A Hand-Eye Calibration Method based on Nonlinear Joint Optimization for RGBD Cameras

Yanling Zhou[1,2], Runhua Wang[1,2], Xuebo Zhang[1,2]

1. Institute of Robotics and Automatic Information System, College of Artificial Intelligence, Nankai University, Tianjin 300350, China

2. Tianjin Key Laboratory of Intelligent Robotics, Nankai University, Tianjin 300350, China
E-mail: wrunhua@nankai.edu.cn

2022 34th Chinese Control and Decision Conference (CCDC) | 978-1-6654-7896-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/CCDC55256.2022.10034272

Abstract: Hand-eye calibration is a basic research topic in the field of robotic vision. At present, most works focus on the hand-eye calibration of RGB cameras, while methods for calibration of RGBD cameras are relatively rare. Two popular kinds of modes for RGBD cameras are based on the stereo vision and the structured light, which both contain at least two cameras for generating the color image and the depth image, respectively. Motivated by this special structure feature, the paper proposes a novel hand-eye calibration method for RGBD cameras by integrating the vision system and the hand-eye robotic system into a whole model. Specifically, the traditional hand-eye calibration method is firstly used to obtain initial values of hand-eye matrices of the two cameras in the RGBD camera separately. Then, a nonlinear joint optimization algorithm is carried out to refine hand-eye matrices, in which the objective is to minimize the joint reprojection error of feature corners in images outputting by the two cameras in the RGBD camera. Thus, the optimized hand-eye calibration matrices can be finally obtained for both cameras. The proposed calibration approach is validated on the robotic hand-eye system, and the experimental results demonstrate that the presented algorithm greatly improves the calibration accuracy compared with the traditional methods.

Key Words: hand-eye calibration, RGBD cameras, reprojection error, nonlinear joint optimization

1. INTRODUCTION

For industrial robots performing grabbing and assembling tasks[1], rescue robots searching for casualties[2], mobile robots guiding customers and many other robotic applications, the computer vision technique is widely used to obtain the environment information based on visual sensors[3][4][5]. As one kind of commonly used visual sensors, the RGBD cameras have many advantages with respect to ordinary RGB cameras, such as providing depth information directly, small and convenient in size and robust to the light intensity, etc. Contribute to above mentioned superiorities, the RGBD cameras have been applied in many areas[6][7][8]. An important prerequisite for robots to accomplish various tasks based on vision information is to obtain the pose transformation matrix between the camera coordinate system and the end-effector coordinate system, which is named hand-eye calibration.

For RGB cameras, hand-eye calibration is a well-studied task. In 1989, Tsai[9] first proposed the formulation of hand-eye calibration as a homogeneous equation: $AX = XB$. Following this hand-eye equation, many domestic and foreign scholars put forward various researches and solutions, using different mathematical tools (quaternion combined with singular value decomposition[10], the dual quaternion[11], shaft angle transformation and matrix direct product[12], matrix product[13], etc.) to express the rotation matrix and translation vector of eye position matrix, and then obtain parameters by solving the hand-eye equation. At present, the hand-eye calibration technique for RGB cameras has achieved high precision and applicability. For RGBD cameras, though they can output both depth images and 2D color images, most methods just chose one kind of images to perform calibration. For example, Reinbacher[14] and Kahn[15][16] chose the 2D image-based method used for RGB cameras to carry out hand-eye calibration between RGBD cameras and robots. In addition, there are some methods that only use depth images. Dong[17] used 3D measurements in the center of the calibration plate for hand-eye calibration. Fuchs[18] applied a calibration plane with known position and direction to estimate the hand-eye matrix, in which the calibration problem is inverted to a least square curve fitting problem between the measured depth value and the calibration plane. In order to solve the shortcomings of less depth image features and no color and texture information, some methods[19] were proposed to use special hollow checkerboard or 3D calibration pieces with special structure to complete hand-eye calibration, However, these methods have low applicability.

