Q1.1 Show that if the image coordinates are normalized so that the coordinate origin (0,0) coincides with the principal point, the F_{33} element of the fundamental matrix is zero.

Solution:

According to the relation of fundament matrix, we have:

$$\tilde{x}_2^T F \tilde{x}_1 = 0$$

Note that $\tilde{x}_2^T$ *and* $\tilde{x}_1$ are the coordinates of two cameras. Since we have the condition that the both of coordinate origins of $\tilde{x}_1$ and $\tilde{x}_2$ are (0, 0), we now have:

$$\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \rightarrow F_{33} = 0$$

Q1.2 Consider the case of two cameras viewing an object such that the second camera differs from the first by a pure translation that is parallel to the x-axis. Show that the epipolar lines in the two cameras are also parallel to the x-axis. Backup your argument with relevant equations.

Solution:

According to the relation among essential matrix E, rotation matrix R and translation matrix t, we have:

$$E = t_{\times}R = \begin{pmatrix} 0 & -t_z & -t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} R$$

Since we know that the pure translation of these two cameras is parallel to the x-axis and there is no rotation, we can get our essential matrix E:

$$E = t_{\times|t_y=t_z=0}R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix}$$

Based on the calculated E, we can get epipolar lines of two camera views:

$$l_1^T = \tilde{x}_2^T E = (x_2 \quad y_2 \quad 1) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix} = (0 \quad t_x \quad -y_2 t_x)$$

$$l_2^T = \tilde{x}_1^T E^T = (x_1 \quad y_1 \quad 1) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{pmatrix} = (0 \quad -t_x \quad y_1 t_x)$$

Based on the previous result, we can get the line equations:

$$Camera\ view\ \#1:\ (0 \quad t_x \quad -y_2 t_x) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \rightarrow t_x y - y_2 t_x = 0$$

$$Camera\ view\ \#2:\ (0 \quad -t_x \quad y_1 t_x) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \rightarrow -t_x y + y_1 t_x = 0$$

Hence, both epipolar lines of two camera views are parallel to x-axis.

Q1.3 Suppose we have an inertial sensor which gives us the accurate positions (R_i and t_i, the rotation matrix and translation vector) of the robot at time i. What will be the effective rotation (R_{rel}) and translation (t_{rel}) between two frames at different time stamps? Suppose the camera intrinsics (K) are known, express the essential matrix (E) and the fundamental matrix (F) in terms of K, R_{rel} and t_{rel}.

Solution:

Assume that the point with respect to the real-world is $w = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$. Then, we can get two points w_1 and w_2 which are points with respect to the position of the camera on the robot at timestamp 1 and 2, respectively.

$$w_1 = R_1 w + t_1$$

$$w_2 = R_2 w + t_2$$

From the first equation, we can express w as the following:

$$w = R_1^{-1}(w_1 - t_1)$$

Therefore, we can re-write w_2 as:

$$w_2 = R_2 w + t_2 = R_2 R_1^{-1}(w_1 - t_1) + t_2 = R_2 R_1^{-1} w_1 + (t_2 - R_2 R_1^{-1} t_1)$$

From the previous equation, we can get R_{rel} and t_{rel} (convert time stamp1 to time stamp2):

$$R_{rel} = R_2 R_1^{-1}$$

$$t_{rel} = t_2 - R_2 R_1^{-1} t_1$$

Now, we can express the essential matrix (E) and fundamental matrix (F) with R_{rel} and t_{rel} :

$$E = t_{rel} \times R_{rel}$$

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

Q1.4 Suppose that a camera views an object and its reflection in a plane mirror. Show that this situation is equivalent to having two images of the object which are related by a skew-symmetric fundamental matrix. You may assume that the object is flat, meaning that all points on the object are of equal distance to the mirror.

