A Novel Approach to Calibrating Multifunctional Binocular Stereovision Sensor

T Xue, J G Zhu, B Wu and S H Ye

State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China

E-mail: snowxue@eyou.com

Abstract. We present a novel multifunctional binocular stereovision sensor for various three-dimensional (3D) inspection tasks. It not only avoids the so-called correspondence problem of passive stereo vision, but also possesses the uniform mathematical model. We also propose a novel approach to estimating all the sensor parameters with free-position planar reference object. In this technique, the planar pattern can be moved freely by hand. All the camera intrinsic and extrinsic parameters with coefficient of lens radial and tangential distortion are estimated, and sensor parameters are calibrated based on the 3D measurement model and optimized with the feature point constraint algorithm using the same views in the camera calibration stage. The proposed approach greatly reduces the cost of the calibration equipment, and it is flexible and practical for the vision measurement. It shows that this method has high precision by experiment, and the sensor measured relative error of space length excels 0.3%.

1. Introduction

With the development of optoelectronics, image processing and computer vision technique, vision inspection has made giant strides and has been widely used in various industrial applications, such as industrial automation, robotics, component quality control and object recognition[1,2], due to non-contact, fast measuring speed, moderate accuracy, well flexibility and low cost.

In order to adapting various three-dimensional (3D) vision inspection tasks, many kinds of vision sensors have been developed, such as structured light stripe vision sensor, passive stereo vision sensor, LCD raster vision sensor and so on. But it is unavoidably bring about measurement complexity because these sensors possess different structures, mathematical models and calibration methods. In the paper, we present a novel multifunctional binocular stereovision sensor which consists of two cameras and a feasible projector for choosing according to various measure tasks. It not only avoids the so-called correspondence problem of passive stereo vision, but also possesses the uniform mathematical model.

The fundamental principle of multifunctional binocular stereovision sensor is stereo vision. The complete calibration parameters include camera intrinsic parameters (e.g. effective focal length, principle point and lens distortion coefficient), extrinsic parameters (including the 3D position and orientation of two camera coordinate frames relative to a certain world coordinate frame) and sensor parameters (the relative 3D position and orientation of two camera coordinate frames).

Conventional approaches to estimating the complete sensor parameters usually use elaborate 3D calibration target, such as the target with two or three planes orthogonal to each other, 3D sphere
target and so on [3]. For calibration targets with orthogonal planes, two cameras have difficulty observing all the feature points on different target planes simultaneously. For 3D sphere target, it is not easy to maintain and usually requires recalibration by CMM before used again. Wu [4] proposed a method using virtual 3D calibration target by 2D one, but target plane must undergo a precisely known translation which orientation must be perpendicular to the plane. A two-step calibration method based on the principle of “vanishing point” of the parallel lines in Ref. [5] is difficult to acquire parallel line vectors with high precision, thus it is unsuitable for on-line sensor calibration.

In this paper, inspired by the work of Zhang [6], we propose a novel approach to calibrating complete primitive parameters of this kind of stereovision sensor with free-position planar reference object. In this technique, the pattern can be printed with a high-quality printer and attached to a hard surface, or ray-echoed for much higher precision. The planar target can be moved freely by hand, and two cameras observe it shown at a few different orientations. All the camera intrinsic and extrinsic parameters with coefficient of lens radial and tangential distortion are estimated, and sensor parameters are calibrated based on the 3D measurement model and optimized with the feature point constraint algorithm using the same views in the camera calibration stage.

2. Mathematical model of the multifunctional binocular stereovision sensor

The multifunctional binocular stereovision sensor we designed consists of two cameras and a variable projector, which can be replaced flexibly. For example, lamp or high brightness LED is usually assembled to inspect feature hole, single structured light strip can be substituted for measuring feature point or feature line, multiple structured light strip or LCD raster is often applied in surface scanning or object reconstruction.

Figure 1 shows the 3D mathematical model of the multifunctional binocular stereovision sensor. It is supposed that $o_l X_l Y_l z_l$ and $o_r X_r Y_r z_r$ are the left camera 3D coordinate frame and its responding 2D image plane coordinate frame, $f_{x_l}$, $f_{y_l}$ are the effective focal length in the $x$ and $y$ directions of left camera. $o_r X_r Y_r z_r$, $o_r X_r Y_r z_r$ are denoted the right camera 3D coordinate frame and its responding 2D image plane coordinate frame. $o_s x_s y_s z_s$ is the sensor measure coordinate frame, which overlaps with left camera coordinate frame.

