Camera Calibration

Theory and Application Of Automated Optical Inspection
Final Project

B09502040 Jim Zeng
B08502097 Benny Cheng
B10502098 Suvanna V. W.

ABSTRACT

Camera calibration can be defined as a technique to determine a specific camera parameter to complete a task with a specific performance measurement. Using camera calibration, our experiments aim to practice using camera calibration in daily life while using a camera that is owned by plenty of students (phones). The object we took a photo of is a chessboard-like paper, and the devices used are iPhone 11 and iPhone 13. There will be 68 pictures taken in total; each phone took 34 photos using different camera settings, distances, and angles. By the end of the experiment, a transformation matrix is obtained by using a calibration board and continuously adjusting the camera, allowing the camera to position itself in any situation accurately.

Keywords: Camera calibration, matrix computations

CHAPTER 1

INTRODUCTION

I. BACKGROUND

A. What can Camera Calibration do?

(1) To transform every 3D image to 2D (RGB or grayscale) images

(2) Reconstruct a world and interact with it (e.g., the hand-eye coordination of robots)

B. How to use Camera Calibration?

(1) Given a set of points and their corresponding image coordinates.

(2) Use linear algebra to find the projection matrix

(3) Determine intrinsic and extrinsic parameters and do certain computations.

C. Why Do We Choose Camera Calibration?

(1) We need some parameters to do calculations and analyze the photos.

(2) The parameter comes from the camera itself, but hard to be precisely measured

What is worse, there exists lens distortion, an inevitable disadvantage:
We should do camera calibration first to avoid effects from lens distortion, and therefore, we don't have to measure most of the parameters.

D. The choice of Camera Calibration method?

We choose Zhang's method as our approach, and our reference paper is "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations," written by Zhengyou Zhang.

E. The math tool of camera calibration

We use **linear algebra** to solve equations related to camera calibration :

(a). 3D rotation matrix:

$$\begin{bmatrix} cos\theta & -sin\theta & 0 \\ sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

What's more, the rotation matrix will always be an orthogonal matrix.

(b). Properties of the orthogonal matrix :

Theorem 6.9

The following conditions about an $n \times n$ matrix Q are equivalent:

- (a) Q is orthogonal.
- (b) $Q^T Q = I_n$.
- (c) Q is invertible and $Q^{-1} = Q^T$.

(d) Dot product of any two (column) vectors in an orthogonal matrix always be zero. Also, each vector's norm (length) = 1.

(c). Transform matrix from the 3D world coordinates to the 3D camera coordinate:

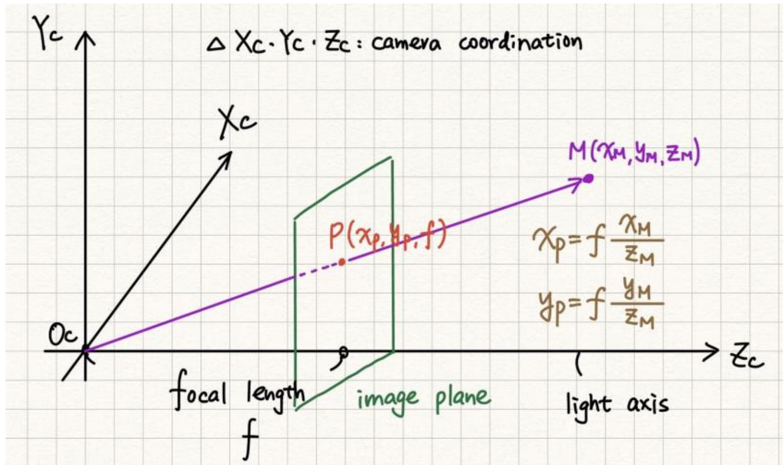
$$\begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix}$$

(1). This transform matrix is the combination of the rotation matrix and translate matrix.

(2). It is also an extrinsic parameter matrix (represents orientation and position) of the camera.

F. The Calculation of Camera Calibration

I. From the 3D camera coordinate to the 2D image coordinate



We can rewrite it as homogeneous equations:

$$z_M \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_M \\ y_M \\ z_M \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_M \\ y_M \\ z_M \\ 1 \end{pmatrix}$$

II. from 2D image coordinates to 2D memory coordinate

Definitions and an assumption:

1. image coordinate (x,y) and memory coordinate (u,v)
2. Each memory(pixel) point represents (d_x,d_y) in image coordinate
3. The memory point we are interested in is located at (x_c,y_c)
4. Assume the origin point in the image coordinate is located at (u_0,v_0) in the memory coordinate.