From the above analysis it is known that, current hand-eye calibration technology for RGBD cameras only uses its 2D color image or depth value. However, the depth camera based on stereo vision or structured light has at least two cameras to output both color images and depth images. The two cameras are two RGB cameras or one RGB camera and one infrared camera. Motivated by this special structural characteristic, this paper considers the two cameras of the RGBD camera at the same time and then put

This work is supported in part by National Key Research and Development Project under Grant 2018YFB1307503, in part by Tianjin Science Fund for Distinguished Young Scholars under Grant 19JCJQJC62100, in part by Tianjin Natural Science Foundation under Grant 19JCYBJC18500, and in part by the Fundamental Research Funds for the Central Universities.

forward a novel hand-eye calibration method, wherein the nonlinear joint optimization approach is used to minimize the total reprojection error of feature corner points in both cameras' images.

2. THE HAND-EYE CALIBRATION

The basis of hand-eye calibration is camera calibration, which is to construct the camera projection model to obtain the corresponding relationship between the 3D space and the 2D image. This projection process mainly involves four important coordinate systems and their mutual conversion relations, including the world coordinate system, the camera coordinate system, the image coordinate system and the pixel coordinate system.

The conversion relation from the world coordinate system $\begin{bmatrix} X_W & Y_W & Z_W \end{bmatrix}^T$ to the pixel coordinate system $\begin{bmatrix} U & V \end{bmatrix}^T$ can be described as:

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \frac{1}{Z_C} M_{3\times3} \begin{bmatrix} \mathrm{R}_{3\times3} & \mathrm{t}_{3\times1} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}, \tag{1}$$

where Z_C is a positive scale factor, $M_{3\times3}$ is the internal parameters matrix of camera, and $\begin{bmatrix} R_{3\times3} & \mathrm{t}_{3\times1} \end{bmatrix}$ is the external parameters matrix. $M_{3\times3}$ and $\begin{bmatrix} R_{3\times3} & \mathrm{t}_{3\times1} \end{bmatrix}$ are the results of camera calibration.

For the eye-in-hand system studied in this paper, the hand-eye calibration is to obtain the pose relationship between the camera coordinate system and the manipulator's end-effector coordinate system, which is represented by matrix ${}^{g}_{c}T$. The main method for hand-eye calibration is to control the robotic arm with the camera to different poses and simultaneously observe the corner points on the calibration plate. With these observation results, the matrix ${}^{g}_{c}T$ just can be calculated. This classical calibration process will be briefly introduced as follows.

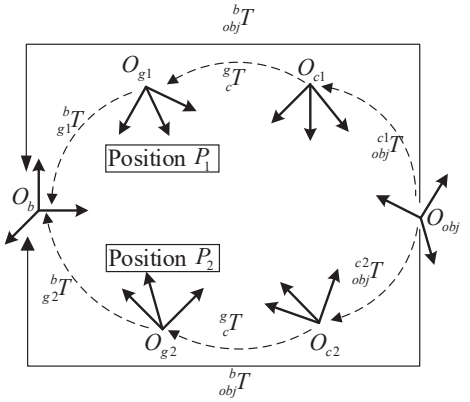

Fig 1. Transformation of coordinate systems in the process of hand-eye calibration

Figure 1 shows relative pose relationships among coordinate systems of robot's base (O_b), manipulator's end-effector (O_{g1} and O_{g2}), camera (O_{c1} and O_{c2}) and calibration plate (O_{obj}) when the manipulator's end-effector moves from pose P_1 to pose P_2. The matrices ${}^{c1}_{obj}T$ and ${}^{c2}_{obj}T$ represent the pose transformation between coordinate systems of the calibration plate and the camera at those two poses, respectively. They are the external parameters of the camera and can be obtained through camera calibration. The matrices ${}^{b}_{g1}T$ and ${}^{b}_{g2}T$ represent the pose transformation between coordinate systems of the manipulator's end-effector and the robot's base at those two poses, respectively. They can be read out through the controller of the robot arm.

When the manipulator's end-effector is at pose P_1 or P_2, according to the transformation relationship between coordinate systems, the pose transformation matrix from the calibration plate coordinate system O_{obj} to the robot base coordinate system O_b can be obtained as follows:

$${}^{b}_{obj}T = {}^{b}_{g1}T\,{}^{g}_{c}T\,{}^{c1}_{obj}T = {}^{b}_{g2}T\,{}^{g}_{c}T\,{}^{c2}_{obj}T\ . \tag{2}$$

