1. Introduction:

Camera calibration is a way to find the extrinsic and intrinsic parameters of the camera. We use five images of a checkerboard pattern with Jean-Yves Bouguet Camera Calibration Toolbox for MATLAB to perform camera calibration. The camera parameters can then be used for various applications such as 3D reconstruction, robotics, and augmented reality.

2. Methodology:

All five images of a checkerboard loaded in Matlab using the Toolbox. The corners of the checkerboard pattern in each images identified manually. The window size of 11X11 has been taken with 13 number of squares along the X direction and 14 number of squares along the Y direction and the size of each square is 30mm. By processing all the images on the toolbox, following camera parameters has been found:

Calibration results after optimization (with uncertainties):

- Focal Length: fc = [647.15106 698.36340] +/- [8.73470 14.25538]
- Principal point: cc = [314.69728 240.37709] +/- [12.88965 10.73796]
- Skew: alpha_c = [0.00000] +/- [0.00000] => angle of pixel axes = 90.00000 +/- 0.00000 degrees
- Distortion: kc = [-0.17712 0.44414 -0.00317 0.00155 0.00000] +/- [0.06620 0.28992 0.00458 0.00587 0.00000]
- Pixel error: err = [1.77855 1.31120]

3. Results:

a. Corners Detection

Below we have result of corner detection for all five images of checkerboard. Here:

- The crosses represent the actual corner points in the image identified by the calibration algorithm.
- The circles represent the predicted corner points after estimating the camera parameters.
- The green line connecting the crosses and circles represents error between the detected and predicted corners points in the images.

Image 1 - Image points (+) and reprojected grid points (o)

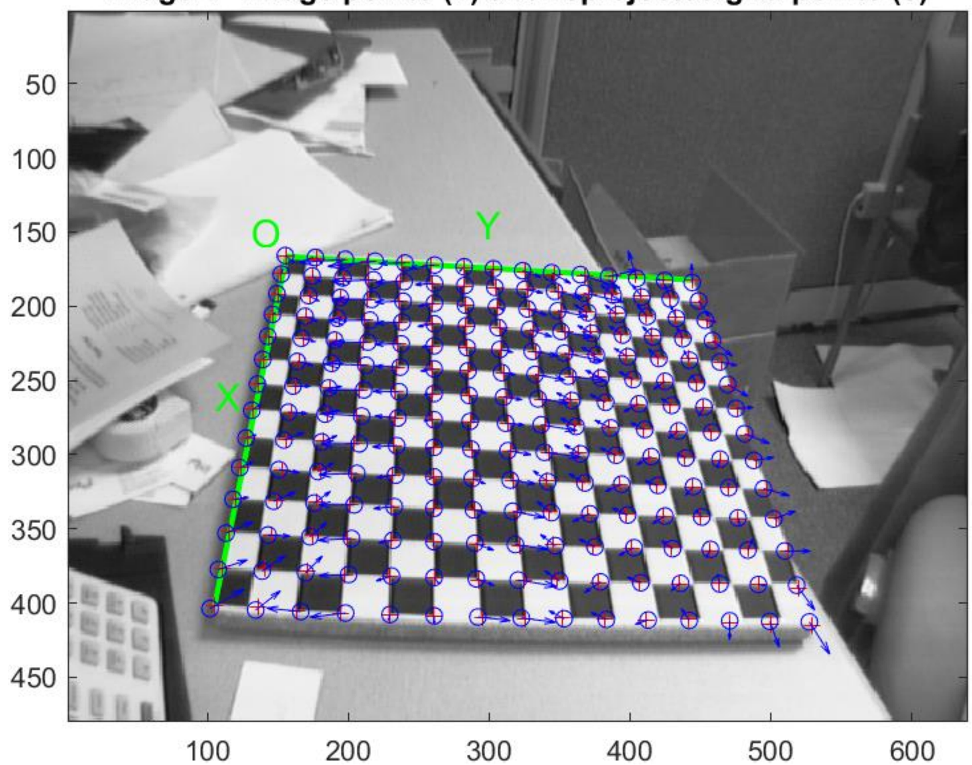


Image 2 - Image points (+) and reprojected grid points (o)

