IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 26, 2021, accepted March 12, 2021, date of publication March 17, 2021, date of current version March 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3066379

Phase-Based Calibration Method for a Binocular Vision Sensor

MINGWEI SHAO, PAN WANG, AND YANJUN WANG
School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266000, China
Corresponding author: Mingwei Shao (smw1987@163.com)

This work was supported by the Natural Science Foundation of Shandong Province under Grant ZR2020QF101.

ABSTRACT In this paper, a phase based method to calibrate a binocular vision sensor is presented. In this method, only a surface plate is utilized. A series of phase shifting patterns are projected onto the surface plate. In this case, fundamental matrix can be calculated out from point correspondences on two image planes. As intrinsic parameters of each camera have been obtained according to related calibration method, essential matrix can be worked out. Then, rotation matrix and translation matrix with a coefficient are deduced based on singular value decomposition of the essential matrix. As size of the pattern can be confirmed from a look-up table, the coefficient is determined. So far, the binocular vision sensor is calibrated. In our proposed method, only one planar target without any pattern is necessary, which is with high precision and easy machined. As the calibration method is based on related phase algorithm, the calibration process is not affected by the ambient light which makes our calibration method robust. Experiment results show the precision of our proposed method. When we utilized a laser displacement sensor with a precision of 0.01 mm to determine size of the pattern, root mean square error of our calibration method is 0.103 mm with regard to the measurement area of about 100×80 mm. Moreover, as planar features are widely existing, our proposed method is very suitable for on-site calibration and auto-calibration.

INDEX TERMS Calibration, binocular vision sensor, target without any pattern, measurement.

I. INTRODUCTION

Binocular vision sensor is widely utilized in the field of vision measurement [1]–[3]. Two cameras in a binocular vision sensor are fixed in the process of measurement. Determining the relationship of two cameras is significant which is known as calibration.

Heretofore, calibration methods can be classified into three categories according to utilized target, i.e. three-dimensional (3D) target method, planar target method and one-dimensional (1D) target method.

In 3D target method, a special target is utilized. Enough non-collinear (or non-coplanar) feature points can be obtained from known size of the target [4], [5]. Then relationship between two cameras can be confirmed. As is known, calibration results of 3D target method are accurate enough. Unfortunately, 3D target method is with some inevitable problems, such as mutual occlusion between different planes, difficulty of producing precisely and so on.

In planar target method, features on a planar target are utilized such as corner points, cross points, center points and so on [6]–[8]. The typical one is Zhang's calibration method [9] using a planar target with checkerboard pattern. Relationships of these feature points (corner points) are known exactly beforehand. Once images of these feature points are captured by two cameras simultaneously, coordinates of points under each camera coordinate system can be deduced according to camera imaging model and relative position constraints of features. When these corresponding points are confirmed, relationship between two cameras can be confirmed.

In 1D target method, geometrical feature is normal utilized, such as co-linearity feature, invariance of double cross ratio feature and so on [10]–[12]. Similar with the planar target method, as relative locations (or relations) of features are known exactly, relationship between two cameras in a binocular vision sensor can be confirmed.

In these calibration methods, a special target with known features is required. These features should be machined precisely as calibration results relay on these features absolutely. As is known, the machining process is time-consuming

The associate editor coordinating the review of this manuscript and approving it for publication was Guangcun Shan.

and difficulty. Moreover, as feature points are extracted depending on the contrast between features and background, these calibration methods are affected by the ambient light inevitably.

In this case, a phase based calibration method for a binocular vision is proposed. In this method, only a surface plate is necessary. Vertical and horizontal phase shifting patterns are projected on the plane. Matching points on image planes of two cameras can be extracted according to absolute phase images. Then fundamental matrix can be calculated out. As each camera has been calibrated, i.e. intrinsic parameters of each camera have been obtained, essential matrix can be worked out. Based on singular value decomposition of essential matrix, rotation matrix and translation matrix with a coefficient are obtained. As the distance from the projector to the plane is known, size of the pattern can be confirmed according to related look-up table. Then the coefficient is determined, i.e. the translation matrix is confirmed. In our proposed method, any one target with a special pattern or features is not necessary. Obtained matching points are nearly not affected by the ambient light. Moreover, the proposed method is robust and precise, which is very suitable for on-site calibration and auto-calibration.

II. MULTI-VIEW VISION SENSOR ON SCHEIMPFLUG CONDITIONS

A. IMAGING MODEL OF A SINGLE CAMERA

A single camera is described by a pin-hole model. Imaging model of a camera is illustrated in Fig. 1. $O_C - X_C Y_C Z_C$ is the camera coordinate system (*CCS*) while *o-xy* is the image coordinate system *(ICS)*. Under *CCS*, camera center is at the origin and the optical axis points in the positive Z_C direction. A spatial point P is projected onto the plane *o-xy*, referred to as the real image plane under *CCS*. f_0 is the effective focal length (*EFL*). Supposing the image point $p = (x, y, 1)^T$ is projection of the spatial point $P = (X, Y, Z)^T$ on the image plane. Under the idealized pinhole imaging model, i.e. the ideal model of the camera, P, p and the camera center O are collinear [13].

