HW1 AutoCalib: Camera Calibration using Zhang's Method - RBE549

Nishanth Adithya Chandramouli
Robotics Engineering Department
Worcester Polytechnic Institute (WPI)
Email: nchandramouli@wpi.edu

***Abstract*—In this homework, we calibrated a monocular camera using multiple images of a planar checkerboard pattern based on Zhang's calibration method. The objective of the calibration process was to estimate the intrinsic camera matrix K, radial distortion parameters k, and the extrinsic parameters $(\mathbf{R}_i, \mathbf{t}_i)$ corresponding to each calibration image.**

I. INTRODUCTION

In this homework, a planar checkerboard based calibration approach proposed by Zhang was implemented to estimate the camera parameters using multiple images of a known calibration pattern.

The intrinsic matrix of a pinhole camera can be expressed as:

$$K = \begin{bmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Here, (c_x, c_y) represents the principal point, f_x and f_y denote focal lengths along the horizontal and vertical axes respectively and γ represents the shear parameter which accounts for the non-orthogonality between image axes.

II. INITIAL PARAMETER ESTIMATION

A. Corner Detection

Checkerboard corners were detected from each calibration image using OpenCV's corner detection routine - `cv2.findChessboardCorners`. For each valid image, 54 inner corner points corresponding to a (6×9) checkerboard were extracted in image coordinates.

The corresponding world coordinates were generated assuming the checkerboard lies on a plane $(Z = 0)$ with square size of $21.5\,mm$.

B. Homography Estimation

Since the calibration checkerboard lies on a planar surface, the mapping between the world coordinate system of the checkerboard and the image plane of the camera can be represented using a planar projective transformation known as a homography.

For each calibration image, the relationship between a world point (X, Y) and its corresponding pixel location (u, v) is given by:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{2}$$

where $\mathbf{H} \in R^{3\times3}$ is the homography matrix and s is a scale factor.

Each correspondence between a checkerboard corner in the world frame and its detected pixel location generates two linear equations:

$$\mathbf{Ah} = 0 \tag{3}$$

where $\mathbf{h}$ represents the vectorized form of the homography matrix. The system was solved using Singular Value Decomposition (SVD), and the homography matrix was obtained as the right singular vector corresponding to the smallest singular value.

C. Intrinsic Matrix Estimation

The homography matrix can be decomposed into intrinsic and extrinsic camera parameters as:

$$\mathbf{H} = \mathbf{K} \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \tag{4}$$

where $\mathbf{K}$ is the intrinsic camera matrix and r_1, r_2, and t correspond to the rotation and translation parameters.

Since r_1 and r_2 belong to a rotation matrix, they must satisfy the orthogonality constraints:

$$r_1^T r_2 = 0 \tag{5}$$

$$\|r_1\| = \|r_2\| \tag{6}$$

Using these constraints across multiple homographies, a symmetric matrix $\mathbf{B}$ was formed such that:

$$\mathbf{B} = \mathbf{K}^{-T}\mathbf{K}^{-1} \tag{7}$$

Stacking these constraints results in a homogeneous linear system:

$$\mathbf{Vb} = 0 \tag{8}$$

Solving this system using SVD yields the elements of matrix $\mathbf{B}$, from which the intrinsic parameters f_x, f_y, c_x, c_y, and γ were obtained:

$$\mathbf{K} = \begin{bmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Here, γ represents the shear parameter accounting for non-orthogonality between the image axes.

D. Extrinsic Parameter Estimation

Once the intrinsic matrix $\mathbf{K}$ was computed, the rotation and translation parameters for each calibration image were obtained from the homography matrix using:

$$r_1 = \lambda \mathbf{K}^{-1} h_1 \tag{10}$$

$$r_2 = \lambda \mathbf{K}^{-1} h_2 \tag{11}$$

$$t = \lambda \mathbf{K}^{-1} h_3 \tag{12}$$

where h_1, h_2, and h_3 denote the columns of the homography matrix $\mathbf{H}$ and λ is a normalization factor ensuring unit norm of the rotation vectors.