Since ${}^{b}_{obj}T$ keeps unchanged, the basic equation of the hand-eye calibration can be deduced:

$$AX = XB, \tag{3}$$

where

$$A = ({}^{b}_{g2}T)^{-1}\,{}^{b}_{g1}T\ ,\ B = {}^{c2}_{obj}T({}^{c1}_{obj}T)^{-1}\ . \tag{4}$$

According to formula (3), the hand-eye calibration is equivalent to solving the parameter X. The reference [9] proposed Tsai-Lenz hand-eye calibration method in 1989, which is the most basic and widely used algorithm in the field of hand-eye calibration at present. With this method, all the pose matrices in the hand-eye equation are represented by rotation matrices and translational vectors, so as to obtain:

$$\begin{pmatrix} R_A & t_A \\ 0 & 1 \end{pmatrix}\begin{pmatrix} R_X & t_X \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_X & t_X \\ 0 & 1 \end{pmatrix}\begin{pmatrix} R_B & t_B \\ 0 & 1 \end{pmatrix}. \tag{5}$$

By expanding formula (5), it can be obtained that:

$$\begin{cases} R_A R_X = R_X R_B \\ (R_A - I)t_X = R_B t_B - t_A \end{cases}. \tag{6}$$

Then the problem of hand-eye calibration is transformed into solving R_X and t_X from above two equations, as shown in formula (6). Tsai-Lenz method is to solve these equations with "two-step method". The first step is to obtain a rotation matrix R_X by solving the first equation in formula (6). Then the second step is to obtain a translation vector t_X through substituting R_X into the second equation.

Theoretically, the unique values of the rotation matrix R_X and translational vector t_X could be solved by two sets of manipulator's non-translational motion data. However, in order to improve the precision of hand-eye calibration, Tsai proposed to control the manipulator for multiple sets of poses. Then, by establishing multiple sets of hand-eye calibration equations, the corresponding rotation matrices and translational vectors can be solved, respectively. Finally, the least square fitting method is used to obtain the hand-eye matrix.

3. OPTIMIZATION OF HAND-EYE MATRICES

Tsai-Lenz method linearizes the nonlinear model and improves calculation's efficiency, while the accuracy is low. On the one hand, the error of each step in the process of hand-eye calibration would transmit and result in a large error in the final. Especially, the precision of camera calibration has great influence on the precision of hand-eye calibration. On the other hand, when the model is linearized to calculate the rotation matrix, the linearization error is bound to occur.

Based on above drawbacks and taking full advantages of structural characteristics of the RGBD cameras, this paper proposed a novel hand-eye calibration for RGBD cameras.

3. 1. Establishment of Overall Model

Firstly, this paper proposes to integrate the vision system of the RGBD camera and the eye-in-hand system of the robot into an overall model by the common parameter ${}^{c}_{obj}T$.

The matrix ${}^{c}_{obj}T$ represents the pose relationship between the calibration plate and the camera. In the vision system of a camera, the matrix ${}^{c}_{obj}T$ represents camera's external parameters which can be directly obtained by the camera calibration. In the eye-in-hand system of the robot, the matrix ${}^{c}_{obj}T$ can be obtained through the pose relationships among the calibration plate, the robot's base, manipulator's end-effector and the camera.

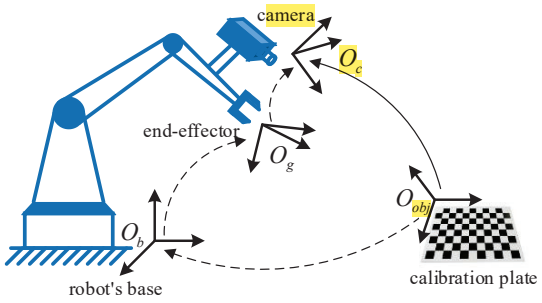

Fig 2. The coordinate systems in the hand-eye system

As shown in Figure 2, in the eye-in-hand system, the conversions between the camera coordinate system and the calibration plate along the dotted line and the solid line are ultimately equal, which can be expressed as:

$$ {}^{c}_{obj}T = ({}^{g}_{c}T)^{-1} \cdot ({}^{b}_{g}T)^{-1} \cdot {}^{b}_{obj}T , \tag{7} $$

where ${}^{g}_{c}T$ represents the hand-eye relationship between the camera and the manipulator's end-effector; ${}^{b}_{g}T$ represents the pose relationship between the manipulator's end-effector and the robot's base; ${}^{b}_{obj}T$ represents the pose relationship between the calibration plate and the robot's base which can be obtained by formula(2).