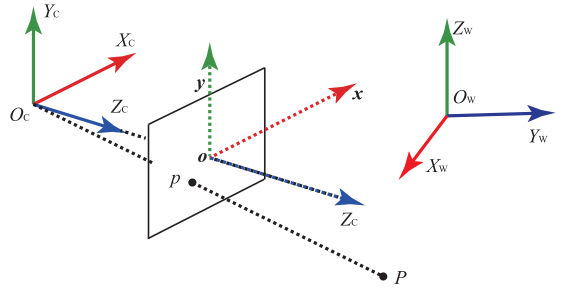


FIGURE 1. Imaging model of a single camera.

When we define the world coordinate system (*WCS*) as O_W-$X_W Y_W Z_W$, coordinate of a spatial point P is defined as (X_W, Y_W, Z_W).Then we can obtain an expression as

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (1)$$

where $R_W^C = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ is rotation matrix from *WCS* to *CCS* and $T_W^C = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$ is the corresponding translation matrix. Eq.(1) can be rewritten as [9], [14]

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (2)$$

where (u, v) is the projection of point P on image plane in pixel unit, and matrix $K_u = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ is the intrinsic parameter matrix, where (u_0, v_0) is the principal point, $f_x = f/dx, f_y = f/dy$. Practically, the radial distortion and the tangential distortion of the lens are inevitable. Normally, the radial distortion is obvious and cannot be neglected. When only considering the radial distortion, we have following equations:

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

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

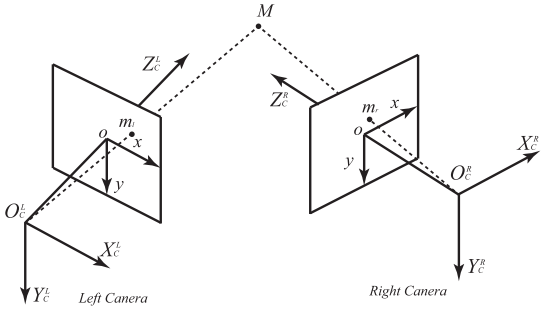


FIGURE 2. Measurement model of a binocular vision sensor.

B. MEASUREMENT MODEL OF A BINOCULAR VISION SENSOR

A binocular vision sensor is consist of two cameras (as illustrated in Fig. 2). Define the left camera coordinate system (*LCS*) as $O_C^L - X_C^L Y_C^L Z_C^L$, while the right camera coordinate system (*RCS*) as $O_C^R - X_C^R Y_C^R Z_C^R$. Traditionally, *LCS* is defined

as the global coordinate system (*GCS*). Measurement model of the binocular vision sensor can be formulated as

$$\begin{cases} s_l m_l = \mathbf{K}_l[\mathbf{I}|\mathbf{0}]\mathbf{M} \\ s_r m_r = \mathbf{K}_r[\mathbf{R}|\mathbf{T}]\mathbf{M} \end{cases} \tag{4}$$

where s_n is a coefficient, $\mathbf{I}$ is the unit matrix, $\mathbf{K}_n$ is the intrinsic parameter matrix of camera n, $\mathbf{R}$ is the rotation matrix while $\mathbf{T}$ is the translation matrix from *GCS (LCS)* to *RCS*. $\mathbf{M}$ is the homogeneous coordinate of a spatial point under *GCS*, while m_{n} is its corresponding image coordinate under *ICS* of camera n.

C. IMPROVED BINOCULAR VISION SENSOR

In the measurement of a binocular vision sensor, feature points should be extracted. For the purpose of extracting more feature points, additive feature pattern is projected onto the object target, such as laser stripes, phase-shifting pattern and so on [15], [16] (two typical improved binocular vision sensors are illustrated in Fig. 3).

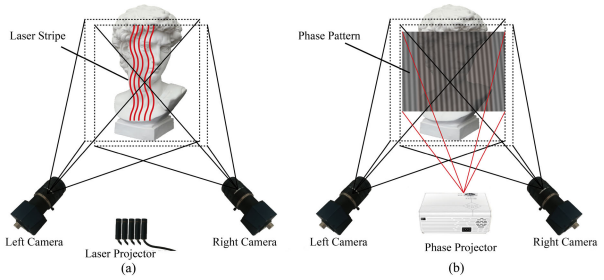


FIGURE 3. Two typical improved binocular vision sensors.

In Fig. 3(a), laser stripes are projected onto surface of the measured target. Images are captured by two cameras of the binocular vision sensor simultaneously. In this sensor, feature points are centerline points (or edge points) of each laser stripe, which are obtained by related extraction algorithms such as C.Steger's algorithm [17]. Moreover, as is known, in order to capture a clear feature image, the contrast between light strips and background should be sharp. Optical filter is normal utilized to reduce effect of the ambient light. In this case, an improved binocular vision sensor based on phase shifting algorithm is present as illustrated in Fig. 3(b). Effect of the ambient light can be removed based on the mathematic description of the phase shifting algorithm.

III. PRINCIPLE

A. PHASE SHIFTING ALGORITHM

Phase shifting algorithm is widely utilized in the binocular vision sensor as its high robustness and denseness. In general, a projector is fixed with these two cameras. Series of fringe patterns are projected onto the measured object. Corresponding images are captured by two cameras separately. Absolute phase can be calculated out according to related phase shifting algorithm [18]. As absolute phase of matching points are identical, one additive constraint can be given. Combined with the epipolar constraint, matching points can be confirmed precisely.

