RBE/CS 549 Computer Vision

HW1 - AutoCalib

Noopur Koshta
Robotics Engineering Department
Worcester Polytechnic Institute
Worcester, MA, USA
nkoshta@wpi.edu

***Abstract*—Estimating the camera intrinsic, extrinsic parameters and distortion coefficients using Zhengyou Zhang's method [1]. The calibration target image is a checkerboard pattern with each square size of 21.5 mm. First, the camera intrinsic matrix (K) is approximated and using that, the extrinsic parameters Rotation (R) and translation (t) are estimated. Then, using these as initial estimates, non linear optimization is done to minimize the geometric error and refine the intrinsic matrix parameters and the distortion coefficients.**

The pipeline used for this homework is given in Zhang's paper. It involves the following steps:

- Corner Detection of Checkerboard pattern
- Estimate Intrinsic Camera Matrix
- Estimate Camera Extrinsics
- Approximate distortion coefficient
- Non-Linear Geometric error minimization

I. CAMERA CALIBRATION MATRIX ESTIMATE (K)

The first step is to find the pixel coordinates of the chessboard corners in each image. Points are found using the cv2.findChessboardCorners function. The patternsize parameter is (9,6) which is the number of inner corners that are to be detected. A total of 54 corner points are found for each image.

Next, a 3D coordinate system for each of the 54 detected points is specified. In the 3D coordinate system the Z value is 0 for the plane of the chessboard. X and Y coordinates vary by the size of the chessboard square which is given to be 21.5 mm. Now, the correspondence between the 3D coordinates and the 2D pixel coordinates are available. Next is to find the homography matrix between the 3D and 2D corresponding points. This is done with the method of direct linear transform

Given n corresponding points, each corresponding pair can be used to create a 2x9 matrix p.

Multiple such matrices are stacked to create a matrix P. Then the system P H = 0 needs to be solved for H. Singular value decomposition (SVD) of P is done i.e P = USV^T.

The last singular vector of V as the solution to H. After obtaining homography matrix, the matrix V which is in the closed form solution.

For n images, stacking n such equations, we get the system V b = 0. This system is solved for b by taking SVD of V matrix and obtaining the right singular vector associated with smallest single value. Once b is obtained, it is essentially the vector.

$$p_i = \begin{bmatrix} -x_i & -y_i & -1 & 0 & 0 & 0 & x_i x_i' & y_i x_i' & x_i' \\ 0 & 0 & 0 & -x_i & -y_i & -1 & x_i y_i' & y_i y_i' & y_i' \end{bmatrix}$$

$$PH = \begin{bmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & x_1x_1' & y_1x_1' & x_1' \\ 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1y_1' & y_1y_1' & y_1' \\ -x_2 & -y_2 & -1 & 0 & 0 & 0 & x_2x_2' & y_2x_2' & x_2' \\ 0 & 0 & 0 & -x_2 & -y_2 & -1 & x_2y_2' & y_2y_2' & y_2' \\ -x_3 & -y_3 & -1 & 0 & 0 & 0 & x_3x_3' & y_3x_3' & x_3' \\ 0 & 0 & 0 & -x_3 & -y_3 & -1 & x_3y_3' & y_3y_3' & y_3' \\ -x_4 & -y_4 & -1 & 0 & 0 & 0 & x_4x_4' & y_4x_4' & x_4' \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & x_4y_4' & y_4y_4' & y_4' \end{bmatrix} \begin{bmatrix} h1 \\ h2 \\ h3 \\ h4 \\ h5 \\ h6 \\ h7 \\ h8 \\ h9 \end{bmatrix}$$

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, \\ h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T$$

$$b = [B11, B12, B22, B13, B23, B33]^T$$

This is used to solve for the intrinsic parameters of the camera matrix. Then, the intrinsic parameter matrix is basically given by

II. COMPUTING CAMERA EXTRINSICS

the principal point obtained is (u0, v0) = (751.6288, 1338.6347)mm and focal lengths (fx, fy) = (2060.7377, 2045.653)mm and skew factor is -4.2 The rotation matrix R and translation vector t for each image can be obtained from the Intrinsic matrix an = d homography matrix of that image by using the following system.

