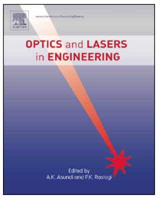
Optics and Lasers in Engineering 68 (2015) 203–213

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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



Calibration for stereo vision system based on phase matching and bundle adjustment algorithm

Huijie Zhao[a], Zhen Wang[a], Hongzhi Jiang[a,*], Yang Xu[a], Chao Dong[b]

[a] Beihang University, School of Instrumentation Science and Opto-electronics Engineering, XueYuan Road No. 37, HaiDian District, Beijing 100191, China
[b] Key Laboratory of Technology and Application for Safeguarding of Marine Rights and Interests, SOA, Guangzhou 510301, China

ARTICLE INFO

Article history:
Received 15 July 2014
Received in revised form
18 November 2014
Accepted 1 December 2014
Available online 30 January 2015

Keywords:
Camera calibration
Extrinsic parameters calibration
3D measurement
Phase matching
Bundle adjustment

ABSTRACT

Calibration for stereo vision system plays an important role in the field of machine vision applications. The existing accurate calibration methods are usually carried out by capturing a high-accuracy calibration target with the same size as the measurement view. In in-situ 3D measurement and in large field of view measurement, the extrinsic parameters of the system usually need to be calibrated in real-time. Furthermore, the large high-accuracy calibration target in the field is a big challenge for manufacturing. Therefore, an accurate and rapid calibration method in the in-situ measurement is needed. In this paper, a novel calibration method for stereo vision system is proposed based on phase-based matching method and the bundle adjustment algorithm. As the camera is usually mechanically locked once adjusted appropriately after calibrated in lab, the intrinsic parameters are usually stable. We emphasize on the extrinsic parameters calibration in the measurement field. Firstly, the matching method based on heterodyne multi-frequency phase-shifting technique is applied to find thousands of pairs of corresponding points between images of two cameras. The large amount of pairs of corresponding points can help improve the accuracy of the calibration. Then the method of bundle adjustment in photogrammetry is used to optimize the extrinsic parameters and the 3D coordinates of the measured objects. Finally, the quantity traceability is carried out to transform the optimized extrinsic parameters from the 3D metric coordinate system into Euclid coordinate system to obtain the ultimate optimal extrinsic parameters. Experiment results show that the procedure of calibration takes less than 3 s. And, based on the stereo vision system calibrated by the proposed method, the measurement RMS (Root Mean Square) error can reach 0.025 mm when measuring the calibrated gauge with nominal length of 999.576 mm.

1. Introduction

The 3-D shape measurement techniques based on digital fringe projection and phase-shifting methods have been widely used in optical noncontact surface measuring field [1–4]. Usually a stereo vision measurement system consists of two stereo cameras and a digital projector. The projector is used to project groups of phase-shifting fringe images onto the objects, and two cameras capture distorted fringe images synchronously.

For a stereo vision system, accurate calibration is quite important since the calibration results determine the mapping relationship between 3D reference coordinate and 2D image coordinate. In many cases, the performance of the vision measurement system strongly depends on the accuracy of the camera calibration. The parameters to be calibrated include the intrinsic parameters and the extrinsic parameters. The intrinsic parameters consist of the principle point, the focal length and the distortion of the camera. The extrinsic parameters include the rotation matrix and translation vector which describe the relationship between two cameras. The intrinsic parameters of camera are stable as the camera is usually mechanically locked once adjusted appropriately after calibrated in lab. While the extrinsic parameters are inevitably changed when the instrument is assembled or moved into industrial measuring field. Especially for the application of vision measurement in large field of view, the system structure parameters tend to change because of the long baseline between the two cameras. Therefore, the fast calibration for the system extrinsic parameters is needed in in-situ 3D measurement and in large field of view measurement.