The third rotation vector (mutually perpendicular) was computed as:

$$r_3 = r_1 \times r_2 \tag{13}$$

The rotation matrix was then formed as:

$$\mathbf{R} = [r_1 \; r_2 \; r_3] \tag{14}$$

To ensure orthonormality, the matrix $\mathbf{R}$ was projected onto the closest valid rotation matrix using SVD

E. Distortion Parameters

In practical cameras, lens imperfections introduce radial distortion which causes image points to deviate from their ideal pinhole projection. This distortion is modeled using two radial distortion coefficients k_1 and k_2.

Given normalized image coordinates (x, y) obtained after perspective projection:

$$x = \frac{X_c}{Z_c}, \quad y = \frac{Y_c}{Z_c} \tag{15}$$

the distorted normalized coordinates are computed as:

$$x_d = x\left(1 + k_1 r^2 + k_2 r^4\right) \tag{16}$$

$$y_d = y\left(1 + k_1 r^2 + k_2 r^4\right) \tag{17}$$

where

$$r^2 = x^2 + y^2 \tag{18}$$

These distorted normalized coordinates are then mapped to pixel coordinates using the intrinsic matrix:

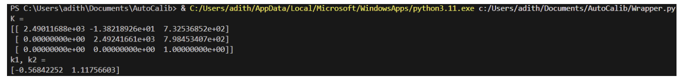

Fig. 1. Obtained Intrinsic matrix and distortion parameters

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} \tag{19}$$

The distortion coefficients were initialized to zero during the homography-based linear estimation stage and subsequently refined during non-linear optimization.

III. Optimization

The initial calibration parameters obtained from homography decomposition minimize only an algebraic error and do not necessarily correspond to the true geometric projection of the scene points. Therefore, a non-linear refinement step was performed to jointly optimize the intrinsic parameters, extrinsic parameters, and radial distortion coefficients.

The optimization minimizes the geometric reprojection error between the detected image corners and the projected world points. The objective function is given by:

$$\sum_{i=1}^{N} \left\| \mathbf{x}_i - \hat{\mathbf{x}}_i(\mathbf{K}, \mathbf{R}_j, \mathbf{t}_j, k_1, k_2) \right\|^2 \tag{20}$$

IV. Results

The final estimated intrinsic matrix is given by:

$$\mathbf{K} = \begin{bmatrix} 2490.12 & -13.82 & 732.54 \\ 0 & 2492.42 & 798.45 \\ 0 & 0 & 1 \end{bmatrix} \tag{21}$$

The estimated radial distortion parameters are:

$$\mathbf{k} = \begin{bmatrix} -0.5684 \\ 1.1176 \end{bmatrix} \tag{22}$$

Refer to Fig. 1 for the terminal output.

V. CONCLUSION

Camera calibration was carried out using Zhang's planar method, where homographies between the world and image planes were first estimated using SVD to obtain initial intrinsic and extrinsic parameters. Radial distortion coefficients were initialized to zero and later refined through joint non-linear least squares optimization by minimizing the reprojection error between detected and reprojected checkerboard corners. The final low reprojection error indicates that the estimated camera model accurately represents the imaging system.

References

[1] Nitin J. Sanket - RBE549 - HW1 Available: https://rbe549.github.io/spring2026/hw/hw1/#coll.

APPENDIX

The following figures show the detected checkerboard corners and the corresponding reprojected points obtained using the optimized calibration parameters.

