

# Computer Vision Assignment 1

Gowri Lekshmy

20171053

## Question 1 : Direct Linear Transform

Direct linear transform (DLT) maps any object point  $X$  to the image point  $x$ .

$$x = KR \begin{bmatrix} I & | & -X_0 \end{bmatrix} X$$

$$= PX$$

We try to find the 11 intrinsic and extrinsic parameters by using DLT. Firstly, 20 image and 3D object points are found out, then we try to find  $M$  matrix by using the following equation.

$$\begin{bmatrix} a_{x_1}^T \\ a_{y_1}^T \\ \vdots \\ a_{x_i}^T \\ a_{y_i}^T \\ \vdots \\ a_{x_I}^T \\ a_{y_I}^T \end{bmatrix} p = \begin{matrix} M \\ 2I \times 12 \end{matrix} \begin{matrix} p \\ 12 \times 1 \end{matrix} \stackrel{!}{=} 0$$

We later perform SVD decomposition on  $M$  matrix, and obtain Projection matrix from the last row of  $V^T$  and resize it to  $(3,4)$ . Then for finding the rest of the parameters the following equations are used.

$$\hat{P} = \hat{K} \hat{R} \begin{bmatrix} I_3 & | & -\hat{X}_O \end{bmatrix} = \begin{bmatrix} \hat{H}_\infty & | & \hat{h} \end{bmatrix}$$

$\begin{matrix} 3 \times 3 & & 3 \times 1 \end{matrix}$

$$\hat{H}_\infty = \hat{K} \hat{R} \quad \hat{h} = -\hat{K} \hat{R} \hat{X}_O$$

$$\hat{X}_O = -\hat{H}_\infty^{-1} \hat{h} \quad \hat{H}_\infty^{-1} = \hat{R}^T \hat{K}^{-1} \quad \hat{\hat{K}} = \frac{1}{\hat{K}_{33}} \hat{K}$$

**Results obtained :**

```

Projection Matrix
[[ 2.43230665e+00  6.06749709e-01 -6.16125164e+00  1.54083681e+03]
 [-1.36811875e+00  6.68999676e+00 -1.24293395e+00  1.59596820e+03]
 [-1.18597072e-03  3.82948419e-04 -1.10526986e-03  1.00000000e+00]]
projection_center
[433.70355384 -72.92804573 414.11841734]
Camera Calibration matrix
[[ 2.70167789e-04 -4.48008173e-07 -4.03902385e-01]
 [-0.00000000e+00  2.73707613e-04 -5.48268951e-01]
 [-0.00000000e+00 -0.00000000e+00  1.00000000e+00]]
Rotation matrix
[[-0.68242233 -0.00375398  0.73094848]
 [-0.16554867 -0.97320881 -0.15955641]
 [ 0.71196447 -0.22989241  0.66351795]]

```

**RANSAC Method:**

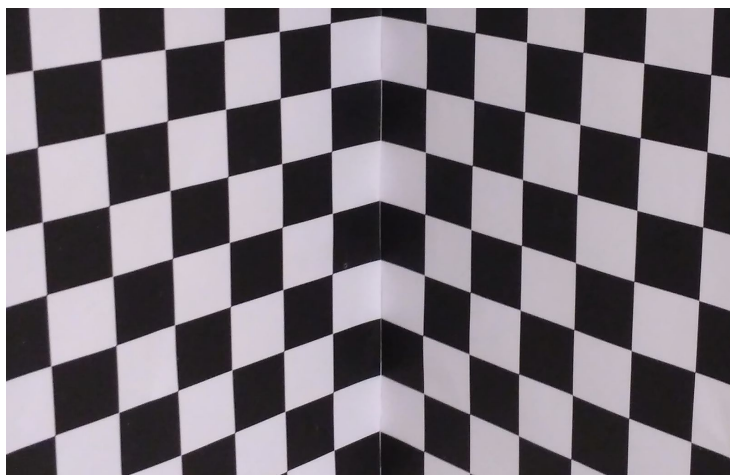
6 points are randomly sampled from 20 points forming different subsets, then a new Projection matrix is estimated using these set of 6 points. For each combination,  $(x - PX)$  is found and the mean square error is found, the Projection matrix corresponding to the least error is chosen.

