Assignment 4

(Using 1 late day)

Question 1

1.1

Given two points x_1 and x_2 from camera 1 and camera 2, we have the following relation with the fundamental matrix defining the pair:

$$x_2^T F x_1 = 0$$

Since both the coordinate axis is normalized, the two points, and in turn the relationship is as follows:

$$x_2^T F x_1 = 0$$
$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T F \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

From solving the above, we get $F_{3,3} = 0$

1.2

We know that for 2 views,

$$x_2^T F x_1 = 0$$

Also, the epipolar lines are given by:

$$l_1 = E^T x_2$$
$$l_2 = E x_1$$

where the essential matrix is given by:

$$E = t_\times R$$

If the two cameras only differ by a translation parallel to the x-axis, the rotation matrix is the identity matrix and the essential matrix is defined by the skew-symmetric translation matrix which only has translations in the x direction.

$$l_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix}^T \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$$
$$l_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

$$l_1 = -t_x y_2 + t_x$$
$$l_2 = t_x y_1 - t_x$$

Therefore, from the equation, both epipolar lines are parallel to the x-axis.

1.3

Given the world coordinates P and the 2-d points from the two frames x_1 and x_2, the following equations define the relationship between the 3-d and 2-d points using the camera matrix K and the corresponding rotation and translation matrices for x_1 and x_2:

$$x_1 = K(R_1 P + t_1)$$
$$x_2 = K(R_2 P + t_2)$$

Assuming that the first frame corresponds to x_1 and the second one to x_2 and we want to know the effective transformations from x_1 to x_2, we can eliminate the world coordinates from the above equations:

$$P = R_1^{-1}(K^{-1}x_1 - t_1)$$

Substituting it back into the other equation:

$$\begin{aligned} x_2 &= K(R_2 R_1^{-1}(K^{-1}x_1 - t_1) + t_2) \\ &= KR_2R_1^{-1}K^{-1}x_1 - KR_2R_1^{-1}t_1 + Kt_2 \end{aligned}$$

The effective rotation and translation matrices are:

$$R_{rel} = KR_2R_1^{-1}K^{-1}$$
$$t_{rel} = Kt_2 - KR_2R_1^{-1}t_1$$

The essential and fundamental matrices in terms of the above defines quantities are as follows:

$$E = t_{rel\times} R_{rel}$$
$$F = K^{-T}EK{-}1 = K^{-T}t_{rel\times}R_{rel}K{-}1$$

1.4

We need to prove that a single camera viewing an object and its mirror reflection is equivalent to two images being related by a skew symmetric fundamental matrix.

$$F = K^{-T}EK^{-1}$$
$$F^T = K^{-T}E^TK^{-1}$$

We need to prove that $F = -F^T$. For this to be true, the essential matrix in the above term should be skew symmetric. In other words, we can instead prove $E = -E^T$. The expression for the essential matrix and its transpose is given by:

$$E = t_\times R$$
$$E^T = R^T t_\times^T$$

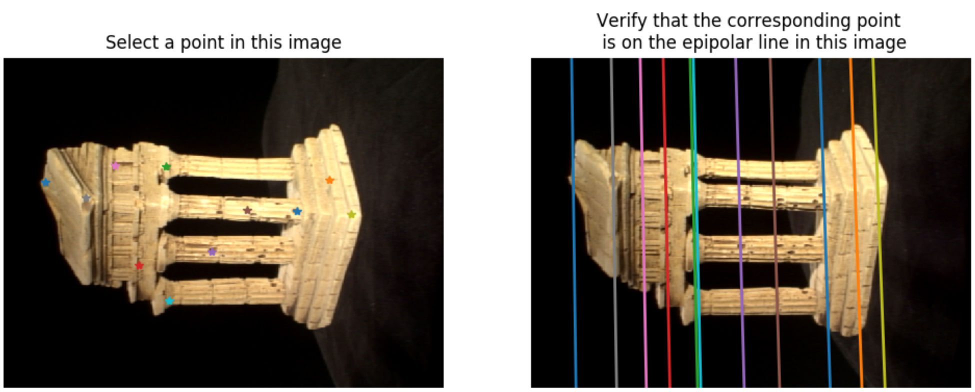
Now, taking a suitable image plane for the camera, we can see that for reflection of an object, there's no 2-D rotation involved. Hence the matrix R can be reduced to an identity matrix. The remaining matrix $t_\times$ is already a skew-symmetric matrix. Hence the E matrix skew symmetric and in turn F is skew-symmetric.

