Optics and Lasers in Engineering 90 (2017) 119–127

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

# Sphere-based calibration method for trinocular vision sensor

Rui Lu[a], Mingwei Shao[b,*]

[a] National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing 100029, China
[b] Qingdao Research Center for Advanced Photonic Technologies, Laser Institute of Shandong Academy of Sciences, Qingdao 266100, China

ARTICLE INFO

*Keywords:*
Sphere
Trinocular vision
Calibration
Feature matching

ABSTRACT

A new method to calibrate a trinocular vision sensor is proposed and two main tasks are finished in this paper, i.e. to determine the transformation matrix between each two cameras and the trifocal tensor of the trinocular vision sensor. A flexible sphere target with several spherical circles is designed. As the isotropy of a sphere, trifocal tensor of the three cameras can be determined exactly from the feature on the sphere target. Then the fundamental matrix between each two cameras can be obtained. Easily, compatible rotation matrix and translation matrix can be deduced base on the singular value decomposition of the fundamental matrix. In our proposed calibration method, image points are not requested one-to-one correspondence. When image points locates in the same feature are obtained, the transformation matrix between each two cameras with the trifocal tensor of trinocular vision sensor can be determined. Experiment results show that the proposed calibration method can obtain precise results, including measurement and matching results. The root mean square error of distance is 0.026 mm with regard to the view field of about 200×200 mm and the feature matching of three images is strict. As a sphere projection is not concerned with its orientation, the calibration method is robust and with an easy operation. Moreover, our calibration method also provides a new approach to obtain the trifocal tensor.

## 1. Introduction

In multi-camera vision system, including stereo vision system, the relationship of cameras is fixed. Therefore, determining the transformation matrix (including the translation matrix and the rotation matrix) between each two cameras is significant. Meanwhile, another task is to determine the multi-focal tensor of multi-camera vision sensor, which is essential to feature matching. The method to finish these two tasks is named as calibration.

For stereo vision system, calibration methods are various, such as planar target-based method, 1-D target-based method, etc. In the planar target-based method [1–3], different features on the planar target are utilized, e.g. centers of circles, corners, crosspoints, parallel lines and so on. Anyway, the relationships of these features are known exactly. When images of these features are captured by the two cameras, the coordinates (or expressions) under each camera coordinate system can be deduced from the camera model and the constraints of features. Once the corresponding features are confirmed, the relationship between the two cameras can be determined. In the 1-D target-based calibration method [4,5], the essential matrix is calculated from the corresponding feature points. When singular value decomposition is conducted to the essential matrix, the rotation matrix and the translation matrix, which needs a factor, can be deduced. As the relationship between two feature points is known, the factor can be confirmed easily.

The sphere target is also widely used to calibrate the stereo vision system. In [6,7], Agrawal and Zhang calibrate the intrinsic parameters of a camera based on the projection of the sphere target and the obtained dual image of the absolute conic. Meanwhile, the relationship between two cameras can be deduced from the 3D points cloud registration method. However, when there are little feature points, calibration results will be with more noise. In [8], two calibration methods are proposed to finish the calibration task. One is using sphere centers and points of tangency to obtain the fundamental matrix, the other is using the homography matrix and epipoles to determine fundamental matrix. In these sphere-based calibration methods mentioned above, the target should be placed in many different positions to obtain enough feature points, meanwhile, a mass of calculation is inevitable.

For multi-camera vision system, which consists of more than two cameras, the relationship of cameras is also fixed. Normally, the calibration method for stereo vision system is valid when each two cameras are treated as a stereo vision system. However, as is known, the relationship between two cameras (or two image plane) can be

* Corresponding author.
*E-mail address:* smw1987@163.com (M. Shao).

http://dx.doi.org/10.1016/j.optlaseng.2016.10.004
Received 8 June 2016; Received in revised form 2 September 2016; Accepted 4 October 2016

represented by fundamental matrix. Similarly, the relationship of cameras is expressed by multi-focal tensor. In trinocular vision sensor, the tensor is trifocal tensor. When each two cameras are calibrated by the method for stereo vision system, the calibration result cannot satisfy the constraint of trifocal tensor as it may be local optimum. In addition, the normal calibration method often meets the problem of self-occlusion as the different shooting angle of each camera. In this case, a new approach to calibrate the trinocular vision sensor needs to be proposed.

In this paper, a new calibration method for trinocular vision sensor is presented. A sphere target with several spherical circles is utilized to determine the trifocal tensor. When the trifocal tensor is determined, the fundamental matrix between each two cameras can be obtained. In this case, the rotation matrix and the translation matrix with a scale factor can deduced from the singular value decomposition of the fundamental matrix. Then the scale factor is determined by the known spherical circle. In our method, the calibration result is precise and the problem of self-occlusion is evitable as the isotropy of a sphere.

## 2. Determination of trifocal tensor

On the sphere target, there are several spherical circles. Therefore, the features include the visible outline of the sphere target and the spherical circles. In this section, we will discuss the relationship between the trifocal tensor and these features.

### 2.1. Outline of the sphere

The projection of a sphere on the image plane is an ellipse, which can be expressed as a matrix form [6]:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} a & c/2 & d/2 \\ c/2 & b & e/2 \\ d/2 & e/2 & f \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0, \tag{1}$$

where $(u, v, 1)^{\mathrm{T}}$is the homogeneous coordinate of the projective point on the image plane and $a, b, c, d, e, f$ are the parameters of the elliptical expression.

