Homework 1: Autocalib

Dushyant Patil

I. Phase1: Initial Parameter Estimation

A closed form solution is used to find an initial estimate of camera intrinsic and extrinsic parameters - A, R, t. A set of 13 images with a checkerboard pattern of grid size 10x7 are used as input images for calibration. The lens distortion parameters k = [k_1, k_2] are not accounted for this section of the calibration. The intrinsic parameters consist of focal lengths f_x, f_y, image principal coordinates u_0, v_0, scale factor λ and skew γ. The extrinsic parameters consists of rotation and translation matrix [R t] between world points and image points. The governing equation between world points and image points with the help of intrinsic and extrinsic parameters:

$$s\hat{m} = H\hat{M} \tag{1}$$

1. Closed-Form Solution - Intrinsic Matrix

The first step in finding the intrinsic parameters is to find the world points of the known pattern (checkerboard in our case) and its corresponding image points. The image points are obtained using the in-built cv2 function *cv2.findChessboardCorners*. The image points obtained using this function are shown in figures below. The homography between these world points and image points is found using *cv2.findHomography* for all 13 images. The homographies are used to find the B matrix[1] which is given as follows:

$$\text{B} = A^{-T}A^{-1}$$

B is a symmetric matrix which is defined by a 6D vector as follows:

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T$$

The b vector is a solution to the following equation:

$$\mathbf{Vb} = \mathbf{0} \tag{2}$$

The vector v is given as follows:

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

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0} .$$

The h_{i1}, h_{j1}, h_{i3},etc are the respective elements of homography matrix found before for each image. Solving equation (2) for b vector using SVD, we are able to obtain the elements of B matrix. The B matrix is used to get the intrinsic parameters as per the appendix B in the paper by Zhang[1]. Using the B matrix, initial estimation of intrinsic camera parameters are as follows:

$$A = \begin{bmatrix} 2035.62 & 0.11187 & 773.202 \\ 0 & 2019.24 & 1360.42 \\ 0 & 0 & 1 \end{bmatrix}$$

The above matrix gives following camera parameters:
f_x = 2035.62
f_y = 2019.24
γ (skew) = 0.11187
u_0 = 773.202
v_0 = 1360.42

2. Closed-Form Solution - Extrinsic Matrix

The extrinsic matrix includes both rotational and translational matrix concatenated in one 3x4 matrix. The extrinsic matrix will be different for each image as the homography between the same checkerboard and different image points will be different. The extrinsic matrices are found using below equations as per section 3.1 in the paper by Zhang[1]

$$\lambda = \frac{1}{\|A^{-1}h_1\|_2} = \frac{1}{\|A^{-1}h_2\|_2}$$

$$r_1 = \lambda * A^{-1} * h_1 \tag{3}$$

$$r_2 = \lambda * A^{-1} * h_2 \tag{4}$$

$$r_3 = r_1 \times r_2 \tag{5}$$

$$t = \lambda * A^{-1} * h_3 \tag{6}$$

II. Non-Linear Geometric Error Optimization

We have the initial estimates of K,R,t,k, now we want to minimize the geometric error. The optimization problem is given as follows:
$argmin_{fx,fy,cx,cy,k1,k2}\Sigma_{i=1}^{N}\Sigma_{j=1}^{M}$ $\|$x$_{i,j}$-$x_{i,j}$(K, R_i, t_i, X_j, k)$\|$

The initial parameter estimates are the intrinsic matrix elements and extrinsic matrices [R t] found in section above. The initial values of K_1 and k_2 are estimated to be zero. The initial parameter estimates were optimized using the function *scipy.optimize*. The loss function which was used in the optimization was the least square function between the image points and world point projections with the help of intrinsic and extrinsic matrices derived in the above section using closed-form solution. The projection coordinates are derived by applying lens distortion equations to eq (1) as shown below:

$$\hat{u} = u + (u - u_0)[k_1 r^2 + k_2 r^4] \tag{7}$$

$$\hat{v} = v + (v - v_0)[k_1 r^2 + k_2 r^4] \tag{8}$$

$r^2 = x^2 + y^2$.
Optimizing the reprojection error using Levenberg-Marquardt method, we obtain following intrinsic matrix:

$$A_{opt} = \begin{bmatrix} 2035.62 & 0.11187 & 773.202 \\ 0 & 2019.24 & 1360.42 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_{lens} = \begin{bmatrix} -1.08e-9 & 2.6389e-23 \end{bmatrix}$$

The error per pixel for all 13 images is 1.5503 after implementation of $scipy.optimize.least\_squares$

III. Output images and observations

The re-projection points were found using equation (1), (7) and (8). The image points obtained initially using cv.findCheesboardCorners are shown in blue circles. The output images with re-projection points are shown in figures below. The chessboard corners highlighted in red using re-projection points are slightly offset 1-2 pixels for majority of points in majority of images. But some points are off from real image points by more distance - 6-10 pixels. This major offset is observed mainly on the images which are either tilted or have skewed perception (namely image3 and image12). The offsetting effect can be minimized if these images are ignored from the calibration input altogether.