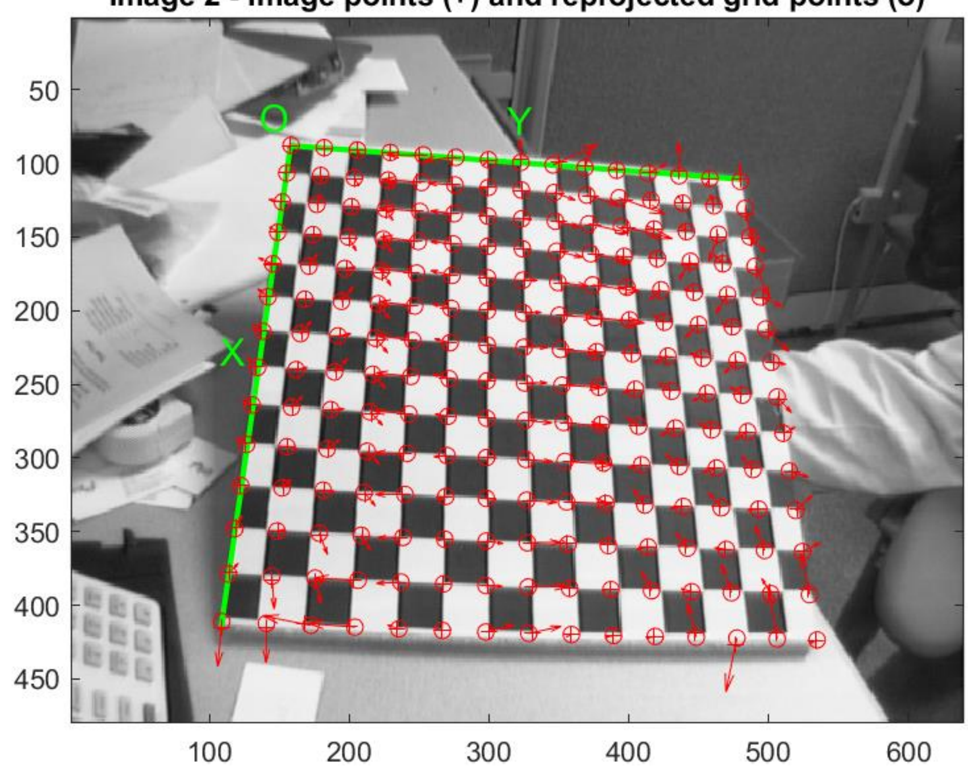


Image 3 - Image points (+) and reprojected grid points (o)