In Fig. 1, *O-XYZ* is the Camera Coordinate System (*CCS*) while *o-xy* is the Image Coordinate System (*ICS*). Under the *CCS*, the projection center of the camera is at the origin and the optical axis points in the positive $Z$ direction. Supposing a spatial point $P$ is projected onto the plane with $Z$=$f_0$, referred to as the image plane under the *CCS*, where $f_0$ is the effective focal length (*EFL*). Supposing $p = (x, y, 1)^{\mathrm{T}}$ is the projection of $P = (X, Y, Z)^{\mathrm{T}}$ on the image plane. Under the ideal pinhole imaging model, the undistorted model of the camera, $P$, $p$ and the projection center $O$ are collinear. The fact can be expressed by the following equation:

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_0 & 0 & 0 & 0 \\ 0 & f_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{2}$$

Practically, the radial distortion and the tangential distortion of the lens are inevitable. When considering the radial distortion, we have the following equations:

$$\begin{cases} \bar{x} = x(1 + k_1 r^2 + k_2 r^4) \\ \bar{y} = y(1 + k_1 r^2 + k_2 r^4) \end{cases}, \tag{3}$$

where $r^2 = x^2 + y^2$,$(x, y)^{\mathrm{T}}$is the distorted image coordinate, and $(\bar{x}, \bar{y})^{\mathrm{T}}$is the idealized one, $k_1$, $k_2$ are the radial distortion coefficients of the lens.

According to Eqs. (1) and (2), we obtain the matrix representation of the right-circular cone under *CCS* as:

$$[X \;\; Y \;\; Z]Q[X \;\; Y \;\; Z]^{\mathrm{T}} = 0, \tag{4}$$

where the matrix $Q$ is defined as [7]:

$$Q = \begin{bmatrix} A & C/2 & D/2 \\ C/2 & B & E/2 \\ D/2 & E/2 & F \end{bmatrix}, \tag{5}$$

and the related definitions are show as follow:

$$A = af_0^2,\ B = bf_0^2,\ C = cf_0^2,\ D = df_0,\ E = ef_0,\ F = f \tag{6}$$

Define the coordinate of the sphere center under the *CCS* is $(X_0, Y_0, Z_0)^{\mathrm{T}}$, the coordinate of its corresponding image point is $(u_0, v_0, 1)^{\mathrm{T}}$, the relationship according to Eq. (2) is

$$K \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} = \lambda \begin{pmatrix} u_0 \\ v_0 \\ 1 \end{pmatrix}, \tag{7}$$

where $K$ is the intrinsic parameter matrix of the camera. So the projection of the sphere center and the outline of the sphere can be deduced from Eqs. (5) and (7):

$$K \begin{pmatrix} e_{3x} \\ e_{3y} \\ e_{3z} \end{pmatrix} = \lambda' \begin{pmatrix} u_0 \\ v_0 \\ 1 \end{pmatrix}, \tag{8}$$

the matrix $K$ in Eq. (8) is defined the same as in Eq. (7), $\lambda$ and $\lambda'$ are scale factors as they satisfies the following equation:

$$\lambda' = \frac{\lambda}{R}\sqrt{\frac{|\lambda_3|}{|\lambda_3| + |\lambda_r|}}, \tag{9}$$

where$\lambda_1$, $\lambda_2$ and $\lambda_3$ are the eigenvalues of matrix $Q$, $\lambda_1$ and $\lambda_2$ must have the same sign while $\lambda_3$ must have the different one. As $Q$ is a spherical matrix, we define $\lambda_{\mathrm{r}} = \lambda_1 = \lambda_2$. $[e_{3x}, e_{3y}, e_{3z}]^{\mathrm{T}}$ is the eigenvector corresponding to $\lambda_3$ (If $e_{3z} < 0$, the eigenvector should be multiplied by scale factor $-1$) and $R$ is the radius of the sphere.

In trinocular vision sensor, the relationship of these corresponding image points $P$ according to trifocal tensor is [8]:

$$P_1^i P_2^j P_3^k \varepsilon_{jqu} \varepsilon_{krv} T_i^{qr} = 0_{uv}, \tag{10}$$

where $T$ is the trifocal tensor,

$$\varepsilon^{ijk} = \begin{cases} 1,\ (i, j, k)\text{is even permutation} \\ -1,\ (i, j, k)\text{is odd permutation} \\ 0,\ \text{others} \end{cases} \tag{11}$$

the indices repeated in the contravariant and covariant positions imply summation over the range (1,2,3) of the index. According to Eqs. (8) and (10), the relationship of the three projections of the sphere outline on the three image planes can be expressed as:

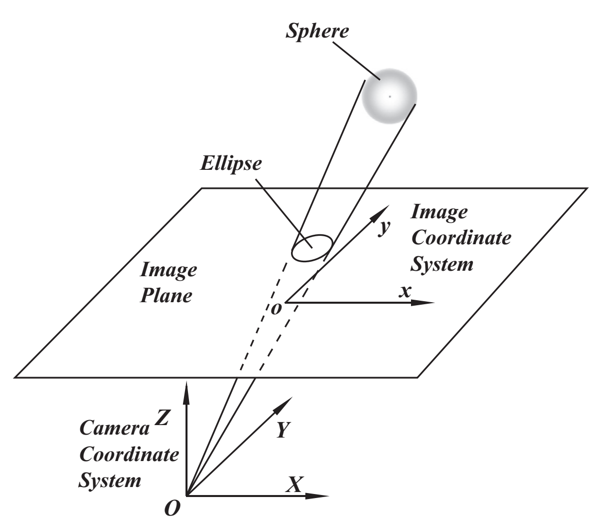


**Fig. 1.** The projection of a sphere under the camera coordinate system.