOI:10.13434/j.cnki.1007-4546.2016.0103

Research on Extraction of Bottom of Shoe Pattern Based on Binocular Stereo Vision

MA Xin-wu, GAN Yi, SUN Fu-jia

College of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, P. R. China

Abstract: In order to quickly and efficiently get the information of the bottom of the shoe pattern and spraying trajectory, the paper proposes a method based on binocular stereo vision. After acquiring target image, edge detection based on the canny algorithm, the paper begins stereo matching based on area and characteristics of algorithm. To eliminate false matching points, the paper uses the principle of polar geometry in computer vision. For the purpose of gaining the 3D point cloud of spraying curve, the paper adopts the principle of binocular stereo vision 3D measurement, and then carries on cubic spline curve fitting. By HALCON image processing software programming, it proves the feasibility and effectiveness of the method.

Keywords: binocular stereo vision; camera calibration; edge; stereo matching; fitting

1 Introduction

In footwear manufacturing, spraying on the sole is a very important process, because it shows the durability of shoes. However, this process is still largely undertaken manually or semi-automatically, which brings many disadvantages, such as low productivity and harmful working condition which is bad to the operator. This situation largely constraints healthy development of footwear industry, thus automated manufacturing becomes the trend of footwear industry, described in Ref[1]. The extraction of sole surface data and automatically generating of spray adhesive trajectory are the key to spraying process, Li[2] invented a machine for shoe sole automatic dispensing,

Received 03 February 2016

Kwon[3] proposed a method of obtaining spray adhesive trajectory based on sole contour, but the spray adhesive trajectory using two methods mentioned above is plane curve, which can't exactly show real information of shoe. Kim[4] proposed a method to extract trajectory based on three view drawing and 3D scanning data, but this method is complex and hard to control precision effectively.

The main content of computer stereo vision is to gain 3D coordinates of the object by great graphic images. In recent years, binocular stereo vision technology based on two images is a hot research, described in Ref[5]. Binocular stereo vision has the advantages of high efficiency, accuracy, simple system structure and low cost, which are very suitable for manufacturing site online and non-contact product testing and quality control. Camera calibration and stereo matching are two key and difficult problems in binocular stereo vision, described in Ref [6]. In order to realize the automation of spray adhesive process, the paper proposes a method of quickly get spray adhesive trajectory 3D data based on binocular stereo vision.

2 Sole pattern three-dimensional reconstruction

The steps of 3D reconstruction based on binocular stereo vision include: calibration plate pattern acquisition, camera calibration, shaping around the image collection, image correction, feature extraction, feature point matching, 3D coordinate calculation, three-dimensional display, etc., the whole process is shown in Figure 1.