III. NON LINEAR GEOMETRIC ERROR MINIMIZATION

For each pixel coordinate in each image the error is calculated as follows:

Here, (u, v) is the detected corner pixel coordinate, A in intrinsic matrix, [R - t] is the transformation matrix and

$$
\begin{aligned}
v_0 &= (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2) \\
\lambda &= B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})]/B_{11} \\
\alpha &= \sqrt{\lambda/B_{11}} \\
\beta &= \sqrt{\lambda B_{11}/(B_{11}B_{22} - B_{12}^2)} \\
\gamma &= -B_{12}\alpha^2\beta/\lambda \\
u_0 &= \gamma v_0/\beta - B_{13}\alpha^2/\lambda \,.
\end{aligned}
$$

$$
\mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
\begin{aligned}
\mathbf{r}_1 &= \lambda \mathbf{A}^{-1}\mathbf{h}_1 \\
\mathbf{r}_2 &= \lambda \mathbf{A}^{-1}\mathbf{h}_2 \\
\mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\
\mathbf{t} &= \lambda \mathbf{A}^{-1}\mathbf{h}_3
\end{aligned}
$$

X, Y are 3D coordinate points corresponding to the pixel coordinates. The error is obtained by taking the l2 norm of the resulting vector. The above process is the same as executing the following functional over all points and all images. The function used is the maximal likelihood estimate which assumes that the image points are corrupted by independent and identically distributed noise. The errors over all points and images is summed up and the total error is divided by (54*11) to get the mean reprojection error. The mean reprojection error before optimization is obtained as **1.24961151844**

To reduce the error, the functional needs to be minimized for the error while dealing with radial distortion as well.

$$
\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}
$$

$$
\sum_{i=1}^{n} \sum_{j=1}^{m} \|\mathbf{m}_{ij} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathtt{M}_j)\|^2
$$

```
u = 751.6288030124533
v = 1338.6346964135987
lamda = -0.6401102612995773
alpha = 2060.737716085073
beta = 2045.6533992962193
gamma = -4.2400863870756185

Initial estimate of Calibration matrix:

[[ 2.06073772e+03 -4.24008639e+00  7.51628803e+02]
 [ 0.00000000e+00  2.04565340e+03  1.33863470e+03]
 [ 0.00000000e+00  0.00000000e+00  1.00000000e+00]]
Initial estimate of Extrinsic parameters:
Rotation Matrix:

 [[-0.10767199  0.67932285  0.71311771]
 [-0.9913797  -0.15255406 -0.02550025]
 [ 0.07465274 -0.70783085  0.68989268]]

Transaltion Vector:

 [-35.32447236  62.30369705 478.01980446]

Mean Reprojection error before optimization:
1.2496115192269146

Calibration matrix after optimization:

[[ 2.91705302e+03 -1.79973987e+03  9.27435860e+02]
 [ 0.00000000e+00  1.40117826e+03  1.39347300e+03]
 [ 0.00000000e+00  0.00000000e+00  1.00000000e+00]]

Distortion coefficients after optimization:
0.015325765990402088, -0.026949984687643192

Mean Reprojection error after optimization:
1.1942568294066958
```