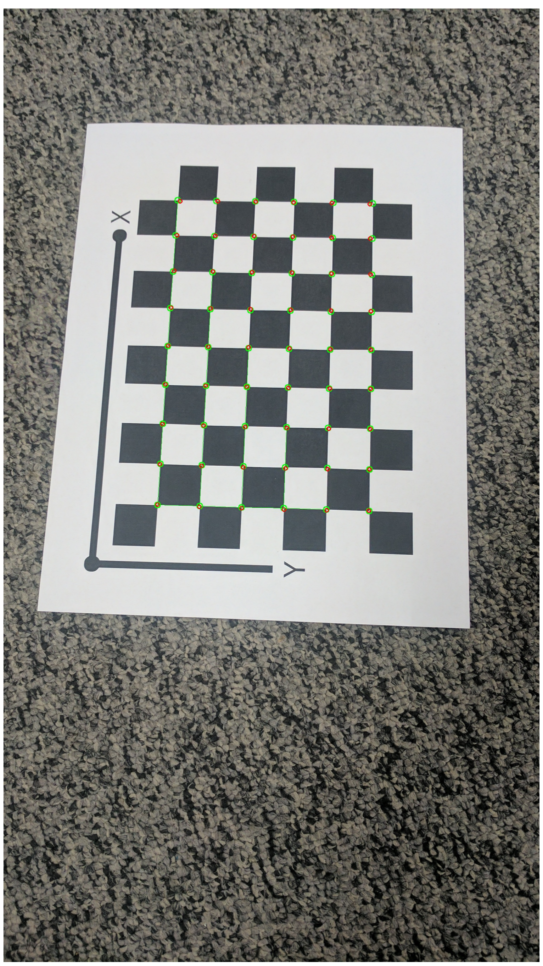

Fig. 2. Calibration Result 1

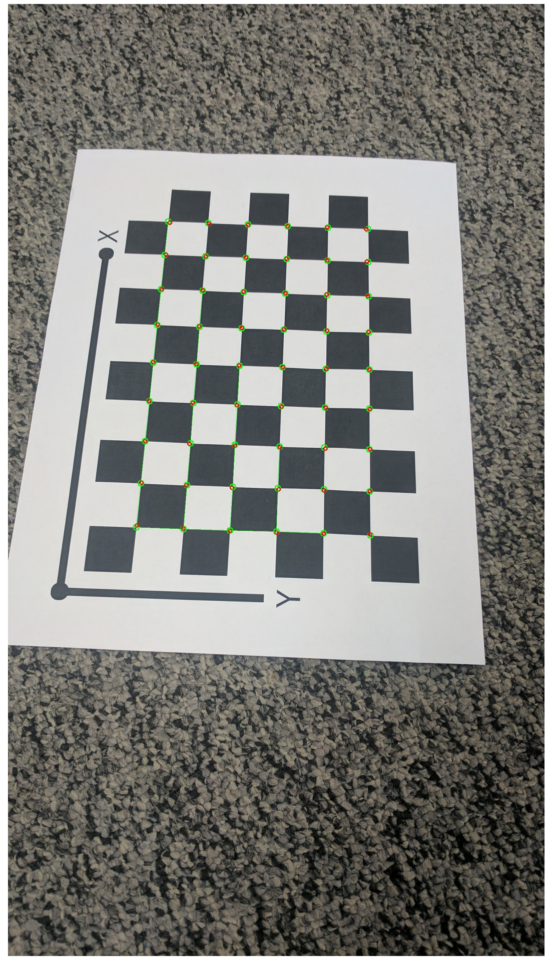

Fig. 3. Calibration Result 2