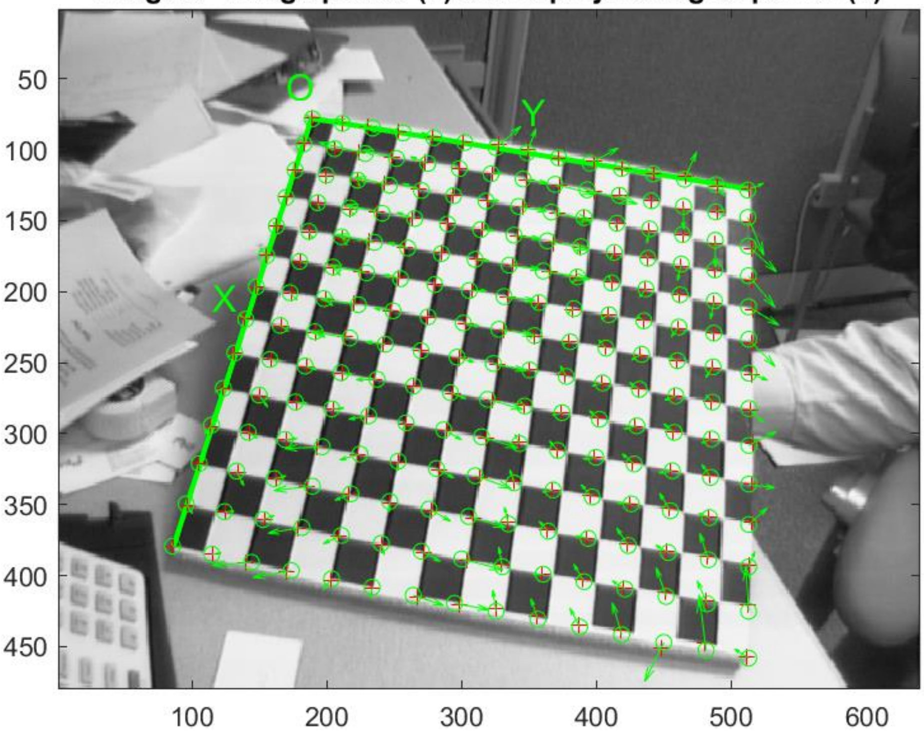
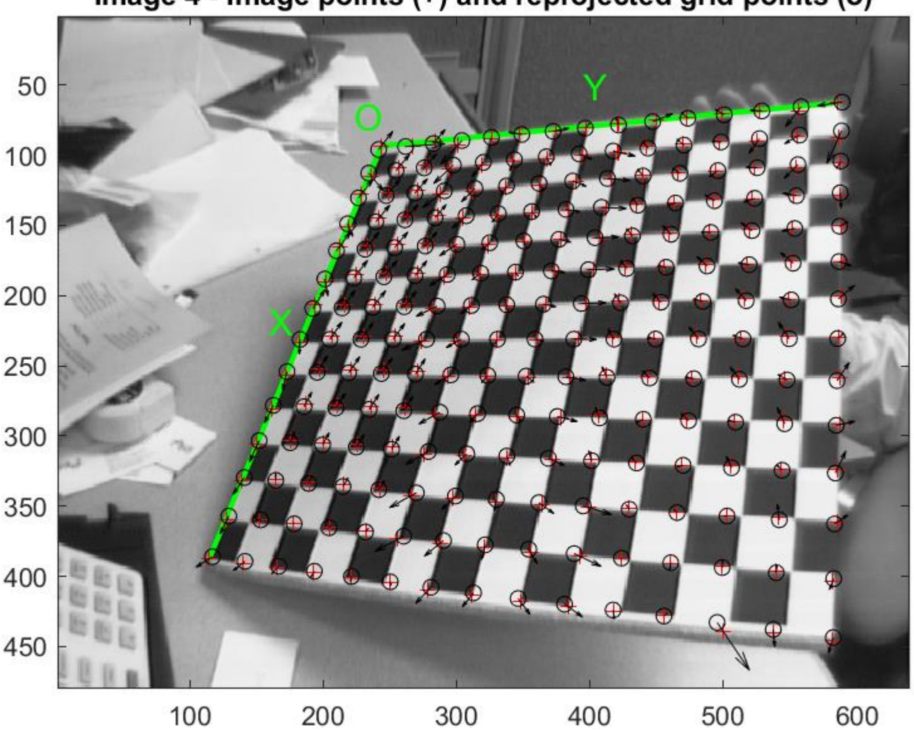
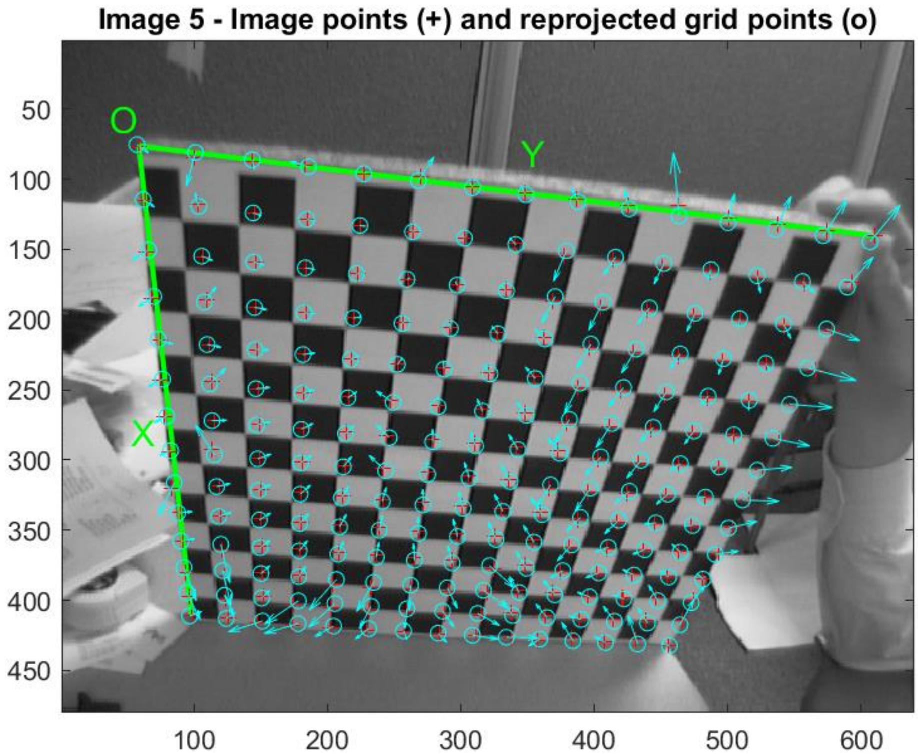
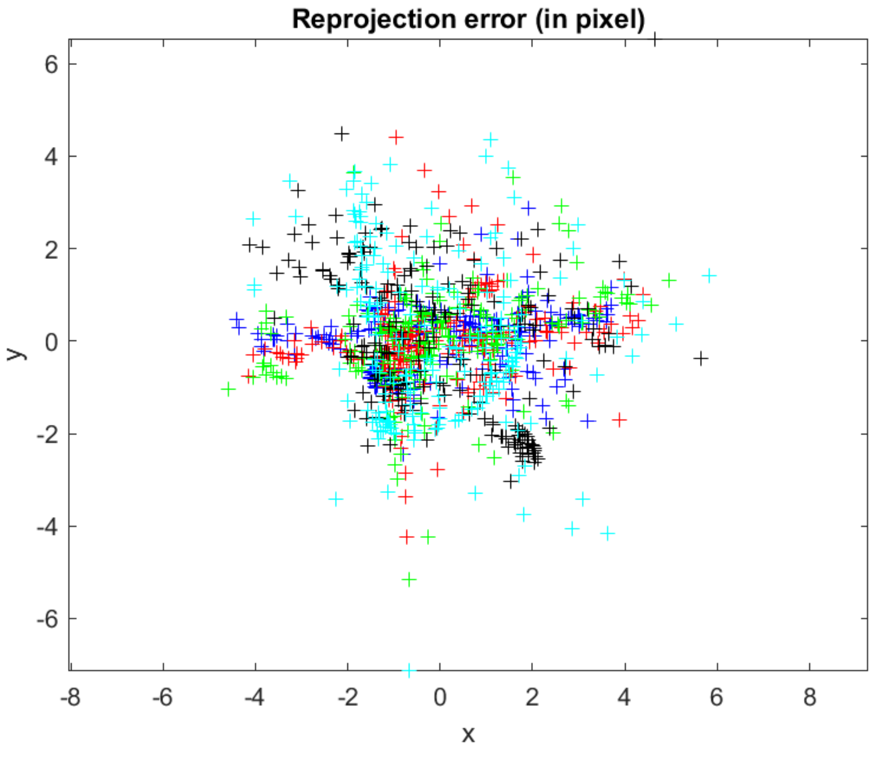