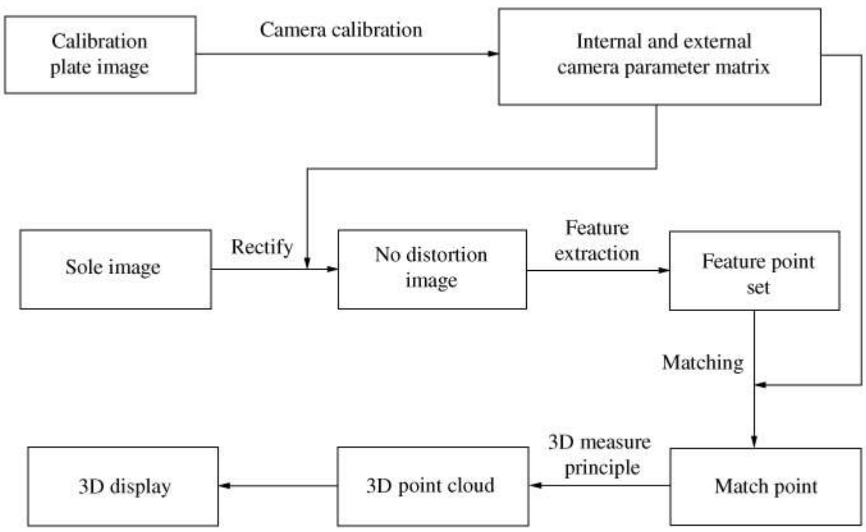

Figure 1 The flow chart of 3D reconstruction

2.1 Camera calibration

Two cameras are adopted to simulate human eye's function of percept distance of object in binocular stereo vision. In order to establish an effective imaging model, we must first calibrate internal parameters, the distortion parameters and external parameters for each camera. Ideally, Optical system imaging model of camera is pinhole model, that is, corresponding coordinates of points in 3D space is projected to image plane through perspective transformation. In image measurement process and machine vision applications, in order to determine the relationship between the three-dimensional geometry of some points on objects surface and the corresponding points in image, we must establish the geometric model of camera imaging. The geometry model parameters are camera parameters. Under most conditions camera parameters are got through experiment and calculation, this process is called camera calibration. Here mainly refers to the internal and external camera parameters. in calibration of binocular vision, internal parameters is left and right camera lens focal length f, lens distortion k, size s_x, s_y, etc. external parameters mainly refers to left and right camera coordinate systems relative to the world coordinate system. The relationship between left camera coordinate system and world coordinate system is showed with the rotation matrix $R_1$and translation vector T_1. In the same way, R_2 and T_2 show the relationship between right camera coordinate system and world coordinate system. RandTshow the relationship between right camera coordinate system and left camera coordinate system, described in Ref[7]. The following is to determine these parameters, the first three steps can get internal and external camera parameters and fourth step obtains position relationship between two cameras.

1) Transform point $p_w = (x_w, y_w, z_w)^T$ in world coordinate system to corresponding point $p_c = (x_c, y_c, z_c)^T$ in camera coordinate system. This transformation is rigid transformation. The transformation relationship between them is

$$p_c == R' * p_w + T' \tag{1}$$

$T' = (t_x, t_y, t_z)$ is a translation vector. $R' = (\alpha, \beta, \gamma)$ is a rotation matrix. (α, β, γ) show the angle between object and axis x, y, z of camera coordinate system, $(\alpha, \beta, \gamma, t_x, t_y, t_z)$ is known as the camera parameters.

2) Project 3D space points p_c from the camera coordinate system onto the imaging plane. It can be expressed mathematically as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{z_c}\begin{pmatrix} x_c \\ y_c \end{pmatrix} \tag{2}$$

where u, v is coordinate in image coordinate system united by mm. In the process of the actual projection, lens distortion causes $(u \quad v)^T$ changed to $(\tilde{u}, \tilde{v})^T$. For most lenses, it can be considered as radial distortion. At this point, the corresponding formula is

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \frac{2}{1 + \sqrt{1 - 4k(u^2 + v^2)}}\begin{pmatrix} u \\ v \end{pmatrix} \tag{3}$$

3) Transform point united by mm to corresponding point united by pixel in image coordinate system. It can be expressed mathematically as

$$\begin{pmatrix} r \\ c \end{pmatrix} = \begin{pmatrix} \dfrac{\tilde{v}}{s_y} + c_y \\ \dfrac{\tilde{u}}{s_x} + c_x \end{pmatrix} \tag{4}$$

Where r, c is coordinate in image coordinate system united by pixel, s_x, s_y is physical size of each pixel, c_x, c_y is light heart's coordinate in image coordinate system united by pixel.

4) Assume that non-homogeneous coordinates of any space point in the world coordinate system, the left and right camera coordinate system are x_w, x_1, x_2, They have the following relationship

$$\begin{cases} x_1 = R_1 x_w + T_1 \\ x_2 = R_2 x_w + T_2 \end{cases} \tag{5}$$

Away x_w to get

$$x_2 = R_2 R_1^{-1} x_1 + T_2 - R_2 R_1^{-1} T_1 \tag{6}$$

so we known the position relations between two cameras

$$\begin{cases} R = R_2 R_1^{-1} \\ T = T_2 - R_2 R_1^{-1} T_1 \end{cases} \tag{7}$$