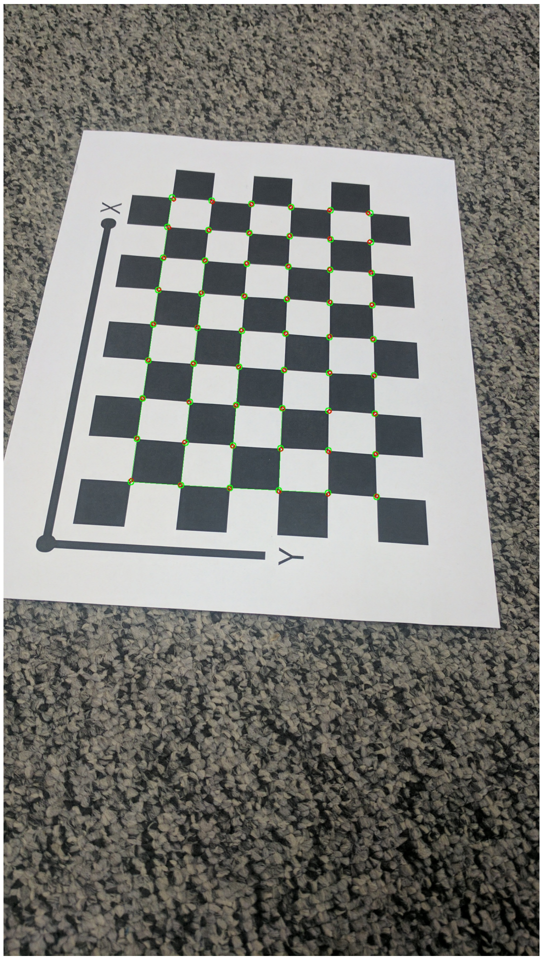

Fig. 4. Calibration Result 3

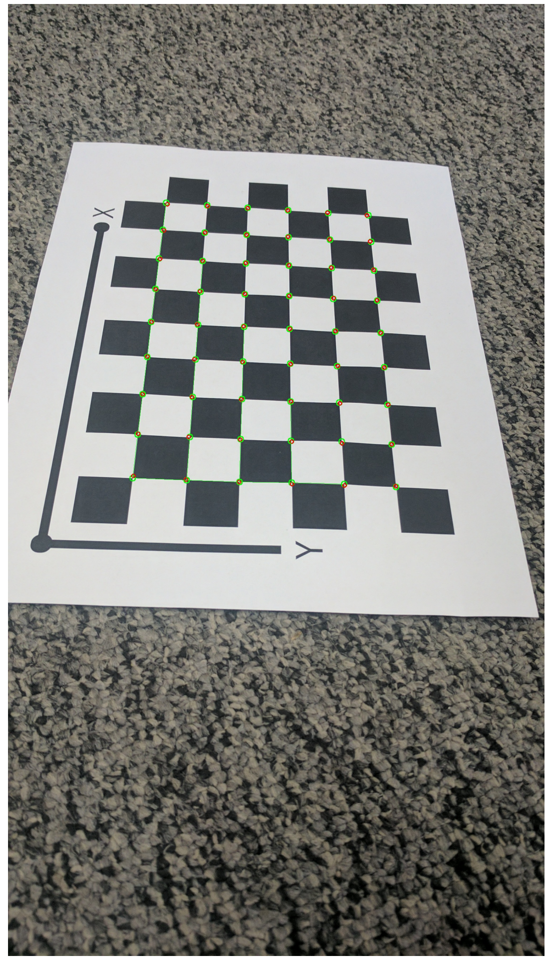

Fig. 5. Calibration Result 4

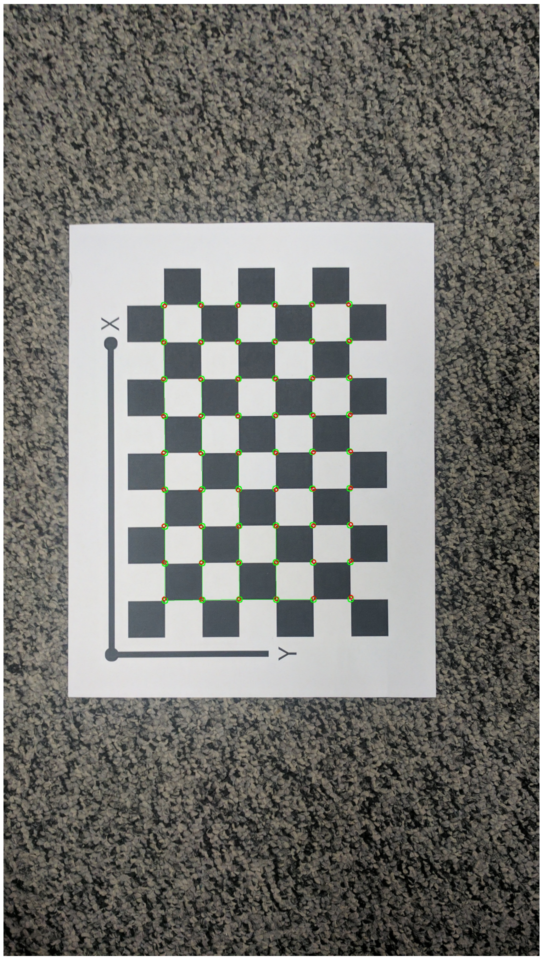

