Study on Binocular Stereo Camera Calibration Method

Guo Yan Xu, Li Peng Chen, and Feng Gao

Transportation Science and Engineering Department, Beijing University of Aeronautics and Astronautics
Beijing, China

***Abstract*—Camera calibration is a key technology in the application of the stereo visual system. On the basis of analyzing the OpenCV camera model and the existing calibration methods of a binocular vision system, a binocular camera calibration algorithm was designed to obtain the internal and external camera parameters. Considering the importance of corner extraction in the camera calibration, a new corner extraction algorithm of gradient threshold is proposed, which can effectively extract the corners. The experiments show that this method is easy to be carried out and the calibration results are stable and accurate, which can satisfy the requirements of binocular stereo visual system application.**

Keywords—binocular vision; camera calibration; camera model; corner extraction; gradient threshold

I. INTRODUCTION

Camera calibration is an essential and fundamental part of computer vision systems [1] that have been widely used in many areas, such as mobile robot navigation and path identification, measurement of mechanical parts processing and complicated diagnosis and surgery, the vehicle identification positioning, and measurement of machine parts. A stereo vision system is designed to obtain position, shape, speed and other information of three-dimensional objects in space by the images taken by camera. In order to obtain the correspondence accurately between the surface point and a corresponding pixel in the image, camera calibration is of great importance.

Currently, camera calibration methods can be divided into two methods: conventional camera calibration and self-calibration method. Establishing the coordination between a known reference point of three-dimensional coordinates and the corresponding point on the image, conventional calibration method calculates the internal and external camera parameters accurately, while calibration process is quite complicated. Self-calibration method gets the internal and external camera parameters through the constraint relation of the image, which is of high degree automation, but is not accurate as conventional methods [2]. Between the two methods, Zhang calibration method takes template plane of different positions to calibrate, which is simpler than conventional method and more precise than self-calibration method. OpenCV calibration algorithm is based on the Zhang method, but uses Brown's method in calculating distortion parameters. It is widely used because it provides a flexible calibration method and can also get precise results. Although the theory on camera calibration has been very mature, and calibration methods are many, so far there hasn't been a method able to adapt to various occasions [3]. The study of binocular vision camera calibration is significant.

In this paper, a binocular vision system is calibrated using planar checkerboard based on the OpenCV calibration algorithm. A new approach of gradient threshold to extract corners is proposed by improving the OpenCV corner extraction method, which is able to adapt to light conditions to extract corners accurately, moreover it is verified on the binocular vision system produced by SRI international company of America.

II. CAMERA MODEL

Camera calibration needs an appropriate camera model, because it directly affects the final calibration results. Based on pinhole camera model, OpenCV camera calibration algorithm takes into account of the lens radial distortion and tangential distortion [4]. Compared to Tsai camera model that only includes first-order radial distortion, it demonstrates the lens distortion more truly and accurately [5]. Camera calibration will get the correspondence between a set of image pixel coordinates and world coordinates. The transformation from image pixel coordinates to world coordinates requires the following four steps as shown in Fig. 1.

$$\boxed{X_wY_wZ_w} \rightarrow \boxed{X_cY_cZ_c} \rightarrow \boxed{x_dy_d} \rightarrow \boxed{xy} \rightarrow \boxed{uv}$$

Figure 1. Coordinate transformation steps

A. From World Coordinate to Camera Coordinate

Considering the object in image as rigid body, the transformation from world coordinate $P_w(X_w, Y_w, Z_w)$ to camera coordinate $P_c(X_c, Y_c, Z_c)$ can be shown in (1) using homogeneous coordinates.

978-1-61284-881-5/11/$26.00 ©2011 IEEE

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{1}$$

In (1) $\boldsymbol{R}$ is a 3×3 orthogonal rotation matrix defining the camera orientation, and $\boldsymbol{T}$ is a translation vector defining the camera position, $\boldsymbol{0} = (0,0,0)^T$.

B. From Camera Coordinate to Image Coordinate

From a standardized projection of a point $P_c(X_c, Y_c, Z_c)$ in space based on pinhole model, we can get the point $p(x, y)$ in image pixel coordinate by $x = \frac{fX_c}{Z_c}$, $y = \frac{fY_c}{Z_c}$. This formula can be expressed in homogeneous coordinates and matrix as

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}. \tag{2}$$

with f is effective focal length of the camera.