2.2 Feature extraction

The accuracy of feature extraction directly affects the accuracy of subsequent processing, such as image feature matching and measurement results. There are all kinds of features in the image, such as feature point, edge line, area, etc. edges contain the main feature information of visual image. In this paper, it extracts edge lines to get original data of sole. Edge detection must meet two conditions: (1) effectively suppress noise; (2) accurately determine position of edge. In numerous edge detection operators, canny edge detection operator based on optimization algorithms was widely used, because of the advantage of high signal-to-noise ratio and precision, described in Ref[8]. The input of canny edge detection algorithm is original image, the output is edge curves. The steps are as follows:

1) Smooth image with Gaussian filter

$$\begin{cases} h(x,y) = \dfrac{1}{2\pi\sigma} e^{-\frac{a^2+b^2}{2\sigma^2}} \\ g(x,y) = f(x,y) \otimes h(x,y) \end{cases} \tag{8}$$

in which $f(x,y)$ is original input image, $g(x,y)$ is image with smoothing, $\otimes$ represents convolution.

2) Calculate the gradient magnitude and direction based on finite difference of a first order partial derivatives,

$$\nabla g(x,y) = \begin{bmatrix} g'_x \\ g'_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial g}{\partial x} \\ \dfrac{\partial g}{\partial y} \end{bmatrix} \tag{9}$$

$$\begin{cases} | g(x,y) | = \sqrt{(g'_x)^2 + (g'_y)^2} \\ \theta(x,y) = \arctan(\dfrac{g'_x}{g'_y}) \end{cases} \tag{10}$$

3) Maximum inhibition on gradient amplitude. If images are smoothed by Gaussian filter, edge blurs. Edge has a certain width by calculating gradient, this thinner method is called inhibition of the maximum point. This method compare adjacent pixel gradient amplitude on the perpendicular direction of edge (the direction of the gradient), and remove the ones which are smaller than field amplitude.

4) Detect and connect edge with double threshold algorithm. Firstly, we should distinguish edges. All edges strength is greater than high threshold must be edge point; All the intensity of edge points must not be less than low threshold; If the edge strength between low threshold and high threshold, we see if the point in the adjacent

pixels over high threshold, if you have, it is the point of edge, otherwise, it is not edge points. Secondly, connect edges. Double threshold algorithm has two thresholds T_1 and T_2 to the image of the maximum inhibitory, they meet the equation $T_2 = 2T_1$. So we can get two threshold edge images $G_1[i,j]$ and $G_2[i,j]$. Due to the use of high threshold value, $G_2[i,j]$ contains a few false edges, but intermittent. It requires to joint edges into outline. When reaching the end of a contour, the algorithm tries to find edge which can connect to the contour through $G_1[i,j]$ eight adjacent points. In this way, the algorithm is constantly collecting edges in $G_1[i,j]$ until it totally links up $G_2[i,j]$, described in Ref[9].

2.3 Stereo matching

In general, stereo matching algorithm can be divided into local and global algorithms. According to different matching primitives, the local algorithm can be divided into regional matching, feature matching and phase matching. Stereo matching algorithm based on region has poor adaptability of in place of texture sparse and depth discontinuities, large amount of calculation and poor matching precision, described in Ref[10]. Stereo matching algorithm based on feature matching has the advantage of high speed and precision, but the matching precision is greatly affected by feature point extracted. If feature points are error or inaccurate, it will directly affect the next matching precision, described in Ref[11].

Combined with the matching algorithm based on region and feature, the paper uses polar constraint stereo matching algorithm based on RANSAC, described in Ref[12]. This algorithm firstly apply initial matching based on region to image feature, it reduces the matching area from the whole image to surrounding area of corresponding characteristics, which greatly reduce search scope and improve calculate speed, compared with region matching algorithm. Then it uses RANSAC method to eliminate errors under the guide of corresponding polar, and estimates fundamental matrix at the same time. The matching algorithm is a combination of region matching and feature matching, thus making it higher precision and robustness.

