Photogrammetric Computer Vision
Assignment 3
Group 4

1. Image acquisition

- The calibration object is a box on which landmarks points drawn by hand, as shown in the above image. To locate these landmarks, I used a ruler to outline a cube of size of 8 centimes, and then subdivided it one more time into even blocks. Theoretically, we needs at least 5 and ½ point correspondences to estimate a camera projection matrix, but there's often noise in measured points, so more points will be better. That's why I tried to create 19 landmark points on the calibration object.

- The camera is the rear camera from Ipad Mini 4, 8MP, ƒ/2.4 aperture.

2. Control Point measurements.

- The object coordinate system is shown in the above picture, where the z axis points down, the x axis points to the right side and the y axis points the left side. The origin is the point closest to the camera.

- The measured locations are quite not exact because the used box is a bit distorted. I tried to find a better object but there was no better one at my place.

3. Computation of the projection matrix.
- please check the function homo_3d_2d in my attached file

4. Interpretation of the projection matrix.

Here is the projection matrix returned by the function homo_3d_2d

```
P =

  389.3300 -220.0195 -116.1411  625.6135
  -96.3653  -98.4565 -454.9744  969.0289
    0.0717    0.0650   -0.0967    0.5398
```

As discussed in the book “Multiple View Geometry in Computer Vision”, this projection matrix has the following form.

$$\mathrm{P} = [\mathrm{M} \mid -\mathrm{M}\tilde{\mathrm{C}}] = \mathrm{K}[\mathrm{R} \mid -\mathrm{R}\tilde{\mathrm{C}}].$$

Where R and C, respectively are the rotation matrix and translation part which represents the orientation and location of object in the camera space. The 3x3 matrix K represents the internal parameters of camera.

$$\mathrm{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}.$$

+ Alpha_x = f*m_x where f is the focal length of the camera. m_x is the number of pixels per one distance unit of the camera in the x direction.
+ Alpha_y = f*m_y where m_y is the number of pixel per one distance unit of the camera in the y direction.
+ S is the skew factor
+ x_0 and y_0 is the principal point of the camera, which is the intersection between the principal axis and the projection plane of the camera.

The RQ decomposition is used to decompose the 3x3 part of the projection matrix into the matrices K and R as shown in below.

```
K =

  424.6857   -8.9719 -181.7602
         0  419.1966 -224.1781
         0         0   -0.1368


R =

    0.6815   -0.7317    0.0140
   -0.5104   -0.4889   -0.7075
   -0.5245   -0.4750    0.7066
```

- It could be seen from the matrix K that alpha_x is almost equal to alpha_y, which could be interpreted that our camera is not CCD, which has different number of pixel per distance unit in x and y dimension.
- The skew parameter and principal coordinates are not equal to zero, which I haven’t found out why.