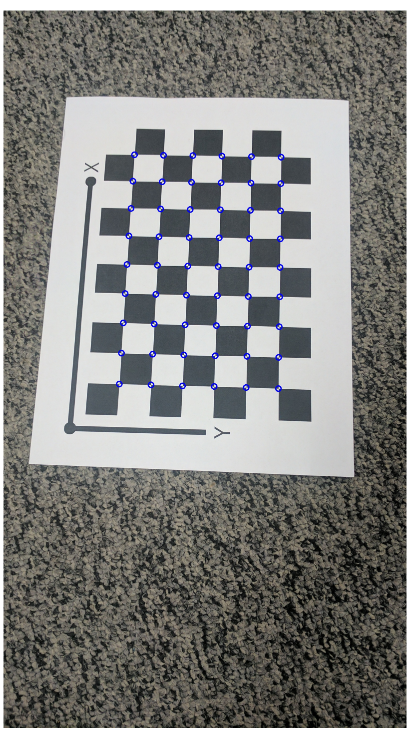

Fig. 1: Image points for image 10 (index 9)

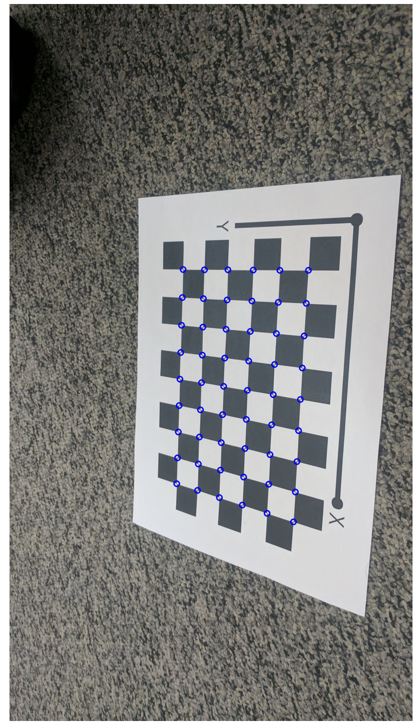

Fig. 2: Image points for image 4

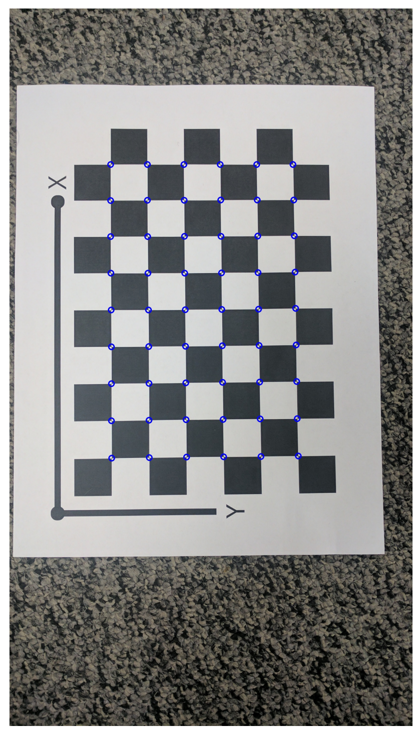

Fig. 3: Image points for image 1

IV. References

[1] Zhengyou Zhang, A Flexible New Technique for Camera Calibration. In MSR-TR-98-71, December 2, 1998

