Practice Report for Distortion-free Tsai Calibration

Kang (Eric) Song

***Abstract*—This article describes a practical experiment to perform a photogrammetric camera calibration using Tsai's technique with the assumption that the radial distortion of lenses is not taken into account. The procedure of the experiment is described in detail, followed by an analysis of the result including the error analysis. Although this experiment assumes distortion-free, the first estimation of the radial distortion factor κ_1 is calculated based on the parameters obtained. The proposal on the potential improvement of the experiment is described at the end.**

***Index Terms*—Calibration, Error Analysis**

I. Introduction

A. What is Camera Calibration

CAMERA CALIBRATION is an essential preparation in many computer vision applications. It involves taking measurements of reference points and their corresponding pixel positions on the image taken in order to calculate the extrinsic and intrinsic parameters of the camera that link a point in the physical world to its location in the image. The extrinsic parameters refer to how the coordinate system of the camera frame relate to a certain world coordinate system defined by the calibration object, normally involves a combination of rotation and translation transforms. The intrinsic parameters refer to internal camera geometric and optical characteristics which normally include effective focal length, radial distortion factor and the uncertainty scale factor which is caused by scanning and acquisition timing error. [1] Camera calibration can be roughly classified into two categories: [2]

1) Photogrammetric calibration: Calibration requires a well designed calibration object, which normally has two or three orthogonal planes. The reference points need to be measured precisely. This is the calibration type to be used in this experiment.

2) Self-calibration: No calibration object required. Instead, by moving the camera inside a static scene, two constraints on the camera's internal parameters will be provided by the rigidity of the scene from one camera displacement. Three images taken by the same camera with fixed internal parameters are sufficient to restore both the external and internal parameters based on their correspondences. However the result is only approximation and not always work since this type of technique is not mature.

Apart from the two categories above, there are also calibration techniques like vanishing points for orthogonal directions and calibration from pure rotation.

Kang (Eric) Song is a postgraduate student in the Computer Science Department of the University of Auckland, New Zealand. Email: kson009@aucklanduni.ac.nz

B. Why and When Camera Calibration is Needed

Camera calibration is an expensive process for not only additional computation is performed, but also requires efforts in precise manual measurements. However it is still an important and necessary step before many computer vision and photogrammetric applications for the two purposes:

1) High accuracy measurements: Applications like automatic assembly of mechanical or electronics components require a ray in 3D space that the object point must lie on, given the image coordinates [1]. However without calibration, the ray won't likely to go through the object point due to radial lens distortion, the rectangular nature of the sensor receptor, scanning and acquisition timing error and reasons like those. There isn't a universal model to rectify those uncertainties therefore each camera needs to undertake its unique calibration to minimise the errors. This is true the other way round as well: for applications that takes some hypothetical points in 3D to predict their positions in 2D image, calibration is necessary for the same reason.

2) Stereo image matching: Stereo images by themselves may not be aligned properly therefore they are not ideal for matching. However by undertaking calibration, one camera knows exactly where its partner is and oriented to. The differences between the cameras' intrinsic parameters are also computed through calibration. The image can then be rectified and to be used for matching. Therefore calibration is an important preparation for stereo matching.

C. The Effect of Radial Lens Distortion

One of the unpredictable effects on images is the lens distortion. A radial lens refracts lights into the focus and it is very common that the lights are not all gathered in the focus due to reasons like the temperature, the material of the lens, the manufacturing quality of the lens and lots of other reasons. Although this effect is inevitable, calibration provides a way to estimate the impact of distortion and carry it through calculation to get reasonable results. The distortion factor is assumed to be zero at the initial stage of the calibration process in this experiment, however an estimation of the first distortion coefficient will be computed at the end of the article. A more specific focus on finding distortion factors will be introduced in the next experiment.

D. The Structure of This Article

This article first introduces the background knowledge related to camera calibration. A brief description of the existing techniques will be covered next. Then there will be a long and elaborate section about the experiment, including the camera specification, the measurements and the procedures in the program. The results and error analysis will be discussed in the section follows by. Then the first estimation on the radial distortion factor κ_1 will be presented. Last but not least, the potential improvements on the experiment and the conclusion will be stated at the end. One thing worth to mention is that the appendix contains all the tables, figures and proofs to keep the body of the article clean and condensed. In order to distinguish the input and calculated values, this article uses a convention such that input values and matrices use upper-case letters whereas calculated results use lower-case letters for naming.

II. Exisiting Techniques

The two most recognized[1] calibration techniques are Tsai [1] calibration and Zhang [2] calibration.

A. Tsai calibration

Although there was a paper with the same title [3] published in 1985, Tsai's calibration technique is mostly referred to his article [1] published in 1987. Both mono-view coplanar and non-coplanar points can be used for Tsai's technique. However in our situation, where the uncertainty scale factor s_x is not known *a priori*, we should use the mono-view non-coplanar points for calibration. One of the advantage for Tsai's technique over Zhang's is only one image required.

B. Zhang calibration

Zhang's technique employs a calibration object with its reference points coplanar. However it needs more images taken from different angle. But the calibration object is easier to construct and it's more accurate than Tsai's technique.

