

## Research Article

# Research on Calibration Method of Binocular Vision System Based on Neural Network

Hao Zhu <sup>1,2</sup>, Mulan Wang <sup>2</sup>, and Weiye Xu <sup>3</sup>

<sup>1</sup>School of Information and Communication, Nanjing Institute of Technology, Nanjing 211167, China

<sup>2</sup>Jiangsu Key Laboratory of Advanced Numerical Control Technology, Nanjing 211167, China

<sup>3</sup>Department of Informatics, University of Leicester, Leicester, LE1 7RH, UK

Correspondence should be addressed to Hao Zhu; zhuhaon@njit.edu.cn

Received 29 January 2021; Revised 15 April 2021; Accepted 14 June 2021; Published 21 June 2021

Academic Editor: Zhaoqing Pan

Copyright © 2021 Hao Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In binocular vision inspection system, the calibration of detection equipment is the basis to ensure the subsequent detection accuracy. The current calibration methods have the disadvantages of complex calculation, low precision, and poor operability. In order to solve the above problems, the calibration method of binocular camera, the correction method of lens distortion, and the calibration method of projector in the binocular vision system based on surface structured light are studied in this paper. For lens distortion correction, on the basis of analyzing the traditional correction methods, a distortion correction method based on radial basis function neural network is proposed. Using the excellent nonlinear mapping ability of RBF neural network, the distortion correction models of different lenses can be obtained quickly. It overcomes the defect that the traditional correction model cannot adjust adaptively with the type of lens. The experimental results show that the accuracy of the method can meet the requirements of system calibration.

## 1. Introduction

With the development of modern electronic technology, the application of 3D detection in the machining field is more and more mature. At present, the common 3D detection methods mainly include contact and noncontact. In the traditional reverse engineering, the common method of 3D object detection is the contact measurement technology represented by coordinate measuring machine (CMM). The advantage of this method is that it is easy to operate. However, it has a large error for the soft measured target. Besides, the cost of special large-scale CMM is also very high [1]. With the development of computer technology, the application of machine vision and noncontact measurement technology in the mechanical manufacturing system has gradually become a research hotspot.

Structured light detection is a representative method of noncontact measurement technology [2]. In the detection process, the projector projects structured light with specific rules to the target surface. The stripe of structured light

changes with the depth of the target surface, resulting in distortion. Due to the different positions of the cameras on both sides of the projector, the distorted images captured by the cameras are also different. The distorted structured light image contains the depth information of the measured target surface and the relative position of the projector and camera. By analyzing and calculating the distortion characteristics, the target depth information can be obtained, and the target 3D coordinates can be achieved. In the process of calculation, in order to determine the relationship between the 3D geometric position of a point on the surface of a space object and its corresponding point in the image, it is necessary to establish the geometric model of the camera and projector. The parameters of these geometric models are the parameters of the camera and projector, including internal parameters, external parameters, and distortion parameters. In most cases, these parameters can only be obtained by experiment and calculation. This process of solving parameters is called system calibration. The accuracy of system calibration directly determines the accuracy of subsequent

measurement and calculation [3–5]. Therefore, it is very important to study the high-precision and high-efficiency system calibration method for the 3D detection system.

## 2. Calibration Principle of Binocular Structured Light System

For a monocular vision system with only one projector and one camera, an equivalent camera can be created by rigid rotation and translation of the projector and camera. However, the premise of the transformation is that the internal parameters of the projector and camera are the same. In engineering practice, the above conditions are generally difficult to meet. Therefore, a binocular vision system composed of one projector and two cameras is usually used for 3D reconstruction [5, 6]. Generally speaking, the projector can be virtual as a pinhole imager, and the camera can be virtual as a linear camera. When the internal parameters of two cameras are the same and the optical center is in the same horizontal plane, the image height is the same. The corresponding points can be determined by searching for feature points at the same height. Therefore, when building a binocular vision system, two cameras of the same model can be selected and placed on a horizontal pan tilt. The projector is located between the two cameras. Two cameras are divided into two angles of view and simultaneously collect the projection image projected by the projector on the 3D target.

The calibration process of the structured light measurement system is the process of solving the functional relationship among the 3D coordinates of the measured point in space, the image information collected by the camera, and the structured light information. The parameters of this function include camera parameters, projector parameters, and the transformation relationship between camera coordinate system and world coordinate system. The calibration of the binocular structured light system includes camera calibration, relative position calculation of two cameras, camera distortion correction, projector calibration, and relative position calculation between the projector and camera [7].

