Camera Calibration Deployed in Mobile Robots

Wei-Ting Weng, Han-Pang Huang, and Yu-Lin Zhao

***Abstract*—Camera calibration is the basis of computer vision, and the quality of the calibration results directly determines the accuracy of a computer vision system. In order to make up for the deficiency of existing calibration methods and to improve accuracy, a novel calibration method is proposed, which can calculate the internal, external, and distortion parameters of the camera. Internal parameters and distortion are inherent properties of the camera, which can be easily calculated by package software and do not change. External parameters, however, are contingent on the real world and therefore subject to change. Thus, we designed a self-made mechanism, which works with the robot navigation system, to calculate them. Overall, this method can obtain the spatial relationship between the robot and the camera while improving the performance of the visual servoing system.**

***Index Terms*—*Calibration, Mobile Robots, Robot vision systems.*

I. INTRODUCTION

The camera projects the three-dimensional real world as two-dimensional images. And the key to finding the relationship between the projection and the real world is camera calibration, which is in turn related to many applications, including 3D reconstruction, visual detection, object localization, and camera localization [1, 2]. Camera calibration aims to find a matrix of camera parameters, including the parameters of the camera [3]. Intrinsic parameters model the internal geometry and optical characteristics of the image sensor. Extrinsic parameters are used to measure the position and orientation of the camera with respect to a world coordinate system. And distortion describes the manufacturing accuracy and the deviation created by the assembly process.

Camera calibration methods can be divided into three main categories: traditional camera calibration methods, active-vision-based calibration methods, and camera self-calibration methods.

Traditional camera calibration methods require shots to be taken of plane calibration boards, such as a checkerboard, which are then analyzed; this is the most common method. For example, direct linear transform (DLT) [4] regards the reprojection error as a cost function and minimizes it. The results are less accurate than other methods due to the usage of linear techniques that ignore distortion. Another example is the two-stage technology [5], which is based on DLT. Tsai [5] considers linear optimization and iterative implementation, thereby effectively removing radial distortion effects

The term active-vision-based calibration refers to motion information pertaining to either the camera or the measured object. The paper [6] describes a system in which odometry and camera are simultaneously calibrated to direct the differential drive used by the mobile robot. The calibration process needs wheel encoders, the camera, and several properly taken camera snapshots of a series of known landmarks. Methods in the same category also include Ma Songde's triple orthogonal motion method [7] and Hartley's camera-based pure rotation calibration [8].

Camera self-calibration methods originate from the idea that camera calibration should not be limited to pre-programmed scenes and camera angles. Faugeras [9] proposes a self-calibration method based on the Kruppa equation. This method establishes an equation for the direct relationship between two images; however, long image sequences can make the results difficult to converge. Hartley [8] gives a different method in which all images are taken from the same point in space. This calibration method is limited; however, its accuracy is not high, and its speed is slow.

In addition, camera calibration methods should also be able to adapt to various new application scenarios. Tang et al. [10] propose a simultaneous calibration of odometry and camera for a differential drive mobile robot that only needs onboard measurements taken by wheel encoders, and snapshots of a set of known landmarks taken by an on-board camera. The two-wheel differential-driven mobile robot with a fixed camera obtains the extrinsic parameter by observing several visual landmarks and recording the measurements from both encoders. Most vision-based industrial robots use eye-in-hand, and paper [11] proposes a method to compute the transformation matrix between the end-effector and camera. It uses the pure translation of the end-effector both to get the rotation matrix and to find the translation matrix between the end-effector and camera. Another eye-in-hand calibration method [12] utilizes a chessboard and transformation to get the extrinsic parameter; a mean filter is then used to decrease the error.

The paper presents a novel camera calibration mechanism, integrated into the mobile robot system, for improving the performance of the visual servoing system with advanced mobile robots.

II. CAMERA MODEL

The concept of the camera model comes from pinhole imaging. The image is obtained by reflecting the light off an object and passing it through a pinhole, although the obtained

High appreciation goes to the Ministry of Science and Technology (MOST) of Taiwan (Grant No. MOST 110-2221-E-002-110-MY3) for the financial support.

W. T. Weng is with the Mechanical Engineering Department, National Taiwan University, Taipei, Taiwan, 10617, R.O.C. (e-mail: r08522802@ntu.edu.tw).

H. P. Huang is with the Mechanical Engineering Department, National Taiwan University, Taipei, Taiwan, 10617, R.O.C. (e-mail: hanpang@ntu.edu.tw). Correspondence addressee.

Y. L. Zhao with the Mechanical Engineering Department, National Taiwan University, Taipei, Taiwan, 10617, R.O.C. (e-mail: f07522844@ntu.edu.tw).

image will be upside down. To make the mathematical formula simpler, the pinhole camera model is organized into another equivalent form. That is, the pinhole position is exchanged with the image plane, so the obtained image is the same as the original one. A schematic diagram of the entire camera model [13] is shown in Fig. 1. The world coordinate system $\{F_{world}\}$ is converted to the camera coordinate system $\{F_{camera}\}$ through extrinsic parameters, and the intrinsic parameters are used to convert the camera coordinate system to the pixel coordinate system $\{F_{pixel}\}$ through the image coordinate system $\{F_{image}\}$. The pinhole point is reinterpreted as the projection center, and the principal point (c_x, c_y) is the intersection of the image plane and the optical axis z.