III. Experiment

A. Camera

The camera used in this experiment is Canon EOS 600D, refer Table I for detailed specification. In order to compare the impact on calibration with different focal lengths, two images were taken under $15mm$ and $85mm$ focal length respectively. The images are taken under the max resolution which is 5184×3456 therefore the image centre is taken as $(2592.5, 1728.5)$, although this pixel doesn't exist on the image it is a conceptual value because there are 2592 pixels on the left half and 2592 pixels on the right half as well, if we think of the pixels as grids, 2592.5 represents the shared edge of the two middle pixels in x direction, which is the true centre of the image. It is similar in case of the y direction that 1728.5 represents the centre position.

B. Calibration object

The calibration object to be used in this experiment is made of a wooden stand with three orthogonal plane surfaces, like a cube cut along diagonal. Two copies of a checker board image are printed and pasted on two surfaces and aligned with the joints. Fig. 1 illustrates the actual calibration object and the world coordinate axes defined[2]. The calibration for the image taken at $85mm$ focal length will also use this world coordinates definition.

As the calibration object is ready for taking images, there are three things to pay attention to:

1) Calibration object should occupy as much space as possible in the image: The larger the object is in the image, the more pixels are used to represent the reference points, thus the smaller the error ratio will be in manual measurements. Another not yet necessary but beneficial reason is some reference points close to the edges of the image can be obtained and measured to be used for either computing or testing the radial distortion factor as the distortion effect is more significant for the points far away from image centre.

2) The camera should not be on the XZ plane: In order to make the calculation work, the translation coefficient t_y for the origin along Y axis must be non-zero. That means the centre of the sensor plane must not lie on the plane formed by the X and Z axes in the world coordinates. Actually, it is better to avoid having t_y value close to zero, therefore as long as there are clearly viewed non-coplanar points from both XZ plane and YZ plane, the condition of non-zero t_y is automatically satisfied.

3) Roughly measure the distance of camera from origin: A laser range finder is used in this experiment to measure the distance between the world coordinates origin and the camera[3] while taking pictures of the calibration object. The measurement will be used to check against the calculated extrinsic parameters to get a rough idea whether the results make sense.

C. Choose reference points

Depending on the images taken, a set of reference points need to be selected and measured for both world and 2D image coordinates. Seven points are the minimum requirement for solving the unknowns however in order to be more robust to errors and outliers 16 points are going to be selected to produce an overdetermined system. The principles of choosing the reference points are:

1) Choose the corner of the squares: This is obvious. The corner of a square is most distinguishable and both easy and accurate to measure.

2) Choose the points close to the centre of the image: At the initial stage of this experiment, the effect of radial distortion is assumed to be negligible in order to simplify the calculation. However radial distortion does happen in real life and will take effect on the pixels in the image. Therefore the reference points used for calibration must be selected as the points that receive least distortion effect.

We can derive the following equations from [1] by only considering the first radial distortion factor κ_1

$$x_d + x_d(\kappa_1 r^2) = x_u$$
$$y_d + y_d(\kappa_1 r^2) = y_u$$

[1]Also the two most cited papers on camera calibration.

[2]We use right hand rule here.

[3]To be more precise, the rough estimation of where the sensor plane is about.

$$r = \sqrt{x_d^2 + y_d^2}$$

where x_d and y_d are the X and Y coordinates respectively for a distorted point in the camera coordinate system, similarly x_u and y_u are for undistorted coordinates. Combining the equations above, we get

$$\delta^2 = \kappa_1^2 r^6 \tag{1}$$

$$\delta = \sqrt{(x_u - x_d)^2 + (y_u - y_d)^2}$$

which implies the less the distance of a point from image centre, the less the point will be distorted. Therefore the reference points will be selected as the points close to the image centre.

D. Measurements

The 16 reference points selected for calibration are marked in Fig.1 and Fig.2 for $15mm$ and $85mm$ focal length scenarios respectively. For each reference point two sets of data need to be measured in the following way:

1) 3D world coordinates: The world coordinate system is defined as shown in Fig.1 for both scenarios. A tape ruler is used in this experiment to measure a point's coordinate values X_w, Y_w and Z_w in the world coordinate system. The measurements are in millimetre (mm) and are presented in TableII and TableIII for each scenario.

2) 2D image coordinates: The image coordinate system is defined as the top left corner being the origin and having X axis going toward right and Y axis going downwards just like the conventional definition in image processing. A software called ImageJ [4] is used in this experiment to measure the image coordinates X_f and Y_f for a point at a zoom level of 3200%. There are blessings and curses for this zoom-in measurement. For obvious reason it is more accurate however it comes with a side effect. Modern digital cameras normally equip a low pass filter for anti-aliasing purpose, this will smooth the edges by interpolating the pixels around the edges. This effect might not be quite distinguishable at original size but it is very significant at a scale of 3200%. A large blurry area formed by interpolated pixels makes the measurement of a corner in a checker board a difficult task as only one pixel can be chosen to represent this blurry area. Although the decision on the choices may lead to inaccurate measurement, still it is better than measuring manually at original size. The potential improvement on the measurement is proposed at the end of the article. The measurements are recorded in TableII and TableIII for each scenario.