Take four step phase shifting algorithm as an example. Four fringe images can be mathematically described as below:

$$\begin{cases} I_1(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + 0] \\ I_2(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + \pi/2] \\ I_3(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + \pi] \\ I_4(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + 3\pi/2] \end{cases} \tag{5}$$

where $\phi(x,y)$ denotes the wrapped phase, $I'(x,y)$ denotes the average intensity of ambient light, while $I''(x,y)$ denotes the intensity modulation. As average intensity is estimated as $I'(x,y)$, phase shifting algorithm can remove effect of the ambient light. The wrapped phase can be solved according to the following equation:

$$\phi(x,y) = \tan^{-1}\left(\frac{I_4 - I_2}{I_1 - I_3}\right) \tag{6}$$

where $\tan^{-1}$ in Eq.(6) denotes *arctan* function. The calculated wrapped phase is ranging from $-\pi$ to $+\pi$ with 2π phase discontinuities. When feature matching is conducted, the absolute phase should be recovered. Heretofore, there are many recovery algorithms, such as space phase unwrapping algorithm, temporal phase unwrapping algorithm, etc. In our calibration method, multi-frequency phase shifting algorithm is used to recover the absolute phase. In this algorithm, the absolute phase of two series of fringes with a close frequency can be calculated according to the following equation:

$$\begin{cases} \Phi_h = \phi_h + 2k_h\pi \\ \Phi_l = \phi_l + 2k_l\pi \\ \phi_{eq} = \phi_h - \phi_l \end{cases} \tag{7}$$

where ϕ_h and ϕ_l are wrapped phases of high and low frequency fringes, while Φ_h and Φ_l are their corresponding absolute phase respectively. The coefficient k_n is defined as

$$k_n = Round\left(\frac{(\lambda_{eq}/\lambda_h)\,\Phi_{eq} - \phi_h}{2\pi}\right), n = h, l \tag{8}$$

where $Round()$ denotes the closest integer value obtained. λ_h and λ_l are the frequency of high and low frequency fringe respectively. And frequency of additional phase is defined as $\lambda_{eq} = \frac{\lambda_h\lambda_l}{\lambda_h - \lambda_l}$.

B. FUNDAMENTAL MATRIX

As is known, projections of one spatial point on two images is with the same phase. For the purpose of determining matching points [19], vertical and horizontal phase shifting patterns are projected onto a plane. Then absolute phase in two different directions can be worked out according to Eq.(7). Feature points can be selected as projections of the same point are with the same absolute phase in two directions (as illustrated in Fig. 4).

Define projection of the spatial point on image plane of camera one as m_l, its corresponding projection on image

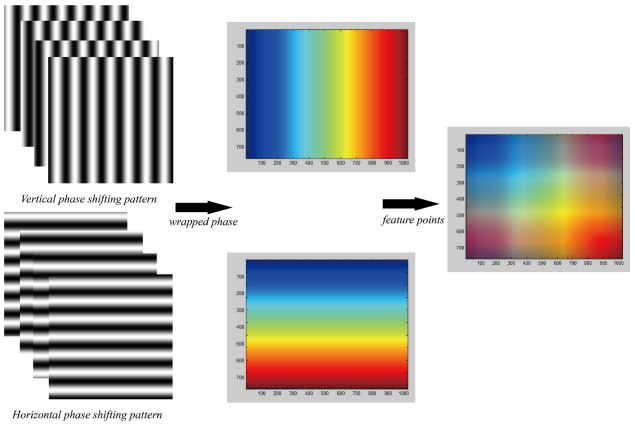


FIGURE 4. Feature points detection in this paper.

plane of camera two as m_r, then we can get the relationship as

$$m_l^{\mathrm{T}}\mathbf{F}m_r = 0 \tag{9}$$

where $\mathbf{F}$ is the fundamental matrix. When we put the plane on several different positions, we can get enough matching point correspondences. In this case, the fundamental matrix can be calculated out according to related algorithm, such as the normalized eight-point algorithm, the algebraic minimization algorithm, the distance minimization algorithm and so on.

C. ROTATION MATRIX

Define intrinsic parameter matrix of left camera as K_l, intrinsic parameter matrix of right camera as K_r, essential matrix can be denoted as

$$\mathbf{E} = \mathbf{K}_r^{\mathrm{T}}\mathbf{F}\mathbf{K}_l \tag{10}$$

Furthermore, essential matrix $\mathbf{E}$ can be decomposed by

$$\mathbf{E} = Udiag(\sigma_1, \sigma_2, 0)V^{\mathrm{T}} \tag{11}$$

where *diag* denotes the diagonal matrix, singular values satisfy $\sigma_1 = \sigma_2 = \sigma$, i.e. essential matrix has two same singular values and one zero singular value. Then rotation matrix R and translation matrix T (with a scale factor κ) can be deduced based on singular value decomposition (*SVD*) of matrix $\mathbf{E}$. Four possible solutions are listed as below [13]:

$$\begin{cases} A : R = UZV^{\mathrm{T}}, T = \kappa u_3 \\ B : R = UZ^{\mathrm{T}}V^{\mathrm{T}}, T = -\kappa u_3 \\ C : R = UZV^{\mathrm{T}}, T = -\kappa u_3 \\ D : R = UZ^{\mathrm{T}}V^{\mathrm{T}}, T = \kappa u_3 \end{cases} \tag{12}$$

matrix Z is defined as $Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, u_3 is the vector with respect to the zero singular value, namely the last column of matrix U. In Eq.(12), A and C, similar with B and D are baseline reversed, while C and D, similar with B and D are rotated 180 degree about the base line. As is known, all points must locate on the positive z-coordinate directions of camera one and camera two simultaneously. In this case, only one solution can be selected from these four possible solutions, i.e. rotation matrix and translation matrix with a coefficient κ are determined.