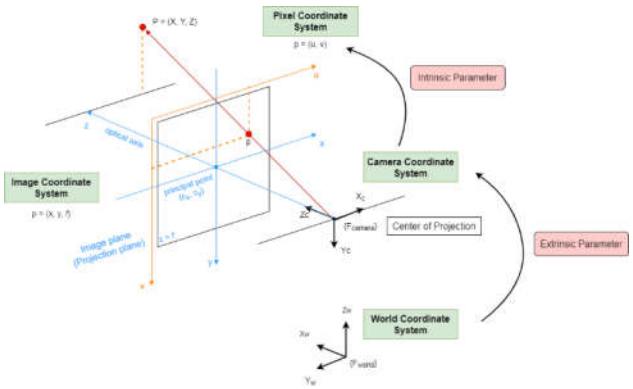

Fig. 1. Camera Model [13].

As shown in Fig. 1. above, a three-dimensional observation point is converted from the world coordinate system to the camera coordinate system through translation and rotation whose process is shown in equation (1). Using homogeneous coordinates to represent three-dimensional coordinate points can combine matrix addition and multiplication.

$$P_{Camera} = \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3\times1} & T_{3\times1} \\ 0 \;\; 0 \;\; 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = M_1 P_{World} \tag{1}$$

The use of similar triangles to project the points of the camera coordinate system onto the image coordinate system is shown in equation (2). Moreover, the equation (3) indicates that the form of matrix multiplication is transformation from the camera coordinate system to the image coordinate system.

$$x = f\frac{X_c}{Z_c}, \quad y = f\frac{Y_c}{Z_c} \tag{2}$$

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{3}$$

However, what we want are the pixels in the pixel coordinate system in camera applications. Therefore, the image needs to be sampled and quantized on the imaging plane. We can transform the points of the image coordinate system to the pixel coordinate system by scaling and translation. The methods for dividing the (x, y) of the image coordinates by the actual length of each pixel (dx, dy) and for moving the original corresponding image origin to the pixel origin (u_0, v_0) are shown in equation (4). Moreover, equation (5) indicates that the form of matrix multiplication is transformation from the image coordinate system to the pixel coordinate system.

$$u = \frac{x}{dx} + u_0, \; v = \frac{y}{dy} + v_0 \tag{4}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

The mathematical formula for the complete camera model is shown in equation (6). When pixel coordinates in the pixel coordinate system, u and v, are non-orthogonal, the term $-\frac{f\cot\theta}{dx}$ is used, otherwise it will be zero.

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \frac{1}{Z_c} \begin{bmatrix} \frac{f}{dx} & -\frac{f\cot\theta}{dx} & u_0 & 0 \\ 0 & \frac{f}{dy\sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3\times1} & T_{3\times1} \\ 0 \;\; 0 \;\; 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \\ &= \frac{1}{Z_c} \begin{bmatrix} f_x & -\frac{f\cot\theta}{dx} & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3\times1} & T_{3\times1} \\ 0 \;\; 0 \;\; 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = M_2 M_1 \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = M \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \end{aligned} \tag{6}$$

With these parameters of the camera model, the pixel position can be transformed to obtain the world coordinate of the target point. The use of a lens, instead of an ideal pinhole, can further enhance the light source, but it also brings image distortion. There are two main types of distortion: radial distortion and tangential distortion.

A. Radial distortion

This type of distortion is mainly caused by greater bending when the light passes through the edge of the lens. Fig. 2. shows the two types of radial distortion, and it follows from equation (7) that the image is increasingly distorted as the distance r is farther from the center.

$$\begin{aligned} x_{distorted} &= x\left(1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) \\ y_{distorted} &= y\left(1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) \end{aligned} \tag{7}$$

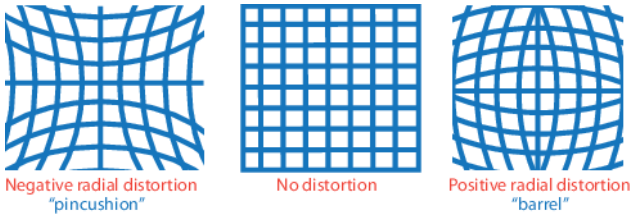

Fig. 2. Type of radial distortion [14].

B. Tangential distortion

This kind of distortion is due to a manufacturing defect in the lens whereby the lens and the image plane are not completely parallel. Fig. 3. shows the type of tangential distortion and equation (8) indicates the correction method.

$$\begin{aligned} x_{distorted} &= x + \left[2p_1xy + p_2\left(r^2 + 2x^2\right) \right] \\ y_{distorted} &= y + \left[p_1\left(r^2 + 2y^2\right) + 2p_2xy \right] \end{aligned} \tag{8}$$

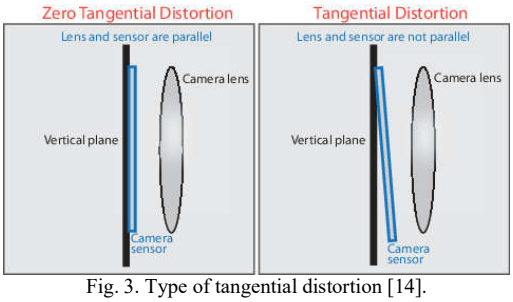

Fig. 3. Type of tangential distortion [14].