In order to use RANSAC algorithm in polar constraints, we must first calculate the basis matrix.

Assumption is that $m_1(u_1, v_1, 1)$, $m_2(u_2, v_2, 1)$ is the projection point of space point P in two images, they are homogeneous coordinates. They meet the equation of $m_1^T F m_2 = 0$, F is basis matrix of two cameras.

By polar constraint, we can get homogeneous equation containing unknown component of basic constraint matrix

$$u^{T}f = 0 \tag{11}$$

Where $u=(u_1u_2, v_1u_2, u_2, u_1v_2, v_1v_2, v_2, u_1, v_1, 1)^{T}$ and $f=(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^{T}$. The rank of third order basic matrix is 2, which means that it has seven degrees of freedom, so we need at least 7 feature points for basic matrix computation.

The calculation steps of polar constraint stereo matching algorithm based on RANSAC algorithm are as follows:

1) Sample repeatedly, it includes seven points Random sampling, estimating fundamental matrix using seven algorithm, calculating how many points are consistent with the basic matrix through compared with distance threshold (each match point to the polar), and determining interior points.

2) Select the point set, which contain the most interior points and whose standard variance is smaller than the residual.

3) Calculate matrixF by using point set accurately.

2.4 Three-dimensional measurementThe principle of three-dimensional binocular stereo vision measurement is parallax, it is showed in Figure 2.

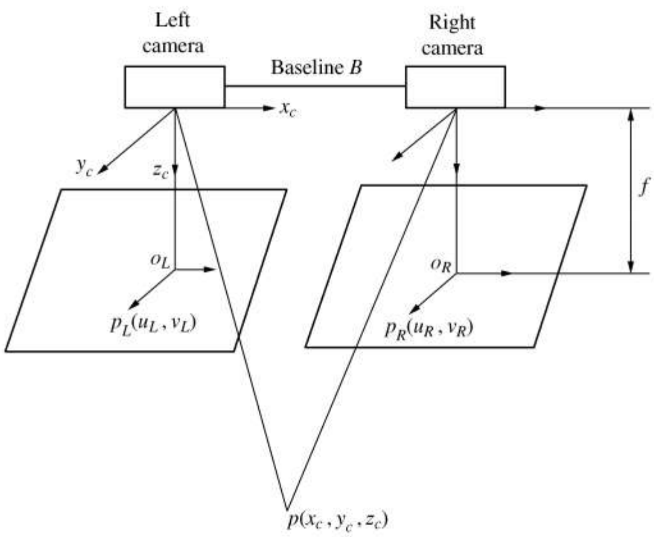

Figure 2 Parallel binocular stereo vision system

In Figure 2, we use baseline B represents the projection center distance of two cameras. In the virtual imaging plane of two cameras, the coordinates of perspective projection of space point P are $p_L(u_{L,} v_L)$ and $p_R(u_{R,} v_R)$. The two cameras are placed in parallels, which means that both imaging plane are in the same plane, so p_L and p_R have the same abscissas. That is v_L is equated to v_R. According to similarity triangular relationship, we can get the following conclusions.

$$\begin{cases} u_L = f\dfrac{x_c}{z_c} \\ u_R = f\dfrac{x_c - b}{z_c} \\ v_L = v_R = f\dfrac{y_c}{z_c} \end{cases} \tag{12}$$

Because the two adjusted image are in the same plane, so we can define the difference between column coordinates in right and left image as the parallax value of corresponding points:

$$disparity = u_L - u_R \tag{13}$$

By (12), (13), we know the 3D coordinates of point P in the camera coordinate system as follows:

$$\begin{cases} x_c = \dfrac{b \times u_L}{disparity} \\ y_c = \dfrac{b \times v_L}{disparity} \\ z_c = \dfrac{b \times f}{disparity} \end{cases} \tag{14}$$

3 Experiments and analysis

3.1 Camera calibration

This experiment use HALCON software to get spray adhesive curve information. Calibration process based on HALCON is shown in Figure 3. In Figure 3, program segment read_image uses for reading the Image, program segment find_caltab uses for determining the effective area of the calibration board, the effect as shown in Figure 4. Procedures section find_marks_and_pose uses for determining two-dimensional coordinates of circular landmark in calibration board[13], the effect as shown in Figure 5. we can determine internal and external parameters of two cameras and relative position of two cameras through procedures section binocular_calibration. We selected 10 images for calibration to get the internal and external parameters of two cameras.