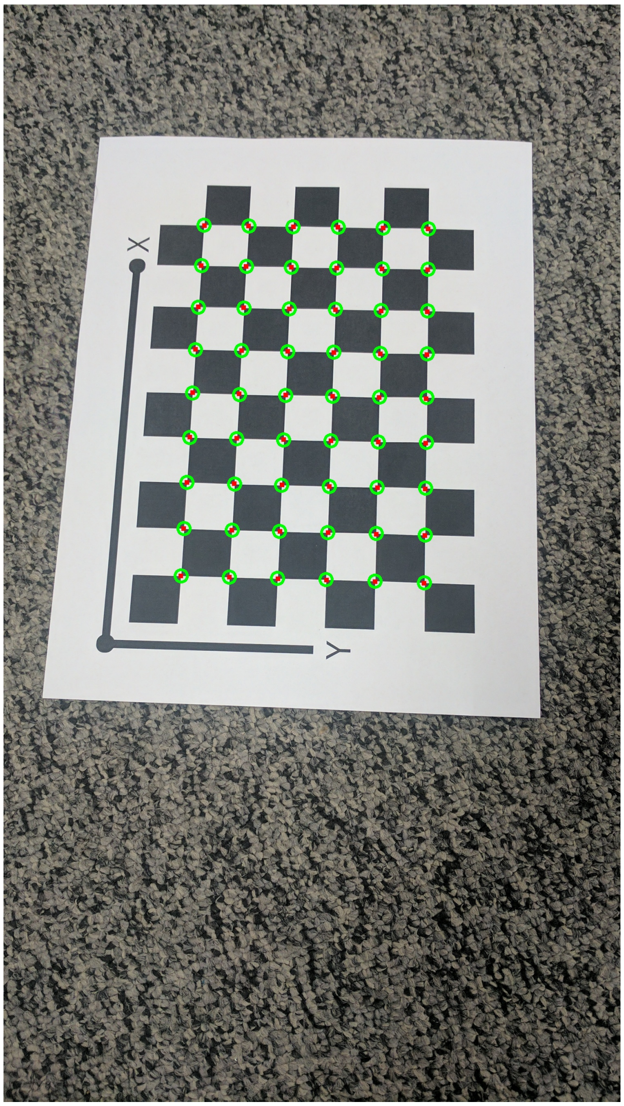
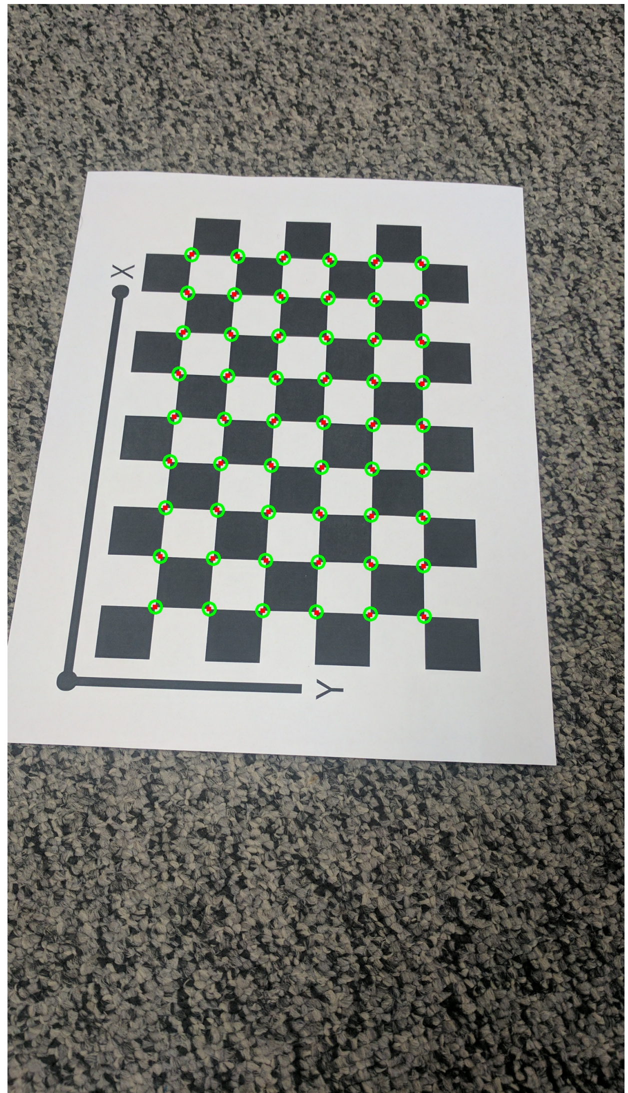
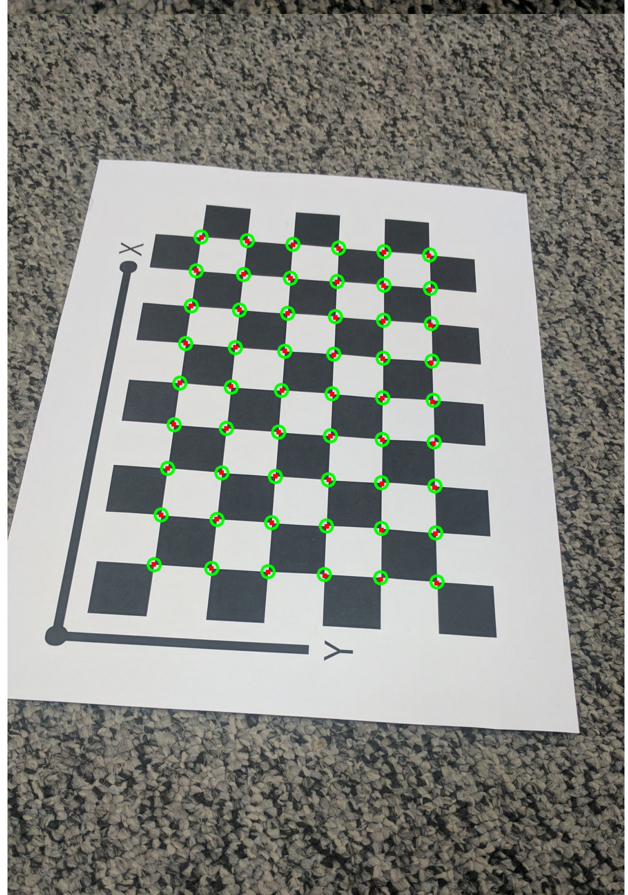
In order to obtain error due to distortion. The distortion coeffs after optimization are **(k1, k2) = (0.0158888379148, -0.0289728808506)**. The mean reprojection error after optimization is **1.1942793426** The reprojected points and the original detected corners are plotted on the original images as shown below. The original detected corners are represented by green circles and the reprojected points are denoted by red squares.