The classical calibration methods include Tsai's and Zhang's methods [5,6]. In these methods, the system takes pictures of a planar board with chess pattern or circular blobs from different views. The procedure of calibration is relatively complicated and time-consuming. And, for the application in large field of view

---

* Corresponding author. Tel.: +86 10 82339563; fax: +86 10 82315884.
E-mail address: jhz1862@buaa.edu.cn (H. Jiang).

http://dx.doi.org/10.1016/j.optlaseng.2014.12.001
0143-8166/

measurement, the size of the planar target must be relative to the field of image, making it a challenge for manufacturing. To achieve rapid calibration, there are some automatic calibration method proposed [7–9], including the method based on dashed lines and two orthogonal vanishing points. Camera calibration methods by using fringe patterns, not-measured calibration target, and Virtual 1D Target are also ways to simplify the procedure [10–12]. There have been a few efforts to solve the problem of calibration for the system with a large field of view without using a big planar target [13–15]. For example, Xiao etc. proposed a calibration method for binocular 3D measurement systems, utilizing a cross target with ring coded points as calibration pattern. Shang etc. proposed a camera calibration method for large field optical measurement, in which the camera's optical center and the control points are approximately coplanar. Only a single image of these control points captured by the camera in measurement state is used in this method. Sun etc. proposed a method based on a baseline ruler. The targets need to be randomly placed in the field of view of the vision sensor for several times. Nearly all the methods above cannot satisfy the requirement for high accurate and speed calibration in in-situ measurement and large field of view measurement.

In this paper, a novel calibration method for fringe projection stereo vision system is proposed based on phase-based matching method and the bundle adjustment algorithm, which is suitable for in-situ calibration in industrial inspection [16]. In this calibration method, the intrinsic parameters of the cameras and extrinsic parameters of the system are calibrated respectively. The cameras are usually locked mechanically before moved into the working field. Thus, the intrinsic parameters are regarded as being stable. We emphasize the extrinsic parameters calibration in the measurement filed. First, matching method based on heterodyne multi-frequency phase-shifting technique is applied to find thousands of homologous pairs of corresponding points between images of two cameras. The large number pairs of matching points can help improve the accuracy of the calibration. Next the bundle adjustment algorithm is used to optimize the extrinsic parameters and the 3D coordinates of the measured objects. The initial extrinsic parameters for the iteration are computed based on the relative orientation method in photogrammetry by two cameras taking a picture of a cross-shaped target respectively. At last, quantity traceability is carried out by measuring a gauge appliance for a few times, which transforms the optimized extrinsic parameters from the 3D metric coordinate system into Euclid coordinate system to obtain the ultimate optimal extrinsic parameters. The advantage of the proposed method is that no planar target is used, which contributes to a simple and rapid calibration method. The method can also successfully applied for scanning high-reflective surfaces based on a fast and high dynamic range digital fringe projector [17,18].

The organization of the paper is as follows. Section 2 introduces the mathematical model of the stereo vision system. Section 3 provides the details of the proposed calibration method and the analysis of the uncertainty. Section 4 describes the experiment process and gives the results of the experiment. The paper ends with some concluding remarks in Section 5.

2. Mathematical model of the stereo vision system

The stereo vision system includes two stereo cameras and a digital projector. The projector projects a group of phase-shifting intensity sinusoidal patterns onto the scene, and the two cameras take picture of images synchronously. Fig. 1 shows the model of the stereo vision system, including two pinhole camera models and the associated coordinate frames.

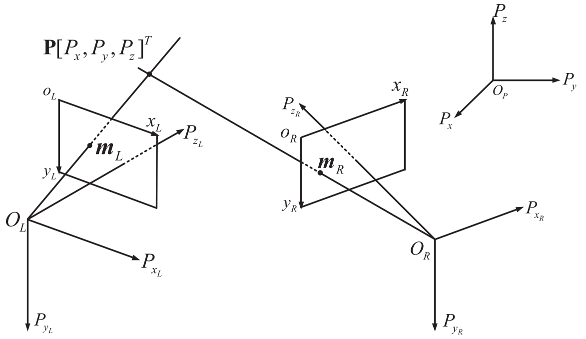


Fig. 1. 3D measurement model of binocular stereo system.