So, we have:

$$u = \frac{x_c}{d_x} + u_0 \text{、} v = \frac{y_c}{d_y} + v_0$$

Again, rewrite them as a matrix:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}$$

III. Intrinsic parameter matrix M (without distortion):

Product of the two previous matrices :

$$M = \begin{bmatrix} f/d_x & 0 & u_0 \\ 0 & f/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

*The complete equation that represents the transformation from the 3D world coordinates to the 2D memory coordinate is shown below(s is a certain size factor corresponding to the target):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}$$

*If consider distortion, M, x_p and y_p will be like:

$$M \text{ w/ distortion: } \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{cases} x_{corr} = x_p(1+k_1r^2+k_2r^4+k_3r^6) \\ y_{corr} = y_p(1+k_1r^2+k_2r^4+k_3r^6) \end{cases}$$

<u>In sum</u>:

(x_w, y_w, z_w) 3D world coordinate

↓

Step 1: Rigid body transformation: R, T

↓

(x, y, z) 3D camera coordinate

↓

Step 2: Pin hole perspective projection: f

↓

(X_u, Y_u) Undistorted image coordinate

↓

Step 3: Radial lens distortion: k_1, k_2

↓

(X_d, Y_d) Distorted image coordinate

↓

Step 4: Scanning, frame grabber sampling: s_x

↓

(X_f, Y_f) Frame memory coordinate

$$\begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Furthermore, define a symmetric matrix B which is also an orthogonal matrix

$$B = (M^{-1})^T M^{-1} = \begin{bmatrix} \frac{1}{f_x^2} & -\frac{\gamma}{f_x^2 f_y} & \frac{v_0\gamma - u_0 f_y}{f_x^2 f_y} \\ -\frac{\gamma}{f_x^2 f_y} & \frac{\gamma^2}{f_x^2 f_y^2} + \frac{1}{f_y^2} & -\gamma\frac{v_0\gamma - u_0 f_y}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} \\ \frac{v_0\gamma - u_0 f_y}{f_x^2 f_y} & -\gamma\frac{v_0\gamma - u_0 f_y}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} & \frac{(v_0\gamma - u_0 f_y)^2}{f_x^2 f_y^2} + \frac{v_0^2}{f_y^2} + 1 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

Then solve the previous matrix equation about B by certain calculations.

Restriction: There should be at least 3 different photos in the experiment.

The differences between each photo can be directly decided by personal settings.

The following is detailed proof of the above statements:

Define $H = s\begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}[r_1\ r_2\ t] = sM[r_1\ r_2\ t] = [h_1\ h_2\ h_3] \Rightarrow \begin{cases} h_1 = sMr_1 \\ h_2 = sMr_2 \\ h_3 = sMt \end{cases} \Rightarrow \begin{cases} r_1 = \lambda M^{-1}h_1 \\ r_2 = \lambda M^{-1}h_2 \\ t = \lambda M^{-1}h_3 \end{cases}$ where $\lambda = s^{-1}$

Property:

① $r_1 \cdot r_2$ are column vectors of a rotation matrix which is also an orthogonal matrix

$\because$ inner product of $r_1 \cdot r_2 = r_1^T r_2 = 0 \Rightarrow h_1^T \underline{(M^{-1})^T M^{-1}} h_2 = 0$ Let $B = (M^{-1})^T M^{-1}$

② $\|r_1\| = \|r_2\| = 1$ $\therefore r_1^T r_1 = r_2^T r_2 = 1 \Rightarrow h_1^T \underline{(M^{-1})^T M^{-1}} h_1 = h_2^T \underline{(M^{-1})^T M^{-1}} h_2$

By ①.② $\Rightarrow \begin{cases} h_1^T B h_2 = 0 \\ h_1^T B h_1 = h_2^T B h_2 \end{cases}$

Then define the i-th column vector $h_i = [h_{i1}\ h_{i2}\ h_{i3}]$ and plug in ①.②:

$$h_i^T B h_j = \underbrace{\begin{bmatrix} h_{i1}h_{j1} \\ h_{i1}h_{j2} + h_{i2}h_{j1} \\ h_{i2}h_{j2} \\ h_{i3}h_{j1} + h_{i1}h_{j3} \\ h_{i3}h_{j2} + h_{i2}h_{j3} \\ h_{i3}h_{j3} \end{bmatrix}^T}_{\text{Let} = V_{ij}} \underbrace{\begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}}_{\text{Let} = b} = V_{ij}^T b \ : \ \begin{cases} V_{12}^T b = 0 \\ (V_{11}^T - V_{22}^T) b = 0 \end{cases} \rightarrow \begin{bmatrix} V_{12}^T \\ V_{11}^T - V_{22}^T \end{bmatrix} b = 0$$

There exists $C_2^3 = 3$ equations $\therefore$ At least 3 photos should be taken!