Question 2

2.1

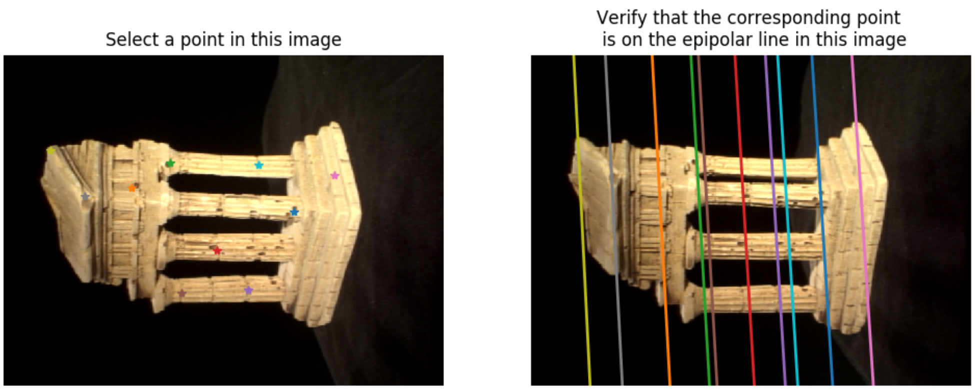
$$F = \begin{bmatrix} 9.80213865e\text{-}10 & -1.32271663e\text{-}07 & 1.12586847e\text{-}03 \\ -5.72416248e\text{-}08 & 2.97011941e\text{-}09 & -1.17899320e\text{-}05 \\ -1.08270296e\text{-}03 & 3.05098538e\text{-}05 & -4.46974798e\text{-}03 \end{bmatrix}$$

Results from the 8-point algorithm shown below



2.2

Results from the 7-point algorithm with manually chosen points shown below



Question 3

3.1

$$E = \begin{bmatrix} 2.26587821e\text{-}03 & -3.06867395e\text{-}01 & 1.66257398e{+}00 \\ -1.32799331e\text{-}01 & 6.91553934e\text{-}03 & -4.32775554e\text{-}02 \\ -1.66717617e{+}00 & -1.33444257e\text{-}02 & -6.72047195e\text{-}04 \end{bmatrix}$$

3.2

$$A = \begin{bmatrix} pts1_x \times C1_z^T - C1_x^T \\ pts1_y \times C1_z^T - C1_x^T \\ pts2_x \times C2_z^T - C2_x^T \\ pts2_y \times C2_z^T - C2_x^T \end{bmatrix}$$

where the two points are given by

$$pts1 = \begin{bmatrix} pts1_x \\ pts1_y \end{bmatrix}$$

$$pts2 = \begin{bmatrix} pts2_x \\ pts2_y \end{bmatrix}$$

and where the two camera matrices $C1, C2$ of shape 3×4 are given by

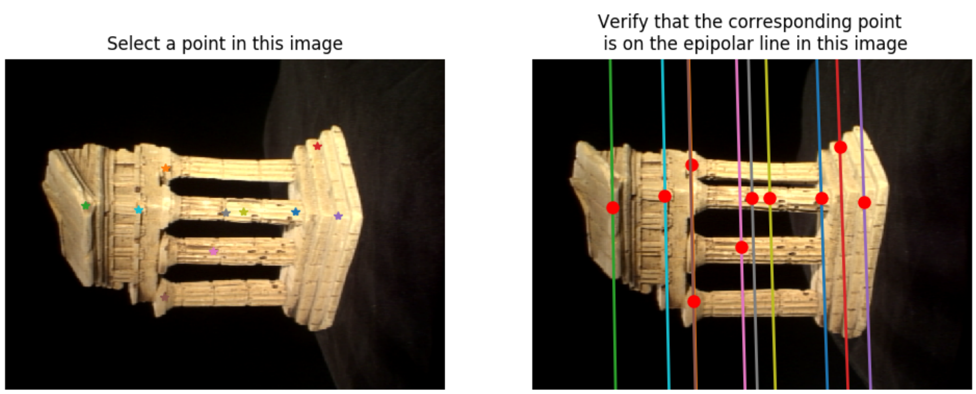
$$C1 = \begin{bmatrix} C1_x \\ C1_y \\ C1_z \end{bmatrix}$$