## 3. Calibration of Camera and Projector

*3.1. Principle of Camera Imaging.* There are two cameras in the binocular vision detection system to obtain the target data. These cameras have their own positions and parameters. The final result of the reconstruction depends on the relationship between the spatial position of the target and the corresponding image points in the camera, that is, the set model and parameters of the camera. Therefore, it is necessary to model the camera and obtain the relevant parameters for 3D reconstruction of the target. Binocular vision detection is to calculate the camera coordinates corresponding to each point according to the coordinates of each point in the distorted structured light plane image obtained by the camera and then obtain the 3D world coordinates corresponding to each point on the target surface [8, 9].

As shown in Figure 1, let the upper left corner of the plane image coordinate be the coordinate origin  $O_0$ , and a known point in the image is  $D(u, v)$ .  $u$  and  $v$  are the number of pixels in the horizontal and vertical directions, respectively. The image coordinate system ( $O_1 - xy$ ) was established. Its origin position is  $(u_0, v_0)$ . If the image coordinates corresponding to point  $D(u, v)$  are  $(x, y)$ , then the corresponding relationship between  $(u, v)$  and  $(x, y)$  is shown in the following equation:

$$\begin{cases} u = \left(\frac{x}{dx}\right) + u_0, \\ v = \left(\frac{y}{dy}\right) + v_0. \end{cases} \quad (1)$$

Its homogeneous form is shown in the following equation where  $dx$  and  $dy$  are the size of each pixel:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (2)$$

The camera coordinate system ( $O_2 - x_c y_c z_c$ ) and world coordinate system ( $O_3 - x_w y_w z_w$ ) are established. According to the imaging principle of the pinhole camera, the relationship between the image coordinate system and the two is shown in Figure 2 where  $O_2$  is the optical center of the camera and the line  $O_1 O_2$  is the focal length  $f$  of the camera.

As can be seen from Figure 2, the world coordinate system can be obtained by rotation and translation of the camera coordinate system. Let the rotation matrix be  $R$  and the translation matrix be  $t$ . The relationship between the world coordinate system and the camera coordinate system can be obtained as shown in the following equation:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \quad (3)$$

According to the similar triangle principle, the relationship between the coordinates of point  $D(x, y)$  on the plane image and its corresponding point  $D(x_c, y_c, z_c)$  in the camera coordinate system is shown in the following equation:

$$\begin{cases} x = \frac{f x_c}{z_c}, \\ y = \frac{f y_c}{z_c}. \end{cases} \quad (4)$$

Equation (5) is the homogeneous form after sorting.

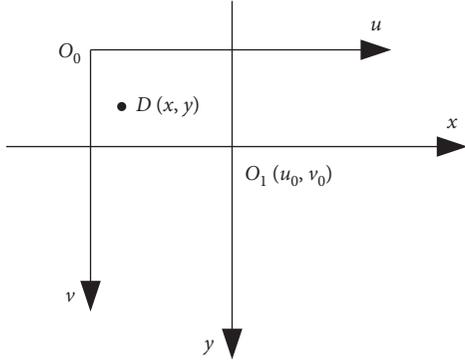


FIGURE 1: Definition of image coordinates.

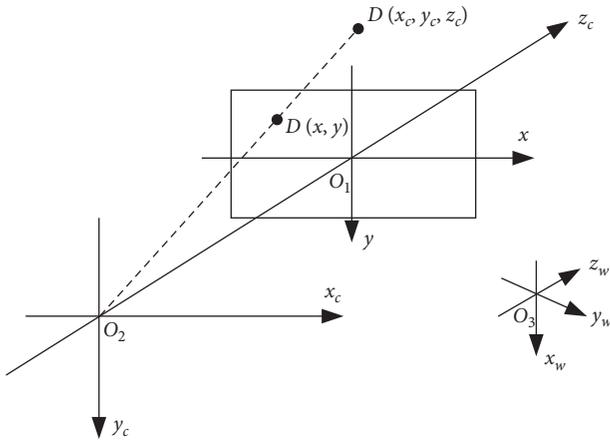


FIGURE 2: The relationship of three kinds of coordinates.