Image 4 - Image points (+) and reprojected grid points (o)







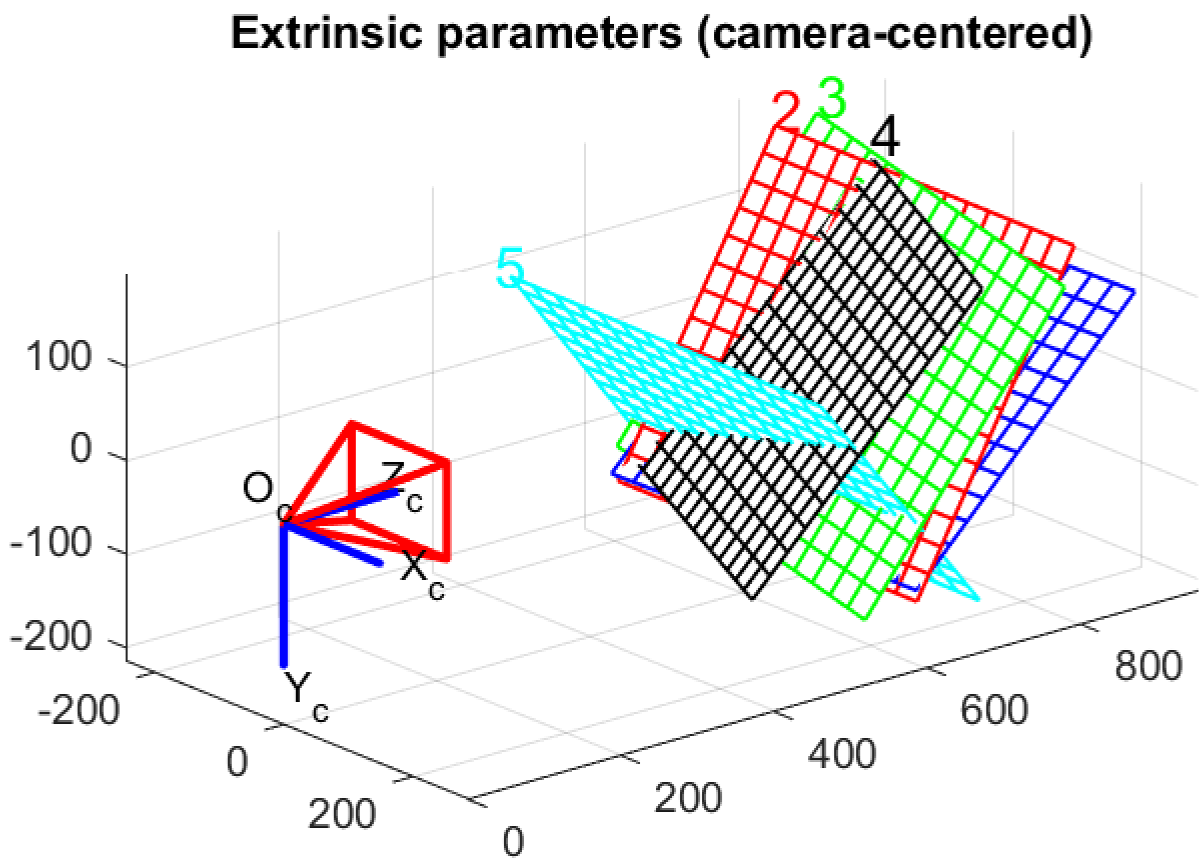
The reprojection error helps us in identifying the points that are contributing most to the error. This plot can be used to further refine the prediction model.