Given an arbitrary 3D point $P$, Let $P = [x_s \ y_s \ z_s]^T$, $\tilde{p}_s = [x_s \ y_s \ z_s \ 1]^T$ be the 3D coordinate and homogeneous coordinate of point $P$ in $o_s x_s y_s z_s$ ($o_l X_l Y_l$), $P_r = [x_r \ y_r \ z_r]^T$, $\tilde{p}_r = [x_r \ y_r \ z_r \ 1]^T$ be the 3D coordinate and homogeneous coordinate in $o_r X_r Y_r z_r$. $p_l, p_r$ are the ideal projection of $P$ in the left camera and right camera image plane, respectively, their corresponding homogeneous coordinates in $o_l X_l Y_l$ and $o_r X_r Y_r$ are denoted by $\tilde{p}_l = [X_l \ Y_l \ 1]^T$ and $\tilde{p}_r = [X_r \ Y_r \ 1]^T$. The position and orientation of two cameras is described by

$$P_r = M_{sr} \tilde{P}_s = \begin{bmatrix} R & T \end{bmatrix} \tilde{P}_s$$  (1)
where \( R \) and \( T \) are the orthogonal rotation matrix and translation matrix between \( o_x y z \) and \( o_r x y z_r \), respectively.

Based on perspective projection model of camera, we have
\[
\rho_l \tilde{P}_l = P_s \quad \text{with} \quad \rho_l = z_s, \quad \rho_r \tilde{P}_r = P_r
\]
(2)

From (1)-(2), the 3D mathematical model of multifunctional stereovision sensor can be represented as follows:
\[
\begin{align*}
x_y &= z_s X_i \\
y_y &= z_s Y_i \\
t_y - X_i t_z &= t_y - Y_i t_z \\
X_r (r_7 X_i + r_8 Y_i + r_9) - (r_1 X_i + r_2 Y_i + r_3) &= Y_r (r_7 X_i + r_8 Y_i + r_9) - (r_4 X_i + r_5 Y_i + r_6)
\end{align*}
\]
(3)

According to Eqs.(3), we can acquire the 3D coordinate of measured feature points if their homogeneous image coordinates of left and right cameras, rotation matrix \( R \) and translation vector \( T \) are known exactly.

3. A flexible novel calibration method

3.1. Camera calibration

We assume calibration target plane is on \( z=0 \) of the world coordinate frame. A 2D point on calibration target plane and its corresponding projection point are related by homography \( H \)
\[
\begin{align*}
s_l \tilde{P}_l = A[r_1 \ r_2 \ r_3][\tilde{P}_w = [h_1 \ h_2 \ h_3][\tilde{P}_w
\end{align*}
\]
(4)

where \( A \) is called the camera intrinsic parameter matrix.

Homography \( H \) can be estimated through linear or nonlinear least-squares algorithm. Using the knowledge that \( r_1 \) and \( r_2 \) are orthonormal, the camera intrinsic parameters \((\lambda, u_0, v_0, f_x, f_y)\) and extrinsic parameters \((r_1, r_2, r_3, t)\) can be estimated.

For vision inspection tasks demanding high precision, complex lens distortion effect must be considered. The radial lens distortion can be given by
\[
\begin{align*}
\delta \begin{bmatrix} u_r \\
v_r 
\end{bmatrix} &= \begin{bmatrix} x_u \\
y_u 
\end{bmatrix} \\
k^r &= k_1 r^2 + k_2 r^4 + \cdots
\end{align*}
\]
(5)

where \( r^2 = x_u^2 + y_u^2 \), \( k_1 \) and \( k_2 \) are the coefficients of the radial distortion. The tangential lens distortion is given by
\[
\begin{align*}
\delta \begin{bmatrix} u_t \\
v_t 
\end{bmatrix} &= \begin{bmatrix} 2p_1 x_u y_u + p_2 (r^2 + 2x_u^2) \\
p_1 (r^2 + 2y_u^2) + 2p_2 x_u y_u 
\end{bmatrix} \\
\end{align*}
\]
(6)

where \( p_1 \) and \( p_2 \) are the coefficients of tangential distortion. Thus actual projection coordinate can be represented as follows:
\[
\begin{align*}
\begin{bmatrix} x_d \\
y_d 
\end{bmatrix} &= \begin{bmatrix} x_u + \delta u_r + \delta u_t \\
y_u + \delta v_r + \delta v_t 
\end{bmatrix}
\end{align*}
\]
(7)

Due to the nonlinear nature of camera model with lens distortion, estimated model coordinate pair \((U_i, V_i)\) and real extraction coordinate pair of feature point pair \((u_i, v_i)\) are often inconsistent. Thus iterative optimizing algorithm must be used to minimize the sum of squared residuals between model and observations. The objective function is expressed by
\[
F(x) = \sum_{i=0}^n ((U_i - u_i)^2 + (V_i - v_i)^2)
\]
(8)
Based on (4)-(8), all the intrinsic parameters with both radial and tangential distortion coefficients and extrinsic parameters of two cameras are estimated through Levenberg-Marquardt least squares fitting algorithm.