E. Processing

Once the input data is ready and stored in a Microsoft® Excel® file, it's time to process the data and generate the parameters. In this experiment the program is implemented in MATLAB® language containing the following steps:

1) Read input information from file: In this step, the world coordinates of all the points are extracted from the file and stored in a matrix P_w in a homogeneous form:

$$P_w = \begin{bmatrix} X_{w1} & X_{w2} & X_{w3} & \dots & X_{wn} \\ Y_{w1} & Y_{w2} & Y_{w3} & \dots & Y_{wn} \\ Z_{w1} & Z_{w2} & Z_{w3} & \dots & Z_{wn} \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

where n is the number of reference points, P_w is the $4 \times n$ matrix and each column represents a vector of a point's homogeneous coordinates for all n points.
Similarly, the image coordinates are stored in P_f in a homogeneous form as well:

$$P_f = \begin{bmatrix} X_{f1} & X_{f2} & X_{f3} & \dots & X_{fn} \\ Y_{f1} & Y_{f2} & Y_{f3} & \dots & Y_{fn} \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

Apart from the points coordinates, the other input parameters are extracted as well, such as image centre (C_x, C_y), pixel width D_x and pixel height D_y.

2) Transform image coordinates to camera coordinate system: It is necessary to transform the image coordinates (X_f, Y_f) to the camera coordinates representation (x_d, y_d) in order to continue on the other operations dealing with lengths units rather than pixels. Since the transformation from (x_d, y_d) to (X_f, Y_f) is:

$$\begin{bmatrix} X_f \\ Y_f \\ 1 \end{bmatrix} = \begin{bmatrix} -D_x^{-1} & 0 & C_x \\ 0 & -D_y^{-1} & C_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$$

the transformation from (X_f, Y_f) to (x_d, y_d) can be done by taking the inverse of the transformation matrix:

$$\begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} = \begin{bmatrix} -D_x & 0 & D_x C_x \\ 0 & -D_y & D_y C_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_f \\ Y_f \\ 1 \end{bmatrix}$$

the calculated results are presented in Table IV and V for each scenario.

3) Construct and solve the overdetermined system of seven unknows: Since x_d and y_d are computed, the next step is to solve an overdetermined system with seven unknowns. Considering distortion-free, for the following equations

$$\frac{x_d}{s_x} = f \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + t_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + t_z}$$

$$y_d = f \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + t_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + t_z}$$

where s_x is the uncertainty scaling factor, f is the effective focal length, $(r_{11} \dots r_{33})$ and (t_x, t_y, t_z) are the corresponding rotation and translation coefficients in the transformation matrix

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

that transforms world coordinate system to camera coordinate system. Assuming t_y is non-zero then x_d can be rewritten as a linear system

$$x_d = ML^T$$

$$M = \begin{bmatrix} y_d X_w & y_d Y_w & y_d Z_w & y_d & -x_d X_w & -x_d Y_w & -x_d Z_w \end{bmatrix}$$

$$L = \frac{1}{t_y} \begin{bmatrix} s_x r_{11} & s_x r_{12} & s_x r13 & s_x t_x & r_{21} & r_{22} & r_{23} \end{bmatrix}$$

where M contains all the knowns and L contains all the unknows. Having more than seven points, an overdetermined linear system can be formed

$$\begin{bmatrix} x_{d1} \\ x_{d2} \\ x_{d3} \\ \vdots \\ x_{dn} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & \cdots & M_{17} \\ M_{21} & M_{22} & M_{23} & \cdots & M_{27} \\ M_{31} & M_{32} & M_{33} & \cdots & M_{37} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & M_{n3} & \cdots & M_{n7} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \end{bmatrix}$$

using Moore-Penrose pseudo-inverse technique, L can be computed by

$$L = (M^T M)^{-1} M^T X$$

however there is a built-in function in MATLAB® namely

```
pinv(...)
```

to compute the pseudo-inverse matrix M_p therefore L can be obtained simply by

$$L = M_p X$$

the results for L are presented in Table VI and VII for each scenario.

4) Find the uncertainty scaling factor: As L is calculated, the absolute value of t_y can be computed next by

$$|t_y| = \frac{1}{\sqrt{L_5^2 + L_6^2 + L_7^2}}$$

furthermore, the uncertainty scaling factor s_x can be calculated by

$$s_x = |t_y| \sqrt{L_1^2 + L_2^2 + L_3^2}$$

the result for s_x can be found in Table VI and VII for each scenario.

5) Determine the sign of t_y: As the absolute value of t_y is found, the next step will be determine the sign of t_y. Assuming t_y is positive, the following parameters can be calculated for temporary use

$$r_{11} = L_1 |t_y| \qquad r_{12} = L_2 |t_y| \qquad r_{13} = L_3 |t_y|$$
$$r_{21} = L_5 |t_y| \qquad r_{22} = L_6 |t_y| \qquad r_{23} = L_7 |t_y|$$
$$t_x = L_4 |t_y|$$

then a vector containing the sum of x_d^2 and y_d^2 for each point is computed and the point i that has the maximum value of $x_d^2 + y_d^2$ in the vector is selected. Compute the following temporary values using the measurements of point i and the temporary parameters above

$$x_t = r_{11} X_{wi} + r_{12} Y_{wi} + r_{13} Z_{wi} + t_x$$
$$y_t = r_{21} X_{wi} + r_{22} Y_{wi} + r_{23} Z_{wi} + |t_y|$$

if x_t and x_{di} have the same sign and y_t and y_{di} have the same sign, our assumption of t_y to be positive is correct, otherwise set t_y to be $-|t_y|$. The result of t_y can be found in Table VI and VII for each scenario.