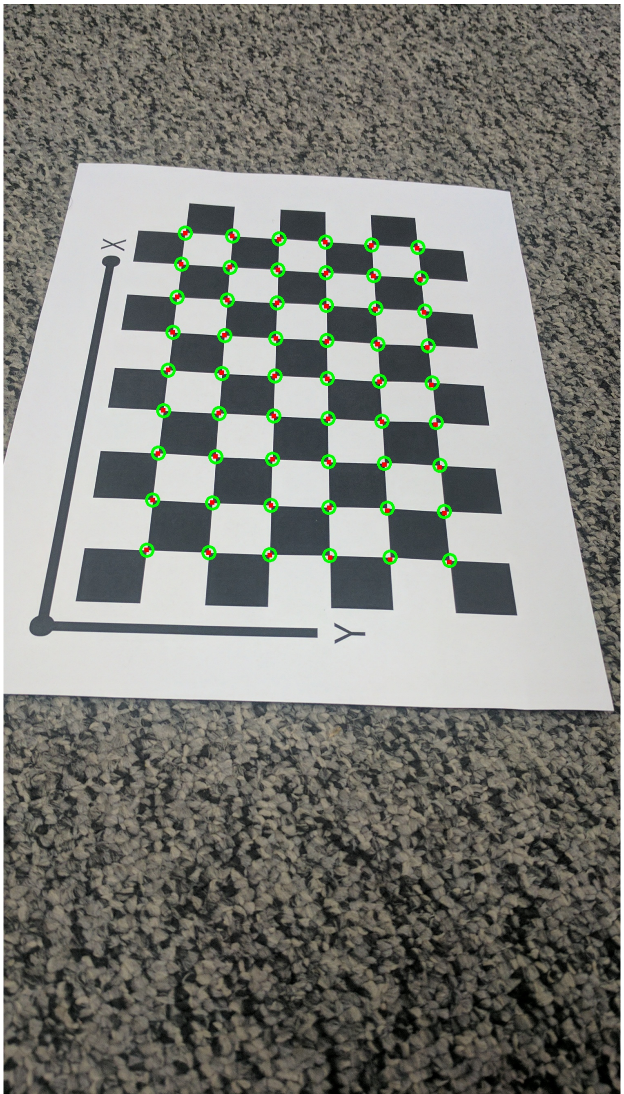
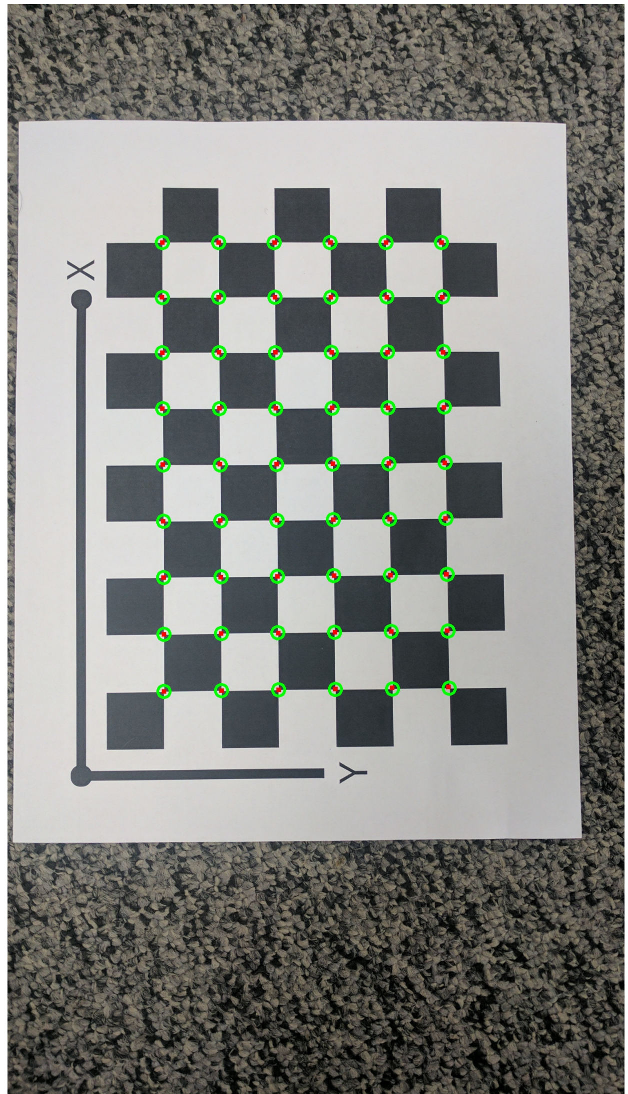
