

# A method for identifying coded feature points on cross target based on multi-information constraint

Ju Huo<sup>1,a</sup>, Junfeng Li<sup>2,b</sup>, Jiashan Cui<sup>2,c</sup> and Wei Li<sup>2,d\*</sup>

<sup>1</sup>Department of Electrical Engineering, Harbin Institute of Technology, Harbin, China

<sup>2</sup>Control & Simulation Center, School of Astronautics, Harbin Institute of Technology, Harbin, China

<sup>a</sup>torch@hit.edu.cn, <sup>b</sup>zdh\_ljunfeng@126.com, <sup>c</sup>cuijiashan@gmail.com, <sup>d</sup>frank@hit.edu.cn

**Keywords:** Stereo vision; Coding recognition; Vision measurement

**Abstract.** In a stereo vision measurement system, Camera calibration is a key step. In order to improve the accuracy of camera calibration, a cross-target calibration scheme with coding points is designed in this paper. Firstly, coding target points of the cross target are identified and positioned. Secondly, a decimal decoding scheme based on median filter is designed to improve the accuracy of the decoding results and realize the matching of the stereo image.

## 1. Introduction

Computer vision technology is commonly used in engineering. As one of the branches, the vision measurement mainly aims at realizing the localization and measurement of the space object [1-3]. The image is the carrier of information in visual measurement technology, which extracts useful information from the image for three-dimensional reconstruction by Image Processing Technology. The advantages of this technology include: non-contact, high precision, high efficiency and low cost. It is widely used in aerospace, facial feature recognition, automated production and testing, robot vision navigation and other fields. The main task of stereo vision is to obtain the geometric information of the objects in the 3D world through captured images by multiple fixed cameras.

Camera calibration is a key technology in visual measurement [4-6]. Based on the image coordinates of the target point, the corresponding three-dimensional coordinates, determined by the camera model, can be calculated. Karara and Abdel-Aziz calibrated the camera by using the direct linear transformation [7], which the linear equation was only thing to solve, speeding up the algorithm of operation greatly. But without considering camera lens nonlinear distortion, inaccurate calibration was resulted. Calibration method proposed by Tsai is based on the radial alignment constraints [8], which did not consider tangential distortion. If the lens tangential distortion is severe, the accuracy of the method will be badly influenced. ZHANG [9] came up with a calibration method between self-calibration and traditional calibration. However, this method requires that all internal parameters of the camera remain unchanged during the calibration process.

In view of the above problems, a cross-target with coding target points is designed in this paper. In order to calibrate the camera, it is necessary to shoot the cross-target under different positions and attitudes. In this paper, a detailed study on coding points identification and decoding of the cross-target is performed. In order to improve the robustness of coding points recognition, a multi-constraint coding algorithm is proposed, and a decimal decoding scheme based on median filter is designed to improve the accuracy of the decoding result and realize the matching of the stereo image.

## 2. The cross-target

The cross target designed in this paper is illustrated in Fig. 1. It consists of two mutually perpendicular target arms, and each arm length of 1.5 meters. There are 33 target feature points on the cross-target, and the adjacent target points are equidistant. The target points are divided into two categories: one is the encoding target points, a total of 17; the other is non-coding target points, a

total of 16. In addition, the coding target points and the non-coding target points are interleaved. The cross-target test report clearly shows the coordinates of each target point.



Fig. 1: The cross-target

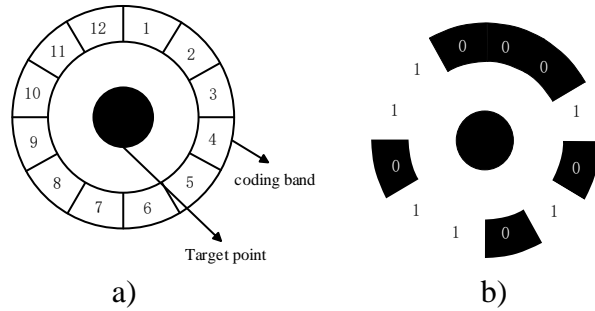


Fig. 2: Encoding target point

Each coding target point can be sorted to two parts: the first part is the center circle, which is the actual target point and used for calibration. The other part is a circular-shaped coding band, used to identify the target point. The coding band is formed by two circles with different radii and concentric with the target point. The encircled area is divided in 12 parts by angle, and the coding target point is called a 12-point code point. Each part corresponds to an angle of  $30^\circ$ , marked by black or white color. The black color corresponds to binary "0", and white corresponds to "1". Any parts of the encoding band can be select as starting point, and reading bits of each part clockwise will result in a 12-bit binary sequence. Since each part can serve as a starting point, one coding band corresponds to twelve different 12-bit binary sequences. Each 12-bit binary sequence is converted to a decimal number, and smallest of them is selected as the value of the encoding point.