b. Views:

i. Camera-Centered View

In this image, we are getting the extrinsic parameters view from a camera-centered perspective. Here:

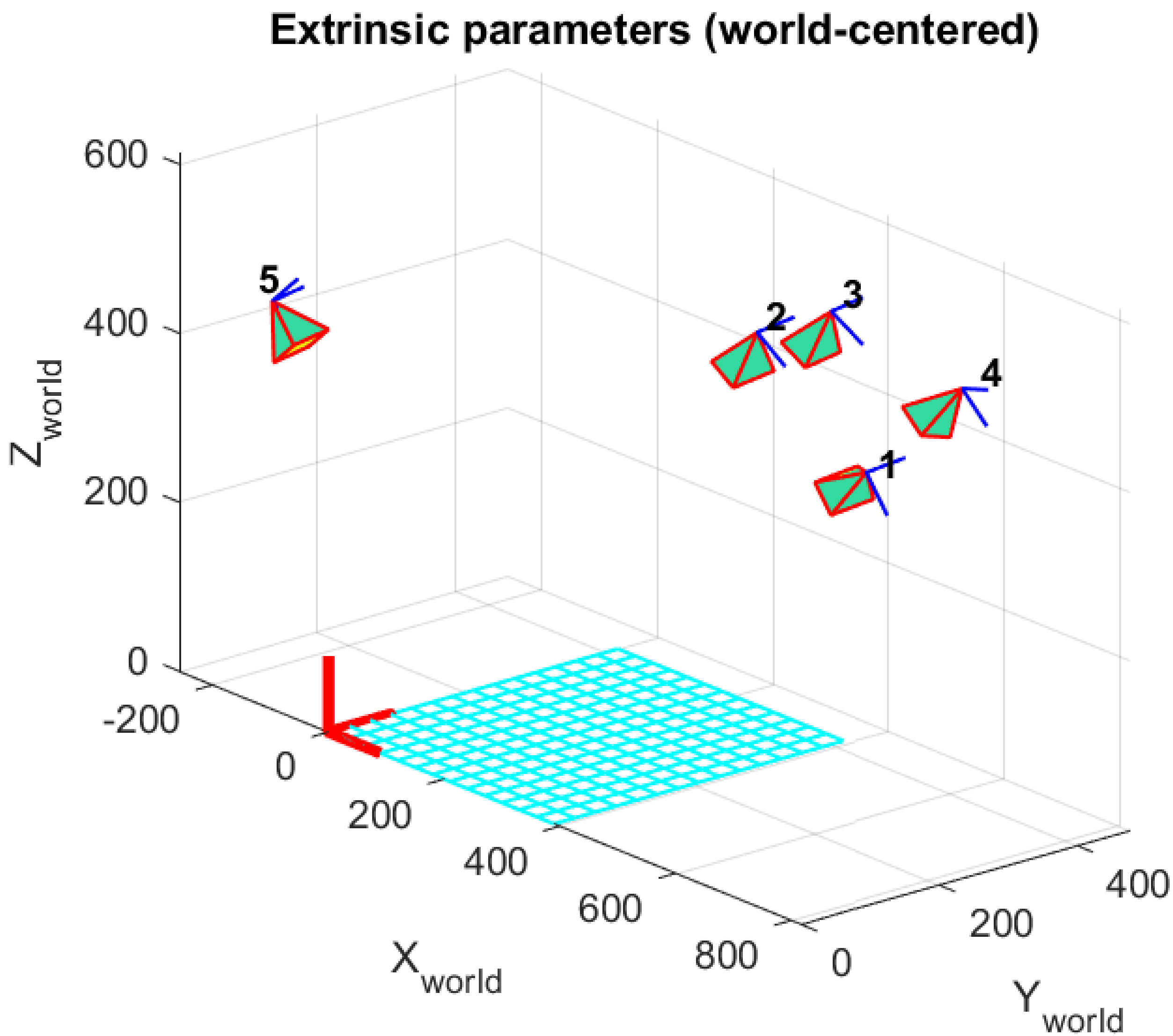
- The origin O represents the position of camera in its co-ordinates.
- The checkerboards numbered from 1 to 5 represents all the images used for camera calibration with their position with respect to the camera co-ordinate system. All these checkerboards have been transformed with respect to the camera.