$$C2 = \begin{bmatrix} C2_x \\ C2_y \\ C2_z \end{bmatrix}$$

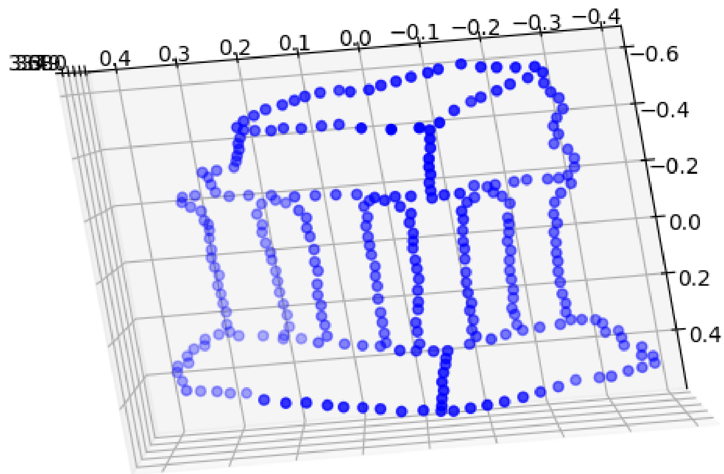
Question 4

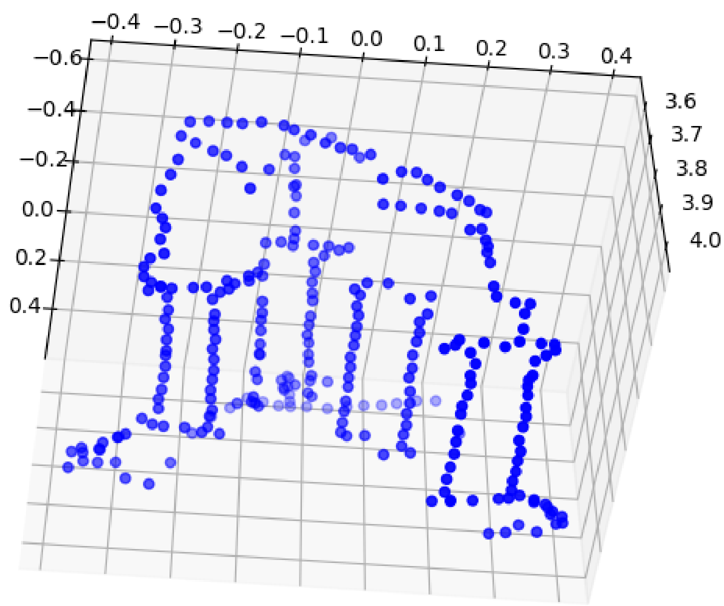
4.1

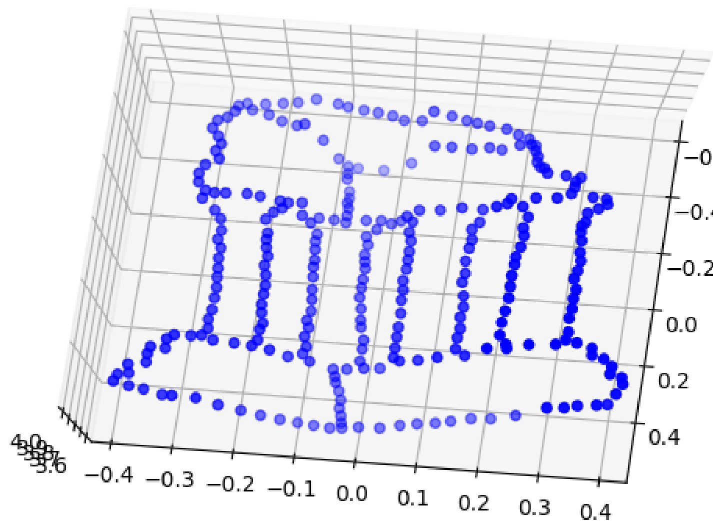
Results from the epipolar correspondence search matching points in image1 to lines in image 2