$$z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}. \quad (5)$$

From equations (1), (3), and (5), the relationship between plane image coordinates and world coordinates can be obtained, as shown in the following equation:

$$\begin{cases} u_i m_{34} = x_{wi} m_{11} + y_{wi} m_{12} + z_{wi} m_{13} + m_{14} - u_i x_{wi} m_{31} - u_i y_{wi} m_{22} - u_i z_{wi} m_{33}, \\ v_i m_{34} = x_{wi} m_{21} + y_{wi} m_{22} + z_{wi} m_{23} + m_{24} - v_i x_{wi} m_{31} - v_i y_{wi} m_{22} - v_i z_{wi} m_{33}. \end{cases} \quad (8)$$

It can be seen from equation (8) that each characteristic point can correspond to two independent equations. Therefore, the 12 unknowns in  $M$  matrix can be obtained from 12 equations obtained from 6 characteristic points. The

$$\begin{aligned} z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= z_c \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_1 M_2 X_w = M X_w. \end{aligned} \quad (6)$$

Among them,  $X = [x_w, y_w, z_w]^T$ ;  $M_1$  is related to  $f$ ,  $u_0$ ,  $v_0$ ,  $dx$ , and  $dy$  and determined by the internal structure of the camera, and is called internal parameter.  $M_2$  is determined by the orientation of the camera relative to the world coordinate system, which is called external parameter.  $M$  is called the projection matrix.

**3.2. Camera Calibration.** Camera calibration is the process of obtaining the internal and external parameters of the camera. For the calibration plate, the 3D coordinates  $(x_w, y_w, z_w)$  of each feature point are known. The plane image coordinates of feature points are also known. Therefore, as long as there are enough characteristic points, the matrix  $M$  can be obtained and then  $M_1$  and  $M_2$  can be obtained. For each characteristic point on the calibration plate, the relationship is shown in the following equation:

$$z_{ci} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix}. \quad (7)$$

The system of equations can be obtained by eliminating  $z_{ci}$  as follows:

solution method is the least square method. The more the number of feature points, the smaller the error. For  $n$  characteristic points,  $2n$  equations are obtained as shown in the following equation:

$$\begin{bmatrix}
x_{w1} & y_{w1} & z_{w1} & 1 & 0 & 0 & 0 & 0 & -u_1x_{w1} & -u_1y_{w1} & -u_1z_{w1} \\
0 & 0 & 0 & 0 & x_{w1} & y_{w1} & z_{w1} & 1 & -v_1x_{w1} & -v_1y_{w1} & -v_1z_{w1} \\
& & & & & & \dots & \dots & & & \\
x_{wn} & y_{wn} & z_{wn} & 1 & 0 & 0 & 0 & 0 & -u_nx_{wn} & -u_ny_{wn} & -u_nz_{wn} \\
0 & 0 & 0 & 0 & x_{wn} & y_{wn} & z_{wn} & 1 & -v_nx_{wn} & -v_ny_{wn} & -v_nz_{wn}
\end{bmatrix}
\begin{bmatrix}
m_{11} \\
m_{12} \\
m_{13} \\
m_{14} \\
m_{21} \\
m_{22} \\
m_{23} \\
m_{24} \\
m_{31} \\
m_{32} \\
m_{33}
\end{bmatrix}
=
\begin{bmatrix}
u_1m_{34} \\
v_1m_{34} \\
\dots \\
u_nm_{34} \\
v_nm_{34}
\end{bmatrix}. \quad (9)$$

It can be seen from equation (6) that the multiplication of  $M$  matrix by any constant other than 0 does not affect the relationship between  $[x_w, y_w, z_w]$  and  $[u, v]$ . Therefore,  $m_{34}=1$  can be specified in equation (9). At this time, the number of unknowns of  $M$  matrix is reduced to 11. Let these 11 unknowns be vector  $m$ , then equation (9) can be abbreviated to the following equation:

$$Km = u, \quad (10)$$

where  $K$  is a  $2n \times 11$  matrix,  $m$  is an 11 dimensional unknown vector, and  $u$  is a  $2n$ -dimensional vector. When  $2n > 11$ , the solution of the equation obtained by the least square method is shown in the following equation:

$$m = (K^T K)^{-1} K^T u. \quad (11)$$

The larger the value of  $2n$ , the smaller the error.