C. From Image Coordinate to Image Coordinate Introduced Distortion

The normalized projection image coordinates introduced distortion is $P_d(x_d,\ y_d)$. The second-order lens distortion model is

$$x_d = (1 + k_1 r^2 + k_2 r^2)x + [2p_1 y + p_2(r^2 + 2x^2)]$$

$$y_d = (1 + k_1 r^2 + k_2 r^2)y + [p_1(r^2 + 2y^2) + 2p_2 x].$$

To simplify the formula using homogeneous coordinates and the equivalent transformation matrix, it is simplified as

$$\begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{3}$$

where $r^2=x^2+y^2$, k_1, k_2 are radial distortion coefficient, p_1、p_2 are tangential distortion coefficient, and $\boldsymbol{A}$ is an equivalent transformation matrix.

D. From Image Coordinate Introduced Distortion to Image Pixel Coordinate

The coordinate transformation can be shown as

$$u = u_0 + \frac{x_d}{dx} + sy_d$$

$$v = v_0 + \frac{y_d}{dy}$$

Simplify the formula in homogeneous coordinate and matrix as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & s & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}, \tag{4}$$

where (u_0, v_0) is a point of camera optical axis and image plane intersection, dx and dy is a pixel's size on the x-axis and y-axis in the image, and s is a camera tilt factor.

From (1)-(4), we can get

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & s & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{A} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \boldsymbol{M_1 M_2} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{5}$$

where $\boldsymbol{M_1}$ is internal camera parameters that includes dx, dy, u_0, v_0, f, k_1, k_2, p_1, p_2 and $\boldsymbol{M_2}$ is external camera parameter that includes rotation matrix $\boldsymbol{R}$ $\boldsymbol{R}$ and translation vector $\boldsymbol{T}$.

III. BINOCULAR CAMERA CALIBRATION

Besides the internal and external camera parameters, binocular vision camera calibration will get the relative position between the two cameras.

In the binocular camera calibration process, single camera calibration can obtain left and right external parameters in the world coordinate. We assume left camera external parameters are rotation matrix $\boldsymbol{R_l}$ and translation vector $\boldsymbol{T}_l$, and right camera parameters are rotation matrix $\boldsymbol{R}_r$ and translation vector $\boldsymbol{T}_r$. Therefore, if a point's x-coordinate is x_w in world coordinate, x_l in left camera coordinate and x_r in right camera coordinate, we can get the following formula.

$$\begin{cases} x_l = \boldsymbol{R}_l x_w + \boldsymbol{T}_l \\ x_r = \boldsymbol{R}_r x_w + \boldsymbol{T}_r \end{cases}. \tag{6}$$

The geometric relationship between the left camera and right camera can be expressed by $\boldsymbol{R}_0$, $\boldsymbol{T}_0$ as

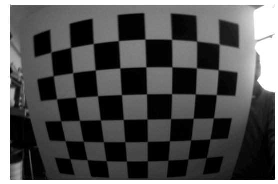
$$\begin{cases} \boldsymbol{R}_0 = \boldsymbol{R}_r \boldsymbol{R}_l^{-1} \\ \boldsymbol{T}_0 = \boldsymbol{T}_r - \boldsymbol{R}_r \boldsymbol{R}_l^{-1} \boldsymbol{T}_l \end{cases}. \tag{7}$$

IV. DESIGN OF CAMERA CALIBRATION PROCEDURE

This paper uses the checkerboard as calibration template. Because the checkerboard's roughness has much greater effects on the calibration's accuracy than random noise does, the calibration template should be printed with high-quality paper and printer. Paste the printed template on a plate, then put calibration template in front of the binocular vision system. We can collect images by varying the angle between template plane and image plane.

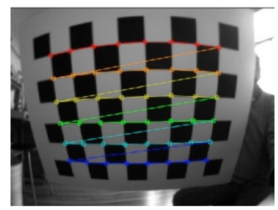
For camera calibration, we should extract the corners of checkerboard, which is essential in camera calibration. For one thing, if corner extraction doesn't work in most of checkerboard images, the accuracy of calibration is poor for the calibration procedure uses the least square method to calculate the calibration parameters. For another, if corner extraction of an image is failed, we need to give up this image to avoid that the memory space is allocated to calculate camera parameters. Although the OpenCV corner extraction algorithm works of the high successful rate, it is not guaranteed to extract all the images corners [7]. For example, when the light is poor because the light is blocked or other factors, the corner extraction is failed because the corner extract number of OpenCV algorithm cannot match the image corner number. So failure of corner extraction must be considered in camera calibration. Therefore, this paper proposed a method, gradient threshold corner extraction method, to extract corners when OpenCV corner extraction algorithm cannot work. This method doesn't change the threshold gradually until all the image corners are extracted successfully.