Therefore, when the intrinsic parameters are corrected, in addition to the geometric model of the camera, a 5×1 distortion vector will also be obtained.

III. Intrinsic Parameters

The geometric model of the camera and the distortion model of the lens can be obtained through the intrinsic parameters of the camera. Any surface that is flat and has a regular pattern can be used as a calibration mark, and we use the calibration marker as shown in Fig. 4. in this paper.

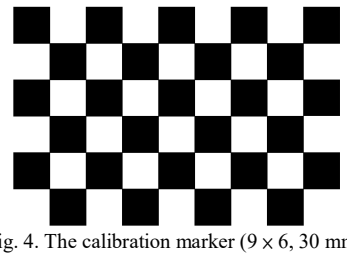

Fig. 4. The calibration marker (9×6, 30 mm).

Planar homography shown in Fig. 5. is used to map the object plane of three-dimensional homogeneous coordinates to the image plane, and $Z = 0$ is selected to define the object plane to increase the generality of the application. Although the planar homography matrix has nine elements, it actually only has eight degrees of freedom plus one constraint, as shown in equation (9). Thus, we can use four points that are not collinear with any three points to provide eight equations to solve H. Furthermore, each new frame provides eight equations and leaves six new unknowns including three rotations and three translations, and the remaining two equations can be used to solve the unknowns of the intrinsic parameters. In addition to the preset mean square error minimization, OpenCV also provides three fitting methods: Random Sample Consensus (RANSAC), Least Median of Squares (LMeDS), and progressive sample consensus (PROSAC). RANSAC calculates a variety of random samples and keeps the sample with the most normal values, LMeDS minimizes the median, and PROSAC is weighted RANSAC.

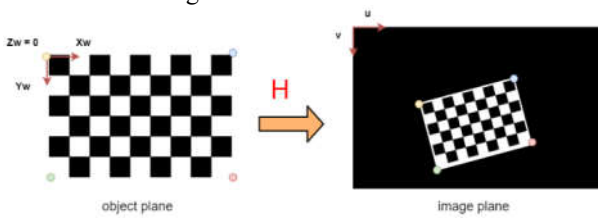

Fig. 5. Planar homography from object plane to image plane.

$$\boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{9}$$
$$h_{11}^2 + h_{12}^2 + h_{13}^2 + h_{21}^2 + h_{22}^2 + h_{23}^2 + h_{31}^2 + h_{32}^2 + h_{33}^2 = 1$$

As discussed above, we can use a planar homography matrix to represent the calculation of the intrinsic and the extrinsic parameter matrices, as shown in equation (10). Thus, where $\boldsymbol{H} = \left[\vec{h}_1 \quad \vec{h}_2 \quad \vec{h}_3 \right] = s \cdot \boldsymbol{M} \left[\vec{r}_1 \quad \vec{r}_2 \quad \vec{t} \right]$ $\boldsymbol{q}$, $\boldsymbol{Q}$, $\vec{r}_i$, $\vec{t}$, $\vec{h}_i \in \mathbb{R}^{3\times1}$.

$$\boldsymbol{q} = s \cdot \boldsymbol{H} \cdot \boldsymbol{Q} = s \cdot \boldsymbol{M} \cdot [\boldsymbol{R} | \boldsymbol{T}] \cdot \boldsymbol{Q}, \; Z = 0$$

$$\Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \cdot \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = s \cdot \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{r}_1 & \vec{r}_2 & \vec{t} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{10}$$

The movement of the camera in space is defined by six extrinsic parameters, while the four intrinsic parameters, belong to the physical properties of the lens, are invariant. And each checkerboard frame can only provide four corner points, which can introduce eight equations. Therefore, in the case of K frames, the solution conditions are $2 \cdot 4 \cdot K \geq 6 \cdot K + 4 \Rightarrow K > 2$. This means that at least two 3×3 checkerboard images are required to solve the calibration problem. Using a larger checkerboard or collecting more images can reduce noise and improve numerical stability.

According to $\boldsymbol{H} = \left[\vec{h}_1 \quad \vec{h}_2 \quad \vec{h}_3 \right] = s \cdot \boldsymbol{M} \left[\vec{r}_1 \quad \vec{r}_2 \quad \vec{t} \right]$, the decomposition equation can be obtained in equation (11).

$$\begin{aligned} \vec{h}_1 &= s \cdot \boldsymbol{M} \cdot \vec{r}_1 \\ \vec{h}_2 &= s \cdot \boldsymbol{M} \cdot \vec{r}_2 \\ \vec{h}_3 &= s \cdot \boldsymbol{M} \cdot \vec{t} \end{aligned} \Rightarrow \begin{aligned} \vec{r}_1 &= \frac{1}{s} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_1 \\ \vec{r}_2 &= \frac{1}{s} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_2 \\ \vec{t} &= \frac{1}{s} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_3 \end{aligned} \tag{11}$$

Since the rotation vector is orthogonal, the inner product of the vectors is zero and the size of the vectors is equal. Through these two constraints, the equation can be listed, as shown in equation (12).