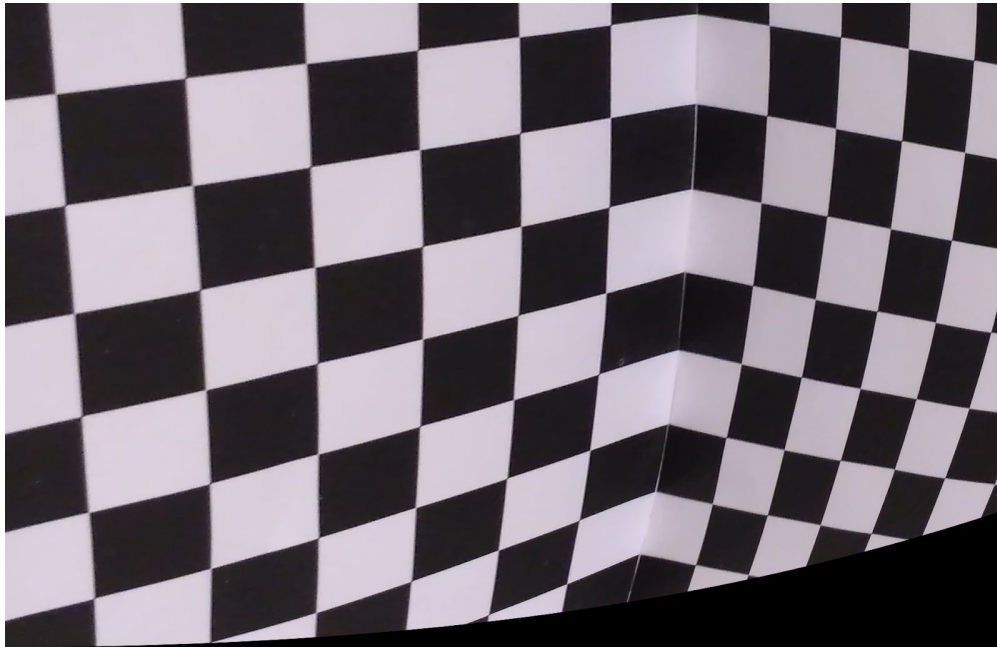
**Best P**

```

[[-1.84404762e+01  1.68719534e-11 -9.22023810e+00  1.54900000e+03]
 [-2.10000000e+01 -7.49692470e-11 -1.05000000e+01  1.76400000e+03]
 [-1.19047619e-02 -7.28219745e-14 -5.95238095e-03  1.00000000e+00]]

```

**Removing Radial Distortion:**



To remove radial distortion from an image, first, the image and the world points are obtained then, the image is converted to grayscale. `cv2.calibrateCamera()` is applied on camera calibration matrix `K` which we already have.

`cv2.getOptimalNewCameraMatrix()` is applied to obtain our new camera matrix, then using the new camera matrix `cv2.undistort()` function is applied.

#### **New Camera Matrix**

```
[[3.73681571e-04 0.00000000e+00 7.25075131e-39]
 [0.00000000e+00 2.90346186e-04 7.66386162e-39]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
```

## **Question 2 : Zhangs Method**

Zhang model is a camera calibration method that uses traditional calibration techniques (known calibration points) and self-calibration techniques (correspondence between the calibration points when they are in different positions).