6) Compute seven extrinsic parameters from knowns: Since t_y, s_x and L are all known, seven more parameters can be calculated by

$$r_{11} = \frac{L_1 t_y}{s_x} \qquad r_{12} = \frac{L_2 t_y}{s_x} \qquad r_{13} = \frac{L_3 t_y}{s_x}$$
$$r_{21} = L_5 t_y \qquad r_{22} = L_6 t_y \qquad r_{23} = L_7 t_y$$
$$t_x = \frac{L_4 t_y}{s_x}$$

the results are shown in Table VI and VII for each scenario.

7) Calculate the remaining three rotation parameters: From the orthonormal characteristics of the coordinate axes, the remaining three rotation parameters in the vector $\vec{r_3}$ can be calculated by taking the cross product of the other two vectors $\vec{r_1}$ and $\vec{r_2}$

$$\vec{r_3} = \vec{r_1} \times \vec{r_2} = \begin{bmatrix} \begin{vmatrix} r_{12} & r_{13} \\ r_{22} & r_{23} \end{vmatrix} \\ -\begin{vmatrix} r_{11} & r_{13} \\ r_{21} & r_{23} \end{vmatrix} \\ \begin{vmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{vmatrix} \end{bmatrix}$$

the results are shown in Table VI and VII for each scenario.

8) Find f and t_z by constructing and solving an overdetermined system: For each point, there is an equation

$$\begin{bmatrix} U_y & -y_d \end{bmatrix} \begin{bmatrix} f \\ t_z \end{bmatrix} = U_z y_d$$

where

$$U_y = r_{21} X_w + r_{22} Y_w + r_{23} Z_w + t_y$$
$$U_z = r_{31} X_w + r_{32} Y_w + r_{33} Z_w$$

for all the n points, an overdetermined system can be formed as

$$\begin{bmatrix} U_{y1} & -y_{d1} \\ U_{y2} & -y_{d2} \\ U_{y3} & -y_{d3} \\ \vdots & \vdots \\ U_{yn} & -y_{dn} \end{bmatrix} \begin{bmatrix} f \\ t_z \end{bmatrix} = \begin{bmatrix} U_{z1} \\ U_{z2} \\ U_{z3} \\ \vdots \\ U_{zn} \end{bmatrix}$$

the overdetermined system can be solved using the same technique as before, the result of f and t_z can be found in the Table VI and VII for each scenario.

So far, apart from the radial distortion factor which assumed to be zero, all the other extrinsic and intrinsic parameters are calculated, the calibration process is finished.

IV. Result Analysis

To check how accurate the result is, a few analyses should be performed.

A. Results at a first glance

Before doing any complicated computation, some values can be checked up quickly to get a rough idea if the results make sense.

1) Uncertainty scaling factor: Normally the uncertainty scaling factor s_x should be some number close to 1. Here we have 1.069772582 and 1.059328402, both make sense.

2) Compare roughly measured distance with calculated distance: The calculated distance between the world coordinate origin and the centre of the sensor plane is $\sqrt{t_x^2 + t_y^2 + t_z^2}$. This is going to be compared with the rough measurement while taking the images. For the $15mm$ scenario, the calculated distance is $400.766827mm$ compared to $460mm$ as the measurement, which is quite an acceptable difference. However for the $85mm$ scenario, the calculated distance is $1118.579439mm$ compared with $1495mm$ as the measurement, the difference is not too bad but not satisfactory either.

3) Effective focal length: Comparing the calculated effective focal length $14.36945686mm$ with the $15mm$ scenario, they are very close. However the $15mm$ is the focal length of the lens used on Canon EOS 600D which has an APS-C type sensor with the crop factor 1.6x, in other words it is equivalent to a full-frame sensor with $15 \times 1.6 = 24mm$ focal length. It is unknown whether the full-frame camera better simulates the pin-hole camera model, the only informative signal is the result is not too far off.
However for the $85mm$ scenario, the calculated focal length is $56.37136471mm$ which is relatively too far from either $85mm$ or $136mm$. So we can't really see much information until we perform the detailed error analysis.

B. Error analysis

1) Calculate estimated pixel positions for reference points: Since the world coordinates of the points and all the parameters of the camera are available, we can calculate the pixel position one point in 3D world should lie on the image by

$$\begin{bmatrix} kx_e \\ ky_e \\ k \end{bmatrix} = M_i M_p M_e \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

where

$$M_i = \begin{bmatrix} -\frac{1}{D_x} & 0 & C_x \\ 0 & -\frac{1}{D_y} & C_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$M_p = \begin{bmatrix} s_x f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$M_e = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

then normalize the coordinates into homogeneous form by dividing k to get the estimated or calculated pixel position (x_e, y_e) on the image for each reference point. The results can be found in Table VIII and IX for each scenario.

2) Compute the error between the estimation and the measurement: The error e_x and e_y will be computed by $x_e - X_f$ and $y_e - Y_f$ respectively. In addition, the magnitude of the error δ is also calculated by $\sqrt{e_x^2 + e_y^2}$. The results can also be found in Table VIII and IX for each scenario.