D. TRANSLATION MATRIX

Now we have obtained the translation matrix with a coefficient κ. If we can calculate out the coefficient κ, the translation matrix is determined. As is known, a projector has functions of autofocus and trapezoidal correction. In this case, the projected pattern can be kept as a rectangle. If we can obtain the distance from the projector to the target plane, size of the pattern can be obtained from a look-up table.

1) LOOK-UP TABLE

Normally, the look-up table is provided by the manufacturer. In other words, if we cannot get the look-up table from the manufacturer, we can get the table with the help of a laser displacement sensor and a planar target easily. This procedure is named as projector calibration which is detailed as follow:

(1) A laser displacement sensor is fixed with the projector but not request to be perpendicular to image plane of the projector. Namely, relationship between the projector and the laser displacement sensor is unchangeable.

(2) A target which is with a checkerboard pattern is utilized. Length of each square is known exactly.

(3) Then a rectangle pattern (or a pattern contains a rectangle) is projected onto the target plane. Homography matrix from image plane to target plane can be calculated out. When four corner points are detected, size of the pattern can be confirmed. Meanwhile, the distance from the laser displacement sensor to the target plane can be read from the sensor easily.

(4) When the target is located in enough different positions, the look-up table can be determined.

This procedure is time-consuming but not complex.

2) TRANSLATION MATRIX

When we get the look-up table, precise size of the pattern can be obtained from the distance. Without loss of generality, four side lengths are assumed to be d_1, d_2, d_3 and d_4. When we project the same pattern which we used to calculate the look-up table onto a target plane, we can confirm size of the projected pattern from reading of the laser displacement sensor. Moreover, as description in *Part A* of this *Section*, the phase shifting algorithm is not affected by the ambient light. We can design special patterns to confirm the size according to related multi-frequency phase shifting algorithm. An example is given as illustrated in Fig. 5.

As illustrated in Fig. 5(b), four sides of the pattern are *AB, BD, AC* and *CD* separately. When the wrapped phase or the absolute phase is obtained, each side can be fitted according to related linear fitting algorithm. The distance from point A to B can be defined as d_1, while the distance from point C to D is d_2, the distance from point A to C is d_3, the distance

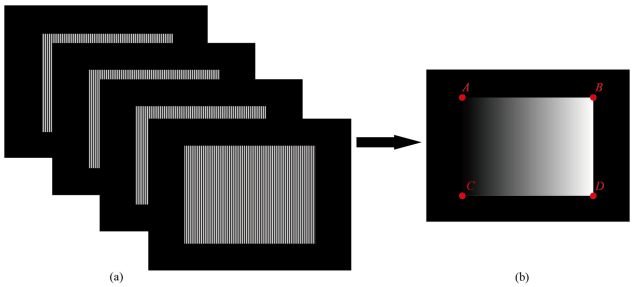


FIGURE 5. Special patterns to confirm the size according to multi-frequency phase shifting algorithm.

from point B to D is d_4. Relationship can be described as

$$\begin{cases} |AB| = d_1 \\ |CD| = d_2 \\ |AC| = d_3 \\ |BD| = d_4 \end{cases} \tag{13}$$

According to Eq. (4), when defined $K_l = \begin{bmatrix} f_x^l & 0 & u_0^l \\ 0 & f_y^l & v_0^l \\ 0 & 0 & 1 \end{bmatrix}$, $K_r = \begin{bmatrix} f_x^r & 0 & u_0^r \\ 0 & f_y^r & v_0^r \\ 0 & 0 & 1 \end{bmatrix}$, we can get the relationship as (14), shown at the bottom of the page, where the spatial point under GCS is $P = (X, Y, Z)^{\mathrm{T}}$, while its projection on image plane of left camera is $(u_l, v_l)^{\mathrm{T}}$, and its projection on image plane of right camera is $(u_r, v_r)^{\mathrm{T}}$. Then point P can be worked out by least square method if the rotation matrix and the translation matrix are confirmed. From *Part C* of *Section* three, the rotation matrix is confirmed completely, while the translation matrix is with a coefficient κ. Combining Eq.(13) with Eq.(14), as real distances are confirmed, the coefficient can be calculated out. Namely, calibration of a binocular vision sensor is finished.

E. OPTIMIZATION

1) OPTIMIZATION OF TRANSLATION MATRIX

When the target is located on several (n) different positions, we can obtain several groups of feature points. Then one objective optimization function can be assumed as

$$f_1(\kappa) = \sum_{i=1}^{n} [(|AB| - d_1)^2 + (|CD| - d_2)^2 + (|AC| - d_3)^2 + (|BD| - d_4)^2] \tag{15}$$

When using a square pattern, we can get the relationship that $d_1 = d_2, d_3 = d_4$. In this case, an additive constraint can be given as

$$f_2(\kappa) = \sum_{i=1}^{n} \left[(|AB| - |CD|)^2 + (|AD| - |BC|)^2\right] \tag{16}$$

By minimizing these two optimization functions, optimal coefficient κ, i.e. the optimal translation matrix is obtained.