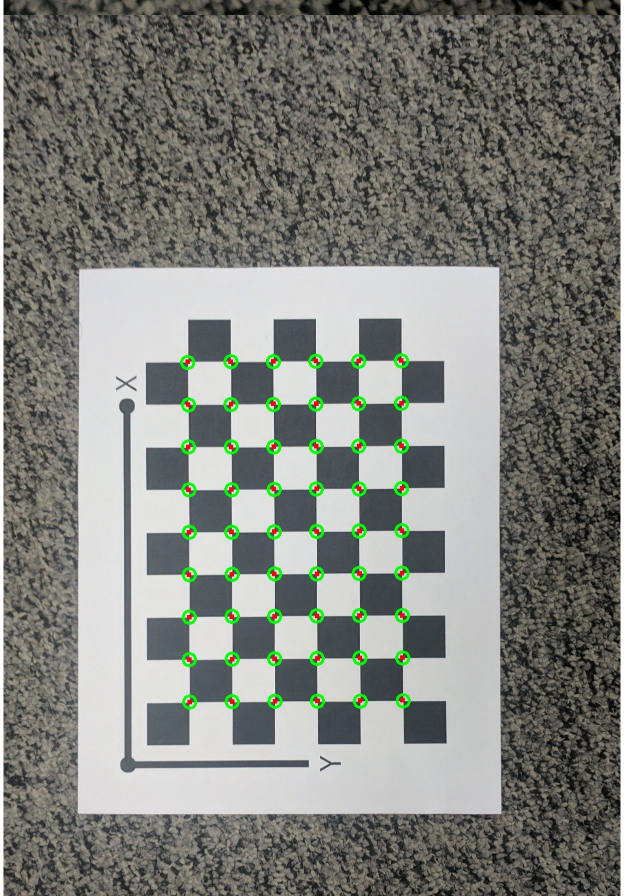
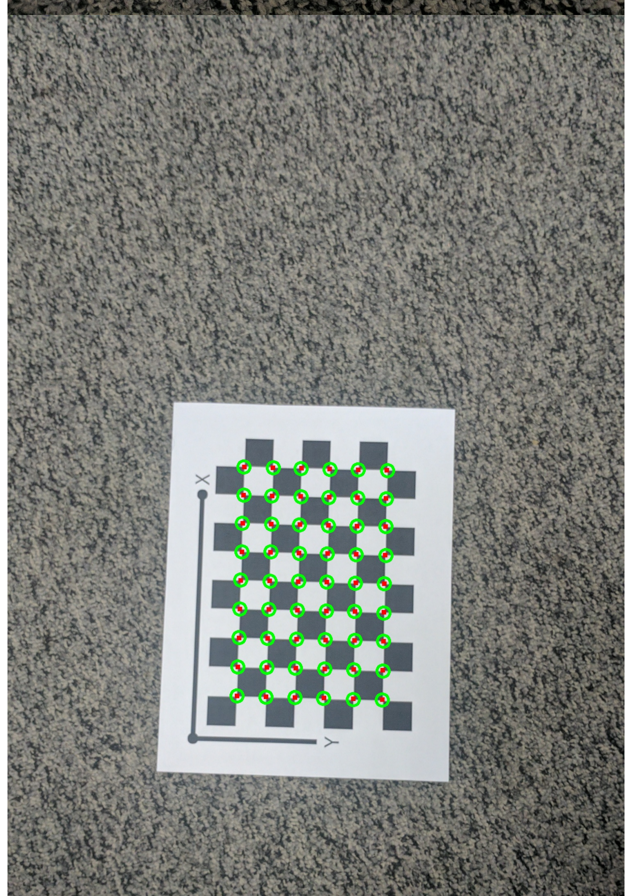
A. Conclusion