3) Statistical figures: The following statistical information is calculated

$$\sum_{i=1}^{n} \delta_i = 84.2457$$

$$\overline{\delta} = 5.265356$$

$$\sigma(\delta) = 3.678504$$

for $15mm$ scenario and

$$\sum_{i=1}^{n} \delta_i = 58.80278$$

$$\overline{\delta} = 3.675174$$

$$\sigma(\delta) = 1.790305$$

for $85mm$ scenario.

4) Improvement and new results: For the $15mm$ scenario, some reference points like 15 and 16 have much higher error than the others so we'll remove those points and redo the calibration process to see if the result can be improved. After several experiments, the best results achieved is by removing points $6, 7, 11, 12, 15, 16$ from calibration. The new parameters are listed in Table X and the new calculated pixel positions and errors are listed in Table XI, and the statistical figures are

$$\sum_{i=1}^{n} \delta_i = 12.15621$$

$$\overline{\delta} = 1.215621$$

$$\sigma(\delta) = 0.84787$$

For the $85mm$ scenario, we'll do the same as above. After experiments, the best result is achieved by removing points $2, 7, 13$ from calibration. The new parameters are listed in Table XII and the new calculated pixel positions and errors are listed in Table XIII, and the statistical figures are

$$\sum_{i=1}^{n} \delta_i = 35.31493$$

$$\overline{\delta} = 2.716533$$

$$\sigma(\delta) = 1.204906$$

5) Analysis: For the $15mm$ scenario, the results got significantly improved by removing some calibration points. Considering the image taken and the points' locations in Fig.1, the image centre is actually a bit to the left of the calibration object, only points $13, 14$ are on the left half of the image, the other points are all on the right side. Therefore points $15, 16$ are the most distant points to the right of the image centre, points $11, 12$ are the most distant points to the top of the image centre and points $6, 7$ may have some more inaccuracy in measurement. By removing those points, the errors achieved

$$-1.327989 < \overline{\delta} \pm 3\sigma(\delta) < 3.759231$$

which means at 99% confidence, the calculated pixel position will lie between 0 and 3.759231 pixels away from the true

value. This is not a perfect result for real applications but not a bad result either for the purpose of this experiment which is to get a first hands-on experience on camera calibration. We can say the experiment is successful, however a few things discussed later can potentially improve the calibration for future experiments.
For the $85mm$ scenario, the results reached optimal when points $2, 7, 13$ are removed. The results become worse if continue removing points with larger error. The errors achieved

$$-0.898185 < \overline{\delta} \pm 3\sigma(\delta) < 6.331251$$

which means at 99% confidence, the calculated pixel position will lie between 0 and 6.331251 pixels away from the true value. This result is worse than the $15mm$ scenario, probably because the squares in the image are larger therefore some points are more distant from the image centre and distorted by the radial distortion effect.

V. κ_1 Estimation

The distorted coordinates are what we've measured in the 2D image whereas the undistorted coordinates are the calculated coordinates based on the parameters obtained through calibration with the assumption of distortion-free. From equation (1) we get

$$\kappa_1 = \frac{\delta}{r^3}$$

where δ is the distance between the calculated coordinates and the measured coordinates in mm for each point, r is the distance between the measured coordinates and the image centre also in mm for each point. Since we have multiple points to work on, the κ_1 can be estimated by taking the average of the above calculation for all the points taking part of the calibration. This step is implemented as a part of the program.
The result for $15mm$ scenario is $0.00192432mm^{-2}$ whereas the result for $85mm$ scenario is $0.011645547mm^{-2}$.

VI. Potential Improvements

For future experiments, some potential improvements may be achieved by

A. A better calibration object

The calibration object used in this experiment is made by taping printed checker-board sheets on the surface. The sheets are taped at the edges which might cause uneven surfaces. Some corners are not taped well and even bended up. A better calibration object is also expensive to make so there is a trade-off. Another thing might be worth mentioning for improvement is to reduce the white margin on the paper around the checker-board, it can make the non-coplanar points closer and actually it makes the points closer to the image centre for less distortion effect.

B. Take image more carefully

Maybe it's because my photography skill is poor, or I was not paying enough attention, the images taken are somehow biased. Take the $15mm$ scenario for instance, I was thinking I aimed at the centre but actually only points 13 and 14 are on the left half of the image, the rest of the points are all on the right side which will incur some distortion especially for points 15 and 16 as they are far from the image centre. For the $85mm$ scenario it is even worse. So next time I'll pay much more attention when taking the images and selecting points.

C. Improve the accuracy of measurements

The result can also be improved by improving the accuracy of measurements. Apart from paying more attention, there isn't much to say about the 3D measurements. However we can do something for the 2D measurements. As mentioned earlier, the pixels are interpolated around the corner thus makes it inaccurate to choose which pixel to measure. We can process the image before taking measurements by choosing a threshold of the grayscale values and convert it into binary image based on the threshold. Then the corner will be either black or white which will help a lot for our measurements.

VII. Conclusion

In this experiment, we use Tsai's calibration method to calculate the extrinsic and intrinsic parameters with the assumption of distortion-free for two scenarios using $15mm$ and $85mm$ focal lengths respectively. Then we analyse the error for both scenarios, it turns out the difference in focal length doesn't seem to have much impact on the result. The first estimate of the radial distortion factor κ_1 is computed as well for both scenarios. Combining the results and error analyses, our calibration is successful, however some improvements can be done in future works.