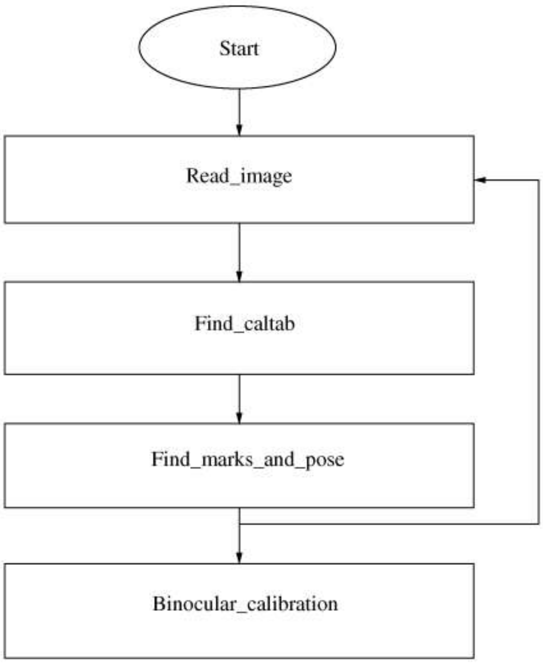

Figure 3 Program flow chart

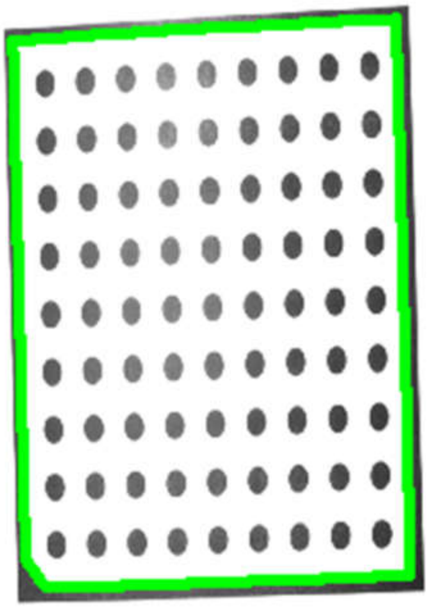

Figure 4 The effective area

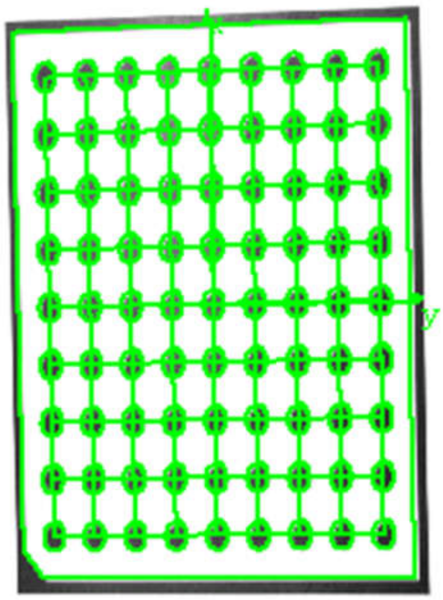

Figure 5 2D coordinates in calibration board

3.2 Correction to stereo images

In order to more accurately match and improve the efficiency of operation, we should first to rectify stereo image and obtain the adjusted parameters after getting camera parameters. We through procedure sections gen_binocular_rectification_map and map_image to implement this feature, within the parameters of camera calibration results are shown in Table 1, outside parameters results are shown in Table 2. You can see from the table of camera calibration the effect is good.