Finding vector  $m$  is to get 11 unknowns in  $M$  matrix. The last unknown number  $m_{34}$  is solved as follows; equation (6) can be written as the following equation:

$$\begin{aligned}
m_{34} \begin{bmatrix} m_1^T & m_{14} \\ m_2^T & m_{24} \\ m_3^T & 1 \end{bmatrix} &= M_1 M_2 = \begin{bmatrix} \alpha x & 0 & u_0 & 0 \\ 0 & \alpha y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1^T & tx \\ r_2^T & ty \\ r_3^T & tz \\ 0^T & 1 \end{bmatrix} \\
&= \begin{bmatrix} \alpha x r_1^T + u_0 r_3^T & \alpha x t x + u_0 t z \\ \alpha y r_2^T + v_0 r_3^T & \alpha y t y + v_0 t z \\ r_3^T & t z \end{bmatrix}, \quad (12)
\end{aligned}$$

where  $\alpha = 1/dx$ . So,  $m_{34} m_3^T = r_3^T$ . Since  $r_3$  is the third row of an orthogonal array of units,  $|r_3| = 1$ . From this, we can get

$$m_{34} = \frac{1}{|m_3|}. \quad (13)$$

After 12 unknowns of  $M$  matrix are obtained, each element in internal and external parameter matrix  $M_1$  and  $M_2$  can be obtained further.

### 3.3. Calculation of Relative Position between Two Cameras.

In binocular vision camera calibration, in addition to calculating the internal and external parameters of each camera, it is also necessary to calculate the relative position between the two cameras. For the two cameras, there are

$$\begin{aligned}
X_{c1} &= R_1 X_w + t_1, \\
X_{c2} &= R_2 X_w + t_2,
\end{aligned} \quad (14)$$

where  $X_w = [x_w, y_w, z_w, 1]^T$ .

After  $X_w$  is eliminated,

$$X_{c1} = R_1 R_2^{-1} X_{c2} + t_1 - R_2^{-1} t_2. \quad (15)$$

Therefore, the relative position between two cameras can be represented by  $R$  and  $t$  as follows:

$$\begin{aligned}
R &= R_1 R_2^{-1}, \\
t &= t_1 - R_2^{-1} t_2.
\end{aligned} \quad (16)$$

**3.4. Calibration of Projector.** The projector can be regarded as a reverse working camera [10]. Therefore, the mathematical model of the projector can be represented by the pinhole camera model shown in equation (6). Although the mathematical model of the projector is the same as that of the camera, the projector cannot directly get the pixel coordinates of each feature point on the projector image plane. The solution is given in reference [11]; the projector projects the horizontal and vertical gray code fringes onto the calibration plate in the order of continuous subdivision. After the camera captures the image, the direct and indirect light components are calculated and the threshold segmentation is performed. Then, the gray code decoding algorithm is used to obtain the coordinates of each point on the image plane of the projector.

## 4. Lens Distortion Correction

**4.1. Traditional Lens Distortion Correction Method.** The ideal pinhole model is only an approximation of the real lens

model. The actual camera and projector are different from the pinhole model because of the different lens structure and the processing error and assembly error in the production process. For ordinary lens, especially for wide-angle lens, lens distortion should be considered [12]. The most important influence on imaging is radial distortion. Let the radial distortion parameters be  $k_1$  and  $k_2$ . Then, there are

$$\begin{cases} x' = x(1 + k_1 r^2), \\ y' = y(1 + k_2 r^2), \end{cases} \quad (17)$$

where  $(x', y')$  is the image coordinate obtained from the single hole camera model and  $(x, y)$  is the actual image coordinate. Formula (17) only considers the radial distortion and ignores the high-order term. Eccentric distortion and thin prism distortion should be considered in real lens. Because of the difference of the optical model and the assembly error, they cannot be expressed by the same mathematical model. In reference [4], a lens distortion correction method based on BP neural network is proposed. However, BP network is slow in calculation and easy to fall into local optimum, so it cannot meet the real-time and accuracy requirements of 3D detection.

**4.2. Lens Distortion Correction Based on RBF Network.** RBF network is an efficient forward neural network. The relationship between input layer and hidden layer is nonlinear. There is a linear weighted relationship between the hidden layer and the output layer [13]. This structure avoids the tedious calculation of BP network. It has not only good nonlinear approximation ability but also fast computing ability. It is especially suitable for nonlinear prediction from  $n$ -dimensional space to  $m$ -dimensional space.