G. Application of Camera Calibration

There are many ways in which camera calibration can be used in daily life. One of the applications includes enabling image correction, which focuses on correcting any imperfectness in one's image.

Fig 1.1

Fig 1.2

In Fig 1.1, the image is taken at an angle that is not straight. Camera calibration enables us to correct the image that is not taken perfectly, such as in Fig 1.1, and change it into a more perfect image, such as Fig 1.2.

Another main application of camera calibration includes changing the visual angle of an image.

Fig 2.1

Fig 2.2

In Fig 2.1, a picture was taken where the floor was below, focusing on many objects. However, using camera calibration, we can change it to focus on the floor, such as in Fig 2.2, by changing the visual angle of the image.

Image correction and changing visual angles are quite similar; however, one focuses on correcting an imperfect image, and one focuses on changing the angle of an image to a desired angle.

CHAPTER 2

EXPERIMENT

II. EXPERIMENT SETTINGS AND PROCEDURE

During the preparation of the experiment, there were a few objects that we needed to prepare. As a device to take a picture of the object, a camera is one of the most important objects needed to prepare; in this experiment, the camera used is iPhone 11 and iPhone 13. Another object is a chessboard that is created using paper (see **Fig 2.1**). This chessboard paper is the object that we are taking pictures of. In every experiment, there are several things that need special attention, and likewise, in this experiment, the focal length needs to remain the same. We took pictures based on several types: types of iPhone, types of camera settings, types of angles, and types of distance.

Fig 2.1

Fig 2.2

The two types of iPhones used in this experiment to take pictures are iPhone 11 and iPhone 13. **Fig 2.1** is one of the pictures taken with iPhone 11, and **Fig 2.2** is one of the pictures taken with iPhone 13.

Fig 2.3

Fig 2.4

Fig 2.5

Fig 2. 6

Fig 2.3 and **Fig 2.4** are pictures taken by iPhone 13, but with different camera settings; they are high-efficiency and high-quality, respectively. **Fig 2. 5** and **Fig 2.6** are pictures taken by iPhone 11, with high-efficiency and high-quality settings, respectively.

The pictures taken by both iPhones are taken from different angles, 8 angles, and 1 front; hence 9 pictures in total from different orientations were taken from each iPhone, each camera setting, and each distance. Left and right bottom (see **Fig 2.7** and **Fig 2.8**), right and left top (see **Fig 2.9** and **Fig 2.10**, respectively), bottom view, top view, and sides view (see **Fig 2.11, Fig 2.12, Fig 2.13** and **Fig 2.14**, respectively), and the last will be the front view (see **Fig 2.15**).

Fig 2.7

Fig 2.8

Fig 2.9

Fig 2.10

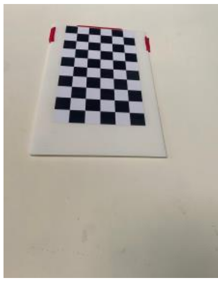




Fig 2.11 **Fig 2.12** **Fig 2.13** **Fig 2.14** **Fig 2.15**

(Fig 2.7 - Fig 2.15 are 8 angles + 1 front (left and right bottom, right left top, bottom, top, sides, front view, respectively) figures taken using iPhone 11 high efficiency)




Fig 2.16 **Fig 2.17** **Fig 2.18** **Fig 2.19**





Fig 2.20 **Fig 2.21** **Fig 2.22** **Fig 2.23** **Fig 2.24**

(Fig 2.16 - Fig 2.24 are 8 angles + 1 front (left and right bottom, right left top, bottom, top, sides, front view, respectively) figures taken using iPhone 11 high quality)


Fig 2.25


Fig 2.26

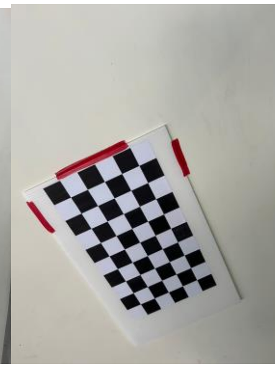


Fig 2.27


Fig 2.28


Fig 2.29


Fig 2.30


Fig 2.31


Fig 2.32


Fig 2.33

(Fig 2.25 - Fig 2.33 are 8 angles + 1 front (left and right bottom, right left top, bottom, top, sides, front view, respectively)figures taken using iPhone 13 high efficiency)