Fig. 6. Calibration Result 5

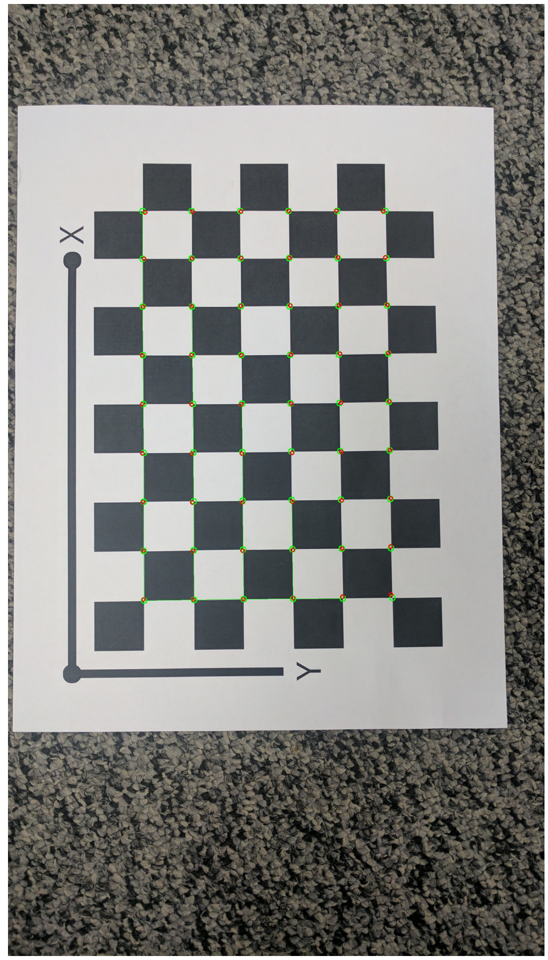

Fig. 7. Calibration Result 6

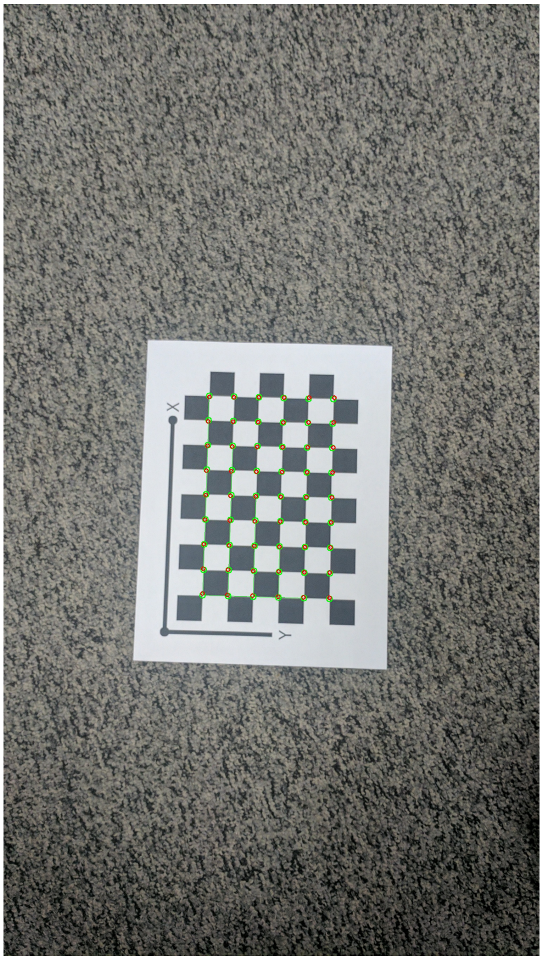