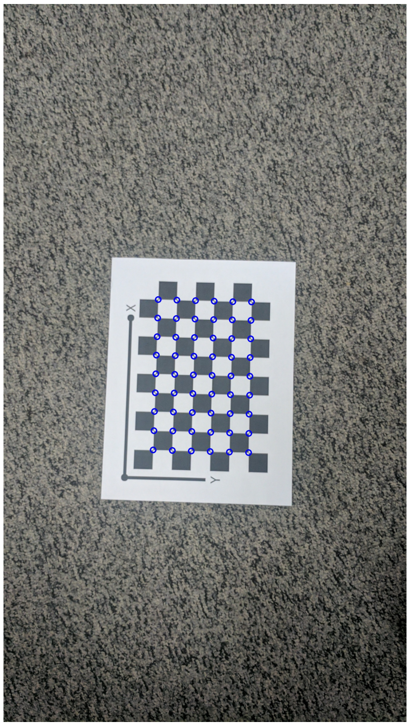

Fig. 4: Image points for image 13

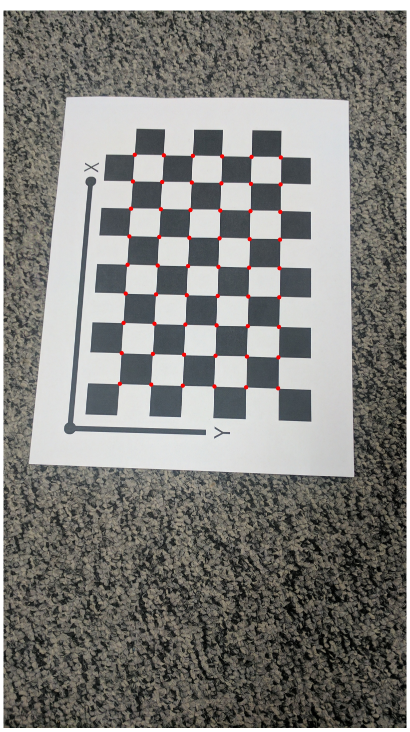

Fig. 5: Calibrated output image 10 (index 9)

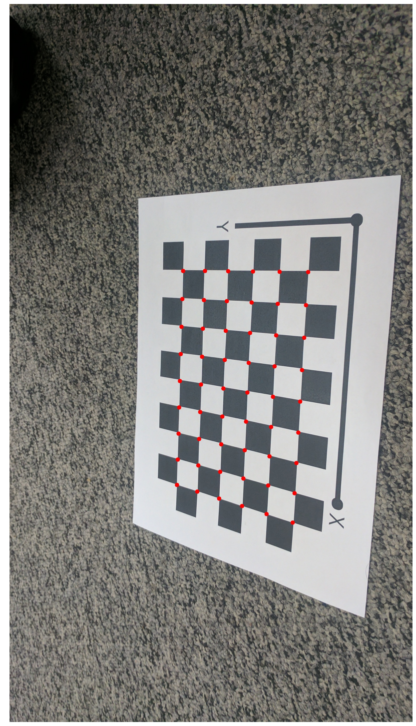

Fig. 6: Calibrated output image 4

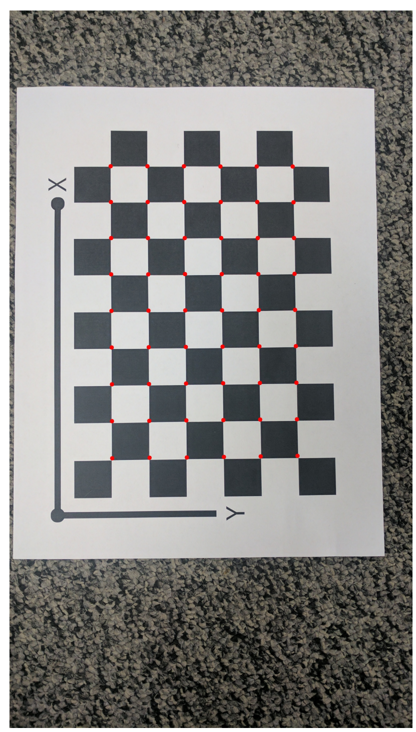

Fig. 7: Calibrated output image 1

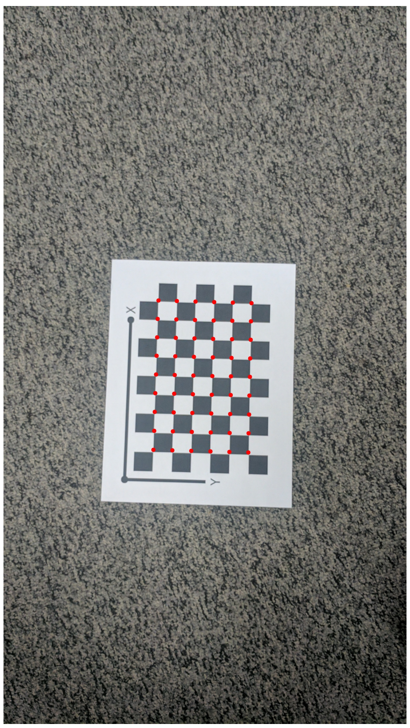

Fig. 8: Calibrated output image 13