$$\begin{aligned}&(1)\ \vec{r}_1^{\,T} \cdot \vec{r}_2 = 0 \\ &(2)\ \left\|\vec{r}_1\right\| = \left\|\vec{r}_2\right\|\end{aligned} \Rightarrow \begin{aligned}&\vec{h}_1^{\,T} \cdot \boldsymbol{M}^{-T} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_2 = 0 \\ &\vec{h}_1^{\,T} \cdot \boldsymbol{M}^{-T} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_1 = \vec{h}_2^{\,T} \cdot \boldsymbol{M}^{-T} \cdot \boldsymbol{M}^{-1} \cdot \vec{h}_2\end{aligned} \quad (12)$$

$$\Rightarrow \begin{aligned}&\vec{h}_1^{\,T} \cdot \boldsymbol{B} \cdot \vec{h}_2 = 0 \\ &\vec{h}_1^{\,T} \cdot \boldsymbol{B} \cdot \vec{h}_1 = \vec{h}_2^{\,T} \cdot \boldsymbol{B} \cdot \vec{h}_2\end{aligned}$$

where

$$\boldsymbol{B} = \boldsymbol{M}^{-T} \cdot \boldsymbol{M}^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x^2} & 0 & -\frac{u_0}{f_x^2} \\ 0 & \frac{1}{f_y^2} & -\frac{v_0}{f_y^2} \\ -\frac{u_0}{f_x^2} & -\frac{v_0}{f_y^2} & \left(\frac{u_0}{f_x^2} + \frac{v_0}{f_y^2} + 1\right) \end{bmatrix}$$

is a symmetric matrix. Therefore, we can write $\vec{h}_i^{\,T} \cdot \boldsymbol{B} \cdot \vec{h}_j$ in the form of a six-dimensional vector inner product as shown in equation (13). Furthermore, we can use equation (13) to rewrite equation (12) into matrix form as shown in equation (14) where $\mathbf{A} \in \mathbb{R}^{2K \times 6}$, K is the number of the frame. When $K \geq 2$, the equation (14) can be solved for $\vec{\boldsymbol{b}}$, so the intrinsic parameters can be obtained by the closed solution of the $\mathbf{B}$ matrix as shown in equation (15) [15].

$$\begin{aligned}&\vec{h}_i^{\,T} \cdot \boldsymbol{B} \cdot \vec{h}_j = \vec{a}_{i,j}^{\,T} \cdot \vec{b} \\ &= \begin{bmatrix} h_{i,1}h_{j,1} & \left(h_{i,1}h_{j,2} + h_{i,2}h_{j,1}\right) & h_{i,2}h_{j,2} & \left(h_{i,3}h_{j,1} + h_{i,1}h_{j,3}\right) & \left(h_{i,3}h_{j,2} + h_{i,2}h_{j,3}\right) & h_{i,3}h_{j,3} \end{bmatrix} \cdot \begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}\end{aligned} \quad (13)$$

$$\mathbf{A} \cdot \vec{\boldsymbol{b}} = \begin{bmatrix} a_{1,2}^{\,T} \\ \left(a_{1,1} - a_{2,2}\right)^T \end{bmatrix} \cdot \vec{\boldsymbol{b}} = \mathbf{0} \quad (14)$$

$$\begin{aligned} f_x &= \sqrt{\frac{1}{s \cdot B_{11}}} \\ f_y &= \sqrt{\frac{B_{11}}{s \cdot \left(B_{11}B_{22} - B_{12}^{\,2}\right)}} \\ u_0 &= s \cdot B_{13} f_x^{\,2} \\ v_0 &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^{\,2}} \\ s &= \frac{1}{B_{33} - \frac{\left(B_{13}^{\,2} + v_0\left(B_{12}B_{13} - B_{11} - B_{23}\right)\right)}{B_{11}}} \end{aligned} \quad (15)$$

Then, using [16], the distortion coefficients of equations (7) and (8) can be found.

Generally, when calibrating intrinsic parameters, there is an image at each angle, as shown in Figure 6. However, because the camera to be calibrated in this thesis is fixed on the chest of the robot, and the robot itself does not have the freedom to move up and down, we cannot get images like number five and number six. Therefore, a self-made mechanism with up and down movement (shown in Fig.7) is used to match the requirements and allow the robot to complete the calibration of the intrinsic parameters. The use of the self-made mechanism and robot navigation system to set up the calibration process is shown in Fig. 8. Because images of the chessboard taken from different views are required, the five points in Fig. 8 are made twice, one is when the putter is in the low state, and the other is when the putter is in the high state. Fig. 9 shows the image captured by the camera on the robot for intrinsic parameter calibration and finally using ten chessboard images.

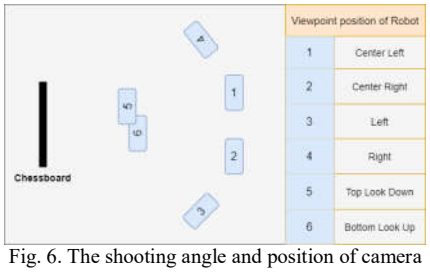

Fig. 6. The shooting angle and position of camera


Fig. 7. A self-made mechanism with up and down movement.