Fig. 8. Calibration Result 7

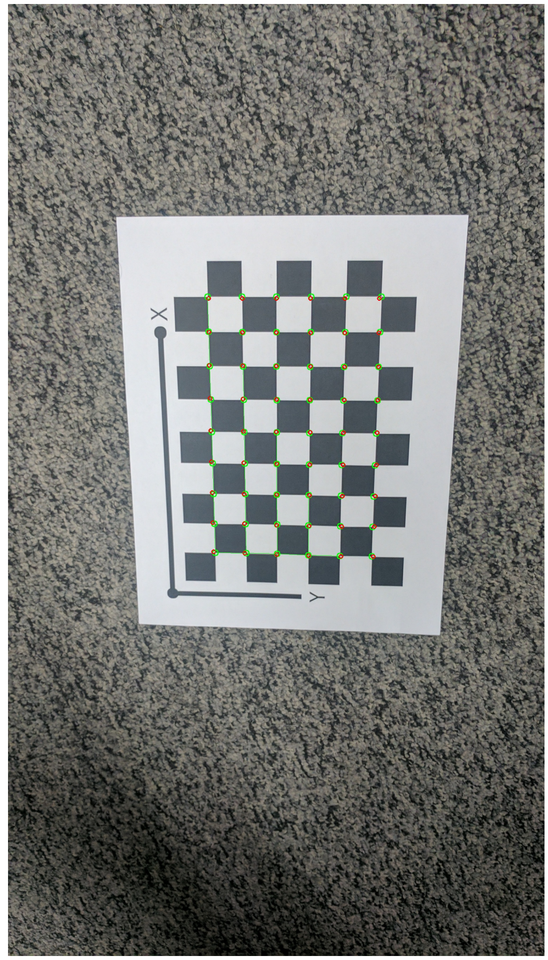

Fig. 9. Calibration Result 8

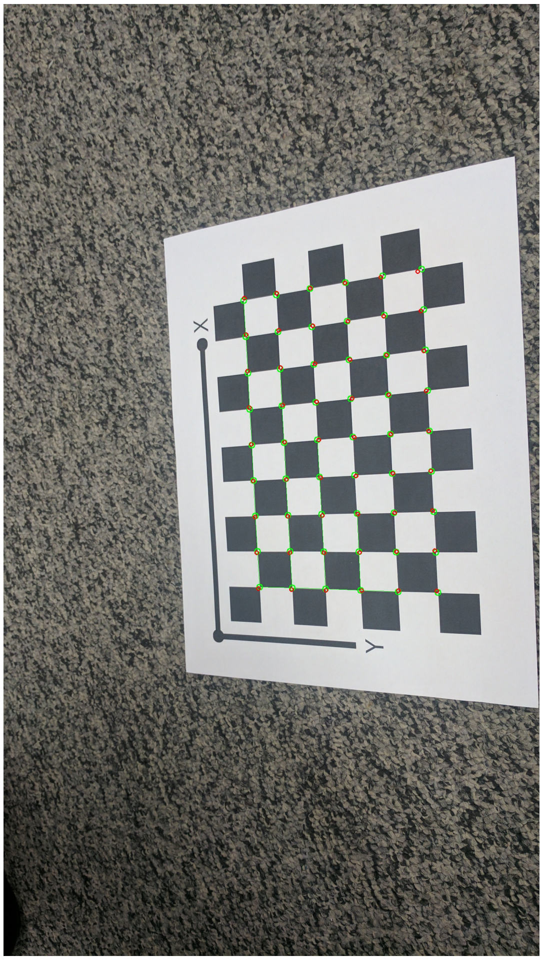