2) OPTIMIZATION OF ROTATION MATRIX

Define a line in the space under GCS as $\mathbf{L}$. If the projection line of $\mathbf{L}$ on image plane of left camera is defined as $l_l = (a_l, b_l, c_l)$, its corresponding projection line on image plane of right camera is defined as $l_r = (a_r, b_r, c_r)$, line $\mathbf{L}$ can be denoted as

$$L = \begin{bmatrix} l_l^{\mathrm{T}} \mathbf{P}_l \\ l_r^{\mathrm{T}} \mathbf{P}_r \end{bmatrix} \tag{17}$$

where

$$\mathbf{P}_l = \mathbf{K}_l [\mathbf{I}|\mathbf{0}] = \begin{bmatrix} f_x^1 & 0 & u_0^1 & 0 \\ 0 & f_y^1 & v_0^1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{18}$$

$$\mathbf{P}_r = \mathbf{K}_r [R|T] = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ m_5 & m_6 & m_7 & m_8 \\ m_9 & m_{10} & m_{11} & m_{12} \end{bmatrix} \tag{19}$$

If two space lines (L_l and L_r) are coplanar, we can get the relation as

$$(\boldsymbol{L}_l | \boldsymbol{L}_r) = 0 \tag{20}$$

Combine with Eq.(17), we can get the definition as

$$(\boldsymbol{L}_l | \boldsymbol{L}_r) = det \left[l_l^{\mathrm{T}} \mathbf{P}_l, l_r^{\mathrm{T}} \mathbf{P}_r, \hat{l}_l^{\mathrm{T}} \mathbf{P}_l, \hat{l}_r^{\mathrm{T}} \mathbf{P}_r \right] = 0 \tag{21}$$

where *det* denotes determinant of the matrix. In the process of calibration, we can get several coplanar lines, such as side lines of the projected pattern, the line with same phase and so on. In this case, we can define an objective function to optimize our calibration result:

$$F_1(R, t) = \sum_{i=1}^{n-1} [(\boldsymbol{L}_i | \boldsymbol{L}_{i+1})] \tag{22}$$

Certainly, the rotation matrix should also satisfies its inherent properties. Then we can get another objective function as

$$F_2(R) = sum(\left| RR^{\mathrm{T}} - \mathbf{I} \right|) \tag{23}$$

By minimizing Eq.(22) and Eq.(23), we can get an optimal rotation matrix.

$$\begin{bmatrix} f_x^l & 0 & u_0^l - u_l \\ 0 & f_y^l & v_0^l - v_l \\ f_x^r r_{11} + (u_0^r - u_r) r_{31} & f_x^r r_{12} + (u_0^r - u_r) r_{32} & f_x^r r_{13} + (u_0^r - u_r) r_{33} \\ f_y^r r_{21} + (v_0^r - v_r) r_{31} & f_y^r r_{22} + (v_0^r - v_r) r_{32} & f_y^r r_{23} + (v_0^r - v_r) r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ (u_r - u_0^r) t_z - f_x^r t_x \\ (v_r - v_0^r) t_z - f_y^r t_y \end{bmatrix} \tag{14}$$

IV. EXPERIMENTS AND DISCUSSION

In this section, we built a binocular vision sensor with a phase projector. The binocular vision sensor is consisted of two cameras (Daheng Imaging MER-231-41GM). Each camera is with a maximum resolution of 1920 × 1200 pixels. The phase projector which is used to generate a phase pattern is GX1600 with a resolution of 1600 × 1200 pixels.

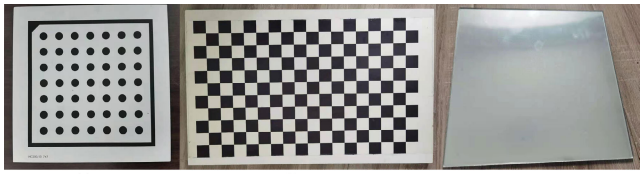
A. CAMERA CALIBRATION

In our experiment, each camera is calibrated by Zhang's calibration method [9], [14]. The planar target to calibrate each camera is illustrated in Fig. 6(a), which is with a precision of 0.01 mm. Obtained intrinsic parameters are listed in Table 1.

TABLE 1. Intrinsic parameters of two cameras.

No.	f_x	f_y	u_0	v_0	kc_1	kc_2
1	1403.327	1408.318	956.417	614.874	-0.095	0.071
2	1414.263	1391.802	956.313	604.829	-0.113	0.189

In Table 1, f_x and f_y represent the scale factor in x-coordinate direction and y-coordinate direction respectively. $(u_0, v_0)^T$ are coordinates of the principal point in terms of pixel dimensions. Planar targets used in our experiments are illustrated in Fig. 6.



(a) (b) (c)

FIGURE 6. (a) The planar target used to calibrate each camera, (b) the target used to calibrate the projector and evaluation, (c) the target used to calibrate the binocular vision sensor.

B. PROJECTOR CALIBRATION

In this section, the phase projector is calibrated according to the procedure detailed in *Section three.* A planar target with checkerboard pattern is utilized (as illustrated in Fig. 6(b)). As the size of the square is known exactly (20 mm with precision of 0.1 mm), homography matrix H from the image plane to the target plane is worked out. When a square pattern (as illustrated in Fig. 7(a)) is projected onto the target plane, four corner points can be extracted precisely according to related sub-pixel corner extraction algorithm [17]. According to properties of homography matrix, four corner points under *CCS* can be calculated out, i.e. size of the projected pattern can be confirmed (as illustrated in Fig. 7(b)).