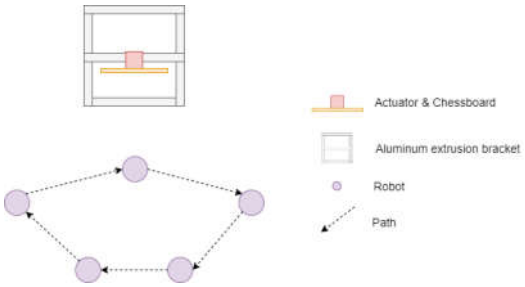

Fig. 8. Top view of equipment setup.

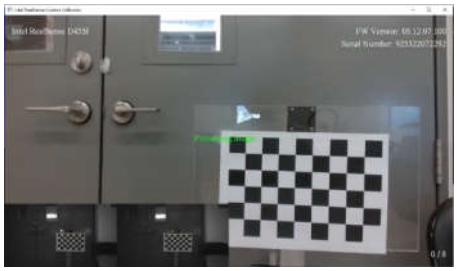
Fig. 9. The image captured by the camera on the robot for intrinsic.

Finally, the average result is calculated through intrinsic parameter results, as shown in equation (16).

$$\boldsymbol{M}_{avg} = \frac{\sum_{0}^{n} \boldsymbol{M}_i}{n}, \text{ where } \boldsymbol{M} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

IV. Extrinsic Parameters

The extrinsic parameters are mainly used to solve the transformation relationship between the camera coordinate system and the robot base coordinate system, and this relationship can then be used to transform the feature points seen by the camera into an orientation that the robot can understand, or transformed into the map coordinate system. In this paper, we use the addition of two space vectors as shown in equation (17) to obtain the relationship between the camera coordinate system and the base coordinate system as shown in Fig. 10. First, the RGB-D camera plus the previously calibrated intrinsic parameters are used to obtain the vector from the camera coordinate system to the corners of the chessboard, and then the geometric relationship of the mechanism, the robot odometry and 2D map (shown in Fig. 11) to obtain the vector from the base coordinate system to the corners of the chessboard.

$$\boldsymbol{V}_{Base\rightarrow Camera} = \boldsymbol{V}_{Base\rightarrow Chessboard} + \boldsymbol{V}_{Chessboard\rightarrow Camera} \quad (17)$$

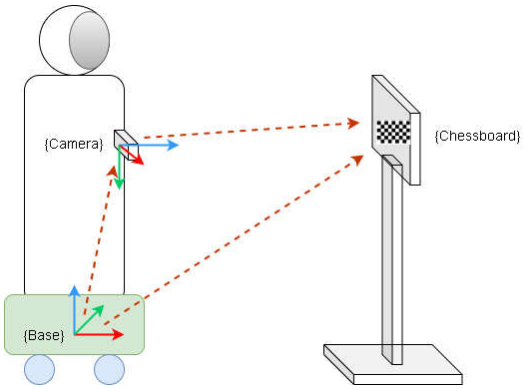

Fig. 10. Schematic diagram of the transformation between camera coordinates and robot coordinates.

Fig. 11. 2D Map for camera calibration.

In order to improve the accuracy, we take six corners on the chessboard to average like number seven, fourteen, seventeen, twenty-four, twenty-seven and thirty-four as shown in Fig. 12.

The translation in the transformation relationship can be obtained in equation (18). Because of the geometric design of the robot, the X axis of the base will be parallel to the z axis of the camera. Therefore, we obtain the rotation relationship between the base and the camera as shown in equation (19). Finally, a more accurate result is obtained by averaging over n iterations of the experiment as shown in equation (20).

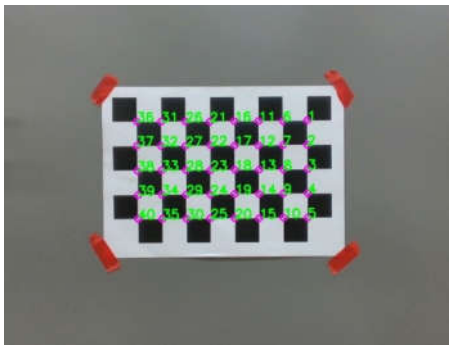

Fig. 12. The corners on the chessboard.

$${}^{Base}t_{Base\rightarrow Camera} = \begin{bmatrix} {}^{Base}X_{Base\rightarrow Chessboard} + {}^{Base}x_{Chessboard\rightarrow Camera} \\ {}^{Base}Y_{Base\rightarrow Chessboard} + {}^{Base}y_{Chessboard\rightarrow Camera} \\ {}^{Base}Z_{Base\rightarrow Chessboard} + {}^{Base}z_{Chessboard\rightarrow Camera} \end{bmatrix} \quad (18)$$

$${}^{Base}R_{Camera} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \quad (19)$$

$$\left({}^{Base}T_{Camera}\right)_{avg} = \begin{bmatrix} {}^{Base}R_{Camera} & \frac{\sum_{i}^{n}\left({}^{Base}t_{Base\rightarrow Camera}\right)_i}{n} \end{bmatrix} \quad (20)$$

V. Experiment Results

A. Introduction to Mobile Robot

Our laboratory developed a novel mobile robot called Mobi which can be divided into four parts: head, upper body, lower body, and hands and arms, as shown in Fig. 13. The hardware architecture of new mobile system is shown in Fig. 14.