Fig. 10. Calibration Result 9

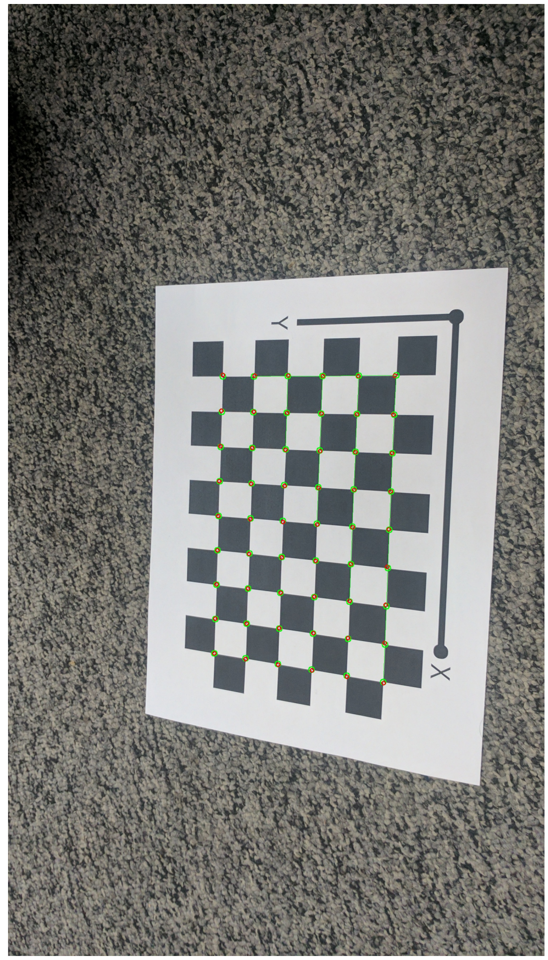

Fig. 11. Calibration Result 10

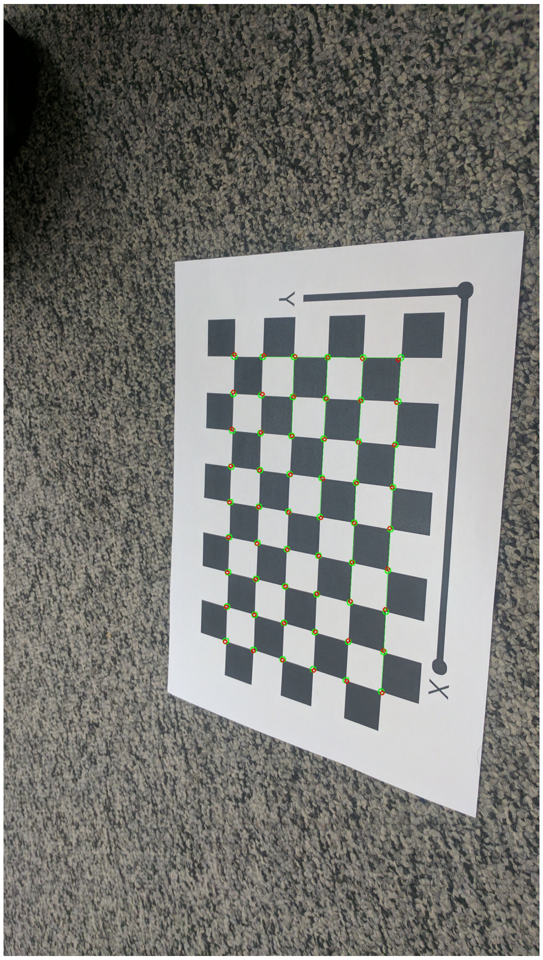

Fig. 12. Calibration Result 11

