Camera Calibration

Sudheer Nadella
*U*93511802
University of South Florida
snadella@mail.usf.edu

I. Introduction

Camera Calibration is a widely used technique used to estimate camera parameters and use them to fix distortion of lens, to find the correspondences in real world. There are many algorithms in use which can be used to find the parameters. In this project, Tsai algorithm has been used. After finding these parameters, triangulation is used to find the world coordinates from pixel locations given as input.

II. Algorithm

There are two types of parameters: intrinsic and extrinsic. Intrinsic parameters are used to map camera coordinates with image coordinates(thus, intrinsic). Extrinsic parameters map camera and world coordinates. To find these parameters, Tsai algorithm is used. With the help of parameters obtained, triangulation is carried out. Triangulation is a technique of mapping the pixel coordinates to world coordinates after making the algorithm learn the camera's orientation with the help of six or more correspondences. The algorithm is as follows:

- Images taken with two or more views are taken. The parameters (3D coordinates and pixel coordinates of the points) for each image are given as input to this algorithm.
- The parameters of each image are K, R and t matrices. With the help of these matrices, v_j and c_j are calculated. c_j is the vector from world coordinate origin to the camera coordinate origin, v_j is a normalised vector of pixels in the image with respect to world coordinates and j represents the camera.
- c_j is calculated using the following formula

$$c_j = -R_j^T t_j \tag{1}$$

- v_j is calculated using the following formula

$$v_j = R_j^T K_j^{-1} x_j \tag{2}$$

- With these matrices calculated, we can calculate least square estimate of 3D location of pixel by using the formula

$$p = (\sum_{j=1}^{N} (I - v_j v_j^T))^{-1} \sum_{j=1} (I - v_j v_j^T) c_j \tag{3}$$

- p is the matrix that gives world coordinates to the pixel locations on the image.

All the equations discussed above tend to minimise the residual. That is the reason why least square estimate is considered.

III. Implementation

To start the project, two images of a Rubik's cube are taken in two different views. Initially, 3D coordinates and the pixel coordinates of the points are taken on these two images. That is, each pixel(x,y) has (X,Y,Z) real world coordinates. At the least 6 points must be considered. These pixels and their respective real world coordinates are passed as input or rather calling it a training set, to Tsai algorithm[1]. The algorithm returns K, R and t matrices. This process is carried out for both the images. With these parameters, we can continue with the triangulation part.

For the Tsai algorithm I have tried implementing the A matrix of size 2i * 12 and then computing P from resizing A and performing rq decomposition on the matrix. With the final projection matrix, I have tried checking my intrinsic and Extrinsic parameters i.e, my K R and T matrices by performing reverse operation, i.e, by computing pixel coordinates from 3D point. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \end{bmatrix} \begin{bmatrix} X_\mathrm{p} \\ Y_\mathrm{p} \\ Z_\mathrm{p} \end{bmatrix}$

$$x = \frac{x_1}{x_3} = \frac{P_{00}X_\mathrm{p} + P_{01}Y_\mathrm{p} + P_{02}Z_\mathrm{p} + P_{03}}{P_{20}X_\mathrm{p} + P_{21}Y_\mathrm{p} + P_{22}Z_\mathrm{p} + P_{23}} \tag{4}$$

$$y = \frac{x_2}{x_3} = \frac{P_{10}X_\mathrm{p} + P_{11}Y_\mathrm{p} + P_{12}Z_\mathrm{p} + P_{13}}{P_{20}X_\mathrm{p} + P_{21}Y_\mathrm{p} + P_{22}Z_\mathrm{p} + P_{23}} \tag{5}$$

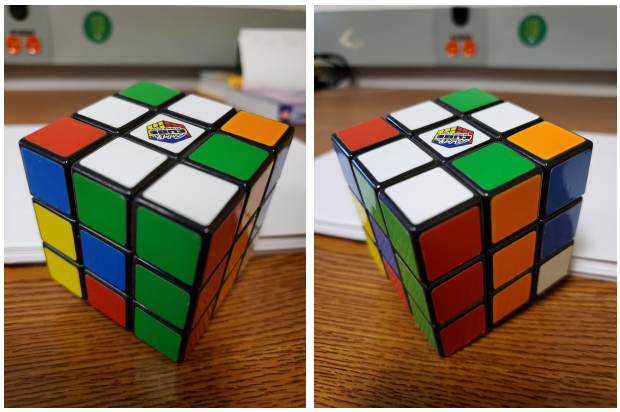
Here x and y are the pixel coordinates of the 3D point in the respective image.

Using the equation (1), (2) and (3), we get a final 3x1 matrix which has 3D coordinates to the specified pixel locations in two images. The input file containing pixel locations and corresponding 3D locations is as follows-

The parameter file that is obtained afterwards-

The 3D location of the pixel location () from camera1 and () from camera2

Rubik's cube from two views:



The image on the left is taken from the camera 1 and the image to the right is taken from camera 2

```
1   3  3  0  2909  1217
2   0  3  3  147   1363
3   0  2  1  125   2481
4   0  1  2  801   2612
5   2  2  0  2539  2059
6   1  1  0  2185  2785
7   3  3  3  1408  759
8   0  3  0  2026  2262
9   0  2  2  744   2190
10  0  1  1  1255  2896
11  1  2  0  2281  2414
12  2  1  0  2417  2454
```

This is the input to Tsai of the image 1 i.e taken from camera 1 to get the Intrinsic and Extrinsic parameters of the camera 1

```
1   3  3  0  2776  1354
2   0  3  3  172   1300
3   0  2  1  759   2352
4   0  1  2  648   2443
5   2  2  0  2232  2183
6   1  1  0  1758  2879
7   3  3  3  1594  828
8   0  3  0  987   2250
9   0  2  2  539   2068
10  0  1  1  869   2741
11  1  2  0  1742  2455
12  2  1  0  2200  2592
```

This is the input to Tsai of the image 2 i.e taken from camera 2 to get the Intrinsic and Extrinsic parameters of the camera 2

The 3D estimates that I got for the points (1 0 1) are (1.12 -0.13 and -0.96) and for (0 3 0) are (-0.53 3.79 and -0.35). My Z axis is considered in opposite way to which it is considered for calibration, so the Z values that I am getting here are actually positive.

```
(base) C:\Users\Dell\Desktop\CV 4>python triangulation.py
[[ 1.12682076]
 [-0.13453903]
 [-0.96312017]]

[[-0.53352956]
 [ 3.79586577]
 [-0.35493325]]
```

IV. Conclusion

The triangulation algorithm works fine when the world coordinates are unknown, but parameters and pixel locations are known. However,this procedure of reducing the residual to a minimal value is not as effective as other procedures which solve with techniques like non-linear least squares or SVD. To get higher accuracy, it is preferred to take more views and them perform Tsai and get the K R and T matrices and then do the triangulation.

V. Issues

There are few issues with the implementation. They are-

- This assignment consists of two parts, one is get the parameters from Tsai and then do the triangulation. We have to be careful when doing the Tsai part.
- The accuracy is directly proportional to the number of images considered. As only two images are used for training, accuracy is expected to be low.

VI. References

[1]Tsai Camera calibration by Alex, https://github.com/alexprz/tsai-camera-calibration.

[2]Computer Vision: Algorithms and Applications by Richard Szeliski, http://szeliski .org/Book/drafts/SzeliskiBook-20100903-draft.pdf.