The experiment is conducted to extract the calibration template corners in poor light conditions using gradient threshold corner extraction method. The results proved this method could adapt to the light condition to extract corners effectively. The results are shown in Fig. 2, in which Fig. 2(a) is the source image, Fig. b is the certain threshold image in gradient threshold, Fig. c is the right-threshold image, and Fig. d is the extraction of corners.

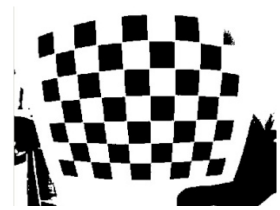


(a) Source image

(b) Certain threshold image



(c) Right-threshold image



(d) Extraction of corners

Figure 2. Gradient threshold corner extraction

According to the camera calibration principles, the calibration algorithm in this paper uses the gradient threshold method to extract corner based on the conventional camera calibration algorithm. The calibration procedure is conducted in Visual C++6.0 and OpenCV1.0, and the procedure flow chart is shown in Fig. 3.

Calibration procedure is as follows:

1) Read into left and right image data.
2) Read the image data into function cvFindChessboardCorners (). If the return value is 1: the number of extracted corners is the same as the set number, corner extraction is successful and save image to continue step (3); If the return value is 0: corner extraction is failed, read the image into function cvThreshold() for changing the threshold gradually. In this paper, change threshold at the rate of 0.5 until the corners are all extracted. Now, the return value of cvThreshold() is 1: corner extraction is successful and save image to continue step (3). Otherwise, corner extraction fails because corners are obscured due to camera angles and special causes, then give up the image, return to step (1).
3) Read the images that extract corner successfully into function cvFindCornerSubPix () to obtain sub-pixel corner coordinates more accurately.
4) Transform the sub-pixel corner coordinates from CvPoint2D32f structure into CvMat structure. The reason is that the input parameter of follow-

up function is CvMat structure, otherwise it will lead incorrect results.

5) Read the corner coordinates into function cvCalibrateCamera2() to get the left and right camera internal and external parameters matrix and distortion coefficient, then save the results.

6) Calculate the relative position of binocular vision camera parameters $\boldsymbol{R_0}$ and $\boldsymbol{T_0}$ according to (7), and save the results.

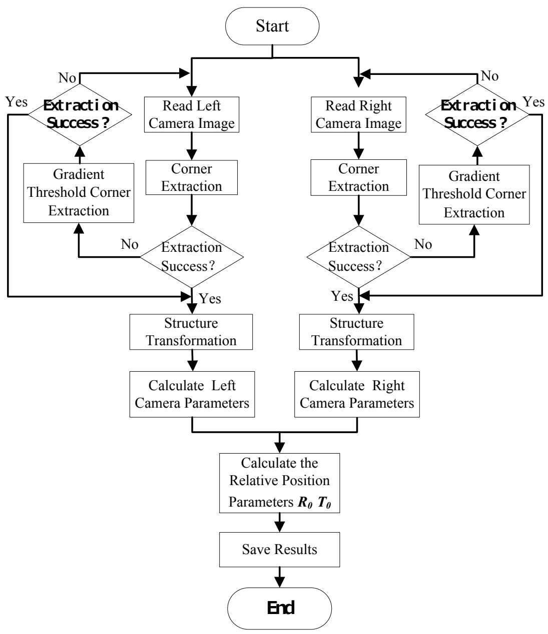


Figure 3. Calibration procedure flow chart

V. EXPERIMENTAL RESULTS AND ANALYSIS

According to the calibration procedure, this paper uses VC++6.0 to develop a camera calibration program based on OpenCV in Windows XP platform, which is experimental verified on binocular vision system. The experimental binocular vision system is STH-MDCS3-VARX produced by SRI International Company of America and it is shown in Fig. 4. The binocular vision system uses low-noise and high sensitivity CMOS imager, IEEE1394 interface to transmit data and supports 1280 × 960 and 640 × 480 resolution imaging.