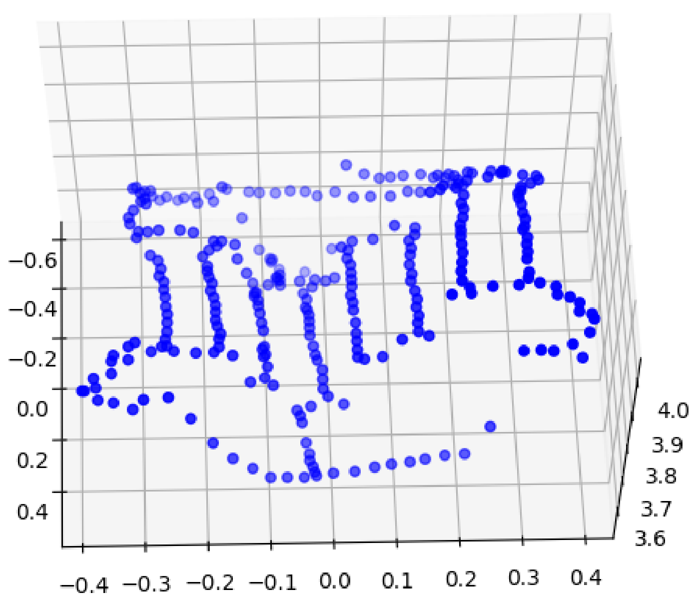


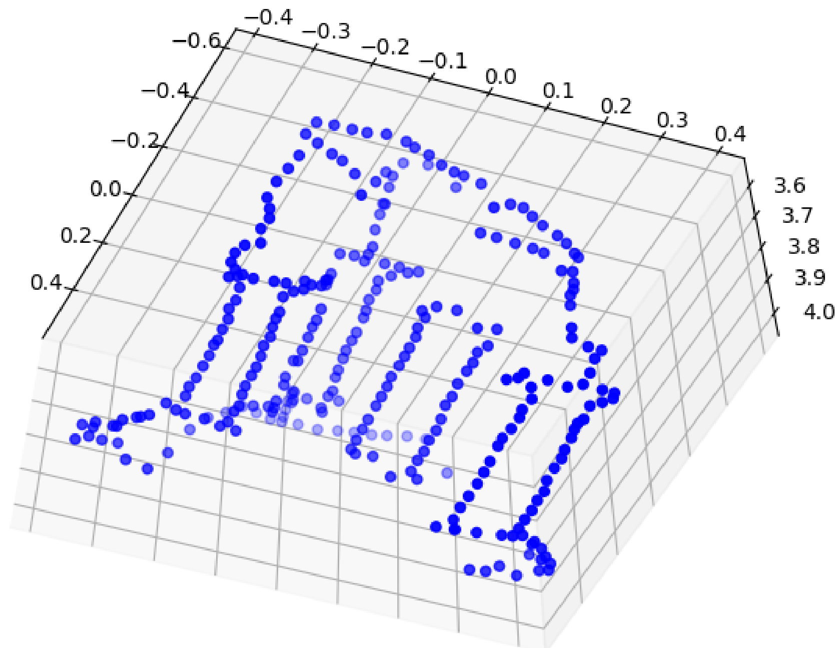
4.2

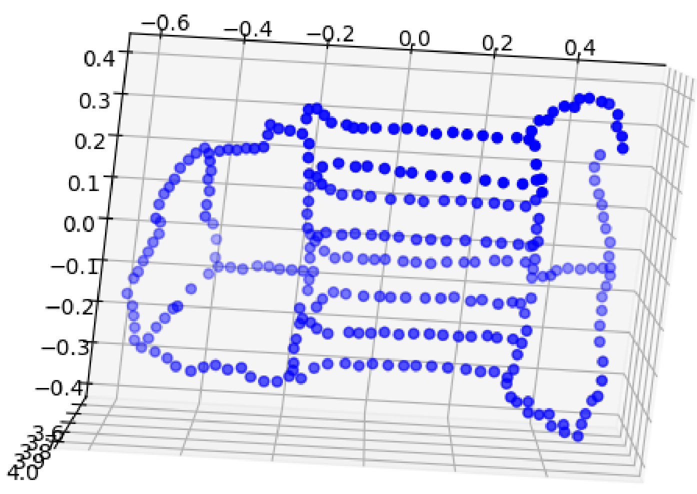

0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.6
−0.4
−0.2
0.0
0.2
0.4


−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
−0.6
−0.4
−0.2
0.0
0.2
0.4
3.6
3.7
3.8
3.9
4.0