Moreover, a laser displacement sensor (ZSY ZLDS112-200) is fixed with the phase projector. Measurement range of

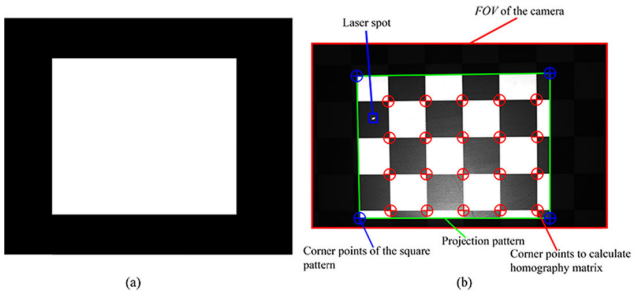


FIGURE 7. (a) Original image of the square pattern, (b) image of the projected square pattern on the target captured by one camera.

TABLE 2. Part results of the obtained look-up table (mm).

No	*dis*	d_1	d_2	d_3	d_4
1	201.02	77.91	78.03	61.08	60.59
2	203.11	79.07	79.24	62.33	62.51
3	206.72	81.87	81.96	63.90	64.02
4	209.25	83.12	83.29	66.15	65.97
5	211.93	85.48	85.77	67.91	68.23
6	216.50	86.73	87.03	69.72	70.10
7	219.32	88.33	88.57	71.02	71.35
8	222.01	90.11	90.37	73.22	73.49
9	227.39	92.15	93.01	75.31	75.53
10	230.07	94.00	94.26	77.93	78.02

the displacement sensor is 200 mm in the distance of 150 mm (i.e. from 50 mm to 250 mm) with a precision of 0.01 mm. In this case, the distance from the laser displacement sensor to the target plane (*dis*) can be read out. When the target is located in enough different position, we can obtain the relationship between the distance and size of the pattern, namely the look-up table as mentioned in *Part D* of *Section three*.

Part results of the look-up table are listed as Table. 1. Without loss of generality, define four sides length of the pattern as d_1, d_2, d_3 and d_4 respectively.

C. CALIBRATION FOR THE BINOCULAR VISION SENSOR

In this paper, three-frequency phase shifting algorithm is utilized to get the absolute phase. Three frequencies are 70, 64, and 59 respectively. Phase images are generated by the projector. Original phase images and ideal absolute phases are illustrated in Fig. 8.

Phase images are projected onto the target plane (as illustrated in Fig. 6(c)). The planar target is not with any special pattern and its machining precision is 0.01 mm. Images of the phase shifting patterns are captured by two cameras as illustrated in Fig. 9.

Based on the calibration procedure detailed above, we can get calibration results of the binocular vision sensor,

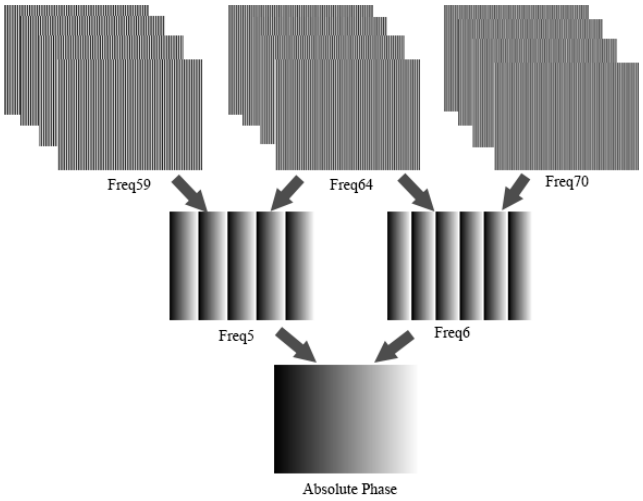


FIGURE 8. Original phase images and ideal absolute phases.

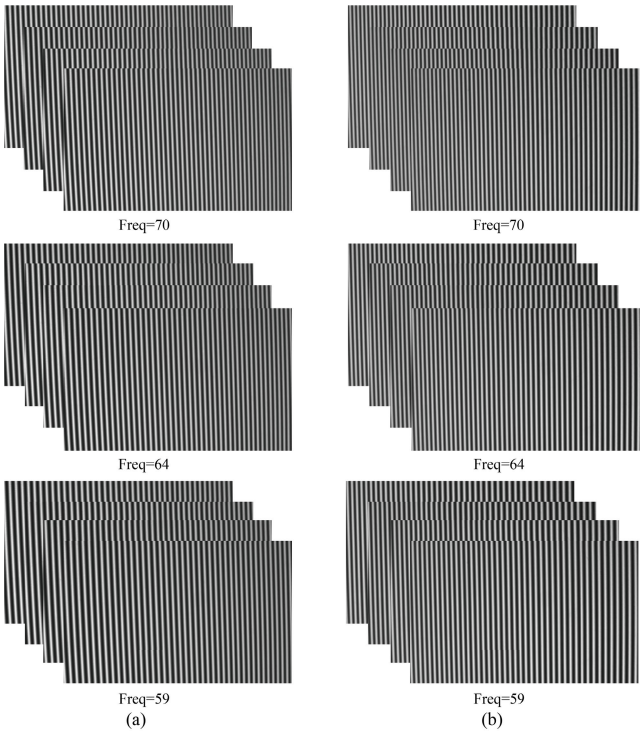


FIGURE 9. (a) Series of phase images captured by left camera, (b) series of phase images captured by right camera.