Table 1 Camera calibration parameters

Initial parameters	[0.005, 0, 3.45×10^{-6}, 3.45×10^{-6}, 1224, 1025, 2 448, 2 050]
Left camera internal parameters	[0.005 022 83, −433.51, $3.416\ 39\times10^{-6}$, 3.45×10^{-6}, 1 267.91, 1 132.19, 2 448, 2 050]
Right camera internal parameters	[0.005 031 43, −470.287, $3.471\ 45\times10^{-6}$, 3.45×10^{-6}, 1 504.3, 908.367, 2 448, 2 050]
Rectified left camera internal parameters	[0.005 032 25, 0.0, 3.45×10^{-6}, 3.45×10^{-6}, 1 404.7, 1 666.82, 3 214, 3 153]
Rectified right camera internal parameters	[0.005 032 25, 0.0, 3.45×10^{-6}, 3.45×10^{-6}, 1 896.91, 1 666.82, 3 226, 3 153]

Table 2 External camera calibration parameter results

Cameras relative position	[−0.079 823 3, 0.086 862 1, 0.009 752 67, 5.463 71, 0.892 027, 179.826, 0]
Rectified cameras relative position	[0.118 372, 0.0, 0.0, 0.0, 0.0, 0.0, 0]

3.3 Feature extraction

In order to better deal with image, we use program rgb1_to_gray to convert color images to gray image. With using program edges_sub_pix based on canny algorithm for image edge extraction, the extracted precision reaches to subpixel level. To choose target from extracted edge, we can adopt the characteristics of the curve, program select_contours_xld can realizes the function. We can use program get_contour_xld to gain curve coordinates. There are 3 620 extracted points in left image and 3 740 points in right image. Extracted experiment edge is shown in Figure 6.

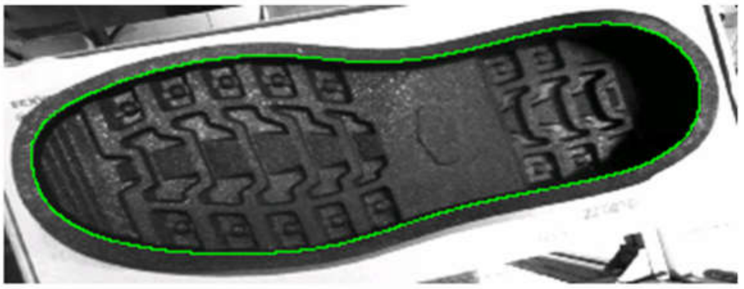

Figure 6 Extracted edge

3.4 Stereo matching

After edge extraction, we can get around coordinates of the curve, getting basic matrix by coordinates of points with RANSAC operator, we achieve rapid accurate match. The basic matrix is as follows:

$$F = [-0.000\,340\,17, -0.025\,744\,3, 0.180\,545, 0.034\,036\,1, 0.043\,948, 0.982\,241, -0.190\,019, -0.980\,198, 0.044\,263\,8]$$

By program match_essential_matrix_ransac in HALCON, We get the matching points up to 2336. The match images as shown in Figure 7.

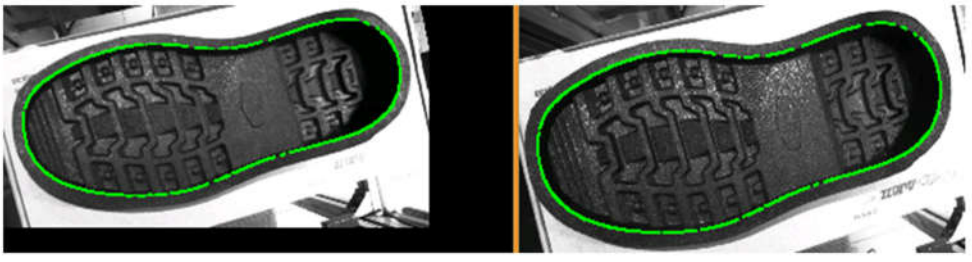

Figure 7 Matching points

3.5 3D display

After stereo matching, 3D measuring principle are used to get the matched point clouds, the image is shown in Figure 8.