−0.6
−0.4
−0.2
0.0
0.2
0.4
−0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4


−0.6
−0.4
−0.2
0.0
0.2
0.4
−0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4
4.0
3.9
3.8
3.7
3.6

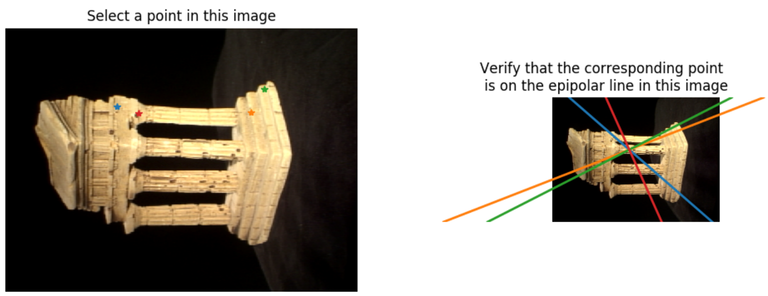

−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
−0.6
−0.4
−0.2
0.0
0.2
0.4
3.6
3.7
3.8
3.9
4.0


−0.6
−0.4
−0.2
0.0
0.2
0.4
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
4.0

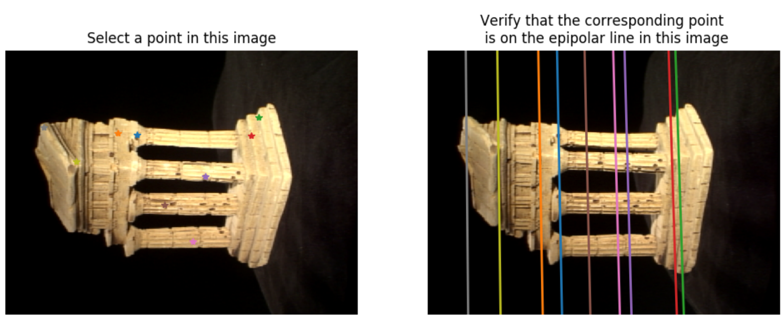
Question 5

5.1

The following are results from directly running the 8-point algorithm on the noisy correspondences.



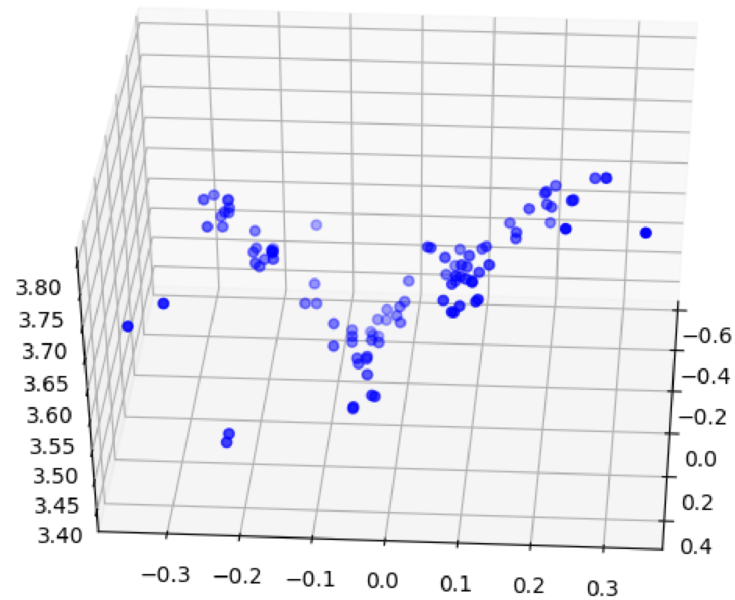
The following are results from running the 7-point algorithm with RANSAC on the noisy correspondences.