Solution:

First, assume that the mirror is the plane of unit normal vector v. Therefore, the points in the real-world w and the reflection w' can be expressed using the household reflection transformation (link), that is:

$$w' = (I - vv^T)w$$

Also, the corresponding points on 2D planes $\tilde{x}_1$ *and* $\tilde{x}_2$ of points (w' and w) can be expressed as:

$$\lambda\tilde{x}_1 = w$$

$$\lambda'\tilde{x}_2 = w'$$

Applying the household transformation and previous equations, we have:

$$\lambda'\tilde{x}_2 = w' = (I - vv^T)w = \lambda(I - vv^T)\,\tilde{x}_1$$

If we take the cross product of v, we have (the cross product of itself equals to 0):

$$\lambda' v_\times \tilde{x}_2 = \lambda v_\times (I - vv^T)\,\tilde{x}_1 = \lambda v_\times \tilde{x}_1$$

If we take the left dot $\tilde{x}_2{}^T$ on both sides, we can get:

$$\tilde{x}_2{}^T \lambda v_\times \tilde{x}_1 = 0$$

Let F = $\lambda v_\times$ and we can get the fact that F is a skew-symmetric fundamental matrix.

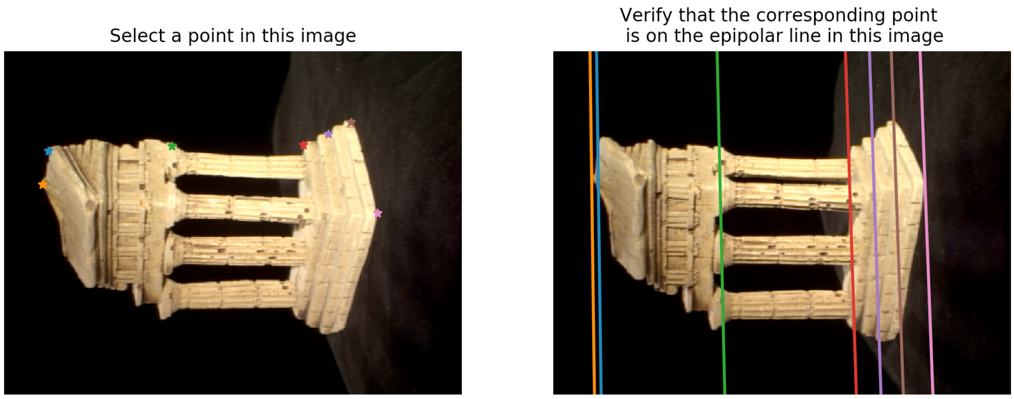
Q2.1 Write your recovered F and include an image of some example output of displayEpipolarF.

Solution:

The recovered F using eight-point algorithm is shown as following:

$$F = \begin{pmatrix} 9.78833287e-10 & -1.32135929e-07 & 1.12585666e-03 \\ -5.73843315e-08 & 2.96800276e-09 & -1.17611996e-05 \\ -1.08269003e-03 & 3.04846703e-05 & -4.47032655e-03 \end{pmatrix}$$

Besides, the following figures are some example outputs using displayEpipolarF:



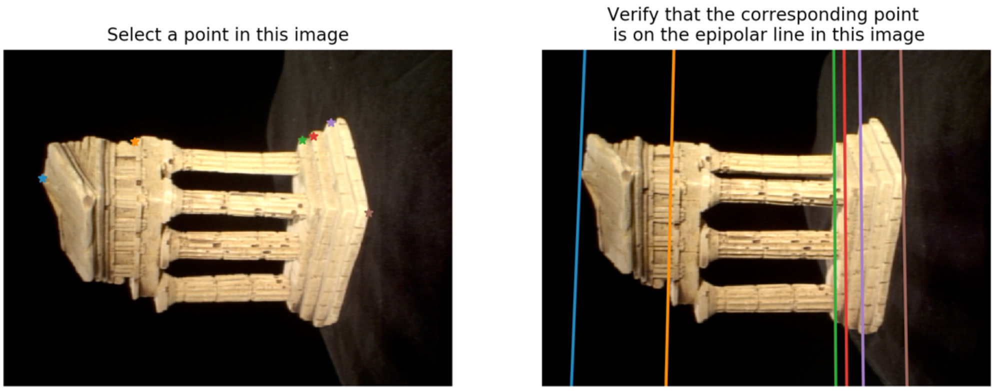
Q2.2 Write your recovered F and print an output of displayEpipolarF. Also, include an image of some example output of displayEpipolarF using the seven-point algorithm.