The RBF network structure of lens distortion correction is shown in Figure 3. The input signal  $(x, y)$  is the coordinates of the real shot image. The output signal  $(x', y')$  is the actual image coordinates, which can be defined as the coordinates of feature points on the calibration board. The number of nodes in the hidden layer is the number of samples  $n$ .

The input vector of the system is  $d = [x, y]^T$ . The output vector is  $d' = [x', y']^T$ . The weight matrix  $w$  is  $2 \times n$  matrix. The element  $w_{ij}$  is the weight between the  $i$ -th node in the hidden layer and the  $j$ -th node in the output layer. The radial basis function  $\phi(d^i, d^p)$  is Gaussian kernel function where  $d^i$  is the  $i$ -th input vector and  $d^p$  is the center point vector.

$$\Phi(d^i, d^p) = \exp\left(-\frac{1}{2\sigma^2}\|d^i - d^p\|^2\right). \quad (18)$$

The system provides  $n$  characteristic point samples on the calibration board. According to the network structure, the system output is

$$d'_j = \sum_{i=1}^n w_{ij} \exp\left(-\frac{1}{2\sigma^2}\|d^i - d^p\|^2\right), \quad j = 1, 2. \quad (19)$$

In order to avoid each radial basis function being too sharp or too flat, the definition of the expansion constant of the radial basis function is shown in the following equation:

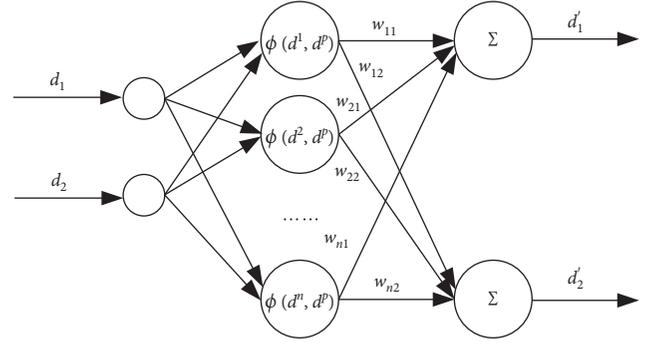


FIGURE 3: The RBF network structure of lens distortion correction.

$$\sigma = \frac{d_{\max}}{\sqrt{2n}}, \quad (20)$$

where  $d_{\max}$  is the maximum distance among samples and  $n$  is the number of samples.

Learning of system is divided into two stages. The first stage is unsupervised learning. The concrete work is to solve the center and variance in the hidden layer. The second stage is supervised learning. The specific work is to solve the weight matrix from implicit layer to output layer. The adjustment of weight can be realized by the least mean square error. The weight adjustment formula is

$$\Delta w_{ij} = \eta(d'_j - w_j^T \Phi) \varphi_i. \quad (21)$$

Among them,  $d'_j$  is the  $j$ -th expected value.

$$\begin{aligned} w_j &= [w_{1j} \ w_{2j} \ \dots \ w_{nj}]^T, \\ \Phi &= [\varphi_1 \ \varphi_2 \ \dots \ \varphi_n]^T. \end{aligned} \quad (22)$$

## 5. Experiment and Analysis

**5.1. Experimental Process.** In this paper, the binocular detection system shown in Figure 4 is used to verify the above algorithm. The camera pixel is 1.3 million, and the measurement format is 200 mm  $\times$  150 mm. The nominal scanning accuracy is 0.01 mm. The calibration board used in the experiment is shown in Figure 5. There are  $11 \times 13 = 169$  regularly arranged feature points on the calibration board, including 17 locating points.

In the calibration experiment, the position of the calibration plate is fixed first. The projector projects the gray code structured light to the calibration plate as shown in Figure 6. The camera takes pictures. The attitude of calibration board is changed, and the above work is repeated. The calibration board is placed in 8 different positions in the measurement space, as shown in Figure 7. Each camera obtains 8 images for system calibration. In calibration, firstly, the edge of the image is extracted at pixel level to identify the mark points and fit the center point and number the marker points according to the locating points. Then, the 3D coordinates of the positioning point are reconstructed by using the location of the positioning point in the first two pictures. Next, the 3D coordinates of the landmarks are reconstructed with the rest of the images.



FIGURE 4: Binocular vision system.