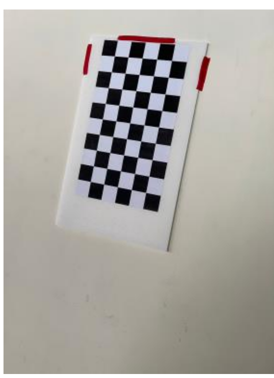


Fig 2.34


Fig 2.35

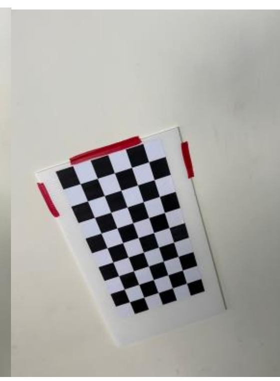


Fig 2.36


Fig 2.37


Fig 2.38


Fig 2.39


Fig 2.40


Fig 2.41


Fig 2.42

(Fig 2.34 - Fig 2.42 are 8 angles + 1 front (left and right bottom, right left top, bottom, top, sides, front view, respectively) figures taken using iPhone 13 high quality)

In the experiment, we also use two types of distance; far and close. We took it from a close-up and a further place for every angle except for the front view. It was also taken by two phones with two types of camera settings. (Examples of the same angles taken with the same iPhone with the same camera settings with different types of distance (**Fig 2.43** is the close-up view, and **Fig 2.44** is the further view)).

Fig 2.43

Fig 2.44

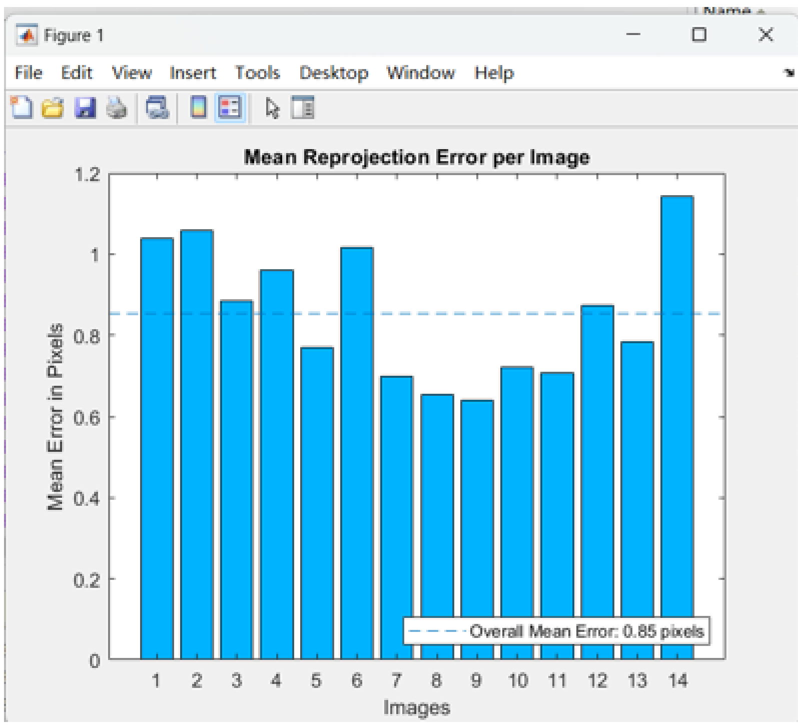
To sum up, this experiment took 68 pictures in total, with 34 photos taken on each phone. In each phone, 17 pictures are taken with high quality, and 17 pictures are taken with high-efficiency settings). In each camera setting, 8 are from a close-up view, 8 angles from a further view, and 1 front view.