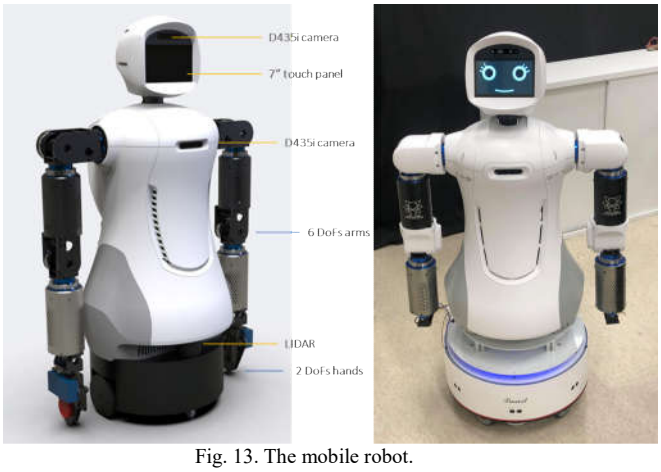

Fig. 13. The mobile robot.

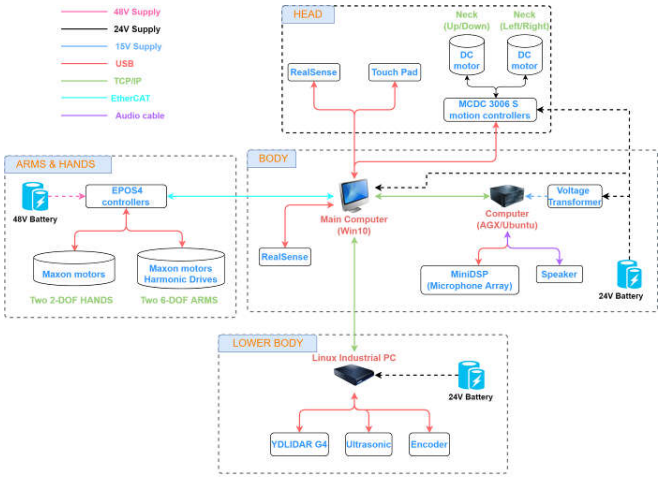

Fig. 14. The hardware architecture of new mobile system.

B. Introduction to the Calibration Mechanism

In the self-made mechanism, we use an Arduino as the control board, an L298N as the motor drive board, and the Wi-Fi module (ESP-01) for communication. The entire circuit design is shown in Fig. 15. The control process is the following: The Wi-Fi module receives the signal and then transmits it to the Arduino, which then controls the L298N, making the linear actuator rise or fall. To achieve this, the main program sends the Arduino ascending and descending commands through HTTP as shown in Fig. 16.

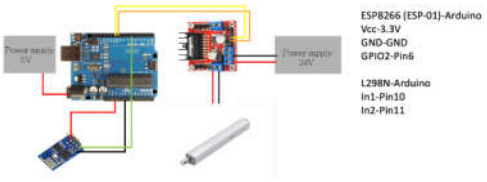

Fig. 15. The entire circuit design for camera calibration.

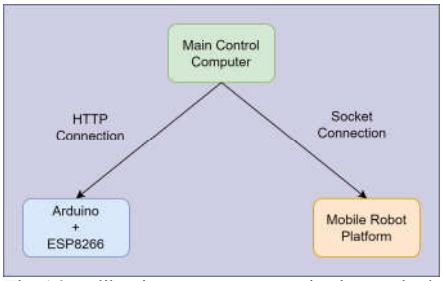

Fig. 16. Calibration system communication method.

C. Deployment

1) Intrinsic parameters

In accordance with the above, the robot is instructed to visit the five points on the map to obtain images of the checkerboard from different heights, and then use ten images to complete the calibration of the internal parameters as shown in Fig. 17. It collects fifteen experimental data points as shown in Fig. 18, and then averages the results by removing outliers through standardization methods. The average and standard deviation of the original data is used to calculate the standardized score of the data, as shown in equation (21). The data with scores higher than 1 or lower than -1 are then removed, and the average of the remaining data is calculated as the results shown in Table I.

$$score = \frac{d_i - \frac{\sum_{i=1}^{n} d_i}{n}}{\sqrt{\frac{\sum_{i=1}^{n} (d_i - \frac{\sum_{i=1}^{n} d_i}{n})}{n-1}}} \tag{21}$$

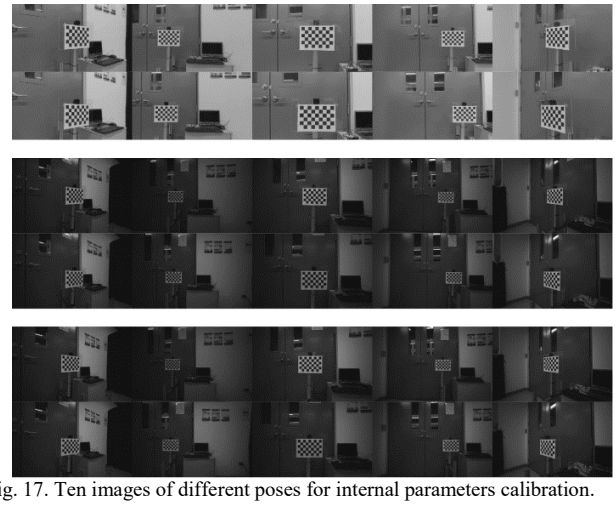

Fig. 17. Ten images of different poses for internal parameters calibration.