Fig. 2 is a schematic diagram of a 12-point coding target. The reason for choosing a circular for a target point is that: circular feature point projection to the image plane is a ellipse, and there are abundant theoretic supports for the ellipse center coordinates extraction. For example, more accurate center coordinates of the ellipse can be extracted after appropriate image processing, which fundamentally improve the accuracy of the system, founded on the least square method of ellipse fitting. The schematic image shows the division of the coding band.

When camera is calibrated, the more points used, the better results calculated. The encoding method of the coding target points has been described in the above paragraph. Non-coding target points can be coded according to the encoding result of coding target points.

The coordinates of the center and the code value of each target point are encapsulated in a structure, which contains following sections: the flag indicating whether the point is a target point, the center of the ellipse, and the encoding value of the point. The value of point is the corresponding value for encoding target points while it is 0 for the non-coding target points.

All non-coding target points are located between the two coding target points. Assuming that the coordinates of a non-coding target point is  $(x, y)$ , and the coordinates of the adjacent two coding target points are  $(x_1, y_1)$  and  $(x_2, y_2)$ . During camera calibration, the cross-target has to be rotated repeatedly, so that the camera system can capture images of the cross-target from different angles, and this causes four positional relationships between the non-coding target point and its adjacent coding target points, namely:

$$\begin{cases} x_1 < x < x_2 \\ y_1 < y < y_2 \end{cases} \quad (1)$$

$$\begin{cases} x_2 < x < x_1 \\ y_1 < y < y_2 \end{cases} \quad (2)$$

$$\begin{cases} x_2 < x < x_1 \\ y_2 < y < y_1 \end{cases} \quad (3)$$

$$\begin{cases} x_1 < x < x_2 \\ y_2 < y < y_1 \end{cases} . \quad (4)$$

If the positional relationship between the non-coding target point and the two adjacent coding target points coincides with one of them, then the point will be coded accordingly.

### 3. Target points recognition and location of the cross-target

After image preprocessing, the Canny operator is used to extract the edges of the image, the pixel of different region boundary is detected according to the gray difference. In order to obtain satisfying ellipse contour, criteria for filter should be as follows:

(1) *Geometric parameter constraints*: The shape of the circular target point projected onto the image plane is elliptical, assuming that the maximum projection angle of the camera set during the experiment is  $\theta_{\max}$  (after simulation we can see that the value can be taken  $70^\circ \sim 75^\circ$ ), so the perimeter  $C$  and area  $S$  of the target ellipse contour in the image should satisfy the following formula:

$$C_{\min} \leq C \leq C_{\max} . \quad (5)$$

$$S_{\min} \leq S \leq S_{\max} . \quad (6)$$

Which,  $C_{\min}, C_{\max}$  represent the maximum and minimum values of the ellipse circumference in the range from 0 to  $\theta_{\max}$ , respectively;  $S_{\min}, S_{\max}$  represent the maximum and minimum values of the ellipse area in the range from 0 to  $\theta_{\max}$ , respectively;

(2) *Target ellipse position constraint*: If the extracted ellipse center point is located near the edge of the image, and it is a coding target point, then the coding band of the point may not be fully projected onto the image, therefore, even if the point is a target code point, normal encoding cannot be performed. For the above reasons, it is necessary to remove the point near the edge of an image according to a certain threshold value. So the center of the ellipse  $(u_0, v_0)$  needs to satisfy the following condition:

$$\min\{u_0, v_0, W - u_0, H - v_0\} \geq \delta . \quad (7)$$

Which,  $\delta$  represents the minimum value of the ellipse center distance from image boundary;

$W$  is the width of the image(Unit: pixels);

$H$  is the height of the image(Unit: pixels);

(3) *Distance constraint*: Using the above constraints, a large number of noise points and image edge points can be eliminated. However, it is not possible to distinguish between the coding target points and the non-coding target points. For the coding target points, there is a possibility that a small portion of the circular arc band on the coding band is misled to target ellipse. As shown in Fig. 3, the recognized target ellipse contains not only a coding target point but also a part of the coding band, the distance between the misrecognized coding ellipse and the target point is much smaller than the distance between the adjacent two target points. So the distance parameter can be used to filter the code points. Assuming that the distance between adjacent contours is  $d$ , satisfying the following formula, then one of the two identified target points are not an encoding target point,

$$d \leq d_{\max} . \quad (8)$$

Which,  $d_{\max}$  represents the maximum value between the misrecognized coded ellipse and the target ellipse center distance.