Inserting above formula (7) into the camera projection models (formula (1)) of both cameras in the RGBD camera, respectively, it can be obtained as:

$$ \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \frac{1}{Z_{1c}} M_1 ({}^{g}_{1c}T)^{-1} ({}^{b}_{g}T)^{-1} {}^{b}_{obj}T \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \tag{8} $$

$$ \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \frac{1}{Z_{2c}} M_2 ({}^{g}_{2c}T)^{-1} ({}^{b}_{g}T)^{-1} {}^{b}_{obj}T \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \tag{9} $$

where Z_{1c} and Z_{2c} represent the depth value of one checkerboard corner in two cameras' coordinate system respectively; M_1 and M_2 represent the internal parameters of the two cameras respectively; ${}^{g}_{1c}T$ and ${}^{g}_{2c}T$ are the hand-eye matrices of the two cameras respectively; $[u_1 \quad v_1 \quad 1]^T$ and $[u_2 \quad v_2 \quad 1]^T$ represent the pixel coordinates of that checkerboard corner in the images of those two cameras respectively; $[x_w \quad y_w \quad z_w]^T$ represents the three-dimensional coordinates of the checkerboard corner in the world coordinate system.

3. 2. Joint Optimization

The models shown in formulas (8) and (9) are nonlinear models, in which the matrices ${}^{g}_{1c}T$ and ${}^{g}_{2c}T$ are the variables that need to be optimized. In this paper, a nonlinear optimization method is used to solve the hand-eye transformation matrix. The objective is to minimize the reprojection error of checkerboard corners in images of two cameras.

Firstly, corner detection is performed on images of the calibration board. In this way, pixel coordinates $[u_{1i} \quad v_{1i} \quad 1]^T$ and $[u_{2i} \quad v_{2i} \quad 1]^T$ of corner points in the two cameras' images can be obtained respectively. Thus, according to formula (8) and (9), the reprojection error can be expressed as:

$$ error_{1i} = \left(\begin{bmatrix} u_{1i} \\ v_{1i} \\ 1 \end{bmatrix} - \frac{1}{Z_{1ci}} M_1 ({}^{g}_{1c}T)^{-1} ({}^{b}_{g}T)^{-1} {}^{b}_{obj}T \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix} \right)^2, \tag{10} $$

$$ error_{2i} = \left(\begin{bmatrix} u_{2i} \\ v_{2i} \\ 1 \end{bmatrix} - \frac{1}{Z_{2ci}} M_2 ({}^{g}_{2c}T)^{-1} ({}^{b}_{g}T)^{-1} {}^{b}_{obj}T \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix} \right)^2 . \tag{11} $$

The error of two cameras' projection models obtained by camera calibrations is different due to different robustness to light intensity and distortion, etc. Thus, it is necessary to use different weights of the corresponding reprojection error to combine formulas (10) and (11). The joint reprojection error of one certain corner point can be written as:

$$ error_i = \alpha_1 error_{1i} + \alpha_2 error_{2i} . \tag{12} $$

In order to reduce the complexity of the model and obtain

the best results, a nonlinear iterative optimization method is adopted to solve the hand-eye matrices in this paper. The nonlinear model may have multiple extreme points, and the result of iterative optimization depends on the initial value. If the initial value is set improperly, the optimization result can only be locally optimal. Therefore, in this paper, the hand-eye calibration results $({}^{g}_{1c}T)_0$ and $({}^{g}_{2c}T)_0$ obtained by Tsai-Lenz method is set as the initial values of this iterative optimization. Then, the nonlinear approach is used to iterate the cycle optimization to obtain the optimal results- ${}^{g}_{1c}T^*$ and ${}^{g}_{2c}T^*$.

At this stage, the objective function of the nonlinear optimization is:

$$ {}^{g}_{1c}T^*, {}^{g}_{2c}T^* = \arg\min \sum_{i}^{n\times m} error_i , \tag{13} $$

where n is the number of calibration plate pictures taken by each camera, and m is the number of inner corner points on the calibration plate.

The constraint condition of this nonlinear optimization is as follows:

$$ {}^{g}_{2c}T = {}^{g}_{1c}T \cdot {}^{1c}_{2c}T , \tag{14} $$

where ${}^{1c}_{2c}T$ is the pose conversion matrix between the two cameras in the RGBD camera.