Figure 4. Binocular vision system

The binocular vision system is divided into large and small baseline visual system according to the size of the distance between the two cameras. For large baseline visual system, calibration template must be large enough to show clearly in two cameras. The experimental calibration template is printed in A0 paper considering the restrictions of printer. There are 9 × 7 black and white squares and each square size is 108mm × 108mm, so we have 8 × 6 corners to extract. The experiment is carried out in large and small baseline binocular vision system. In both cases, two cameras collect ten images of 640 × 480 pixel to calibrate and the calibration results are shown in Table 1.

TABLE 1. CAMERA CALIBRATION RESULTS

Parameters	Binocular Vision System			
	Large baseline vision system		*Small baseline vision system*	
	Left camera	*Right camera*	*Left camera*	*Right camera*
f_x	1100.46	1115.86	1107.92	1119.86
f_y	1102.89	1122.34	1107.58	1124.18
u_0	251.897	252.777	255.890	253.345
v_0	223.517	231.372	225.379	229.867
k_1	-0.44880	-0.40315	-0.38217	-0.36164
k_2	0.46640	0.39840	0.41014	0.38650
p_1	0.00309	0.00150	0.00152	0.00465
p_2	0.00884	0.00145	0.00428	0.00114
$\boldsymbol{R_0}$	$\begin{bmatrix} 0.99995 & 0.00562 & -0.00185 \\ -0.00765 & 0.99987 & -0.00789 \\ 0.00055 & 0.00794 & 0.99994 \end{bmatrix}$		$\begin{bmatrix} 0.99982 & 0.00312 & -0.00511 \\ -0.00245 & 0.99996 & -0.00789 \\ 0.00143 & 0.00487 & 0.99976 \end{bmatrix}$	
$\boldsymbol{T_0}$	$\begin{bmatrix} -582.06809 & -4.09763 & 32.81186 \end{bmatrix}^T$		$\begin{bmatrix} -162.12504 & 1.80315 & 5.85282 \end{bmatrix}^T$	

From the experimental data, we can get the following conclusions. There are certain differences between the internal parameters of left and right camera, which are mainly due to the errors caused by manufacturing and installation. For large and small baseline vision system, camera internal parameters are slightly different, but the difference is rather small. Two cameras rotation matrix $\boldsymbol{R_0}$ are approximately the unit matrix, which means that there is not rotation, but translation between two cameras. From the translation vector $\boldsymbol{T_0}$, we know the translation of y-axis and z-axis is very small, which is consistent with the actual system. The distance between two cameras of large baseline vision system and small baseline system are 600mm and 160mm respectively and the x-axis displacement of two vision systems are -582.06809mm and -162.12504mm according to the translation vector T0. The experimental data is very similar as the actual system. Thus, this paper's algorithm can accurately calibrate the system for binocular vision, but also for large baseline and small baseline visual system with good stability.

VI. CONCLUSION

By studying the binocular vision camera model, a binocular vision system calibration procedure was developed that is effective and has good cross-platform portability. This paper proposed a gradient threshold corner extraction method, successfully extracting the calibration template corners and conducted a calibration experiment on the applied binocular vision system. The experimental results showed that this method is suitable for large and small baseline visual system with stability and accuracy.

ACKNOWLEDGMENTS

This project was supported by National Natural Science Foundation of China (No. 51105021).

REFERENCES

[1] M. L. Qiu, D. S. Ma, and Y. Li, "Overviw of camera calibration," *Acta Automatic Sinica*, vol. 26, pp. 43-45, January 2000.

[2] D. A. Forsyth, and J. Ponce, "Computer vision: A modern approach," *Electronic Industry Press*, 2004.

[3] Z. Z. Liu, and L. Fu, "Reasearch on camera calibration method," *Journal of Shenyang Institute of Aeronautical Engineering*, vol. 27, pp. 43-47, January 2010.

[4] G. Bradski, and A. Kaehler, *Learning OpenCV (Chinese)*, China: Tsinghua University Press, 2010.

[5] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330-1334, 2000.

[6] G. F. Xu, B. Li, and Z. K. Shen, "Measurement of motion parameters based on binocular vision model," *Infrared and Laser Engineering*, vol. 32, pp. 199-202, 2003.

[7] W. S. Yin, Y. L. Luo, and S. Q. Li, "Camera calibration based on OpenCV," *Computer Engineering and Design*, vol. 28, pp. 197-199, 2007.