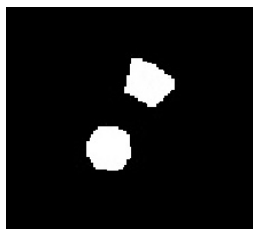


Fig. 3: Contour of mistaken identification

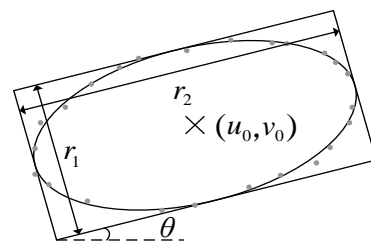


Fig. 4: Schematic of ellipse fitting

(4) *Encoding constraint*: Decoded result of the coding point is known and stored in the form of an array, so it need to be compared with the actual value of code points. If the result is not the same with all the actual value's, the point should be discarded.

(5) *Roundness Constraint*: The outline of the target point projected onto the image plane should be elliptical after contour extraction and polygon fitting, so that the irregularities are removed by roundness constraints.

The roundness of the ellipse is defined as  $R = 4\pi S/C^2$ . Under the specified projection angle,  $R$  should be satisfied:

$$R_{\min} \leq R \leq 1. \quad (9)$$

Which,  $R_{\min}$  represents the minimum value of ellipse roundness, generally it takes 0.75 ~ 0.8, in this paper, it is 0.8.

When the ellipse region of the target point is extracted, the least squares fitting method is used to figure out the edge pixels of the target point, and the part of edge pixels with the least fitting error is set as the central ellipse, so five parameters  $(u_o, v_o, r_1, r_2, \theta)$  of the ellipse can be got. Accordingly,  $(u_o, v_o)$  is the ellipse center coordinates,  $r_1, r_2$  as the major and minor axes of the ellipse respectively, and  $\theta$  as the rotation angle of the ellipse. The result of the elliptic fitting is represented by a rectangle which completely encloses the ellipse, as showed in Fig. 4.

#### 4. Decoding of the cross-target coding band

The division of a coding target point is shown in Fig. 5. The radius of circle  $B$  is the half of the total of circle  $A$  and  $C$ , and it is referred to the coding band median. The pixel acquisition at the time of coding is carried out on a circle, whose radius is substantially the same as circle  $B$ .

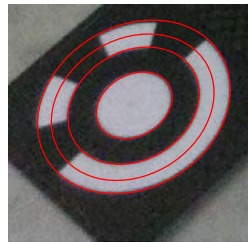


Fig. 5: Coding target point division

The geometric parameters of the coding target point are given in Fig.5( $\phi_{\min}, \phi_{\text{mid}}$  and  $\phi_{\max}$ ). We can get the ratio of the radius of the circle in the coding band to the radius of the target circle, namely:

$$n = \frac{\phi_{\max} + \phi_{\text{mid}}}{2\phi_{\min}}. \quad (10)$$

In practice, the value of  $n$  can be set according to the size of the image and the distance between the object and the camera. In order to increase the decoding speed and maintain the robustness of the algorithm, operations should be executed on the threshold-segmented image rather than the source image. As the image preprocessing for each step of the image have been stored before, there is no need to carry out the extra image processing but directly manipulate on the image which after the threshold segmentation.

The center position information that including some non-target points and all target points can be obtained by coordinate extraction process of circular feature points and positioning process of the target points. And the elliptic parameters after ellipse fitting to each point have also been obtained, so the center of the sampled ellipse can be set to the center of the target ellipse, and the length of the long axis and the minor axis of the sample ellipse are respectively  $n$  times of the fitted ellipse. In addition, the rotation angle of the sampling ellipse is the same with the rotation angle of the target fitting ellipse.

Then set 60 sampling points around the sampling ellipse to read the pixels corresponding to each sample point.

For each sampling point on the sampled ellipse, to compute the corresponding coordinates, and coordinates are calculated as follows:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = n \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} r_1 \cos \theta \\ r_2 \sin \theta \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}. \quad (11)$$

Which,  $[x_1, y_1]^T$  is the coordinates of sampling point;  $\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$  is a rotation transformation matrix with rotation angle  $\alpha$ ;  $r_1$  and  $r_2$  are the long and short axes of the target fitted ellipse, respectively;  $\theta$  is the angle ranged from sampling point to the starting point; The whole  $360^\circ$  is divided into 60 parts, and each part corresponding to a sampling point;  $[x_0, y_0]^T$  is the center coordinates of the fitted ellipse.