Figure 8 Point clouds

From Figure 8, we draw conclusion that this experiment can get evenly sole contour point cloud information. There are local noises in 3D point cloud data, so denoising process is need[14], then data fitting with cubic spline curve[15-17]. The fitting error is shown in Figure 9. The fitting result is shown in Figure 10.

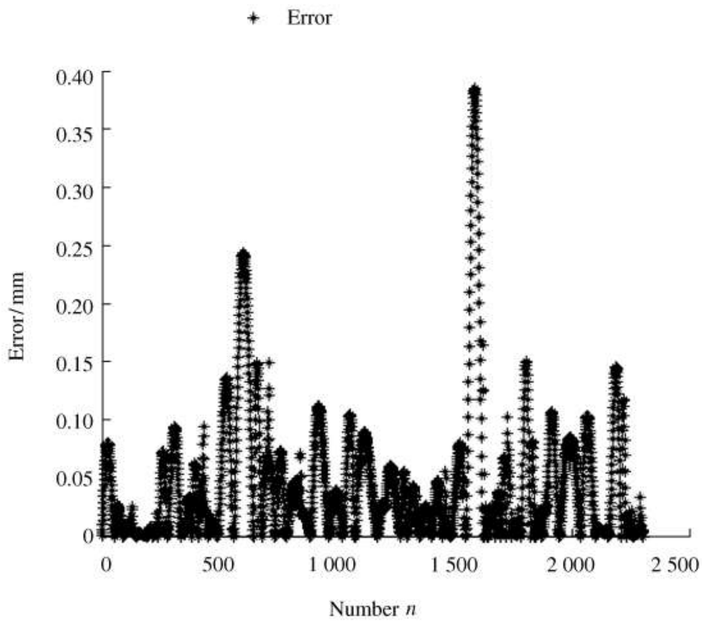

Figure 9 Fitting error

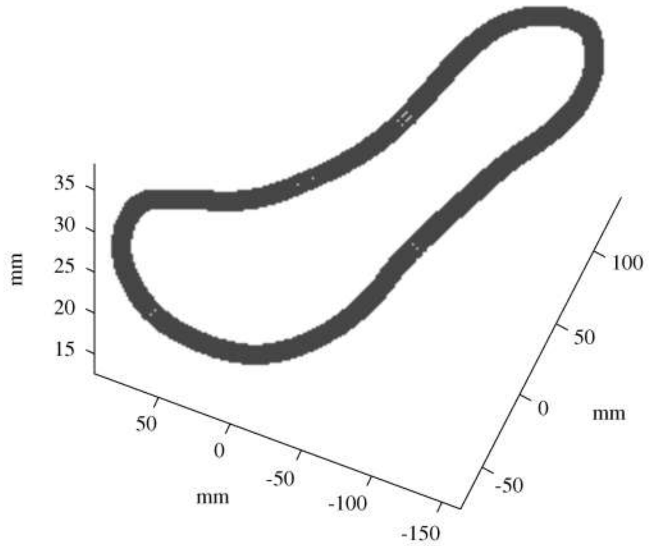

Figure 10 Fitting path chart

It is shown in Figure 9 that the maximum fitting error requiring not more than 0. 5 mm is less than 0. 4 mm. If there is a higher accuracy, it can be further local optimization. It is shown in Figure 10 that the minimum contour height is 16. 8 mm, and maximum height is 34. 1 mm, height difference is 17. 3 mm. The actual minimum height of shoe model is of 17 mm, and maximum height is 34. 5 mm, height difference is 17. 5 mm. Conclusion is drawn that this method can get good results and satisfy the practical production requirements by comparing experimental results with real data.