References

[1] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 4, pp. 323–344, 1987.

[2] Z. Zhang, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000.

[3] R. Tsai, "A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses." IBM, Research Report 11413, Oct 1985.

[4] Imagej. [Online]. Available: http://rsb.info.nih.gov/ij/

Appendix

TABLE I
Camera Specification

Camera Model	Canon EOS 600D
Lens	Canon EF-S 15-85 f/3.5-5.6 IS USM
Sensor	CMOS
Sensor size	APS-C $22.3 \times 14.9mm$
Crop factor	1.6x
Mega pixels	17.9MP
Light sensitivity	6400 ISO (12800 ISO boost)
Max resolution	5184×3456
Sensor resolution	5196×3464
Pixel width	$4.292\mu m$
Pixel height	$4.301\mu m$
Focal length	15-85mm

TABLE II
Points and measurements for $15mm$ scenario

Point ID	$X_w(mm)$	$Y_w(mm)$	$Z_w(mm)$	$X_f(pix)$	$Y_f(pix)$
1	0	15.5	124.5	2710	2366
2	23	0	125	3016	2322
3	0	15.5	148.8	2702	2148
4	23	0	149	3023	2105
5	0	15.5	173.5	2696	1910
6	23	0	173.9	3029	1868
7	0	15.5	198	2693	1654
8	23	0	198.3	3032	1617
9	0	16	222.5	2677	1383
10	23	0	222.9	3034	1346
11	0	16	246.5	2670	1092
12	23	0	247	3038	1060
13	0	40	173.5	2496	1831
14	0	40	198	2482	1589
15	47.5	0	173.5	3188	1775
16	47.5	0	198.2	3195	1534

TABLE III
Points and measurements for $85mm$ scenario

Point ID	$X_w(mm)$	$Y_w(mm)$	$Z_w(mm)$	$X_f(pix)$	$Y_f(pix)$
1	0	16	75	2098	2544
2	23.5	0	75.5	2452	2539
3	0	15.5	99.5	2090	2264
4	23	0	100	2441	2257
5	0	15.5	124.5	2077	1981
6	23	0	125	2433	1974
7	0	15.5	148.8	2061	1696
8	23	0	149	2425	1689
9	0	15.5	173.5	2051	1408
10	23	0	173.9	2415	1400
11	0	15.5	198	2045	1118
12	23	0	198.3	2404	1111
13	47.5	0	100	2690	2211
14	47.5	0	124.5	2684	1930
15	47.5	0	149	2675	1647
16	47.5	0	173.5	2665	1360

TABLE IV
Calculated camera coordinates for $15mm$ scenario

Point ID	$x_d(mm)$	$y_d(mm)$
1	-0.5043	-2.7419
2	-1.8177	-2.5526
3	-0.4700	-1.8043
4	-1.8477	-1.6193
5	-0.4442	-0.7806
6	-1.8735	-0.6000
7	-0.4313	0.3204
8	-1.8863	0.4796
9	-0.3627	1.4860
10	-1.8949	1.6451
11	-0.3326	2.7376
12	-1.9121	2.8752
13	0.4142	-0.4409
14	0.4743	0.6000
15	-2.5559	-0.2000
16	-2.5859	0.8365

TABLE V
CALCULATED CAMERA COORDINATES FOR $85mm$ SCENARIO

Point ID	$x_d(mm)$	$y_d(mm)$
1	2.1224	-3.5075
2	0.6030	-3.4860
3	2.1567	-2.3032
4	0.6502	-2.2731
5	2.2125	-1.0860
6	0.6846	-1.0559
7	2.2812	0.1398
8	0.7189	0.1699
9	2.3241	1.3785
10	0.7618	1.4129
11	2.3499	2.6258
12	0.8090	2.6559
13	-0.4185	-2.0752
14	-0.3927	-0.8667
15	-0.3541	0.3505
16	-0.3112	1.5849

TABLE VI
THE CALCULATED SEVEN UNKNOWNS AND THE EXTRINSIC AND INTRINSIC PARAMETERS FOR $15mm$ SCENARIO

L_1	0.0043
L_2	-0.0043
L_3	-0.00021538
L_4	0.1564
L_5	-0.0019
L_6	-0.0016
L_7	-0.0051

r_{11}	-0.702261117
r_{12}	0.711034927
r_{13}	0.035477548
r_{21}	0.341191808
r_{22}	0.279459914
r_{23}	0.897491118
r_{31}	0.628232979
r_{32}	0.642377763
r_{33}	-0.438853124
t_x	-25.75799897
t_y	-176.2099822
t_z	359.0273211
f	14.36945686
s_x	1.069772582

TABLE VII
THE CALCULATED SEVEN UNKNOWNS AND THE EXTRINSIC AND INTRINSIC PARAMETERS FOR $85mm$ SCENARIO

L_1	0.0059
L_2	-0.0044
L_3	-0.00019364
L_4	-0.2044
L_5	-0.0011
L_6	-0.0014
L_7	-0.0067

r_{11}	-0.804065225
r_{12}	0.593954506
r_{13}	0.026403774
r_{21}	0.157471322
r_{22}	0.203191982
r_{23}	0.966393192
r_{31}	0.568628555
r_{32}	0.781200996
r_{33}	-0.256910408
t_x	27.87493037
t_y	-144.4443254
t_z	1108.863737
f	56.37136471
s_x	1.059328402