CHAPTER 3

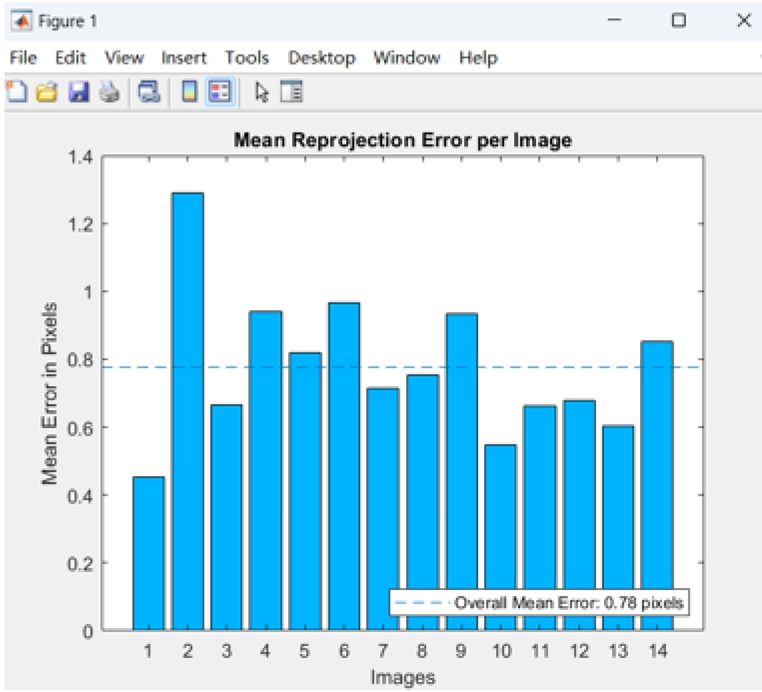
RESULT ANALYSIS AND APPLICATION

III. RESULT ANALYSIS

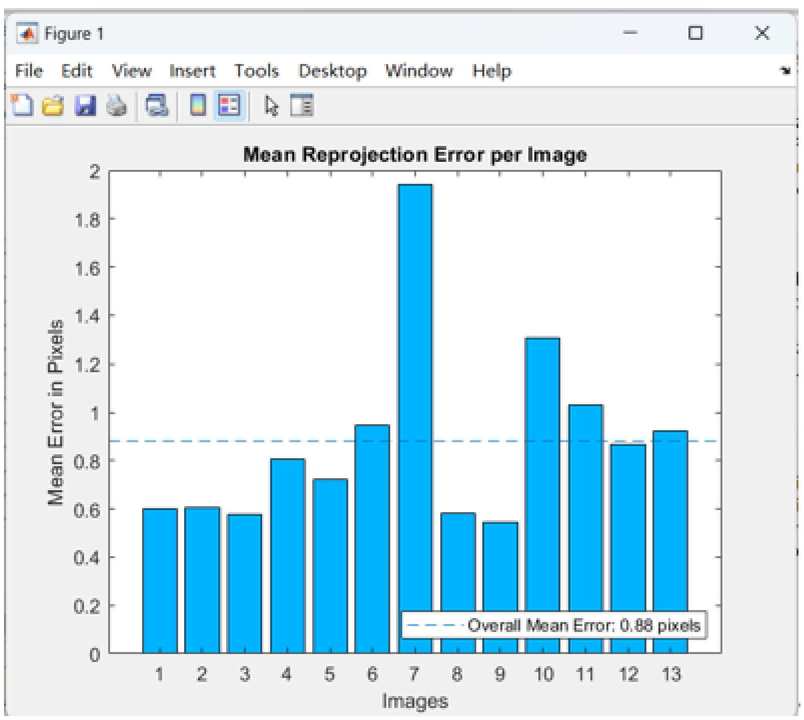
As we shown in the document named “AOI project code”, the mean error in iphone11 or iphone13/high-efficiency mode or high-quality mode is shown orderly below, and they are briefly concluded by the following chart :



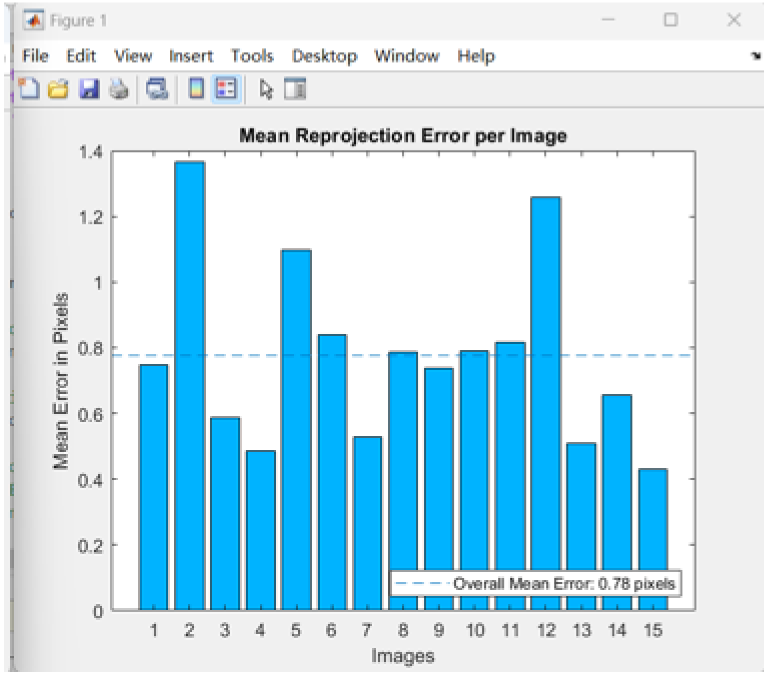
Iphone 11 high-efficiency mode



Iphone 11 high-quality mode



Iphone 13 high-efficiency mode



Iphone 13 high-quality mode

	iphone11	iphone13
High-efficiency mode	0.85	0.88
High-quality mode	0.78	0.78

Analysis conclusion:

Obviously, the error in high-efficiency mode is both larger than that of high-quality mode. This means that high-quality mode can indeed enhance the quality of the photos taken. Also, when it comes to the comparison between iphone11 and iphone13, there's almost no difference under the same shooting mode.

IV. APPLICATION

When it comes to camera calibration, we cannot overlook the crucial aspect of robotic vision. In the era of Industry 4.0, unmanned automation has become a popular trend in manufacturing. However, this has also given rise to a series of preliminary issues, with robotic vision being of utmost importance. If the positioning is inaccurate, it can lead to deviations in the movement of the robotic arm. In the case of processing extremely delicate chips, such deviations in positioning could potentially damage the chips or even the robotic arm itself. Therefore, the calibration between the robotic arm, camera, and objects is exceptionally critical. The following image explains how camera calibration connects a robotic arm and a target object perfectly.

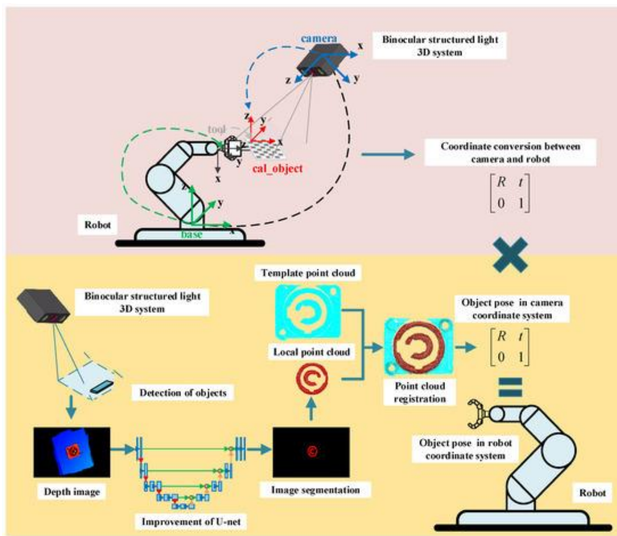


Fig 3.1 Object pose in robot coordinate system

Through the use of a calibration board, the robotic arm and the camera can obtain a coordinate conversion matrix:

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

By aligning the camera with the target of the robotic arm, a transformation matrix can also be obtained.

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

Finally, by multiplying these two transformation matrices, the calibration can be completed, enabling the robotic arm to accurately grasp the target object.

The paper proposes a validation method to detect poor calibration data in the calibration process between the first camera and the robotic arm. Through continuous data updating and debugging, the minimum reprojection error is calculated and stored in the database. By repeating this process, the calibration becomes more accurate. The following figure illustrates the calibration flowchart.

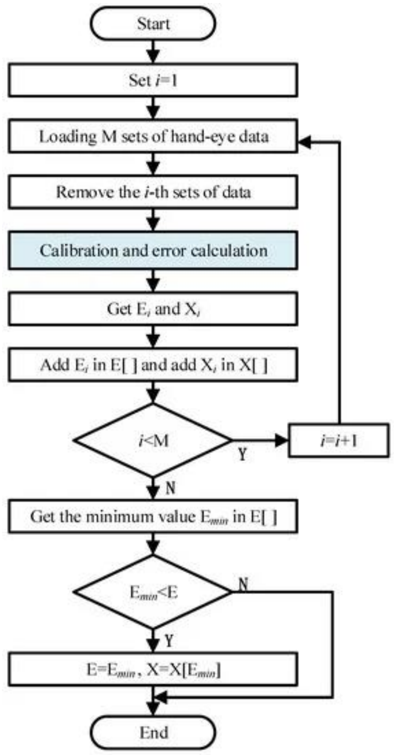


Fig 3.2 Hand–eye calibration optimization process based on data filtering.

Regarding the correction of the camera to the object, this study achieves precise positioning of the target object through point cloud segmentation. Point cloud segmentation involves dividing data points in the original point cloud data into groups with similar and different attributes based on their relevant functional definition. In addition to point cloud segmentation, the study utilizes the U-net architecture to contract the pathway, enabling the capture of more accurate points within a reduced range in the image and enhancing precision.