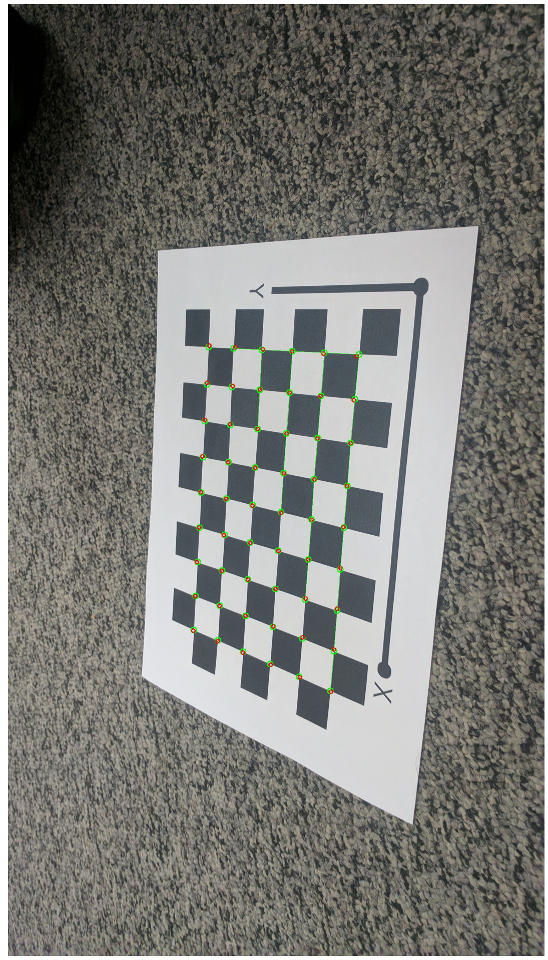

Fig. 13. Calibration Result 12

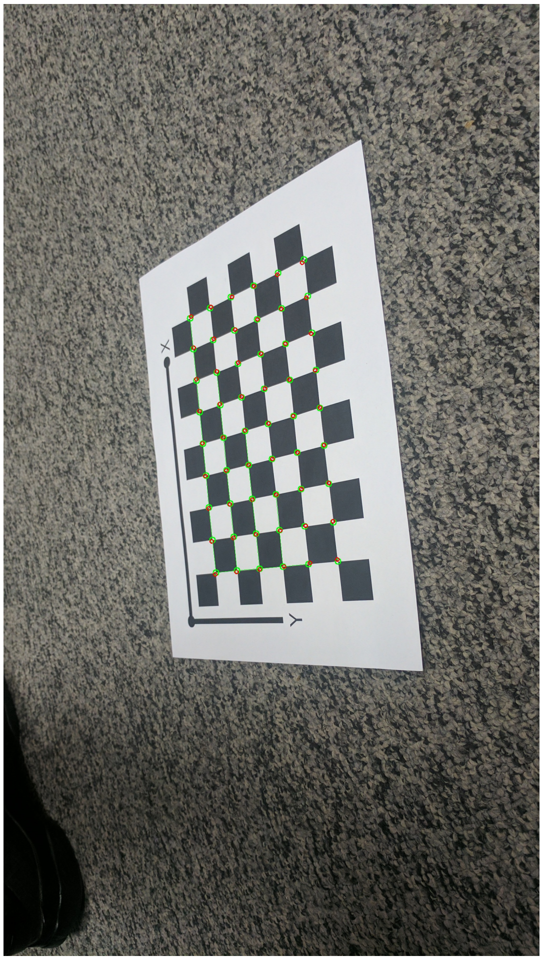

Fig. 14. Calibration Result 13