4 Conclusion

In order to effectively gain spray adhesive track information, this paper adopted the method of binocular stereo vision. First of all, two cameras were calibrated and the canny operator was adopted for image edge extraction, and then, matching points were got based on RANSAC algorithm stereo matching algorithm with polar constraint. Then, point cloud of the outline was gained with principle of three dimensional measuring. In the end, spray adhesive curve was obtained using the method of cubic spline curve fitting.

This method can effectively obtain 3D information of spray adhesive trajectory, and spray adhesive curve is space curve. Compared with plane curve, information is more rich and accurate. It doesn't need to modeling in the actual production. The method is simple in operation and suitable for different type soles. Only two CCD cameras and a computer are needed, so it is low cost and more economical.

References

[1] Wu C Y, He L Y, Li Q C, et al. Method for extraction of information of shoe upper and for automatic generation of spraying trajectory[J]. Chinese Journal of Mechanical Engineering, 2008,44(8):86-87 (in Chinese)

[2] Zuo L, Jiang H D, Zhang Y M. Automatic dispensing machine for shoe sole: China,15505997A[P]. 2004-06-23 (in Chinese)

[3] Kwon D S, Song S K. A method for generating a cementing trajectory of a shoe sole:UK, 226025[P]. 2005-10-26

[4] Kim J Y. CAD-based Automated Robot Programming in Adhesive Spray Systems, 2004, 21(11):625-634

[5] Xue G S. Acquisition of point clouds of cattle in complex situation based on binocular stereo vision[J]. Computer Engineering and Design, 2015,36(5):1390-1391 (in Chinese)

[6] Chen S Y, Liu S. Realization of Computer Vision Based on OpenCV[M]. Beijing:Science Press, 2008 (in Chinese)

[7] Li Z. Binocular vision system calibration based on HALCON[J]. Development & Innovation of Machinery & Electrical Products, 2013,26(3):95-96 (in Chinese)

[8] Wang X H, Qian Y J. An improved canny edge detection algorithm[J]. Mechanical & Electrical Engineering Magazine, 2008,25(12):60-61 (in Chinese)

[9] Zhao P. Research and development of machine vision[M]. Beijing:Science Press, 2012:37-42 (in Chinese)

[10] Yu N G, Qing Y G, Yuan X G. Review of stereo correspondence algorithms[J]. Computer Measurement & Control, 2009,17(5):817 (in Chinese)

[11] Zhu X F. Three-dimensional measurement of the workpiecesize based on binocular vision[D]. Zhejiang:

Zhejiang Ocean University, 2014 (in Chinese)

[12] Zhang Y P, Hua X J, Xia L C, et al. Stereo matching with epipolar line constraints based on RANSAC algorithm[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2013, 11: 20-22 (in Chinese)

[13] Luo Z X. Camera calibration based on HALCON[J]. Video Application & Project, 2010,34(4):100-102 (in Chinese)

[14] Wu S X, Wang C Y. Scattered noise segmentation of point cloud data[J]. Journal of Mechanical Engineering, 2007,43(2):230-233 (in Chinese)

[15] Zhou H M, Wang Y M. Non-uniform rational b-splines curve fitting based on the least control points[J]. Journal of Xi'an Jiaotong University, 2008,42(1):73-77 (in Chinese)

[16] Han J, Jiang B C. A B-spline curve fitting algorithm based on contour key points[J]. Applied Mathematics and Mechanics, 2015,36(4):423-430 (in Chinese)

[17] Gan Y, Qi C Q. Study on a class of modifiable C2-continuous quartic parametric curves and surfaces[J]. Journal of Tongji University, 2001,29(2):154-159 (in Chinese)

Brief Biographies

MA Xin-wu is now a master in the College of Mechanical Engineering, University of Shanghai for Science and Technology, research direction is machine vision. mxw1406860994@sina.com

GAN Yi is now a doctor in the College of Mechanical Engineering, University of Shanghai for Science and Technology, associate professor, research direction is networked manufacturing.

SUN Fu-jia is now a doctor in the College of Mechanical Engineering, University of Shanghai for Science and Technology, lecturer, research direction is precision manufacturing.