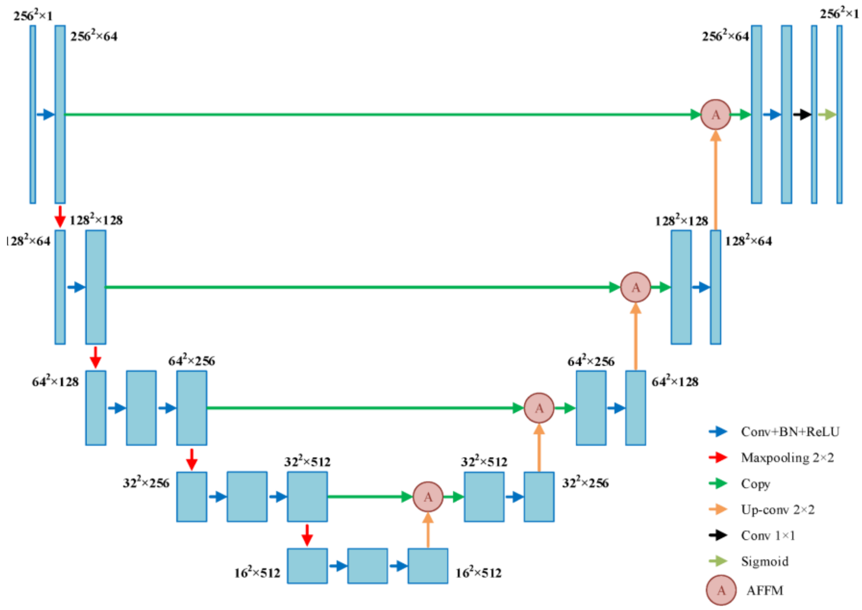


Fig 3.3 Improved U-net structure.

The main reasons for using camera calibration instead of other methods, such as a theodolite or a coordinate measuring machine (CMM), are as follows. Firstly, although the theodolite provides accurate positioning, it is relatively slow in measurement speed. As for the CMM, it has complex calibration mechanisms and a large size, which contrasts with the quick, cost-effective, and flexible nature of optical calibration. Despite the higher precision of the former methods, many people still prefer camera calibration to overcome technical limitations.

Below, the positioning methods between the robotic arm and the camera will be introduced. On the left is the Eye-to-Hand configuration, which involves using multiple cameras positioned around the robotic arm for calibration. On the right is the Eye-in-Hand configuration, where a camera is mounted on the robotic arm itself and moves with the arm for calibration. Previous experimental results have shown that the Eye-to-Hand configuration is often more accurate than the Eye-in-Hand configuration. In the Eye-to-Hand calibration, the camera is fixed in a specific position within the working environment while the arm moves. In this configuration, the camera's perspective remains unchanged as the arm moves, and the position and orientation of the camera are fixed. Therefore, it is possible to accurately measure the position and pose of the arm's end-effector and correlate it with the objects observed by the camera. This allows for obtaining a relatively accurate relationship between the arm and the objects. In contrast, in the Eye-in-Hand calibration, the camera is mounted on the end-effector of the arm and its perspective and position change as the arm moves. In this configuration, the camera's observations of object position and pose are influenced by the arm's motion, leading to potential uncertainties. This can result in decreased accuracy in calibration, particularly for applications requiring precise positioning.