The ideal pinhole camera model [6] is represented by a relationship between the object space (object point $\mathbf{P}[P_x, P_y, P_z]^T$) and the image space (image point $\mathbf{m}[x, y]^T$):

$$s\begin{bmatrix}\mathbf{m}\\1\end{bmatrix} = \mathbf{A}[\mathbf{R}_w|\mathbf{T}_w]\begin{bmatrix}\mathbf{P}\\1\end{bmatrix} \tag{1}$$

where

$$\mathbf{A} = \begin{bmatrix}\alpha_x & \gamma & x_\mathbf{o}\\ 0 & \alpha_y & y_\mathbf{o}\\ 0 & 0 & 1\end{bmatrix} \tag{2}$$

s is arbitrary scale factor, $\mathbf{P}[P_x, P_y, P_z]^T$ and $\boldsymbol{m}[x, y]^T$ are a spatial point and its corresponding image point, $[\mathbf{R}_w|\mathbf{T}_w]$ are the camera extrinsic parameters, including rotation matrix and translation vector, and $\boldsymbol{A}$ is the camera intrinsic matrix consisting five parameters: the effective focal length in pixels of the camera along the x and y direction α_x and α_y, the principle point coordinate $[x_\mathbf{o}, y_\mathbf{o}]^T$, and the skew factor γ usually set to zero.

Considering the impact of lens distortion [19], we have the follow formula:

$$\begin{bmatrix}\check{x}\\ \check{y}\end{bmatrix} = \begin{bmatrix}x\\ y\end{bmatrix} + \begin{bmatrix}\delta_x(x,y)\\ \delta_y(x,y)\end{bmatrix} \tag{3}$$

where $[x, y]^T$ is the ideal (distortion-free) pixel image coordinate, and $[\check{x}, \check{y}]^T$ is the corresponding real observed image coordinate. $\delta_x(x,y)$ and $\delta_y(x,y)$ are the distortion values calculated as:

$$\begin{cases}\delta_u(x,y) = u\left(k_1 r^2 + k_2 r^4\right) + 2k_3 xy + k_4\left(r^2 + 2x^2\right)\\ \delta_v(x,y) = v\left(k_1 r^2 + k_2 r^4\right) + k_3\left(r^2 + 2y^2\right) + 2k_4 xy\end{cases}$$
$$r^2 = x^2 + y^2 \tag{4}$$

where $\mathbf{k} = \begin{bmatrix}k_1 & k_2 & k_3 & k_4\end{bmatrix}$ are the lens distortion coefficients.

According to the above model, the relationship between the object space (object point $\mathbf{P}[P_x, P_y, P_z]^T$) and the image space of the left and right camera (image point $\mathbf{m}_L[x_L, y_L]^T$ and $\mathbf{m}_R[x_R, y_R]^T$) are described as follows:

$$s_L\begin{bmatrix}\boldsymbol{m}_L\\1\end{bmatrix} = \mathbf{A_L}[\mathbf{R_L}|\mathbf{T_L}]\begin{bmatrix}\mathbf{P}\\1\end{bmatrix} \tag{5}$$

$$s_R\begin{bmatrix}\boldsymbol{m}_R\\1\end{bmatrix} = \mathbf{A_R}[\mathbf{R_R}|\mathbf{T_R}]\begin{bmatrix}\mathbf{P}\\1\end{bmatrix} \tag{6}$$

In a measurement system, the left camera coordinate system is usually defined as the world coordinate system. The relationship of the left camera coordinate system ($O_L\_P_{x_L}P_{y_L}P_{z_L}$) and the right camera coordinate system ($O_R\_P_{x_R}P_{y_R}P_{z_R}$) is described as follows:

$$\begin{bmatrix}P_{x_L}\\ P_{y_L}\\ P_{z_L}\end{bmatrix} = \mathbf{R}\begin{bmatrix}P_{x_R}\\ P_{y_R}\\ P_{z_R}\end{bmatrix} + \mathbf{T} \tag{7}$$