ii. World-Centered View

In this image, we are getting the extrinsic parameters view from a world-centered perspective. We assume that world remains stationary and the camera moves Here:

- The large grid at the bottom represents the stationary reference plane of the checkerboard pattern.
- The camera numbered from 1 to 5 represents the orientation of the camera for all the five calibrated images of the checkerboard.



Both these views remain very useful for 3D reconstruction and augmented reality applications.

c. Distortion Models:

It represents the bending or warping of images that can occur because of imperfections in camera lenses. It can make straight lines look curved or objects appear stretched or compressed in photos. The following camera parameters found by calibration as:

- Focal Lengths: [647.151, 698.363] pixels
- Principal Point: [314.697, 240.377] pixels
- Skew: 0 (indicating the x and y pixel axes are orthogonal)

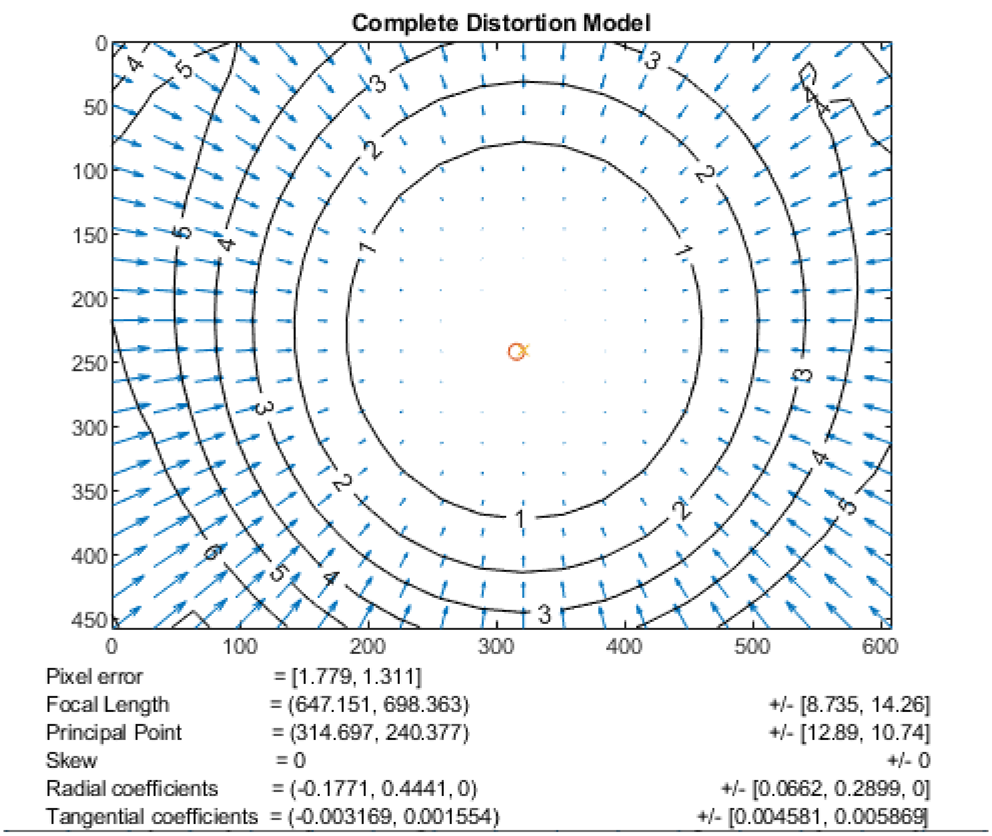
The distortion coefficients are as:

- Radial: [-0.1771, 0.4441, 0]
- Tangential: [-0.003169, 0.001554]

The error got in the distortions are in the acceptable range representing a successful Calibration process.

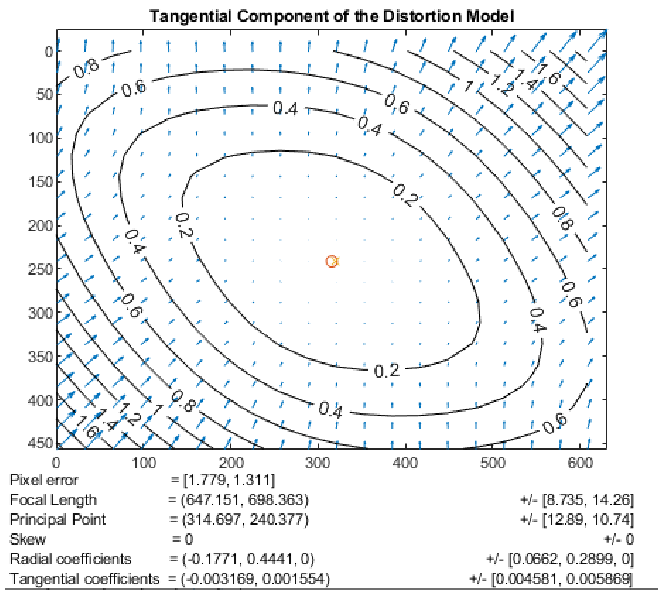
i. Complete Distortion Model

It combines both the radial and tangential components of distortion. The blue arrows on the plot represents the distortion direction and magnitude. The concentric circles indicate the pixel error and we can see it increases as one we move towards the edges of images.



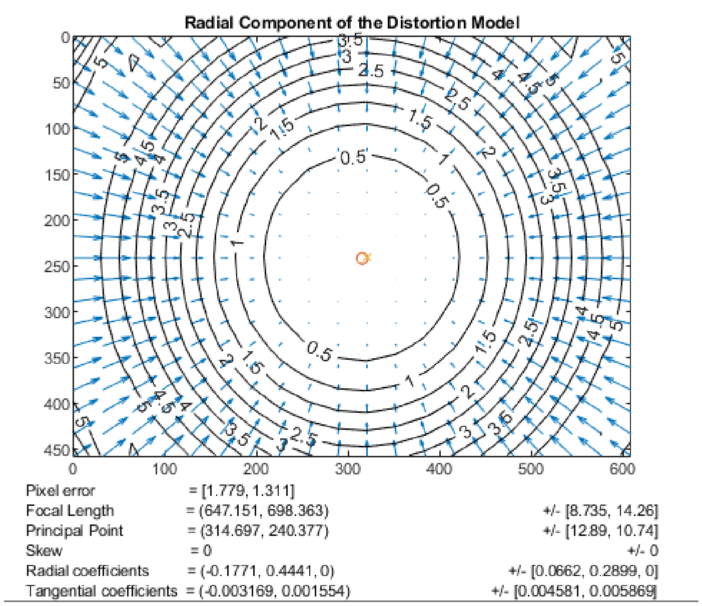
ii. Tangential Component

It represents the skewed or tilted effect in photos especially towards the origin. Tangential distortion occurs when the lens plane is not parallel to the image plane.



iii. Radial Component

It represents that the straight lines in real world appear to be curved in images. It is more noticeable towards the edges of the images.



4. Discussion:

a. If the initial selection of corners is not accurate, how would that impact the final result. Consider the fact that the initial corner estimate is used only in the closed form, or linear regression, portion of the calibration routine. You should consider two distinct cases: the initial corner selection is only slightly incorrect and if the initial selection is significantly in error.