Solution:

The recovered F using seven-point algorithm is shown as following:

$$F = \begin{pmatrix} 4.13913858e-08 & -8.14615550e-08 & 7.89199802e-04 \\ -9.77328767e-08 & -5.40466077e-09 & 3.73569488e-05 \\ -7.67523686e-04 & -1.15930010e-05 & -3.17234766e-03 \end{pmatrix}$$

Besides, the following figures are some example outputs using displayEpipolarF:



Q3.1 Write your estimated E using F from the eight-point algorithm.

Solution:

The recovered E using F from the eight-point algorithm is shown as below:

$$E = \begin{pmatrix} 2.26268684e-03 & -3.06552495e-01 & 1.66260633e+00 \\ -1.33130407e-01 & 6.91061098e-03 & -4.33003420e-02 \\ -1.66721070e+00 & -1.33210351e-02 & -6.72186431e-04 \end{pmatrix}$$

Q3.2 Write down the expression for the matrix A_i

Solution:

First, we make the following notation to express $C_1\ and\ C_2$:

$$C_1 = \begin{pmatrix} c_{00}^1 & c_{01}^1 & c_{02}^1 & c_{03}^1 \\ c_{10}^1 & c_{11}^1 & c_{12}^1 & c_{13}^1 \\ c_{20}^1 & c_{21}^1 & c_{22}^1 & c_{23}^1 \end{pmatrix} C_2 = \begin{pmatrix} c_{00}^2 & c_{01}^2 & c_{02}^2 & c_{03}^2 \\ c_{10}^2 & c_{11}^2 & c_{12}^2 & c_{13}^2 \\ c_{20}^2 & c_{21}^2 & c_{22}^2 & c_{23}^2 \end{pmatrix}$$

And, the coordinates with homogeneous term:

$$\widetilde{w}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} x_{1i} = \begin{pmatrix} x_{1i} \\ y_{1i} \\ 1 \end{pmatrix} x_{2i} = \begin{pmatrix} x_{2i} \\ y_{2i} \\ 1 \end{pmatrix}$$

Then, we could project 3D points back to 2D images and form the following equation:

$$C_1\widetilde{w}_i = \lambda_{1i}x_{1i} \rightarrow \begin{pmatrix} c_{00}^1 & c_{01}^1 & c_{02}^1 & c_{03}^1 \\ c_{10}^1 & c_{11}^1 & c_{12}^1 & c_{13}^1 \\ c_{20}^1 & c_{21}^1 & c_{22}^1 & c_{23}^1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \lambda_{1i} \begin{pmatrix} x_{1i} \\ y_{1i} \\ 1 \end{pmatrix}$$

$$C_2\widetilde{w}_i = \lambda_{2i}x_{2i} \rightarrow \begin{pmatrix} c_{00}^2 & c_{01}^2 & c_{02}^2 & c_{03}^2 \\ c_{10}^2 & c_{11}^2 & c_{12}^2 & c_{13}^2 \\ c_{20}^2 & c_{21}^2 & c_{22}^2 & c_{23}^2 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \lambda_{2i} \begin{pmatrix} x_{2i} \\ y_{2i} \\ 1 \end{pmatrix}$$

Further, we could get:

$$\begin{pmatrix} c_{00}^1 - c_{20}^1 x_{1i} & c_{01}^1 - c_{21}^1 x_{1i} & c_{02}^1 - c_{22}^1 x_{1i} & c_{03}^1 - c_{23}^1 x_{1i} \\ c_{10}^1 - c_{20}^1 y_{1i} & c_{11}^1 - c_{21}^1 y_{1i} & c_{12}^1 - c_{22}^1 y_{1i} & c_{13}^1 - c_{23}^1 y_{1i} \\ c_{00}^2 - c_{20}^2 x_{2i} & c_{01}^2 - c_{21}^2 x_{2i} & c_{02}^2 - c_{22}^2 x_{2i} & c_{03}^2 - c_{23}^2 x_{2i} \\ c_{10}^2 - c_{20}^2 y_{2i} & c_{11}^2 - c_{21}^2 y_{2i} & c_{12}^2 - c_{22}^2 y_{2i} & c_{13}^2 - c_{23}^2 y_{2i} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = 0$$