### 5. Experimental results and analysis

In this section, in order to verify the extraction and decoding algorithm of the target point center coordinates, a cross-target image with 17 encoding target points and 16 non-coding target points has been tested.

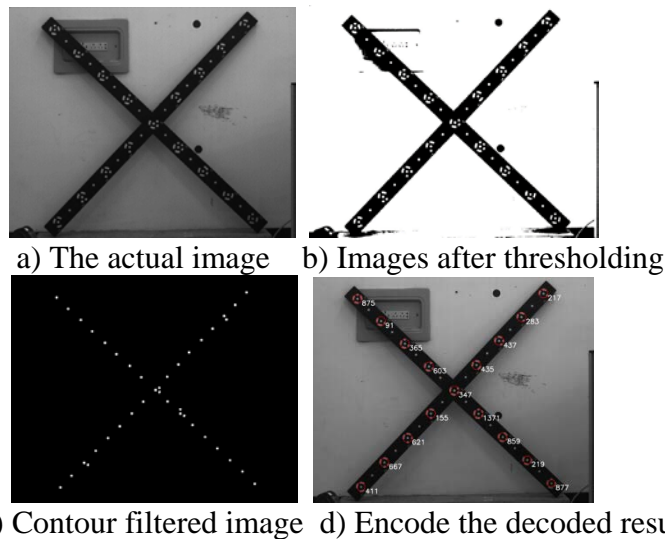


Fig. 6: The target recognition and positioning

Fig. 6 c) is the filtered image under geometric parameters constraint, roundness constraint, the target ellipse position constraint, distance constraint and other constraints. As shown from the image, the extracted contours include rounded target points and small coding blocks (pseudo contours) on the coding band.

For the points shown in Fig. 6 c), the pseudo-contours are eliminated by using the constraint conditions. The coded target points are identified and decoded, and the decoding results are shown in Fig. 6 d). According to the decoding results and center coordinates of the code points, the non-coding target points are filtered and decoded, and the decoding results are shown in Fig. 7.

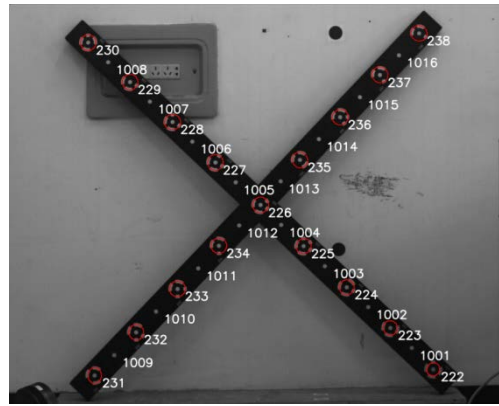


Fig. 7: Target decoding results

## 6. Conclusion

In a certain range of projection angle, the coordinates of the extracted target points are less affected by the projection angle, performing better robustness. By using the five constraints in section 3, it is effective to exclude the non-coding target region. The system can automatically extract the encoding target points and non-coding target points. In addition, it is accurate to obtain the identity information and center coordinates of all target points.

## 7. Acknowledgements

This work was supported by the National Natural Science Foundation of China [grant number 61473100].

## 8. References

- [1] Li Jing, Yuan Feng, Hu Ying-hui. Attitude measurement of space objects based on multi-linear CCD and multi-point cooperation target[J]. Optics and Precision Engineering, 2013, 21(6): 1635-1641.
- [2] Zhang H P, Jiang Z G, Elgammal A. Vision-Based Pose Estimation for Cooperative Space Objects[J]. Acta Astronautica, 2013, 91(10-11): 115-122.
- [3] Huo Ju, Yang Ning, Yang Ming. Attitude measurement of spatial moving object based on vectors of light beams[J]. Acta Photonica Sinica, 2015, 44(7): 0712001.
- [4] Chen X, Hu Y, Ma Z, et al. The location and identification of concentric circles in automatic camera calibration[J]. Optics & Laser Technology, 2013, 54: 185-190.
- [5] Heikkila J, Silvén O. A four-step camera calibration procedure with implicit image correction[C]. Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. IEEE, 1997: 1106-1112.
- [6] Cui J S, Huo J, Yang M. Novel method of calibration with restrictive constraints for stereo-vision system [J]. Journal of Modern Optics, 2016, 63(9): 835-846.
- [7] Abdel-Aziz Y I, Karara H M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry[C]. Proceedings of the Symposium on Close-Range Photogrammetry, 1971: 1-18.
- [8] Tsai R Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses[J]. IEEE Journal of Robotics and Automation, 1987, 3(4): 323-344.
- [9] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.