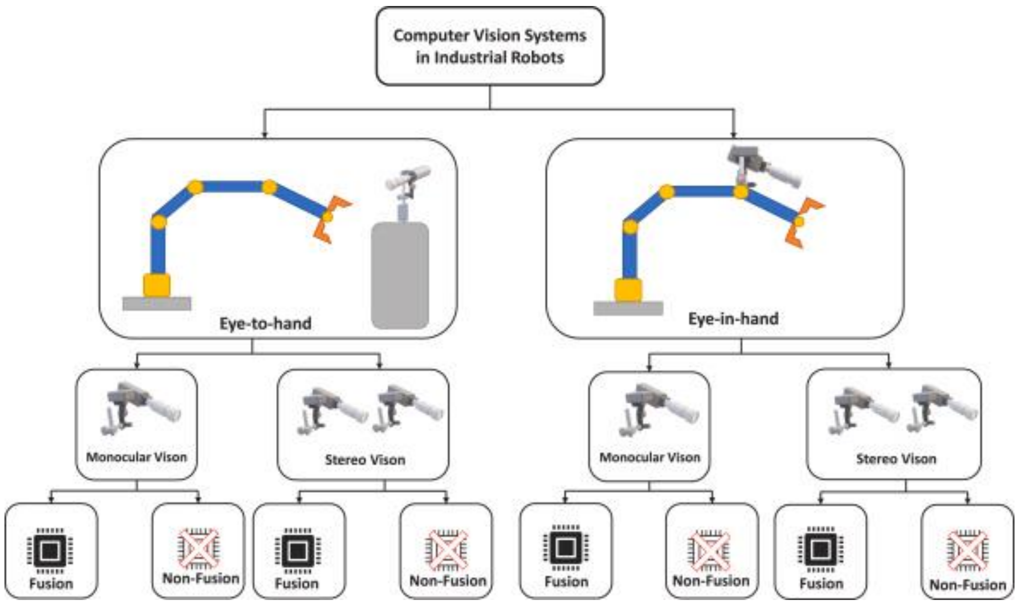


Fig 3.4 Application of computer vision systems in industrial robots.

Finally, we present an experimental result image demonstrating the application in part 4. The results show that regardless of the x, y, and z axes, through camera calibration for error analysis, the errors of the automated assembly system can be kept within the range of 0.7-1.5mm. This level of accuracy enables the system to perform a large number of simple manufacturing processes and demonstrates the importance of camera calibration in robot vision.

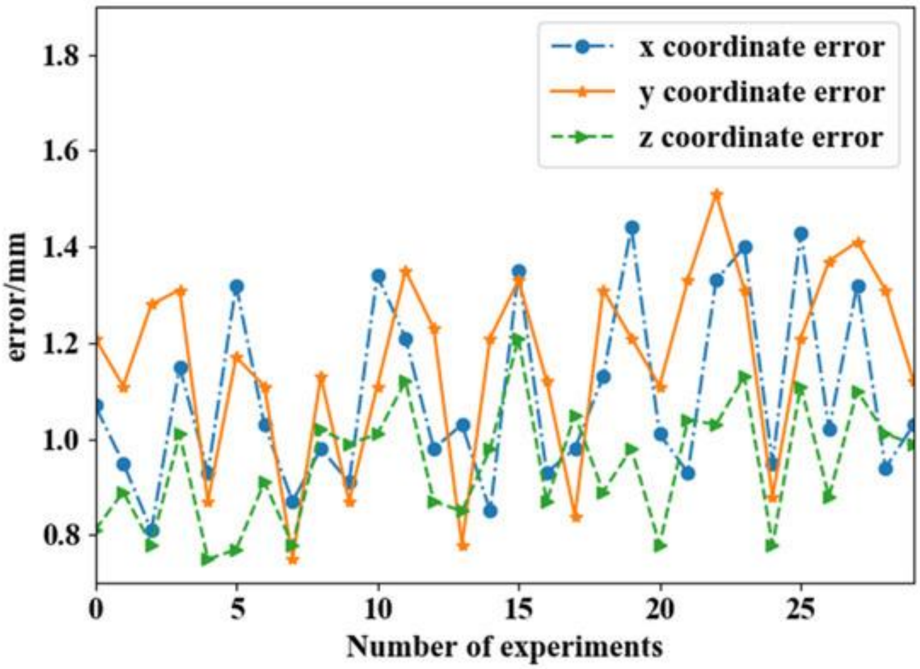


Fig 3.5 Assembly error analysis diagram

CHAPTER 4

CONCLUSION AND REFERENCES

V. CONCLUSION

Through this research on camera calibration, we have come to understand the importance of this technique and its crucial role in the fields of robot vision and automation. Camera calibration allows us to achieve precise alignment between a robotic arm and the target object. By using a calibration board, we can obtain accurate reference points to determine the geometric relationship between the robotic arm and the camera. Simultaneously, by aligning the camera with the robotic arm's target, we can obtain a transformation matrix that converts the observed object position by the camera into the robotic arm's movement position. Ultimately, by multiplying these two transformation matrices, we can complete the calibration process, enabling the robotic arm to accurately grasp the target object. However, camera calibration also faces certain challenges that require further research and optimization. I believe that through continuous effort and innovation, camera calibration techniques will find wider applications in the future, bringing significant breakthroughs and advancements to the fields of automation and robot vision.

VI. REFERENCES

https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-for-camera-calibration-in-computer-vision/

https://www.mdpi.com/2079-9292/11/3/354

https://www.sciencedirect.com/science/article/pii/S0736584521001289#sec3

https://www.mathworks.com/help/vision/camera-calibration.html