Finally, the expression for matrix A_i is:

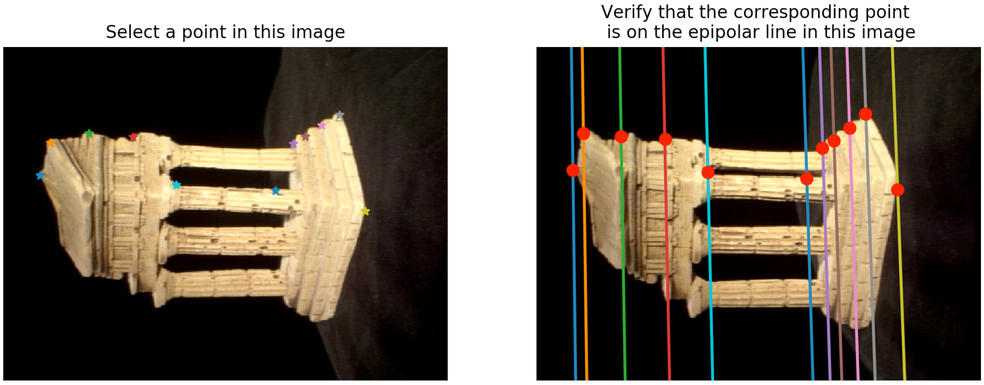
$$A_i = \begin{pmatrix} c_{00}^1 - c_{20}^1 x_{1i} & c_{01}^1 - c_{21}^1 x_{1i} & c_{02}^1 - c_{22}^1 x_{1i} & c_{03}^1 - c_{23}^1 x_{1i} \\ c_{10}^1 - c_{20}^1 y_{1i} & c_{11}^1 - c_{21}^1 y_{1i} & c_{12}^1 - c_{22}^1 y_{1i} & c_{13}^1 - c_{23}^1 y_{1i} \\ c_{00}^2 - c_{20}^2 x_{2i} & c_{01}^2 - c_{21}^2 x_{2i} & c_{02}^2 - c_{22}^2 x_{2i} & c_{03}^2 - c_{23}^2 x_{2i} \\ c_{10}^2 - c_{20}^2 y_{2i} & c_{11}^2 - c_{21}^2 y_{2i} & c_{12}^2 - c_{22}^2 y_{2i} & c_{13}^2 - c_{23}^2 y_{2i} \end{pmatrix}$$

And the projection error is 352.23.

Q4.1 Include a screenshot of epipolarMatchGUI with some detected correspondences.

Solution:

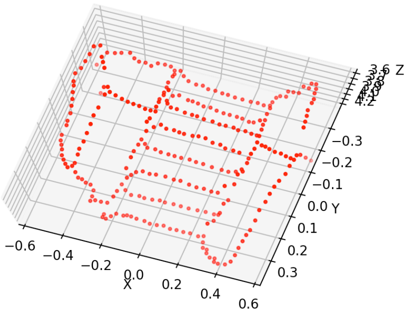
The following image is the screenshot:

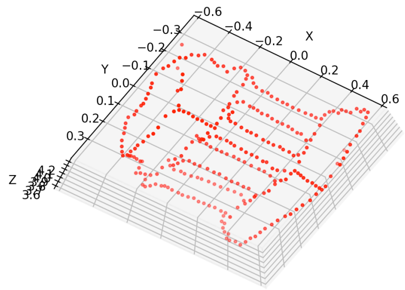


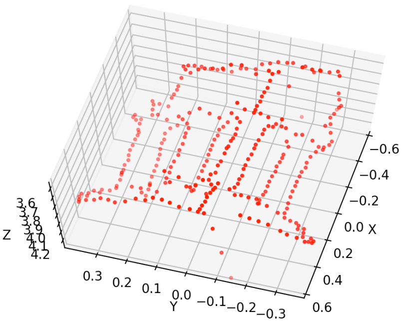
Q4.2 Take a few screenshots of the 3D visualization so that the outline of the temple is clearly visible, and include them with your homework submission.