i.e. rotation matrix and translation matrix as following:

$$R = \begin{bmatrix} 0.731 & -0.015 & 0.682 \\ 0.007 & 0.999 & 0.015 \\ -0.682 & -0.007 & 0.731 \end{bmatrix} \quad (24)$$

$$T = [-437.69, 3.410, 198.59]^{\mathrm{T}} \quad (25)$$

D. EVALUATION

The planar target as illustrated in Fig. 6(b) is utilized to evaluate our calibration method. The distance between each two adjacent feature points (i.e. corner points of checkerboard) is known as D_{Real} exactly. Then feature points can be reconstructed based on our calibration results. Feature matching is conducted according to related algorithm in our previous work [8]. In this case, the measurement distance (D_{Mea}) can also be calculated out. All measurement values are compared with their corresponding true value (20.00 mm). When the planar target is moved into more different positions randomly, we can obtain enough distances to evaluate our calibration results (as illustrated in Fig. 7). Part of evaluation results are listed in Table 3.

TABLE 3. Evaluation of our calibration results.

	Position 1	Position 2	Position3
D_{Mea} (mm)	20.089	19.885	19.893
	19.919	19.706	20.144
	20.033	19.925	20.137
	19.829	19.990	19.976
	20.032	20.031	19.914
	19.997	19.984	20.063
	20.109	20.111	19.914
	20.008	19.879	19.889
	19.999	20.153	19.923
	20.037	19.977	20.112
RMS error (mm)		0.103	

As listed in Table. 3, Root mean square (*RMS*) error of measurement is about 0.103 mm. As the measurement area of our binocular vision sensor is about 100×80 mm, the relative measurement precision is about 1.03% with regards to the measurement area.

E. COMPARISON WITH TRADITIONAL CALIBRATION METHOD

Heretofore, the traditional calibration method is the planar target based calibration method for a binocular vision sensor. Normally, a planar target with a special pattern is located on the overlapping view fields of two cameras. For comparison, we calibrate the binocular vision sensor using a planar target as illustrated in Fig. 5(a). The obtained rotation matrix and translation matrix are:

$$R = \begin{bmatrix} 0.733 & -0.015 & 0.681 \\ 0.007 & 0.999 & 0.015 \\ -0.681 & -0.007 & 0.733 \end{bmatrix} \quad (26)$$

$$T = [-436.87, 3.356, 197.72]^{\mathrm{T}} \quad (27)$$

For the purpose of comparison, we evaluate above calibration results based on the procedure detailed in section 4.4. Evaluation results in three positions are listed in Table 4.

As listed in Table. 4, *RMS* error of the calibration results based on traditional calibration method is 0.114 mm. It is with the similar precision as the calibration results based on our proposed method.

In the tradition calibration method, the limitation is the precision of the target which is with precision of 0.01 mm (of course, the target used for evaluation is with a precision of 0.1 mm). In our proposed calibration method, rotation

TABLE 4. Evaluation of our calibration results.

	Position 1	Position 2	Position3
D_{Mea} (mm)	19.885	20.010	20.072
	20.259	19.933	20.019
	19.992	19.807	19.956
	19.821	20.084	19.911
	20.010	19.946	20.030
	19.940	20.049	20.074
	20.171	19.981	19.786
	19.916	20.135	19.893
	20.096	20.012	20.144
	19.804	19.980	19.879
RMS error (mm)		0.114	

matrix is decomposed from the essential matrix. So the limitation for determining rotation matrix is the planar target. For the translation matrix, the limitation is the precision of the laser displacement sensor, which confirm the coefficient κ.

As the precision of limitation is with the same precision, traditional calibration results and our calibration results are with a similar precision.

F. APPLICATION IN SELF-CALIBRATION

As is known, planar features are widely existing in many scenes, such as ground plane, floor plane, wall plane and other plane in some industrial scenes (as illustrated in Fig. 10). In this case, we can calibrate a binocular vision sensor using our proposed method with the help of those planar features.

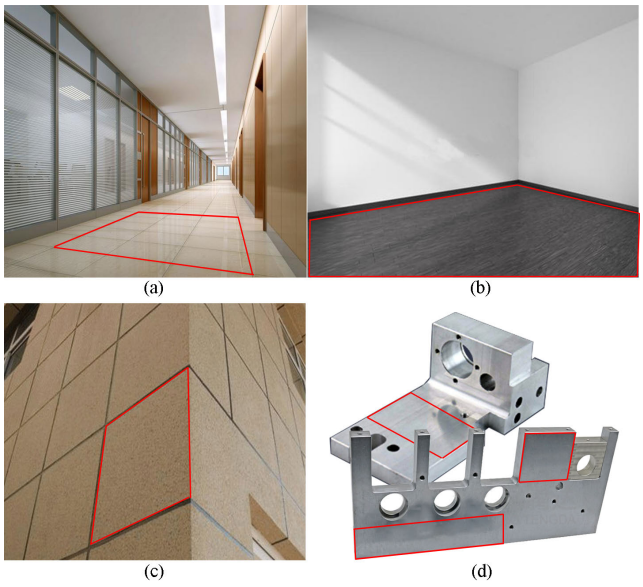


FIGURE 10. (a) Ground plane, (b) floor plane, (c) wall plane and (d) other planes in industrial scenes.