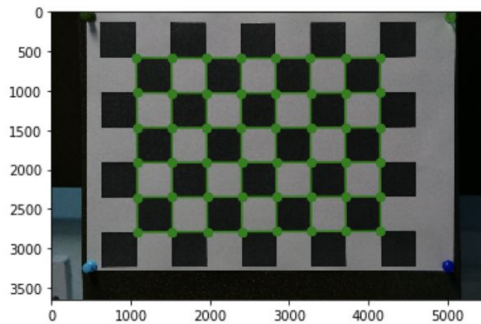
---

```

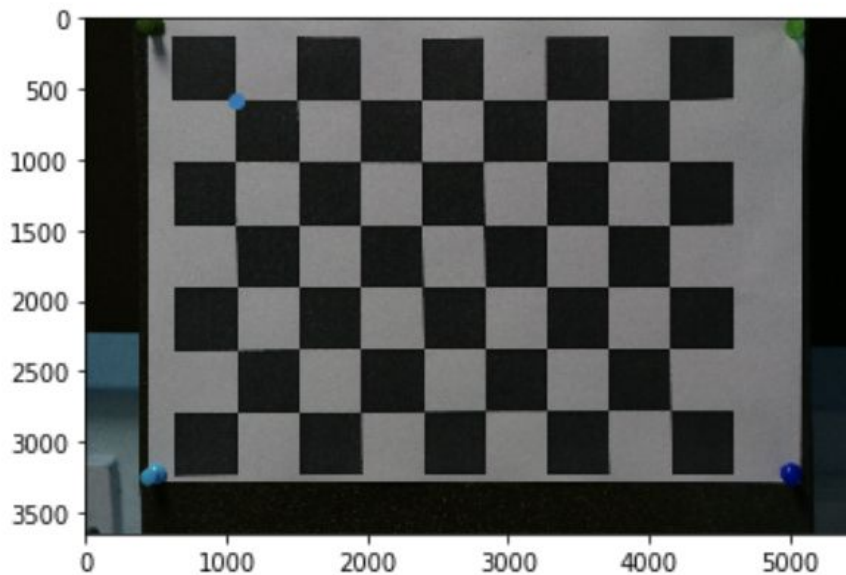
Reprojection Error: 2.48108498382283
Camera Calibration Matrix:
[[1.36634771e+04  0.00000000e+00  3.33653576e+03]
 [0.00000000e+00  1.36813826e+04  1.49660066e+03]
 [0.00000000e+00  0.00000000e+00  1.00000000e+00]]
Distortion Parameters:
[[ 9.52129447e-02  1.01092784e+01 -1.52798046e-02  2.87213059e-02
 -1.60590017e+02]]
Rotation Vectors for the images are:
[array([[ -0.01318089],
        [-0.05189252],
        [-0.00310183]]), array([[ -0.00723373],
        [ 0.22652171],
        [ 0.03532858]]), array([[ -0.00906864],
        [-0.41619657],
        [-0.03572075]]), array([[ 0.01633958],
        [-0.61839253],
        [-0.07755517]]), array([[ 0.00456622],
        [ 0.30778504],
        [ 0.06543949]]), array([[ -0.32566618],
        [-0.22232068],
        [-0.01742856]]), array([[ -0.19221716],
        [-0.38502421],
        [-0.00381478]]), array([[ 0.24427917],
        [-0.32262058],
        [ 0.02472158]]), array([[ 0.31114495],
        [ 0.06632946],
        [ 0.07053941]]), array([[ 0.34957369],
        [-0.56388409],
        [-0.06027059]]), array([[ 0.22457423],
        [-0.4058921 ],
        [-0.05699645]]), array([[ -0.11332321],
        [-0.29806932],
        [ 0.00488493]]), array([[ -0.38084113],
        [ 0.00764409],
        [ 0.06908158]]), array([[ -0.37472662],
        [-0.48684463],
        [ 0.02851414]]), array([[ -0.4414999 ],
        [-0.42570037],
        [ 0.05239045]])]
Translation Vectors for the images are:
[array([[ -5.14439949],
        [-2.05121046],
        [30.87688891]]), array([[ -5.07790437],
        [-1.96194072],
        [32.36119257]]), array([[ -4.53131783],
        [-1.76826385],
        [29.02891504]]), array([[ -4.15553694],
        [-1.7355166 ],
        [28.36758317]]), array([[ -4.81212655],
        [-2.124083 ],
        [32.69759787]]), array([[ -5.1541922 ],
        [-1.85743674],
        [35.99342414]]), array([[ -3.59098663],
        [-2.03074363],
        [31.96546331]]), array([[ -5.65039259],
        [-1.208156 ],
        [35.01639953]]), array([[ -5.31374423],
        [-1.68874033],
        [30.52753119]]), array([[ -2.1672435 ],
        [-1.31735498],
        [29.36434775]]), array([[ -4.69373888],
        [-0.88967538],
        [38.73627633]]), array([[ -4.18738466],
        [-1.93094275],
        [38.61303064]]), array([[ -4.25996064],
        [-2.25825105],
        [35.09459955]]), array([[ -5.17943613],
        [-2.03799116],
        [34.74922656]]), array([[ -4.5053191 ],
        [-2.2730252 ],
        [33.34095519]])]

```

The 3D and 2D points are found then using `cv2.calibrateCamera()`, camera calibration is performed using Zhang's method. After that, using `cv2.projectPoints()`, the points are projected onto the chessboard and the wireframe is found out.



To find the projection of the world origin, again `cv2.projectPoints()` is used for the point  $[0,0,0]$ . The results are as follows:



## Question 3 : Hands On

By using DLT method on an image of a cube, we get the following results:

```

Projection Matrix
[[-1.28659556e+02  1.78069145e+02  6.83621294e+01  5.27150931e+02]
 [-4.78280484e+01 -4.24957389e+01  1.74666956e+02  5.41653424e+02]
 [ 5.75537096e-02  9.57109214e-02  1.49687309e-01  1.00000000e+00]]
projection_center
[-1.8592989  -2.66853575 -4.2594288 ]
Camera Calibration matrix
[[ 9.15616187e-04  1.85096461e-05 -5.31896099e-01]
 [ 0.00000000e+00  1.20819201e-03 -6.69418257e-01]
 [ 0.00000000e+00  0.00000000e+00  1.00000000e+00]]
Rotation matrix
[[-0.79942514  0.59620979 -0.07384665]
 [-0.51570446 -0.6179784  0.59342363]
 [ 0.30816935  0.51248082  0.80149552]]
1046.145858161171