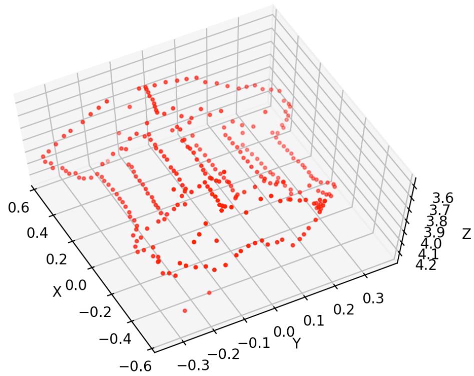
Solution:

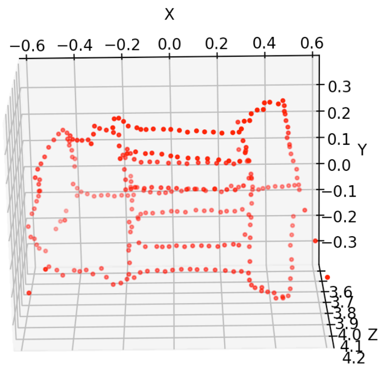
The following is the screenshots of the 3D visualization of the temple.

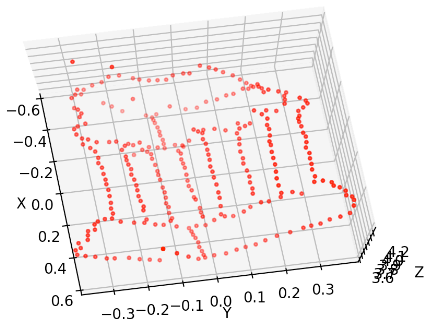








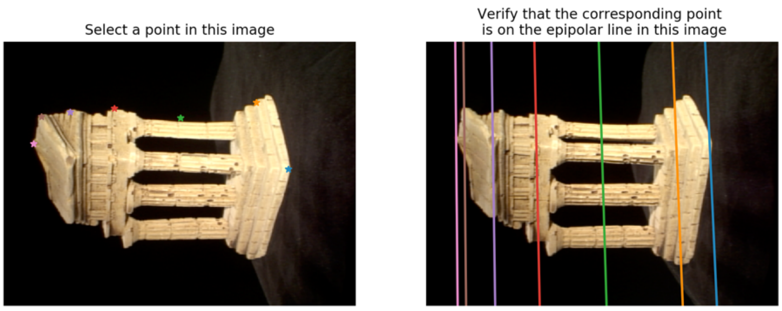




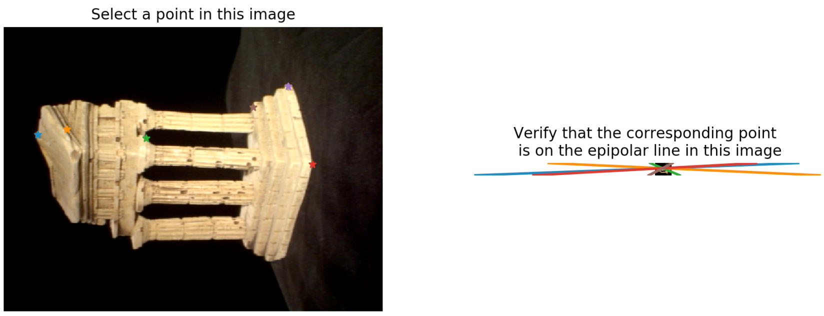
Q5.1 Compare the result of RANSAC with the result of the eightpoint when ran on the noisy coorespondances. Briefly explain the error metrics you used, how you decided which points were inliers and any other optimizations you may have made.

Solution:

The following two figures are the results of RANSAC and eightpoint, respectively.