TABLE VIII
THE CALCULATED PIXEL POSITIONS AND ERRORS FOR EACH REFERENCE POINTS IN $15mm$ SCENARIO

Point ID	$x_e(pix)$	$y_e(pix)$	e_x	e_y	δ
1	2710.082	2367.69	0.08249	1.69032	1.692331
2	3013.754	2317.547	-2.24578	-4.45288	4.987153
3	2704.044	2150.205	2.044076	2.204728	3.006505
4	3018.257	2104.103	-4.74287	-0.89669	4.826889
5	2697.456	1912.906	1.455557	2.906031	3.250179
6	3023.266	1866.663	-5.73372	-1.33687	5.887507
7	2690.42	1659.513	-2.57982	5.512749	6.086533
8	3028.546	1616.417	-3.45441	-0.5827	3.503213
9	2678.033	1384.724	1.033321	1.724361	2.010267
10	3034.283	1344.439	0.283377	-1.56128	1.586786
11	2669.805	1095.271	-0.19513	3.27132	3.277135
12	3040.358	1056.506	2.357753	-3.49449	4.215503
13	2489.913	1829.372	-6.08723	-1.6278	6.301123
14	2475.757	1586.353	-6.24331	-2.64662	6.781119
15	3199.033	1774.317	11.0332	-0.68298	11.05431
16	3210.415	1530.631	15.4153	-3.36896	15.77914

TABLE IX
THE CALCULATED PIXEL POSITIONS AND ERRORS FOR EACH REFERENCE POINTS IN $85mm$ SCENARIO

Point ID	$x_e(pix)$	$y_e(pix)$	e_x	e_y	δ
1	2095.623	2545.673	-2.37704	1.672879	2.906688
2	2454.066	2534.044	2.066341	-4.95649	5.369969
3	2088.146	2268.585	-1.85373	4.584615	4.945201
4	2439.923	2256.746	-1.07703	-0.25442	1.106673
5	2076.738	1980.993	-0.26197	-0.00716	0.262065
6	2430.597	1969.306	-2.40331	-4.69393	5.273415
7	2065.52	1698.184	4.519515	2.18365	5.019397
8	2421.539	1690.159	-3.46054	1.15859	3.64934
9	2053.983	1407.357	2.982969	-0.64277	3.051434
10	2412.033	1397.166	-2.96701	-2.83437	4.103267
11	2042.405	1115.478	-2.59533	-2.52172	3.618673
12	2402.607	1106.661	-1.39274	-4.33932	4.557351
13	2688.721	2204.57	-1.27868	-6.4305	6.556396
14	2681.117	1926.169	-2.88339	-3.83093	4.794783
15	2673.425	1644.576	-1.57532	-2.42441	2.891261
16	2665.644	1359.734	0.644024	-0.26619	0.696869

TABLE X
IMPROVED RESULTS FOR THE SEVEN UNKNOWNS AND THE EXTRINSIC AND INTRINSIC PARAMETERS FOR THE $15mm$ SCENARIO

L_1	0.0044
L_2	-0.0042
L_3	-0.0002119
L_4	0.1538
L_5	-0.0018
L_6	-0.0015
L_7	-0.0051

r_{11}	-0.722964168
r_{12}	0.690007117
r_{13}	0.034828008
r_{21}	0.322445369
r_{22}	0.265081638
r_{23}	0.908713767
r_{31}	0.617786701
r_{32}	0.668197623
r_{33}	-0.414134125
t_x	-25.2713934
t_y	-177.8969787
t_z	353.8760046
f	14.1588775
s_x	1.082377149

TABLE XI
IMPROVED CALCULATED PIXEL POSITIONS AND ERRORS FOR EACH REFERENCE POINTS IN $15mm$ SCENARIO

Point ID	$x_e(pix)$	$y_e(pix)$	e_x	e_y	δ
1	2709.441	2367.092	-0.55946	1.091711	1.226713
2	3016.326	2320.583	0.326227	-1.41711	1.454172
3	2703.343	2148.108	1.343282	0.108179	1.347631
4	3020.334	2105.455	-2.66602	0.45484	2.704542
5	2696.716	1910.071	0.715512	0.071335	0.719059
8	3029.44	1616.636	-2.55951	-0.36441	2.58532
9	2677.461	1383.701	0.460538	0.700788	0.83857
10	3034.489	1345.653	0.488792	-0.347	0.599437
13	2495.685	1831.198	-0.31538	0.197751	0.372253
14	2482.192	1589.242	0.191536	0.24186	0.308516

TABLE XII
IMPROVED RESULTS FOR THE SEVEN UNKNOWNS AND THE EXTRINSIC AND INTRINSIC PARAMETERS FOR THE $85mm$ SCENARIO

L_1	0.0059
L_2	-0.0044
L_3	-0.00020087
L_4	-0.2022
L_5	-0.0011
L_6	-0.0014
L_7	-0.0067

r_{11}	-0.797491971
r_{12}	0.602712984
r_{13}	0.027269325
r_{21}	0.157176383
r_{22}	0.205689415
r_{23}	0.965912754
r_{31}	0.576559147
r_{32}	0.77459376
r_{33}	-0.258767904
t_x	27.44815016
t_y	-144.4014754
t_z	1058.924375
f	53.64870208
s_x	1.063690633