Remark: For currently widely used RGBD cameras, there are three main principles to measure the depth value, including stereo vision (using two RGB cameras), structured light (using one infrared projector and one or two infrared cameras) and time of flight. The paper focuses on the first two kinds of RGBD cameras, which contains at least two cameras (one for color image and the other for the depth image). In other words, the proposed calibration method is applicable for both the stereo vision based and structured based RGBD cameras, but not the time of flight based ones.

4. EXPERIMENT

In order to verify the effectiveness of this novel hand-eye calibration method, an eye-in-hand system is built for the experiments as shown in Figure 3. The robot is Kinova-Jaco-2-S, which is a 7-DOF manipulator. The camera is RealSense_D435i RGBD camera based on structured light as shown in Figure 4. This RGBD camera contains one RGB camera to output color images and two infrared cameras used for binocular distance measurement and outputting the depth images. Moreover, in this camera, the coordinate system of the camera that outputs depth images is consistent with the coordinate system of the left infrared camera. Therefore, the RGB camera and the left infrared camera in this RGBD camera are selected in this experiment to verify the method proposed in this paper. The calibration plate in this experiment is a checkerboard board with 8×6 grids, and the checkerboard's side length is 35mm.

In this experiment, the α_1 is set to value 0.35, while the α_2 is set to value 0.65.

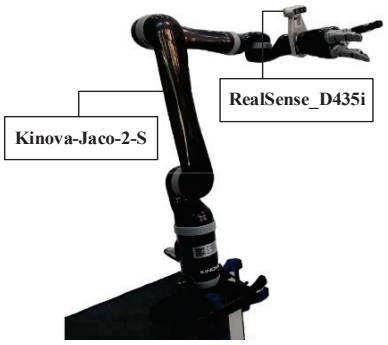

Fig 3. The hand-eye system of the robot in this experiment

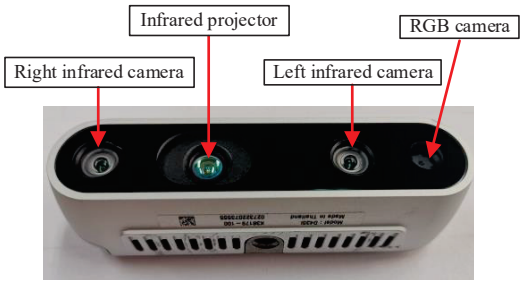

Fig 4. The structural composition of RealSense_D435i camera

During this experiment, the RGBD camera is fixed at the manipulator's end-effector, and the calibration plate is also fixed at an appropriate position outside the robot. In the process of experiment, by manually control the manipulator to form 20 different poses, the corresponding images at each pose for the calibration plate can be taken and the poses of the manipulator can simultaneously be recorded. It should be noted that the images saved at each pose include both the color image from the RGB camera and the infrared image from the left infrared camera.

In this way, a total of 20 sets of experimental data are obtained. The N=1-15 sets of experimental data are taken as the experimental groups to obtain the final results of hand-eye calibration. While, the N=16-20 sets of experimental data are taken as the verification groups to test the effectiveness of the method proposed in this paper. Six sets of experimental images are shown in Figure 5.

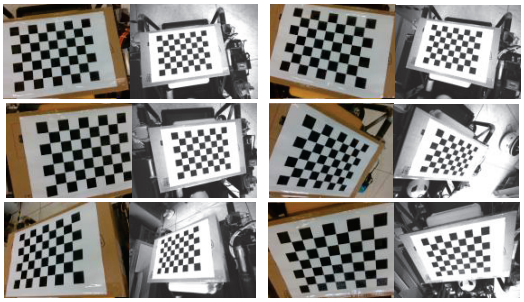

Fig 5. Six sets of calibration images, the left of which are color images and the right of which are infrared images.

The specific calibration steps are as follows:

Firstly, all 20 sets of color and infrared images were processed using Zhang's method[20] for camera calibration,

respectively. Thus, the internal and external parameters of the RGB camera and left infrared camera can be obtained. In addition, the pose transformation matrix between RGB camera and left infrared camera can be obtained according to the factory settings of the RGBD camera.