```



By Zhang's method, we try to capture images at two different focus, the following results are obtained:



For first focus :

```

Reprojection Error: 7.0221769747195255
Camera Calibration Matrix:
[[3.96427211e+03 0.00000000e+00 7.12999598e+02]
 [0.00000000e+00 3.89383471e+03 9.70953774e+01]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
Distortion Parameters:
[[ 0.33522811 -1.54345096 -0.05022154 -0.04568572  2.50388462]]
Rotation Vectors for the images are:
[array([[ -0.78744631,
          -0.31037701,
          [ 2.21208199]], array([[ -1.14527378,
          -0.49620465,
          [ 0.62567945]], array([[ -0.82378429,
          -0.73108059,
          [ 1.41230556]], array([[ -0.81750913,
          -0.65045304,
          [ 1.36072661]], array([[ -0.45665153,
          -0.14677019,
          [ 1.30717843]], array([[ -1.03139328,
          -0.41638315,
          [ 1.65612428]], array([[ -1.08928162,
          -0.32416477,
          [ 0.50040699]], array([[ -1.03275491,
          -0.61029319,
          [ 1.91567675]])]
Translation Vectors for the images are:
[array([[ 5.8533659 ],
          [ 6.7436951 ],
          [21.24956541]], array([[ 5.13099262],
          [ 1.89333972],
          [19.5915101 ]], array([[ 6.57908634],
          [ 0.55207989],
          [19.87378775]], array([[ 5.48734152],
          [ 0.25342952],
          [23.67719903]], array([[ 6.35752296],
          [ 0.76885188],
          [18.88660968]], array([[ 5.77594371],
          [ 3.7024151 ],
          [20.6789661 ]], array([[ 2.2761313 ],
          [ 0.2638275 ],
          [26.29489842]], array([[ 3.83354368],
          [ 5.5279065 ],
          [22.07609674]])]

```

For second focus :

```

Reprojection Error: 6.2155604983834145
Camera Calibration Matrix:
[[3.97513079e+03 0.00000000e+00 1.49470320e+03]
 [0.00000000e+00 4.04235256e+03 4.74789680e+02]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]
Distortion Parameters:
[[ -2.94077203e-01 3.13011031e+00 2.32278553e-03 1.36001178e-03
   -1.23929323e+01]]
Rotation Vectors for the images are:
[array([[ -0.70734786,
          -0.43523084,
          [ 1.40234406]], array([[ -0.89613991,
          -0.5366577 ],
          [ 2.14124214]], array([[ -0.93625293,
          -0.68864613,
          [ 1.34054301]], array([[ -1.07932552,
          -0.35150101,
          [ 0.55792969]], array([[ -0.88770888,
          -0.47412262,
          [ 1.28113684]], array([[ -0.9409666 ],
          -0.67762316,
          [ 1.31052297]], array([[ -1.06456634,
          -0.37416522,
          [ 0.81439179]], array([[ -1.23182631,
          -0.11702556,
          [ 1.24204716]], array([[ -0.76156957,
          -0.70623079,
          [ 2.04792656]], array([[ -1.0188159 ],
          -0.27243871,
          [ 1.55460194]], array([[ -1.05381712,
          -0.18236114,
          [ 0.40436896]])]
Translation Vectors for the images are:
[array([[ 3.1845352 ],
          [-1.27928859],
          [19.73580849]], array([[ 1.68790632],
          [ 5.06753337],
          [23.38892753]], array([[ 3.42138169],
          [-1.10403519],
          [19.67172744]], array([[ 2.09652154],
          [-0.66288244],
          [20.21051034]], array([[ 0.74025448],
          [-1.74710493],
          [24.61711409]], array([[ 3.95118092],
          [-0.60377466],
          [18.24683134]], array([[ 2.34705565],
          [-1.3025161 ],
          [19.51505489]], array([[ 2.25663493],
          -0.94239359,
          [19.39510989]], array([[ 3.53642984],
          -2.92131882,
          [19.65359554]], array([[ 2.50652667],
          -0.96159022,
          [21.73906155]], array([[ -2.82407174],
          [-2.08359513],
          [26.888168 ]]])]

```



**Link to all the input:**

<https://drive.google.com/open?id=19G3kZAjc7g1ckL1cFCyLZf6itTeE1ivz>