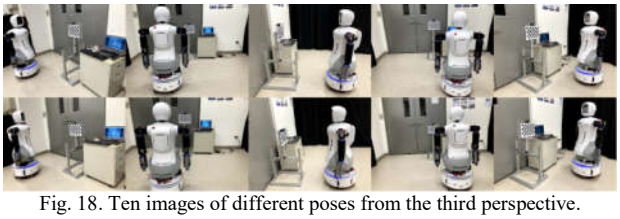

Fig. 18. Ten images of different poses from the third perspective.

Table I
RESULTS OF INTRINSIC PARAMETERS

	RGB	Left	Right
f_x	913.9624	638.4091	638.5296
f_y	913.9624	638.4091	638.5296
u_0	636.0864	641.2085	652.1973
v_0	365.1636	414.5412	408.8523
k_1	-0.0356	-0.0986	-0.0563
k_2	1.088	0.4083	0.0061
p_1	0.0041	0.0038	0.0029
p_2	-0.0014	-0.0026	-0.0008
k_3	-3.5439	-0.8484	0.1913
Transformation			

$$^{Left}T_{RGB} = \begin{bmatrix} 0.984146 & 0.000972 & 0.177357 & 15.8159 \\ 0.013564 & 0.996644 & -0.080725 & -0.0732 \\ -0.176841 & 0.081851 & 0.980830 & -0.5748 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$ {}^{Left}T_{Right} = \begin{bmatrix} 0.993642 & 0.000452 & -0.112589 & -50.4338 \\ -0.007121 & 0.998242 & -0.058841 & 0.1937 \\ 0.112364 & 0.059269 & 0.991898 & 3.4046 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

2) Extrinsic parameters

This set of intrinsic parameters is used to further solve the extrinsic parameters which specify the relationship between the camera and the base through the vector transformation with corner detection, depth camera, known mechanism geometry, and map. The standardization method is used to remove outliers in the data, just as in the method for solving extrinsic parameters, and the corner images are shown in Fig. 19. The result obtained by averaging away the outliers is $T_{\text{measure}} = [0.156408 \quad 0.041839 \quad 0.798142]^T$, and the data are shown in Fig. 20. Finally, the relationship transformation from the base to the camera is shown in equation (22), which verifies the repeatability.

To validate the accuracy of the calibration results, we also use the ruler to indirectly calculate the camera extrinsic parameter $T_{\text{measure}} = [0.1505 \quad 0.0330 \quad 0.8030]^T$. The absolute error is less than 3 mm and the relative error is about 0.29%. The old method was to use CAD to measure parameters, however, the relative error is over 1%. Thus, the mobile robot has a more stable grasp of regular directional objects.

$$ T = \begin{bmatrix} 0 & 0 & 1 & 0.156408 \\ -1 & 0 & 0 & 0.041839 \\ 0 & -1 & 0 & 0.798142 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{22} $$

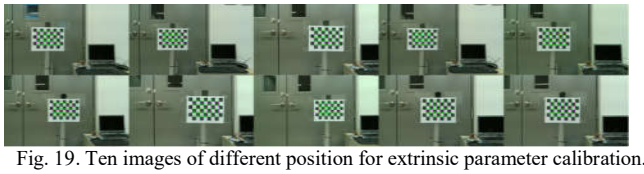

Fig. 19. Ten images of different position for extrinsic parameter calibration.

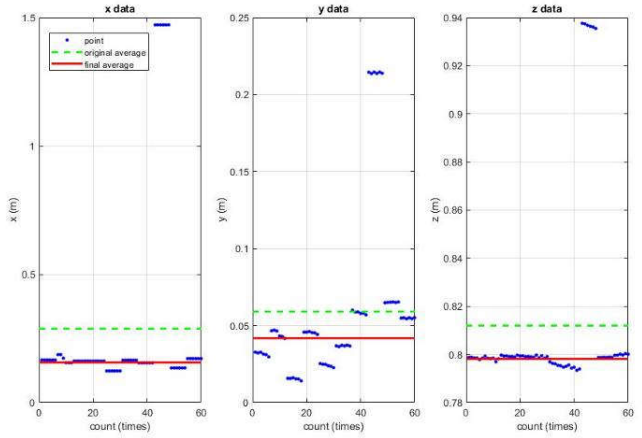

Fig. 20. Translation data (blue is data point, green is original data average, and red is final average).

VI. Conclusions and Future Work

This paper provides solutions for the camera's internal and external parameters. This method can calculate the parameters by inputting multiple checkerboard images and adding robot odometry. Without the need for an expensive professional calibration board, the method uses data filtering and standardization to obtain the relevant transformation matrix with less error. Thus, the experimental results proved that this method improves the accuracy of robot tracking and grasping compared with the past methods. In the future, follow-up work related to the data process can be added; for example, the maximum likelihood method could be used to improve the estimation accuracy.

Acknowledgments

High appreciation goes to the Ministry of Science and Technology (MOST) of Taiwan (Grant No. MOST 110-2221-E-002-110-MY3) for the financial support.