Then, the Tsai-Lenz method is adopted for hand-eye calibration using the RGB camera's external parameters and the manipulator's poses of the first 15 sets of experimental data. Thus, the hand-eye matrix between the RGB camera and the manipulator's end-effector can be obtained as:

$$({}^{g}_{1c}T)_0 = \begin{bmatrix} -0.9857 & -0.1677 & -0.0148 & 0.0541 \\ 0.1678 & -0.9857 & -0.0064 & 0.1387 \\ -0.0135 & -0.0088 & 0.9999 & -0.1795 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

It's noted that the unit of the translation vectors in all the hand-eye matrices is mm in this paper.

Similarly, the hand-eye transformation matrix between the left infrared camera and the manipulator's end-effector can be obtained as:

$$({}^{g}_{2c}T)_0 = \begin{bmatrix} -0.9856 & -0.1669 & -0.0252 & 0.0403 \\ 0.1669 & -0.9859 & 0.0023 & 0.1410 \\ -0.0252 & -0.0019 & 0.9997 & -0.1712 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thirdly, the hand-eye matrices obtained by above traditional method are taken as initial values. The nonlinear joint optimization algorithm aimed at minimizing the reprojection error is used for cyclic iterative optimization.

Thus, the hand-eye pose matrix between the RGB camera and the manipulator's end-effector after optimization was obtained as follows:

$${}^{g}_{1c}T^{*} = \begin{bmatrix} -0.9853 & -0.1692 & -0.0217 & 0.0572 \\ 0.1691 & -0.9856 & 0.0076 & 0.1333 \\ -0.0227 & 0.0038 & 0.9997 & -0.1788 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

After optimization, the hand-eye pose matrix between the infrared camera and the manipulator's end-effector is:

$${}^{g}_{2c}T^{*} = \begin{bmatrix} -0.9853 & -0.1684 & -0.0286 & 0.0420 \\ 0.1681 & -0.9857 & 0.0113 & 0.1377 \\ -0.0301 & 0.0063 & 0.9995 & -0.1709 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

a. Comparison of the Reprojection Error

The reprojection error of corner points in the N=1-15 sets of pictures is calculated by formulas (10)-(12). The comparison of the reprojection error of checkerboard corners in RGB images is shown in Figure 6. The average reprojection error is 4.2116 pixels before optimization, and it is 1.6263 pixels after optimization. The comparison of the reprojection error of checkerboard corners in infrared images is shown in Figure 7. The average reprojection error is 2.3468 pixels before optimization, and it is 1.0936 pixels after optimization. The joint error is reduced from 2.9995 pixels to 1.28 pixels, which demonstrates that the presented approach improves the accuracy of hand-eye matrices.

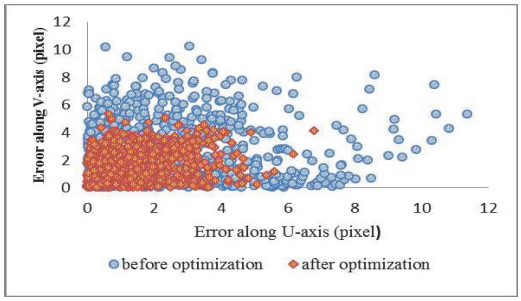

Fig 6. The reprojection error of corner points in color images.

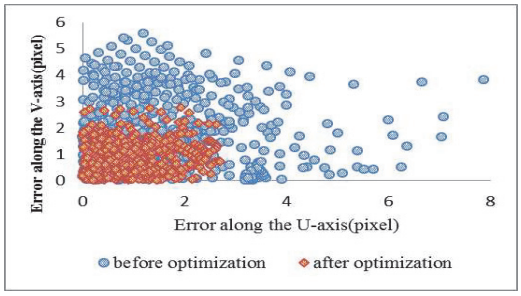

Fig 7. The reprojection error of corner points in infrared images

b. Comparison of the Translation and Rotation Estimation Error

The N=16-20 sets of experimental data are used as the verification groups with the comparison method proposed by Zhao[21]. In this way, the theoretical value of transformation matrix A_N can be calculated through formula (4) by using N=16-20 sets of cameras' external parameters. The estimated value of the transformation matrix A_N (N=16-20) is defined as:

$$\hat{A}_N = XB_NX^{-1}, \tag{15}$$

where B_N can be obtained by formula (4), and X is the hand-eye transformation matrix.