4. Sensor calibration
The extrinsic parameters of left and right camera described in section 2 are denoted by \((R_l, T_l)\) and \((R_r, T_r)\). Given an arbitrary point \(P\) in 3D space, its 3D world coordinate, left and right camera coordinate are defined by \(x_w, x_l\) and \(x_r\), we have

\[
x_l = R_l x_w + T_l, \quad x_r = R_r x_w + T_r
\]

If \(x_w\) is eliminated from (9), we have

\[
x_l = R_l R_r^{-1} x_r + T_l - R_l R_r^{-1} T_r
\]

For calibration reference object with planar pattern, because all the views of planar pattern are acquired by one camera in different positions, thus each view has common intrinsic parameters, but various extrinsic parameters. If only use one group of extrinsic parameters to estimate the geometric parameters of stereovision sensor, parameters will be very unstable because of the noise. Thus we should use nonlinear least squares algorithm to optimize calibration parameters. From (3), we have

\[
(t_x - X_r \cdot t_z) (r_4 \cdot Y_l + r_5) - (r_1 \cdot X_l + r_2 \cdot Y_l + r_3) (t_y - Y_r \cdot t_z) =
(Y_r \cdot t_x - X_r \cdot t_y) (r_7 \cdot X_l + r_8 \cdot Y_l + r_9)
\]

It is supposed that \(T'=\alpha T\), with \(\alpha=1/t_x, t_x \neq 0\), we have \(T' = [1, t_y, t_z]\). Rotation matrix can be transformed by rotation vector through Rodrigues formula, thus Eqs.(11) is actually a nonlinear equation with five unknown parameters. We can denote it by simplified function \(f(x) = 0\).

Using feature point constraint described in Eqs. (11), the objective function for refining multifunctional binocular stereovision sensor parameters is defined as:

\[
F(x) = \sum_{i=1}^{N} f_i^2(x)
\]

where \(x = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T\).

All the more accurate parameters according to (12) can be estimated by Levenberg-Marquardt nonlinear iterative optimizing algorithm. The real translation vector can be obtained based on the accurate distance of feature points on planar pattern. Then all the optimized sensor parameters are acquired according to sensor’s 3D measure model and distance formula of two feature points in 3D space.

5. Experiments
We designed a multifunctional binocular stereovision sensor, which consists of two off-the-shelf WATEC CCD cameras (902H) with 25 mm AVANIA lens and a variable projector. The distance between two cameras is about 300 mm. Each camera took five images of the free-position planar pattern. The complete parameters can be estimated with those feature points in the previous four images. The left camera intrinsic parameters are

- \(f_x=3455.066503\) pixels, \(f_y=3442.459433\) pixels, \(u_0=361.278673\) pixels, \(v_0=254.913139\) pixels
- \(k_1=-1.31777\times10^{-1}\), \(k_2=3.20810\), \(p_1=1.951\times10^{-3}\), \(p_2=1.915\times10^{-3}\)

The right camera intrinsic parameters are

- \(f_x=3491.784928\) pixels, \(f_y=3480.806601\) pixels, \(u_0=324.195347\) pixels, \(v_0=301.981734\) pixels
- \(k_1=-1.22840\times10^{-1}\), \(k_2=-1.513795\), \(p_1=1.951\times10^{-3}\), \(p_2=1.915\times10^{-3}\)

The sensor parameters can be represented by
We evaluate the sensor calibration accuracy with the fifth image taken by the left camera, which possesses 30 feature points. We compare distance between two adjacent points in sensor measure coordinate frame with actual accurate distance in planar pattern. The relative measurement error of sensor is shown in Fig. 2, the horizontal axis is represented as 29 distances in the fifth image, and the vertical axis is defined as relative errors between measured distance and the ground truth. As we can see from figure 2, the measured relative error of the 29 distances excels 0.3%.

![Figure 2. Relative measurement error of sensor.](image)

6. Conclusion
A novel multifunctional binocular stereovision sensor was presented to solve a large amount of vision inspection tasks. And a feasible calibration method with free-position planar pattern was presented. All the camera intrinsic and extrinsic parameters with coefficient of lens distortion are estimated. The sensor parameters can be acquired with the same views of planar pattern in the camera calibration stage, and refined through optimizing algorithm based on feature point constraint. It shows the sensor measured relative error of space length excels 0.3% by experiment. Compared with conventional methods, the proposed technique gains considerable flexibility and simplifies the calibration procedure. It advances vision inspection one step from laboratory environments to real world use.

Acknowledgements
This work was funded by the special funds of national natural science funds of China [50327501] and the major project of natural science funds of Tianjin in China [05YFJZJC01700].

References
[1] Yamada H, Togasaki T, Kimura M and Sudo H 2003 High-density 3-D packaging technology based on the sidewall interconnection method and its application for CCD micro-camera visual inspection system *IEEE Trans. Advanced Packaging* **26** 113–121.
[2] Yang C C, Marefat M M and Ciarallo F W 1998 Error analysis and planning accuracy for dimensional measurement in active vision inspection *IEEE Trans. Robotics and Automation* **14** 476–487.
[3] Faugeras O 1993 *Three-dimensional computer vision: a Geometric Viewpoint*(MIT press).
[4] Wu B 2002 *Study on the key technologies of 3D digital measurement of large-scale objects* Tianjin University (Ph.D dissertation).
[5] Zhang J X, Duan F J and Ye S H 1998 “Two step” method for binocular sensor self-calibraiton *Chinese J. Opto-Electronic Engineering* **25** 37–41.
[6] Zhang Z Y 2000 A flexible new technique for camera calibration *IEEE Trans. Pattern Anal. Mach. Intell* **22** 1330–34.