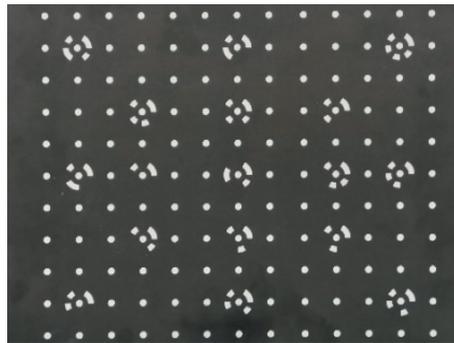


FIGURE 5: Calibration board.

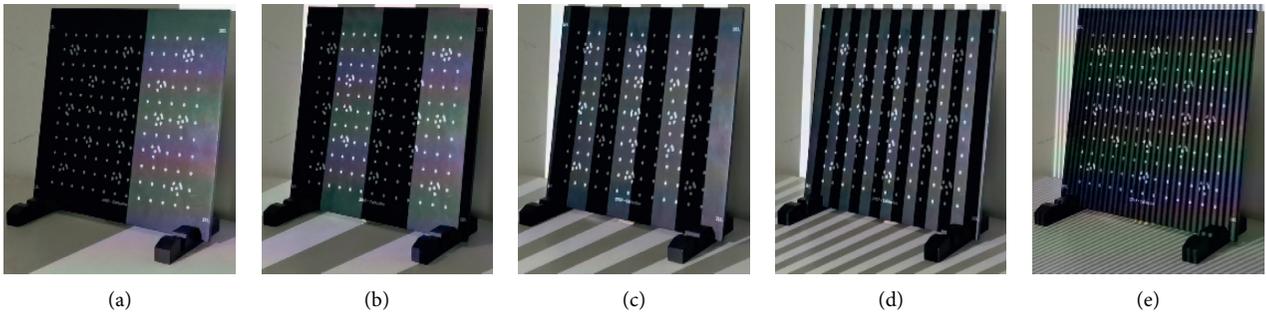


FIGURE 6: Projected gray code structured light.

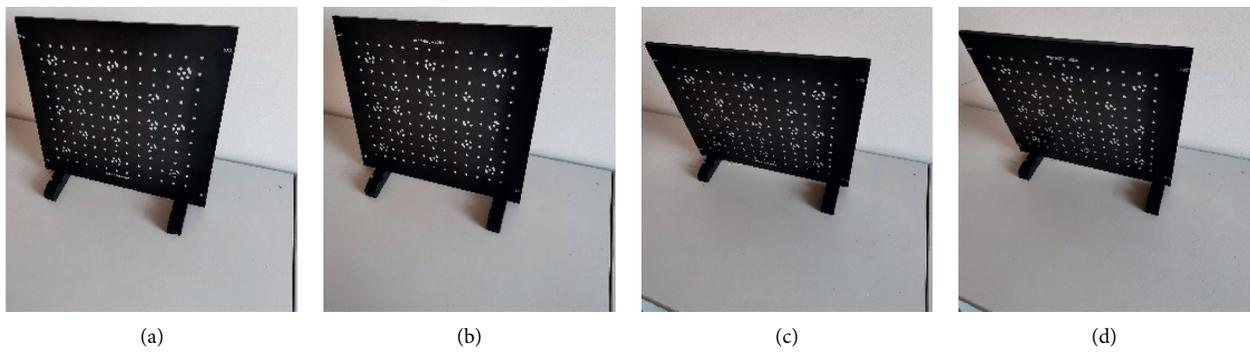


FIGURE 7: Continued.

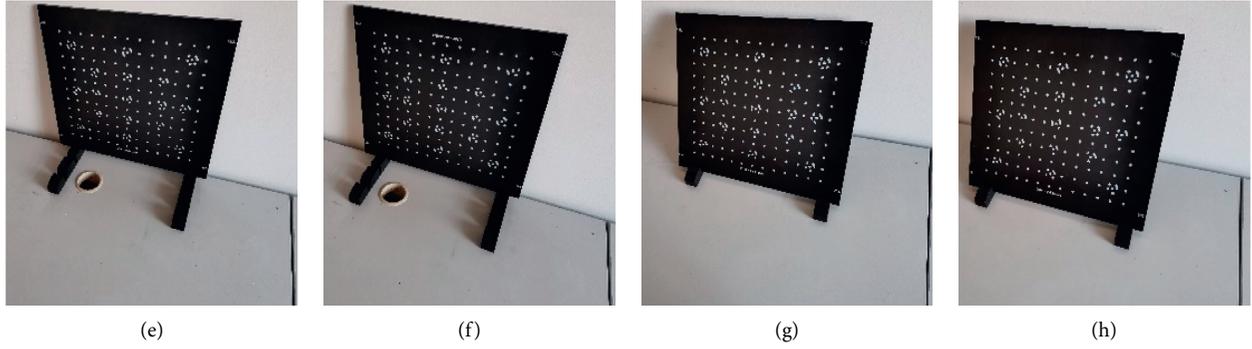


FIGURE 7: Eight attitudes of calibration board.

