Assignment-1 camera calibration

Guruvu Surya Sai Prakash 2019EE10481

Mukku Vamsi Krishna Reddy 2019CS10374

Abstract:

In this study, we calibrate the camera on a mobile phone by utilising a checkerboard object. To calibrate a camera, its intrinsic and extrinsic properties must be estimated. Included among the intrinsic characteristics are the focal length, skew, and optical centre of the camera. For each shot the mobile phone takes, the extrinsic parameters compute the rotation and translation. In addition to the XYZ coordinates of the spots highlighted on the wall checkerboard, a dataset was made that also included the XY coordinates of the corresponding points on the picture. The report goes over the estimation process for the projection matrix and breaks it down into its intrinsic and extrinsic characteristics.

Theory:

The process of estimating a camera's characteristics, including its intrinsic and extrinsic parameters, which characterise the camera's intrinsic qualities and its position in relation to the scene, is known as camera calibration. The intrinsic parameters, which include the focal length, distortion, and primary point, characterise the image characteristics of the camera. The camera's location and orientation in relation to the scene are described by the extrinsic parameters. The 2D and 3D correspondences between the locations on the picture and the scene may be used to estimate the intrinsic and extrinsic parameters.

Procedure:

The following steps were followed to perform camera calibration:

0) **Data preparation:** We have taken the 15 images in different views and taken 16 random points(corners) in each image for estimating the parameters. (We need at least 6 fiducial points, because we have 11 unknowns, each point constitutes 2 equations).

1) **Normalization of the data:** The data was normalized such that the centroid of 2D and 3D points was at the origin, and the average Euclidean distance of 2D and 3D points from the origin was √2 and √3, respectively. The transformation matrices T and U were found to achieve this for 2D and 3D points, respectively. (T and U matrices could be seen in the code).

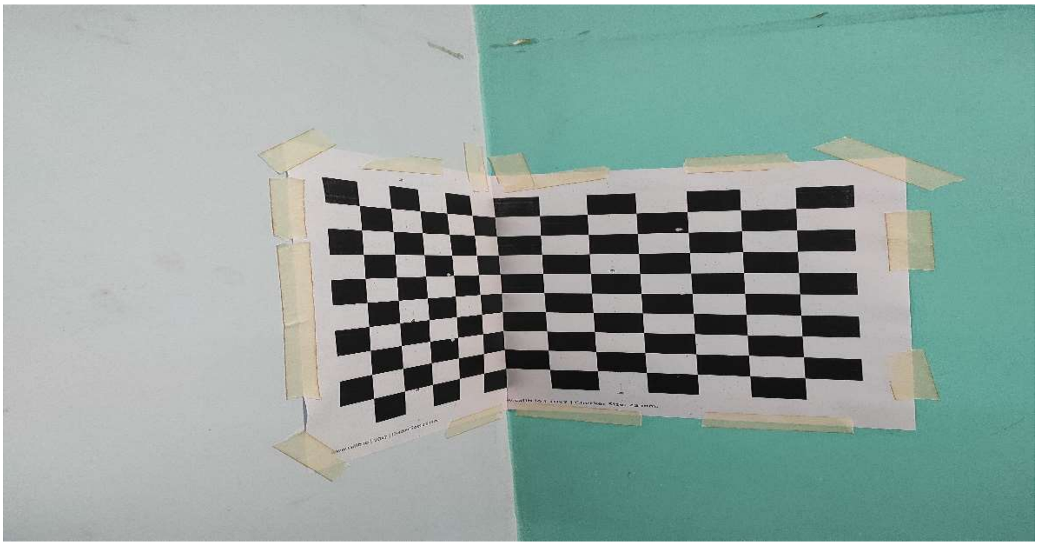
2) **Estimation of the normalized projection matrix:** The normalized projection matrix was estimated using the DLT method. The DLT method is a linear method for estimating the projection matrix that minimizes the sum of squared distances between the observed and estimated points.

3) **Denormalization of projection matrix:** We perform denormalization $P = T^{-1} \hat{P} U$ so that we can get the actual image coordinates not the normalized ones. First, we apply U to a 3-d un-

normalized world coordinate to make it normalized and we can apply P^ which takes normalized world coordinates and maps them to normalized image coordinates then on applying T^{-1} we de-normalize the image coordinates.

4) **Decomposition of the projection matrix:** The projection matrix was decomposed into the intrinsic matrix K, rotation matrix R, and the camera center Xo. The K and R matrices were estimated using the RQ decomposition.
5) **Verification of the projection matrix:** The projection matrix was verified by computing the RMSE between the 2D points marked on the image and the estimated 2D projections of the marked 3D points. The results were visualized on the image and included in the report.

Results and observations:



We are sampling 16 points from the corners of the checkerboard whose world coordinates are predetermined. (Here we have taken 16 points to reduce reprojection error) for an image.

s.no	wc_x	wc_y	wc_z	X_img	Y_img	X_cal	Y_cal
1	0	120	96	1322.212	2357.581	1321.459	2356.601
2	0	144	72	1211.916	2158.508	1203.145	2161.494
3	0	120	24	1302.314	1762.499	1302.393	1763.761
4	72	0	96	2312.202	2322.738	2312.188	2322.32
5	0	48	96	1597.479	2337.636	1617.002	2334.734
6	0	48	72	1592.238	2167.207	1611.747	2155.128
7	96	0	24	2480.373	1777.721	2486.238	1782.023
8	0	0	144	1822.881	2657.927	1796.157	2659.876
9	0	144	48	1205.786	1946.583	1196.358	1957.322
10	96	0	216	2512.508	3212.718	2515.719	3210.449
11	0	120	48	1312.509	1962.924	1308.773	1962.137
12	0	144	168	1233.489	2989.787	1308.773	2970.114
13	72	0	168	2328.021	2848.256	1230.036	2970.114
14	0	72	192	1532.335	3054.856	2324.023	2849.475
15	0	96	120	1428.589	2537.316	1546.543	3074.978
16	48	0	72	2137.164	2152.862	2127.945	2147.818

K: [[6.04025155e+03 6.35634405e+01 1.71892791e+03]

[0.00000000e+00 5.49836664e+03 2.44468401e+03]

[0.00000000e+00 0.00000000e+00 1.00000000e+00]]

R: [[0.89709367 -0.44157863 0.01520739]

[0.01013799 0.05498078 0.99843594]

[-0.44172409 -0.89553639 0.05379963]]

t: [[7.47696843]

[-113.19964183]

[773.47681656]]

RMSE: 14.049258025135176.

The camera calibration was successful in estimating the intrinsic and extrinsic parameters of the mobile phone camera. The RMSE between the 2D points marked on the image and the estimated 2D projections of the marked 3D points was low, indicating that the projection matrix was estimated accurately.

K_avg = [[6.00660893e+03 1.06360723e+02 2.19607755e+03]
[0.00000000e+00 5.93629439e+03 2.06138853e+03]
[0.00000000e+00 0.00000000e+00 1.00000000e+00]]

Skew_angle: 90.602

Effective focal length along x_axis = 6006

Effective focal length along y axis = 5936

Principal point = (2196,2061)

Conclusion:

A checkerboard object and a collection of 2D and 3D correspondences were used to correctly calibrate the camera of a mobile phone. The DLT approach and RQ decomposition were used to precisely estimate the camera's intrinsic and extrinsic characteristics. By calculating the RMSE between the 2D points marked on the picture and the estimated 2D projections of the marked 3D points, the findings were confirmed. It was discovered that normalising the data before using the DLT approach was a good idea since it increases the precision and stability of the calculated parameters.