Then the error between A_N and $\hat{A}_N$ is used as the basis to judge the accuracy of hand-eye calibration. Accordingly, the translation error and rotation error defined by translation vector and rotation vector of A_N and $\hat{A}_N$ are used:

$$e_t = \left\| \hat{t}_N - t_N \right\|_2, \quad e_r = \left\| \hat{r}_N - r_N \right\|_2, \tag{16}$$

where the unit of e_t is mm.

As shown in Table 1 and 2, the translation and rotation estimation error of N=16-20 sets of data are calculated by using three pairs of hand-eye transformation matrices, respectively. The first pair- $({}^{g}_{1c}T)_0$ and $({}^{g}_{2c}T)_0$ are obtained by traditional Tsai-Lenz method. The second pair are obtained by individual optimization, which don't consider the two cameras in the RGBD camera simultaneously. $({}^{g}_{1c}T)_*$ represents the hand-eye matrix of RGB camera by optimization through setting α_1 to value 1 and α_2 to value 0. Similarly, $({}^{g}_{2c}T)_*$ represents the hand-eye matrix of infrared camera by optimization through setting α_1 to value 0 and α_2 to value1. The third pair- ${}^{g}_{1c}T^{*}$ and ${}^{g}_{2c}T^{*}$ are obtained by the

proposed refined method base on nonlinear joint optimization where α_1 is set to value 0.35 and α_2 is set to value 0.65.

Table1. The comparison of rotation and translation error (RGB camera)

N	$({}^{g}_{1c}T)_0$		$({}^{g}_{1c}T)_*$		${}^{g}_{1c}T^*$	
	$(e_{r1})_0$	$(e_{t1})_0$	$(e_{r1})_*$	$(e_{t1})_*$	$(e_{r1})^*$	$(e_{t1})^*$
16	0.0376	0.0242	0.0302	0.0242	**0.0163**	**0.0131**
17	0.0185	0.0074	0.0155	0.0054	**0.0098**	**0.0035**
18	0.0130	0.0068	0.0104	0.0049	**0.0081**	**0.0040**
19	0.0049	0.0057	0.0040	0.0052	**0.0037**	**0.0033**
20	0.0232	0.0112	0.0214	0.0098	**0.0137**	**0.0059**
Averge	0.0194	0.0111	0.0163	0.0099	**0.0103**	**0.0060**

Table2. The comparison of rotation and translation error (infrared camera)

N	$({}^{g}_{2c}T)_0$		$({}^{g}_{2c}T)_*$		${}^{g}_{2c}T^*$	
	$(e_{r2})_0$	$(e_{t2})_0$	$(e_{r2})_*$	$(e_{t2})_*$	$(e_{r2})^*$	$(e_{t2})^*$
16	0.0208	0.0197	0.0177	0.0242	**0.0106**	**0.0113**
17	0.0142	0.0079	0.0125	0.0057	**0.0086**	**0.0038**
18	0.0043	0.0043	0.0038	0.0038	**0.0032**	**0.0034**
19	0.0039	0.0043	0.0036	0.0039	**0.0036**	**0.0032**
20	0.0132	0.0116	0.0102	0.0101	**0.0068**	**0.0057**
Averge	0.0112	0.0096	0.0096	0.0095	**0.0075**	**0.0055**

According to the data in the first and last two columns of the Table 1 and Table 2, the calibration error after nonlinear joint optimization is all reduced to a certain extent for both cameras. According to the data in the last four columns of the Table 1 and Table 2, it can be clearly seen that joint optimization is also better than separate optimization.

5. CONCLUSION

In this paper, a novel hand-eye calibration method is proposed for the RGBD camera, based on its special structure. Firstly, the hand-eye matrices obtained using the traditional hand-eye calibration algorithm are used as the initial values of cyclic iterative optimization. Then a nonlinear joint optimization approach is used to obtain the improved matrices. The optimization goal is to minimize the joint reprojection error of corners in both images generated from the RGBD camera. The proposed calibration algorithm is finally validated on the robotic hand-eye system. Comparative experimental results show that the presented approach can greatly improve the accuracy of the hand-eye calibration.

REFERENCES

[1] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, D. Quillen, Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection[J], International Journal of Robotics Research, 2016:421-436.