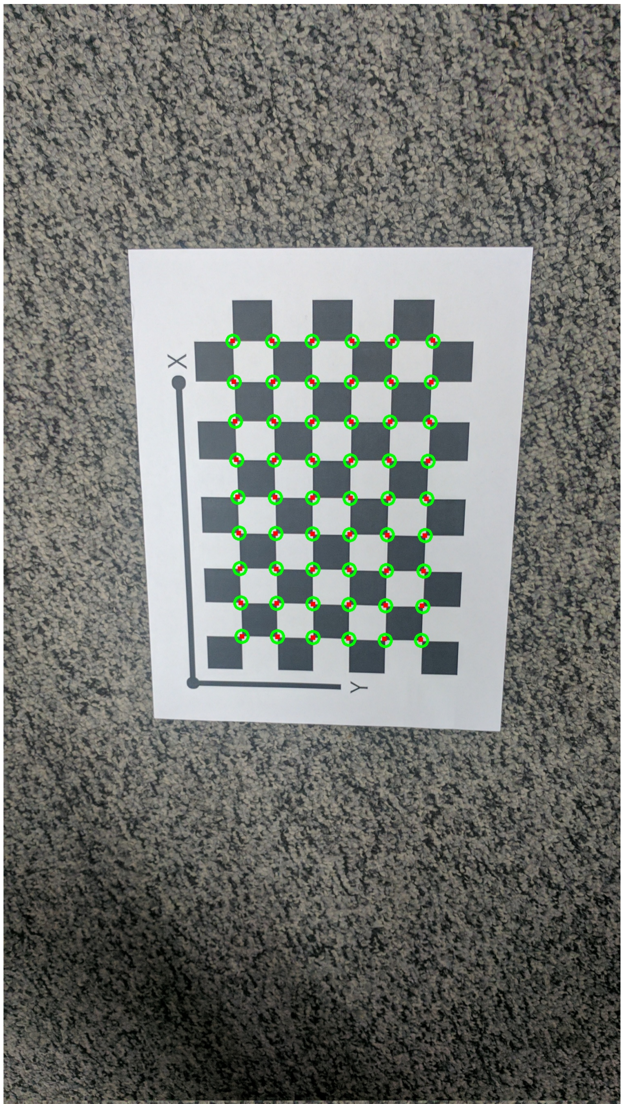
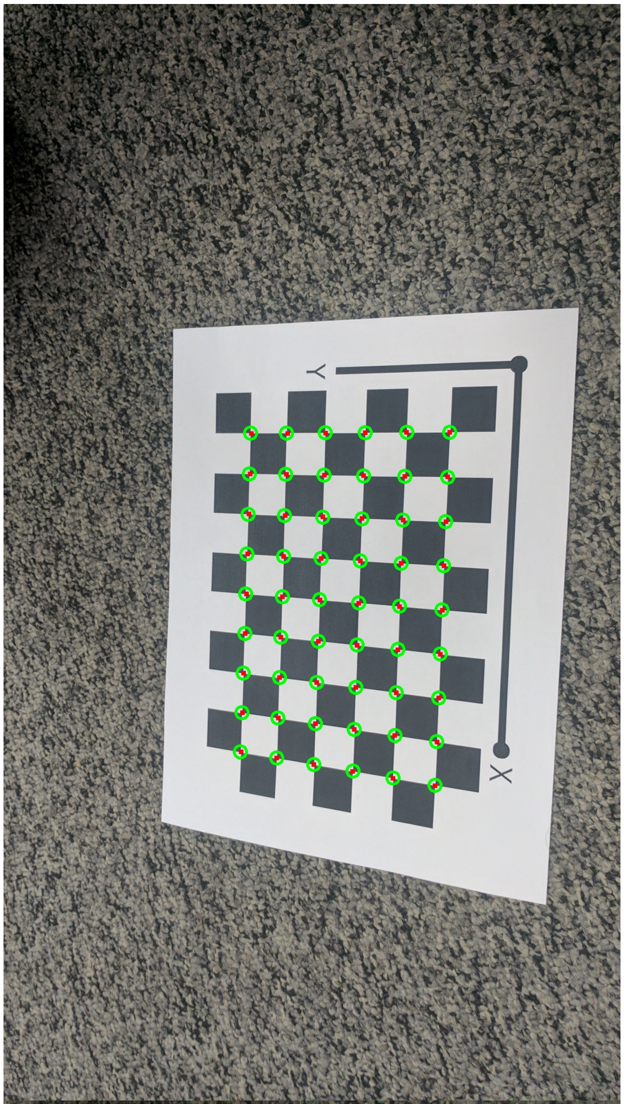
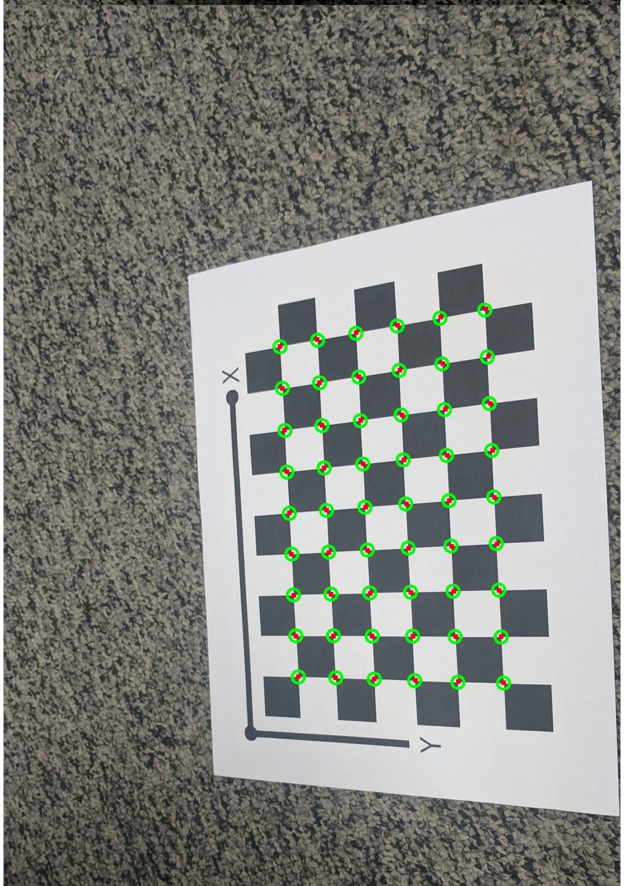
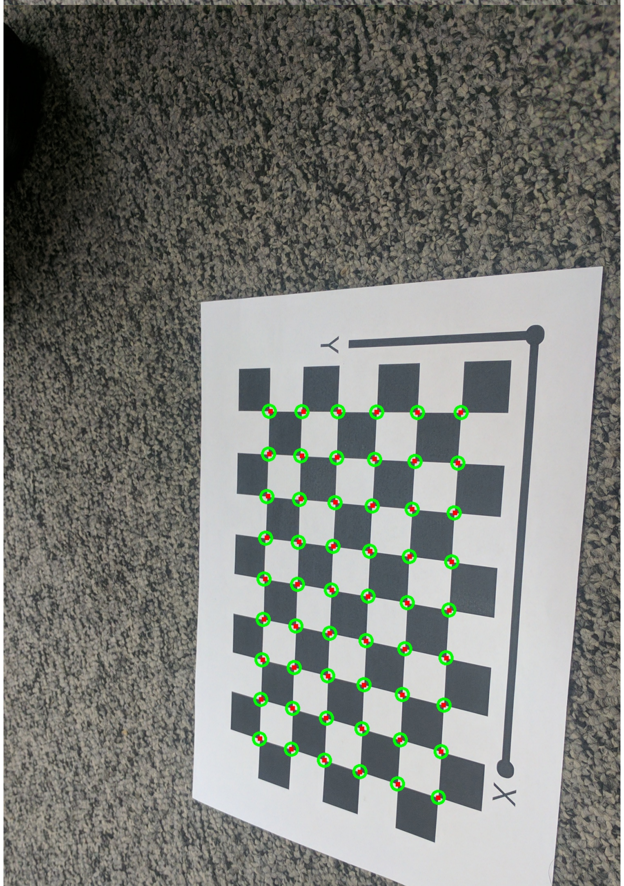
In this report, procedure for calibrating a camera using checkerboard images is discussed and described in detail. The final re-projection error and K intrinsic camera matrix is reported. Non-Linear geometric error minimization is also performed to get the optimal parameter