TABLE XIII
IMPROVED CALCULATED PIXEL POSITIONS AND ERRORS FOR EACH REFERENCE POINTS IN $85mm$ SCENARIO

Point ID	$x_e(pix)$	$y_e(pix)$	e_x	e_y	δ
1	2097.824	2542.752	-0.17616	-1.24755	1.259922
3	2089.975	2266.797	-0.02538	2.796766	2.796881
4	2442.137	2255.371	1.136843	-1.62864	1.986172
5	2078.118	1980.167	1.118488	-0.83312	1.394673
6	2432.484	1968.981	-0.51554	-5.01943	5.045838
8	2423.105	1690.673	-1.89549	1.67275	2.52804
9	2054.447	1407.888	3.446778	-0.11245	3.448612
10	2413.253	1398.371	-1.74708	-1.62861	2.388444
11	2042.39	1116.411	-2.60984	-1.58922	3.055634
12	2403.479	1108.36	-0.52148	-2.63985	2.69086
14	2681.27	1925.875	-2.73018	-4.12494	4.946619
15	2673.329	1645.339	-1.67109	-1.66138	2.356427
16	2665.291	1361.387	0.291312	1.386534	1.416806

Fig. 1. Calibration object and points for $15mm$ scenario

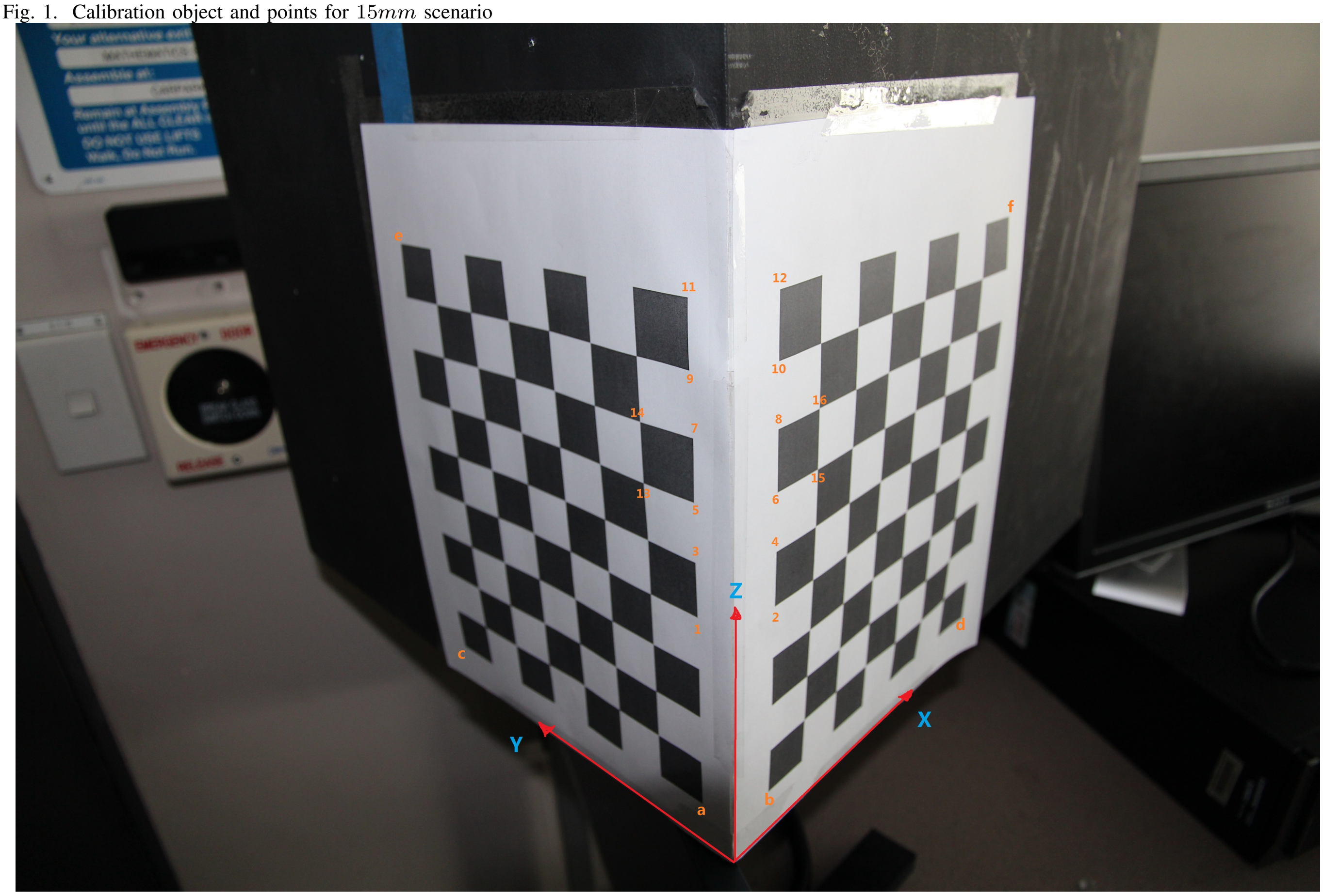

Fig. 2. Calibration object and points for $85mm$ scenario