TABLE 1: Internal parameters of left and right cameras.

Camera	Internal parameters			
	$ax$	$\alpha y$	$u_0$	$v_0$
Left	2785.35	2678.67	229.21	145.54
Right	2820.91	2732.28	257.82	107.85

TABLE 2: External parameters of left and right cameras.

Camera	External parameters			
	$R$			$t$
Left	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$			$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$
Right	$\begin{bmatrix} 0.9006 & -0.02775 & -0.4338 \\ 0.01377 & 0.9992 & -0.3533 \\ 0.4345 & 0.02584 & 0.9003 \end{bmatrix}$			$\begin{bmatrix} -279.1 \\ -15.82 \\ -127.8 \end{bmatrix}$

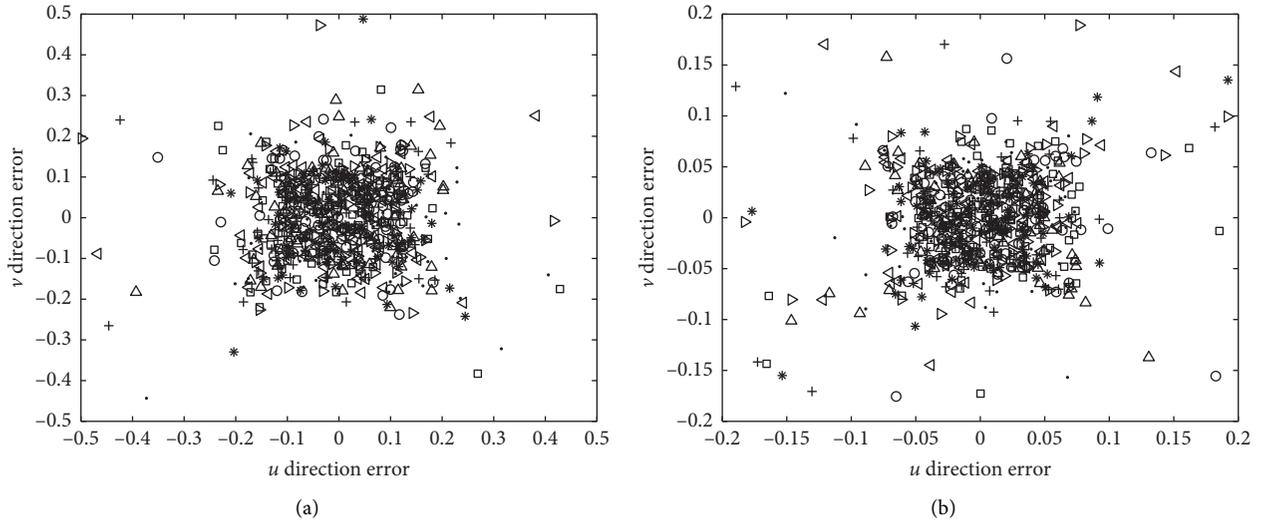


FIGURE 8: Reprojection error.

5.2. Data Analysis. After obtaining the 3D coordinates of the marker points, the internal and external parameters of the camera can be calculated according to the method mentioned above, as shown in Tables 1 and 2.

In the experiment, the reprojection method is used to verify the accuracy of the calibration data. According to the parameters obtained from the calibration, the positioning

points are reprojected to the image plane of the camera and compared with the actual image points. The traditional method and the method proposed in this paper are used to calibrate, respectively. The calculated error is shown in Figure 8. On the left is the result of the traditional calibration method. On the right is the calibration result of the method.

TABLE 3: Residuals corresponding to eight postures of the traditional algorithm.