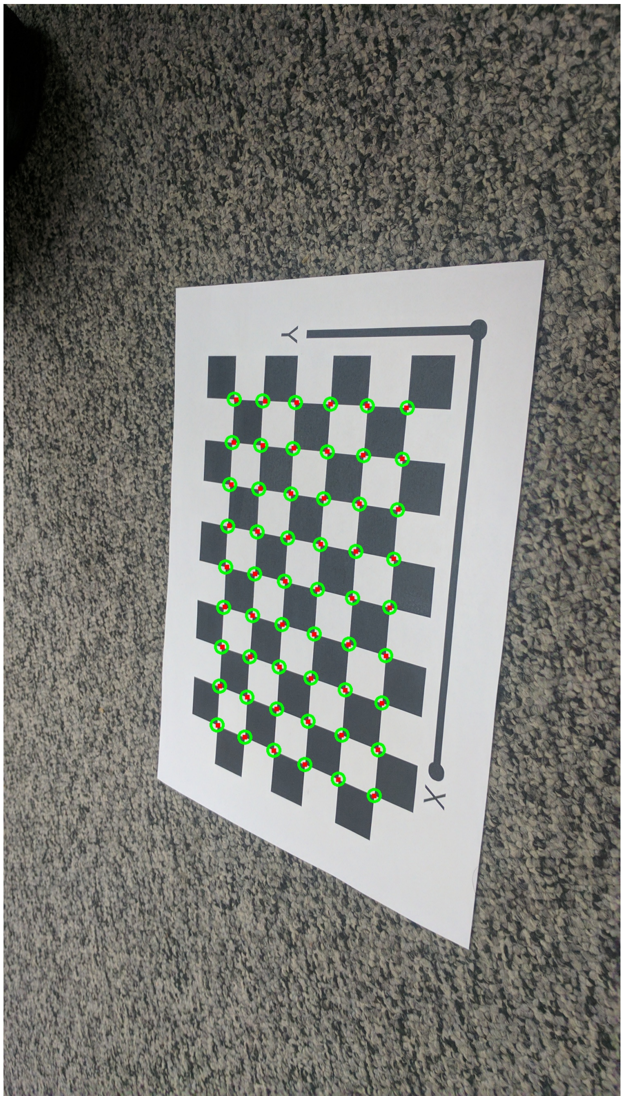
References

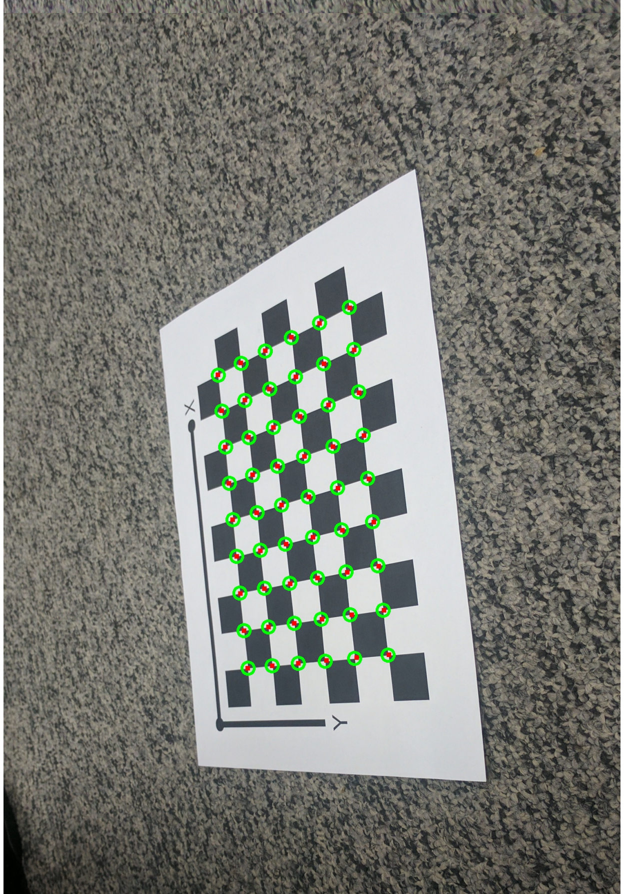
[1] Zhengyou Zhang, A Flexible New Technique for Camera Calibration, Microsoft Research, Redmond, WA, 1998.


X
Y

X
Y

X
Y


X
Y

X
Y

X
Y

X
Y


X
Y

Y
X

X
Y

Y
X


Y
X


X
Y