- **Slightly Incorrect:** Its impact on the final result will be limited. The optimization process can be used to converge it to more accurate solution. The calibration process will require more iterations to improve this.
- **Significant Incorrect:** Its impact on the final result will be significant. The linear regression mostly uses initial selection to predict better. But if the initial points are too far away from the actual points then it may fail to converge to a meaningful solution. The calibration algorithm will also struggle for very large errors.

b. How could the "significant error" case be identified as part of the sequence of steps followed above?

It can be very difficult to accurately identify the error in camera calibration. The error metrics can be used to identify it. We can also inspect by visually looking at the calibration. We can also test it on different scenes to see if calibration is giving distortions or inaccuracies in capturing new images or not.

c. How do you expect the error in the calibration model to change for the following conditions?

i. Only two images are used, with only a slight rotation between them

It will lead to the very high error in the calibration model. It will not be able to capture enough information about the cameras extrinsic and intrinsic parameters correctly. The two images with slight transformation are not sufficient to capture variations in cameras behavior.

ii. A modest number of images are used, with many variations in image orientations, such as rotations and translations.

It will lead to more accurate calibration model of camera as we are giving diverse set of images having many variations. It will help the camera to capture large distortions and characteristics of the camera. It may require more processing power and time to learn camera parameters.

iii. **A very large number images are used, with many different variations in image orientations, rotations and translations. Keep in mind that several of these images may be very similar to other images and that numerical methods are used as part of the error reduction approach (that is, to find the solutions of the camera model equations). Will the error always goes to zero as the number of images used increases ? Why or why not?**

The error in the calibration model of the camera cannot be zero as the number of images increases because of many factors like:

- Noise in images
- Redundancy
- Lens Distortions
    - Etc.

iv. **If the entered square size is incorrect, the values of dX or dY, how could this be identified during the calibration sequence used?**

If the entered square size is incorrect, then the corners will be detected correctly leading to incorrect camera parameters. It can easily identify visually as well as by error metrics used during the calibration.

d. A) How important is even lighting for this calibration approach?

It is important during the calibration process to maintain even lighting as it helps in reducing the variations in brightness and shadows which can effect the accuracy of measurements. If the lighting is not good, it may lead to incorrect identification of corners that will cause significant errors in the calibration.

B) Why is a regular pattern of squares of known size used, why not just use the outer corner points of each image ?

The regular pattern of squares gives us geometric constraints that is very important for accurately estimating the cameras intrinsic parameters like focal length, principal point, camera center etc. Using this we can easily identify relationship between the 3D world coordinate and its corresponding 2D images on camera.

e. **Increasing wintx and winty should make it easier to find corner points, and therefore reduce error. Why is this statement not true for large values of wintx and winty ? What could increase the error for large wintx and winty values depending on the indicated (clicked) point ?**

The increase in window size can help us easily detecting the corner points but there are certain limitations to it. As the larger window will capture more information that will become challenging for the corner detecting algorithm to accurately detecting the corners. A larger window will increase the chances of overlapping features if multiple edges or corners present in the window, so the algorithm will struggle to differentiate between them. Larger window sizes will also lead to the longer processing time.

f. **Read this blog post, from Cleve Molar of Mathworks, on one definition of a matrix condition number, particularly the part on matrix inverses: https://blogs.mathworks.com/cleve/2017/07/17/what-is-the-condition-number-of-a-matrix/**

How would the condition number of the matrix used in determining camera calibration potentially be significant in finding the numerical solution?

In order to determine the accuracy of the numerical solution, the condition number of the matrix used in camera calibration is essential. A high condition number indicates potential challenges in finding accurate calibration. A lower condition number is preferred as it indicates a more stable and precise calibration procedure.

g. **Floating point machine arithmetic is not always exact. In particular, the minimum difference between floating point numbers that a computer can resolve is often called the "machine epsilon". In MATLAB this is assigned the variable: eps. Type "eps" at the command prompt in MATLAB and report the result. How is this significant in determining the stopping condition of any numerical algorithm, such as gradient descent or Newton's method used in camera calibration?**

The eps of the Matlab are as:

```
>> eps
ans =
   2.2204e-16
```

It ensures the convergence, accuracy and precision of the iterative solutions. That's why it is necessary for the numerical algorithms to use it for camera calibration.

5. Conclusion:

The five images of a checkerboard pattern captured from different perspectives of the camera has been used to estimate the intrinsic and extrinsic parameters of the camera. The Jean-Yves Bouguet Camera Calibration Toolbox for MATLAB is used to perform camera calibration on these images.