The result of RANSAC



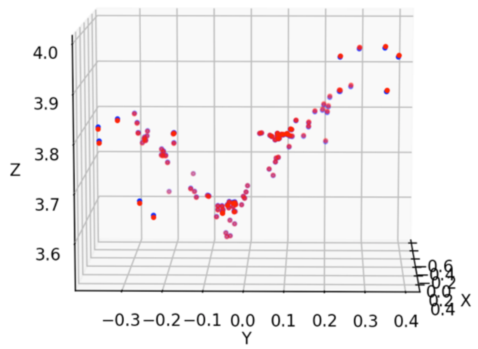
The result of eight-point algorithm

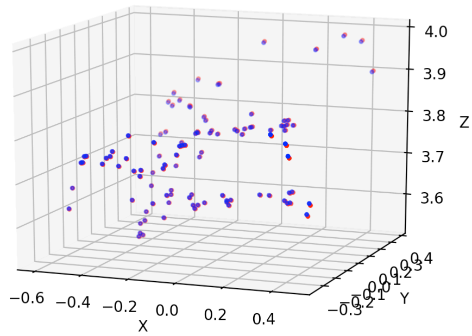
In Q5.1, I adopted the seven-point algorithm to apply RANSAC on the noisy correspondences in some_corresp_noisy.npz. For each iteration, randomly (but with the same seed for each iteration to make the output predictable) pick seven point to calculate the fundamental matrices. Then, generate the corresponding epipolar lines on both image1 and image2 using pts1, F and pts2. Finally, calculate the error defining by the distance between points to their corresponding epipolar lines. For each point pair pts1[i] and pts2[i], we can get two errors and sum them up. If the error is lower than a threshold, then the pair pts1[i] and pts2[i] is regarded as an in-lier. Note that I set this threshold equals to 2 in this assignment.

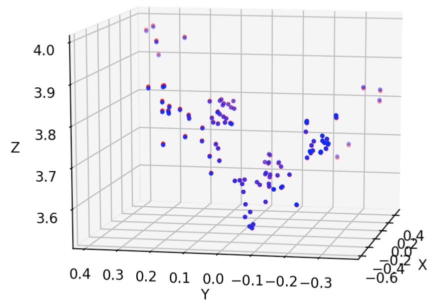
Q5.3 Include an image of the original 3D points and the optimized points as well as the reprojection error with your initial M_2 and w, and with the optimized matrices.

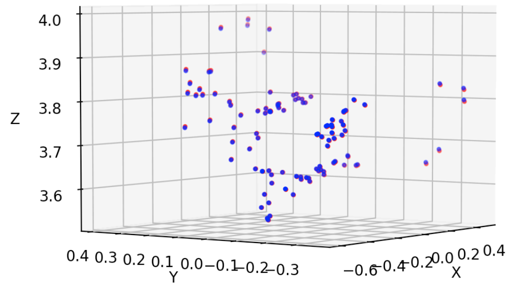
Solution:

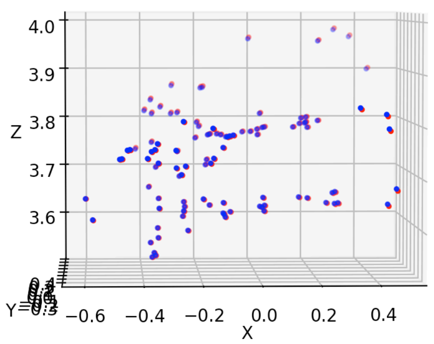
The following figures are the 3D points and the optimized points. Note that the initial points are in blue and the optimized ones are in red.

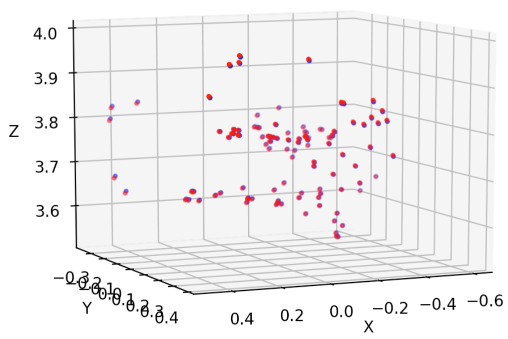












The projection error with the initial M_2 and w is 458.5 and the optimized error is 12.1. Note that the 3D points generated from initial and optimized M_2 are really close.