Attitude	Average error $u$	Average error $v$	Standard error $u$	Standard error $v$
1	0.0129	0.0146	0.0992	0.1187
2	-0.0107	0.0090	0.1085	0.1007
3	-0.0029	0.0063	0.1118	0.1191
4	0.0145	-0.0072	0.1059	0.1109
5	-0.0148	0.0168	0.1104	0.1133
6	0.0081	-0.0034	0.1162	0.1109
7	0.0151	0.0018	0.1218	0.1030
8	0.0013	-0.0198	0.1175	0.1228

TABLE 4: Residuals corresponding to eight postures of the algorithm in this paper.

Attitude	Average error $u$	Average error $v$	Standard error $u$	Standard error $v$
1	0.0083	0.0027	0.0462	0.0485
2	-0.0087	-0.0010	0.0426	0.0494
3	0.0067	0.0076	0.0463	0.0495
4	0.0022	0.0035	0.0488	0.0456
5	-0.0030	0.0012	0.0443	0.0516
6	-0.0031	-0.0011	0.0492	0.0510
7	0.0021	-0.0097	0.0501	0.0478
8	-0.0096	-0.0089	0.0501	0.0495

When the two algorithms are used, the residual data corresponding to the eight attitudes of the calibration board are shown in Tables 3 and 4.

## 6. Conclusion

In this paper, the calibration algorithm of the camera and projector in the binocular vision structured light detection system is introduced. The ideal image and the actual image coordinates are taken as the input and output system, and the image distortion correction system based on RBF neural network is constructed. The actual camera distortion correction calculation is completed by using the good nonlinear fitting ability of neural network. Experimental results show that the algorithm can overcome the shortcomings of traditional methods, and the detection results can meet the actual needs.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the Creation Foundation Project of Nanjing Institute of Technology (Grant No. CXY201933).

## References

- [1] L. Y. Fu, D. Zhang, and Q. L. Ye, "Recurrent thrifty attention network for remote sensing scene recognition," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12, 2020.
- [2] C. F. Jiang, L. Beatrice, and Z. Song, "Three-dimensional shape measurement using a structured light system with dual projectors," *Applied Optics*, vol. 57, no. 14, p. 3983, 2018.
- [3] Z. Wang, Z. Wu, X. Zhen, R. Yang, J. Xi, and X. Chen, "A two-step calibration method of a large FOV binocular stereovision sensor for onsite measurement," *Measurement*, vol. 62, pp. 15–24, 2015.
- [4] D. Zhang, G. Zhang, and L. Li, "Calibration of a six-axis parallel manipulator based on BP neural network," *Industrial Robot: The International Journal of Robotics Research and Application*, vol. 46, no. 5, pp. 692–698, 2019.
- [5] W. G. Li, H. Li, and H. Zhang, "Light plane calibration and accuracy analysis for multi-line structured light vision measurement system," *Optik*, vol. 207, Article ID 163882, 2019.
- [6] X. Chen, F. Zhou, and T. Xue, "Omnidirectional field of view structured light calibration method for catadioptric vision system," *Measurement*, vol. 148, Article ID 106914, 2019.
- [7] J. E. Ha, "Calibration of structured light vision system using multiple vertical planes," *Journal of Electrical Engineering and Technology*, vol. 13, no. 1, pp. 438–444, 2018.
- [8] X. Chen, R. Fan, J. Wu et al., "Fourier-transform-based two-stage camera calibration method with simple periodical pattern," *Optics and Lasers in Engineering*, vol. 133, Article ID 106121, 2020.
- [9] J. Jiang, L. Zeng, B. Chen, Y. Lu, and W. Xiong, "An accurate and flexible technique for camera calibration," *Computing*, vol. 101, no. 4, pp. 1971–1988, 2019.
- [10] S. Yang, M. Liu, J. Song et al., "Projector calibration method based on stereo vision system," *Optical Review*, vol. 24, no. 5, pp. 1–7, 2017.
- [11] S. K. Nayer, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 935–944, 2006.
- [12] M. Zhang, F. L. Wu, L. X. Jin, G. N. Li, S. Han, and Y. Zhang, "Correction optimization of lens radial distortion with

bending measurement function,” *Transactions of Tianjin University*, vol. 22, no. 98, pp. 94–99, 2016.

- [13] Q. Ye, H. Zhao, Z. Li et al., “L1-norm distance minimization-based fast robust twin support vector  $k$ -plane clustering,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4494–4503, 2018.