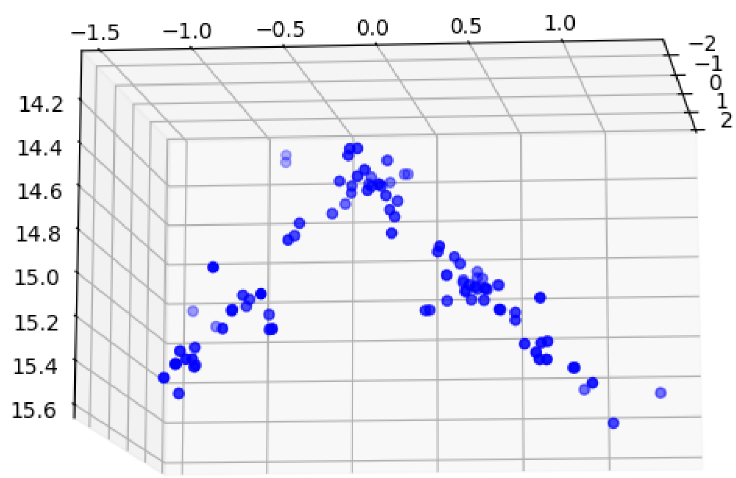


The error metric used for RANSAC was the distance from the epipolar constraint $x_2 F x_1 = 0$. 200 iterations of RANSAC was run with a threshold for the epipolar constraint being set to 10^{-3} for each point to decide if it's an inlier or not.

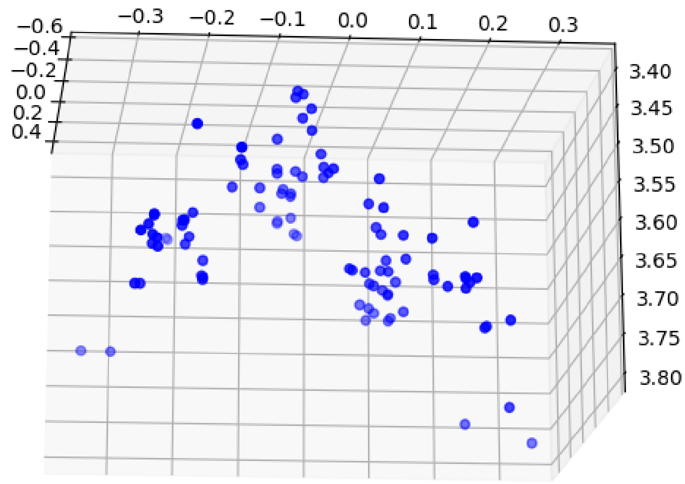
5.3

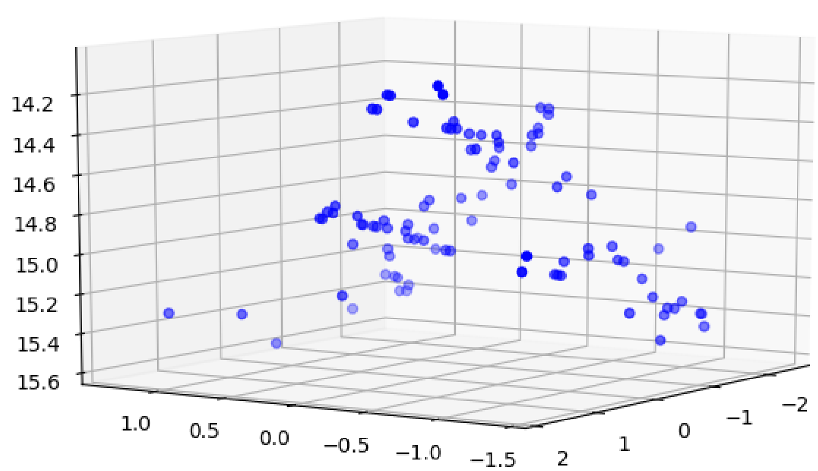
The following images show the inlier points from the arch of the temple's structure before and after the least squares non-linear optimization for M2 and P matrices.





Two more 3-D points of the same execution for before and after optimization are given below





Please run the code 'run_ba.py' to see the inlier points before and after bundle adjustment. The reprojection error before and after the optimization is also printed. The results and errors vary due to the randomness in RANSAC, but the initial error values are in the range 5k-500k and the final error values are usually in the range 10-50.