[2] S. H. Qian, S. R. Ge, Y. S. Wang, Y. Wang, C. Q. Liu, Research Status of the Disaster Rescue Robot and Its Applications to the Mine Rescue[J], Robot, 2006:350-354.

[3] Y. Xu, H. Zhang, L. Cao, X. Shu, D. Zhang, A Shared Control Strategy for Reach and Grasp of Multiple Objects Using Robot Vision and Noninvasive Brain-Computer Interface[J], Transactions on Automation Science and Engineering, IEEE, 2020:1-13.

[4] L. Wu, H. Ren, Finding the Kinematic Base Frame of a Robot by Hand-Eye Calibration Using 3D Position Data[J], Transactions on Automation Science and Engineering, IEEE, 2017:314-324.

[5] A. Hietanen, J. Latokartano, A. Foi, R. Pieters, J. K. Kmrinen, Benchmarking Pose Estimation for Robot Manipulation[J], Robotics and Autonomous Systems, 2021, 143:103810.

[6] D. Wang, W. Li, X. Liu, N. Li, C. Zhang, UAV Environmental Perception and Autonomous Obstacle Avoidance: A Deep Learning and Depth Camera Combined Solution[J], Computers and Electronics in Agriculture, 2020, 175:105523.

[7] L. Xia, J. K. Aggarwal, Spatio-temporal Depth Cuboid Similarity Feature for Activity Recognition Using Depth Camera[C], Conference on Computer Vision and Pattern Recognition, IEEE, 2013:2834-2841.

[8] H. Tian, Zhao, N. Song, Li, King, Ngi, et al, 3-D Reconstruction of Human Body Shape from a Single Commodity Depth Camera[J], Transactions on Multimedia, IEEE, 2018:114-123.

[9] R. Y. Tsai, R. K. Lenz, A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration[J], Transactions on Robotics & Automation, IEEE, 1989:345-358.

[10] J. C. K. Chou, M. Kamel, Finding the Position and Orientation of a Sensor on a Robot Manipulator Using Quaternions[J], International Journal of Robotics Research, 1991:240-254.

[11] A. Malti, J. P. Barreto, Robust Hand-Eye Calibration for Computer aided Medical Endoscopy[C], International Conference on Robotics & Automation, IEEE, 2010:5543-5549.

[12] A. Tabb, K. Yousef, Parameterizations for Reducing Camera Reprojection Error for Robot-World Hand-Eye Calibration[C], RSJ International Conference on Intelligent Robots & Systems, IEEE, 2015:3030-3037.

[13] J. Angeles, G. Soucy, F. Ferrie, On-Line Hand-Eye Calibration[C], 3-D Digital Imaging and Modeling, Second International Conference on Proceedings, 1999:430-436.

[14] C. Reinbacher, M. Rüther, H. Bischof, RoNect: Hand Mounted Depth Sensing Using a Commodity Gaming Sensor[C], International Conference on Pattern Recognition, IEEE, 2014:461-464.

[15] S. Kahn, A. Kuijper, Fusing Real-Time Depth Imaging with High Precision Pose Estimation by a Measurement Arm[C], International Conference on Cyberworlds, IEEE, 2012:256-260.

[16] S. Kahn, D. Haumann, V. Willert. Hand-eye Calibration with a Depth Camera: 2D or 3D?[C], International Conference on Computer Vision Theory & Applications, IEEE, 2014:481-489.

[17] W. K. Dong, J. E. Ha, Hand/Eye Calibration Using 3D-3D Correspondences[J], Applied Mechanics and Materials, 2013:532-535.

[18] S. Fuchs, Calibration and Multipath Mitigation for Increased Accuracy of Time-of-Flight Camera Measurements in Robotic Applications, Master's thesis, TU Berlin, 2012.

[19] N. Andreff, R. Horaud, B. Espiau, On-line Hand-Eye Calibration[C], Second International Conference on 3-D Digital Imaging and Modeling, IEEE, 1999:430-436

[20] Z. Y. Zhang, A Flexible New Technique for Camera Calibration [J], Transactions on Pattern Analysis & Machine Intelligence, IEEE, 2000:1330-1334.

[21] Z. Zhao, Hand-Eye Calibration using Convex Optimization[C], International Conference on Robotics & Automation, IEEE, 2011:2947-2952.