V. CONCLUSION

A phase based calibration method for a binocular vision sensor is proposed. With the help of only one surface plate, a binocular vision sensor can be calibrated. As phase shifting patterns are projected onto the surface plate, matching points on image planes of two cameras can be confirmed. In this case, the fundamental matrix can be worked out easily. Then essential matrix is determined as the intrinsic parameters of each camera have been obtained. According to the singular value decomposition, rotation matrix and translation matrix with a coefficient are worked out from the essential matrix. As the size of the pattern is known from the look-up table, the coefficient can be confirmed. In this proposed calibration method, any one target with a special pattern is not necessary. Moreover, the calibration process is not affected by the ambient light according to the phase shifting algorithm. Experiments show the robustness and precision of this method, which is suitable for on-site calibration and auto-calibration.

ACKNOWLEDGMENT

The authors would like to thank the Electronic Information Laboratory for the use of their equipment in Qingdao University of Technology.

REFERENCES

[1] F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," *Opt. Eng.*, vol. 39, no. 1, pp. 10–22, 2000.
[2] P. Graebling, A. Lallement, D. Zhou, and E. Hirsch, "Optical high precision three-dimensional vision-based quality control of manufactured parts by use of synthetic images and knowledge for image-data evaluation and interpretation," *Appl. Opt.*, vol. 41, pp. 2627–2643, May 2002.
[3] M. Shao, P. Wang, and J. Wang, "Improved sensors based on scheimpflug conditions and multi-focal constraints," *IEEE Access*, vol. 8, pp. 161276–161287, 2020.
[4] J. Yang, Z. Jia, W. Liu, C. Fan, P. Xu, F. Wang, and Y. Liu, "Precision calibration method for binocular vision measurement systems based on arbitrary translations and 3D-connection information," *Meas. Sci. Technol.*, vol. 27, no. 10, Oct. 2016, Art. no. 105009.
[5] R. Lu and M. Shao, "Sphere-based calibration method for trinocular vision sensor," *Opt. Lasers Eng.*, vol. 90, pp. 119–127, Mar. 2017.
[6] Z. Z. Wei and X. K. Liu, "Vanishing feature constraints calibration method for binocular vision sensor," *Opt. Exp.* vol. 23, no. 15, pp. 18897–18914, 2015.
[7] Y. Cui, F. Zhou, Y. Wang, L. Liu, and H. Gao, "Precise calibration of binocular vision system used for vision measurement," *Opt. Exp.*, vol. 22, no. 8, pp. 9134–9149, Apr. 2014.
[8] M. Shao and M. Hu, "Parallel feature based calibration method for a trinocular vision sensor," *Opt. Exp.*, vol. 28, no. 14, p. 20573, Jul. 2020.
[9] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
[10] J. Sun, Q. Liu, Z. Liu, and G. Zhang, "A calibration method for stereo vision sensor with large FOV based on 1D targets," *Opt. Lasers Eng.*, vol. 49, no. 11, pp. 1245–1250, Nov. 2011.
[11] Z. Liu, G. Zhang, Z. Wei, and J. Sun, "Novel calibration method for non-overlapping multiple vision sensors based on 1D target," *Opt. Lasers Eng.*, vol. 49, no. 4, pp. 570–577, Apr. 2011.
[12] M. W. Shao, "Calibration methods for a camera with a tilted lens and a three-dimensional laser scanner in the Scheimpug condition," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 37, no. 7, pp. 1076–1082, 2020.
[13] R. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ., 2003, Ch. 12–15.
[14] J. Y. Bouguet. *Camera Calibration Toolbox for MATLAB*. Accessed: 2015. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/
[15] Y. T. An, J. S. Hyun, and S. Zhang, "Pixel-wise absolute phase unwrapping using geometric constraints of structured light system," *Opt. Exp.*, vol. 24, no. 16, pp. 18445–18459, 2016.
[16] H. Hu, J. Gao, H. Zhou, L. Zhang, H. Deng, X. Chen, and Y. He, "A combined binary defocusing technique with multi-frequency phase error compensation in 3D shape measurement," *Opt. Lasers Eng.*, vol. 124, Jan. 2020, Art. no. 105806.

[17] C. Steger, "Unbiased extraction of lines with parabolic and Gaussian profiles," *Comput. Vis. Image Understand.*, vol. 117, no. 2, pp. 97–112, Feb. 2013.
[18] S. Zhang, "Composite phase-shifting algorithm for absolute phase measurement," *Opt. Lasers Eng.*, vol. 50, no. 11, pp. 1538–1541, Nov. 2012.
[19] Y. Li, J. Zhang, Y. Zhong, and M. Wang, "An efficient stereo matching based on fragment matching," *Vis. Comput.*, vol. 35, no. 2, pp. 257–269, Feb. 2019.

MINGWEI SHAO received the Ph.D. degree from Beihang University, Beijing, China, in 2016. He completed his postdoctoral research with the Ocean University of China, Qingdao, China, in 2019. He is currently a Lecturer with the Qingdao University of Technology. His current research interests include computer vision, machine learning, calibration, and 3D reconstruction.

PAN WANG received the Ph.D. degree in control theory and control engineering from Southeast University, Nanjing, China, in 2019. She is currently a Lecturer with the Qingdao University of Technology. Her current research interests include state feedback, output feedback, sampling control, and tracking control of nonlinear control systems.

YANJUN WANG received the Ph.D. degree from the Harbin Institute of Technology, in 2018. He is currently an Associate Professor with the Qingdao University of Technology. His research interests include complex multiphase flow logging technology, well multiphase flow metering technology, and intelligent water injection/oil production technology.