References

[1] L. R. Ramírez-Hernández, J. C. Rodríguez-Quiñonez, M. J. Castro-Toscano, D. Hernández-Balbuena, W. Flores-Fuentes, R. Rascón-Carmona, L. Lindner, and O. Sergiyenko, "Improve Three-Dimensional Point Localization Accuracy in Stereo Vision Systems Using a Novel Camera Calibration Method," *International Journal of Advanced Robotic Systems,* vol. 17, no. 1, p. 1729881419896717, 2020.

[2] I. Enebuse, M. Foo, B. S. K. K. Ibrahim, H. Ahmed, F. Supmak, and O. S. Eyobu, "A Comparative Review of Hand-Eye Calibration Techniques for Vision Guided Robots," *IEEE Access,* vol. 9, pp. 113143-113155, 2021.

[3] L. Long and S. Dongri, "Review of Camera Calibration Algorithms," *Advances in Computer Communication and Computational Sciences*, Singapore, S. K. Bhatia, S. Tiwari, K. K. Mishra, and M. C. Trivedi, Eds.: Springer Singapore, pp. 723-732, 2019.

[4] Y. I. Abdel-Aziz, H. M. Karara, M. J. P. e. Hauck, and r. sensing, "Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry," vol. 81, no. 2, pp. 103-107, 2015.

[5] R. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3d Machine Vision Metrology Using Off-the-Shelf Tv Cameras and Lenses," *IEEE Journal on Robotics and Automation,* vol. 3, no. 4, pp. 323-344, 1987.

[6] G. Antonelli, F. Caccavale, F. Grossi, and A. Marino, "Simultaneous Calibration of Odometry and Camera for a Differential Drive Mobile Robot," *2010 IEEE International Conference on Robotics and Automation*: IEEE, pp. 5417-5422, 2010.

[7] M. Sang De, "A Self-Calibration Technique for Active Vision Systems," *IEEE Transactions on Robotics and Automation,* vol. 12, no. 1, pp. 114-120, 1996.

[8] R. I. Hartley, "Self-Calibration of Stationary Cameras," *International Journal of Computer Vision,* vol. 22, no. 1, pp. 5-23, 1997.

[9] O. D. Faugeras, "What Can Be Seen in Three Dimensions with an Uncalibrated Stereo Rig?," *Computer Vision — ECCV'92*, Berlin, Heidelberg, G. Sandini, Ed.: Springer Berlin Heidelberg, pp. 563-578, 1992.

[10] H. Tang, Y. Liu, and L. Li, "Simultaneous Calibration of Odometry and Camera Extrinsic for a Differential Driven Mobile Robot," *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*: IEEE, pp. 2246-2251, 2015.

[11] H. Sung and S. Lee, "A Robot-Camera Hand/Eye Self-Calibration System Using a Planar Target," *IEEE ISR 2013*: IEEE, pp. 1-4, 2013.

[12] K.-H. Lee, H.-S. Kim, S.-J. Lee, S.-W. Choo, S.-M. Lee, and K.-T. Nam, "High Precision Hand-Eye Self-Calibration for Industrial Robots," *2018 International Conference on Electronics, Information, and Communication (ICEIC)*: IEEE, pp. 1-2, 2018.

[13] K. Lelowicz, "Camera Model for Lens with Strong Distortion in Automotive Application," *2019 24th International Conference on*

Methods and Models in Automation and Robotics (MMAR): IEEE, pp. 314-319, 2019.
[14] "What Is Camera Calibration?" MATLAB. 2021. https://www.mathworks.com/help/vision/ug/camera-calibration.html?w.mathworks.com (accessed.
[15] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 22, no. 11, pp. 1330-1334, 2000.
[16] C. B. Duane, "Close-Range Camera Calibration," *Photogramm. Eng,* vol. 37, no. 8, pp. 855-866, 1971.

Wei-Ting Weng was born in New Taipei, Taiwan, in 1996. He received the B.E. degree in department of mechanical engineering from the National Chung Hsing University, Taiwan, in 2019. He completed the M.S. degree in department of mechanical engineering the National Taiwan University (NTU), Taiwan, in 2021. He had worked on robotics laboratory, and the research direction is the vision application in wheeled robots.

Han-Pang Huang (S'83–M'86) received Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 1986. Since 1986, he has been with National Taiwan University, Taipei, Taiwan, where he is currently a distinguished professor and Zhong Zhuo-Zhang Chair Professor with the Department of Mechanical Engineering and the Graduate Institute of Industrial Engineering His current research interests include machine intelligence, humanoid robots, intelligent robotic systems, prosthetic hands, manufacturing automation, nanomanipulation, and nonlinear systems. He has authored more than 390 papers on these topics that have been published in refereed technical journals and conference proceedings. He is the Fellows of the Chinese Institute of Automation Engineers, the Chinese Society of Mechanical Engineers, and Robotics Society of Taiwan.

Yu-Lin Zhao was born in Shenyang, China, in 1996. He received the B.E. degree in department of automation engineering from the National Formosa University, Taiwan, in 2018. He is currently direct pursuing the Ph.D. degree in department of mechanical engineering the National Taiwan University (NTU), Taiwan. He is working on robotics, Internet of Things technology and image recognition. He has published 3 papers as the first author and holds a patent.