Homework 4

Thomas Zhang

November 2023

Part 1

Q1.1

$$\begin{bmatrix} x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

Since both the points are normalized the coordinates for the point P in both images is $(0, 0, 1)$ in homogeneous coordinates. So the above equation becomes:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

$$0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + f_{33} = 0$$

Hence f_{33} must be equal to 0.

Q1.2

$$E = \left[t_x\right] \times R$$

Since the relationship between the two cameras is purely translational. R is the identity matrix (no rotation).

$$E = \left[t_x\right]$$

The epipolar line is represented by the equation:

$$l'_r = Fp_l$$

where p_l is a point on the left image while l'_r is the epipolar line on the right image. The fundamental matrix, F, can be related to the essential matrix by the equation:

$$F = K_r^{-T} E K_l^{-1}$$

In this case since $E = \left[t_x\right]$ we can say $F = K_r^{-T}\left[t_x\right] K_l^{-1}$ where $\left[t_x\right]$ is the skew symmetric matrix with just the terms for x translation since we are assuming the two cameras are related via pure translation parallel to the x-axis.

This means the epipolar line can be represented as:

$$l'_r = K_r^{-T} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} K_l^{-1} p_l$$

$$l'_r = \begin{bmatrix} f_r^{-1} & 0 & 0 \\ 0 & f_r^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} f_l^{-1} & 0 & 0 \\ 0 & f_l^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ 1 \end{bmatrix}$$

$$l'_r = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x f_r^{-1} \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} f_l^{-1} & 0 & 0 \\ 0 & f_l^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ 1 \end{bmatrix}$$

$$l'_r = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x f_r^{-1} \\ 0 & t_x f_l^{-1} & 0 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ 1 \end{bmatrix}$$

$$l'_r = \begin{bmatrix} 0 \\ -t_x f_r^{-1} \\ y_l t_x f_l^{-1} \end{bmatrix}$$

This shows that the epipolar line is parallel to the x-axis since the x value of l'_r in homogenous coordinates is 0.

Q1.3

The projection of a point P onto an image can be represented by the equation:

$$
\begin{aligned}
p_{image} &= K \begin{bmatrix} R & t \end{bmatrix} P \\
K^{-1} p_{image} &= \begin{bmatrix} R & t \end{bmatrix} P \\
\hat{p}_{image} &= \begin{bmatrix} R & t \end{bmatrix} P \\
\hat{p}_{image} &= RP + t
\end{aligned}
$$

Suppose the two time points are denoted by i and j. Let p be the real world coordinates of a point. Let p_i and p_j be the image coordinates at the times i and j.

$$
\begin{aligned}
\hat{p}_i &= R_i p + t_i \\
\hat{p}_j &= R_j p + t_j
\end{aligned}
$$

$$p = R_i^{-1}(\hat{p}_i - t_i)$$

Plug expression for p into equation for p_j.

$$
\begin{aligned}
\hat{p}_j &= R_j R_i^{-1}(\hat{p}_i - t_i) + t_j \\
\hat{p}_j &= R_j R_i^{-1}\hat{p}_i - R_j R_i^{-1} t_i + t_j
\end{aligned}
$$

So:

$$
\begin{aligned}
R_{rel} &= R_j R_i^{-1} \\
t_{rel} &= -R_j R_i^{-1} t_i + t_j
\end{aligned}
$$

$$
\begin{aligned}
E &= t_{rel} \times R_{rel} \\
F &= K_i^{-T} t_{rel} \times R_{rel} K_j^{-1}
\end{aligned}
$$

Q1.4

Let X be a point on the object in 3D and X' be the reflection of X.

The position of the object and it's reflection in the image is:

$$X_l \equiv KMX$$
$$X'_l \equiv KMX'$$

Where $M = \begin{bmatrix} R_{cam} & t_{cam} \end{bmatrix}$ is the extrinsic matrix of the camera and K is the intrinsic matrix. Assuming the world coordinate system is at the camera, M can be simplified to $\begin{bmatrix} I & 0 \end{bmatrix}$. Hence the equations for X_l and X'_l become:

$$X_l \equiv KX$$
$$X'_l \equiv KX'$$

Since X' is a reflection of X, we can represent X' using the equation $X' = \begin{bmatrix} R & t \end{bmatrix} X$ where R represents some rotation matrix and t represents the translation. Using this substitution to update our equations we get:

$$X_l \equiv KX$$
$$X'_l \equiv K \begin{bmatrix} R & t \end{bmatrix} X$$

If we let $X'_l = X_r$, then we see that this is the same exact scenario as having two cameras with the same intrinsic matrix take a picture of X from two different positions where the first camera has an extrinsic matrix of $\begin{bmatrix} I & 0 \end{bmatrix}$ and the second camera has a extrinsic matrix of $\begin{bmatrix} R & t \end{bmatrix}$

Now we can normalize the coordinates using $\hat{X}_l = K^{-1}X_l$ and $\hat{X}'_l = K^{-1}X'_l$.

$$K^{-1}X_l \equiv K^{-1}KX$$
$$\hat{X}_l \equiv X \qquad (1)$$

$$K^{-1}X'_l \equiv K^{-1}K \begin{bmatrix} R & t \end{bmatrix} X$$
$$\hat{X}'_l \equiv \begin{bmatrix} R & t \end{bmatrix} X \qquad (2)$$

Now we can substitute equation (1) into equation (2).

$$\hat{X}'_l \equiv \begin{bmatrix} R & t \end{bmatrix} \hat{X}_l$$

Take the cross product of both sides with t to get:

$$t \times \hat{X}'_l \equiv t \times \begin{bmatrix} R & t \end{bmatrix} \hat{X}_l$$
$$t \times \hat{X}'_l \equiv t \times (R\hat{X}_l + t)$$
$$t \times \hat{X}'_l \equiv t \times R\hat{X}_l + t \times t$$
$$t \times \hat{X}'_l \equiv t \times R\hat{X}_l$$

Now multiply both sides by ${\hat{X}'_l}^T$.

$${\hat{X}'_l}^T (t \times \hat{X}'_l) \equiv {\hat{X}'_l}^T (t \times R\hat{X}_l)$$

Based on the property that $q(p \times q) = 0$ for any two 3 coordinate vectors p and q, we can say that ${\hat{X}'_l}^T (t \times \hat{X}'_l) = 0$ which means ${\hat{X}'_l}^T (t \times R\hat{X}_l) = 0$. Now we can convert the cross product to matrix multiplication using the skew-symmetric matrix to get ${\hat{X}'_l}^T [t_\times] R\hat{X}_l = 0$ where

$$[t_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

Now converting back to non-normalized coordinates:

$${\hat{X}'_l}^T [t_\times] R\hat{X}_l = 0$$
$$(K^{-1}X'_l)^T [t_\times] RK^{-1}X_l = 0$$
$$X_l^{'T} K^{-T} [t_\times] RK^{-1}X_l = 0$$

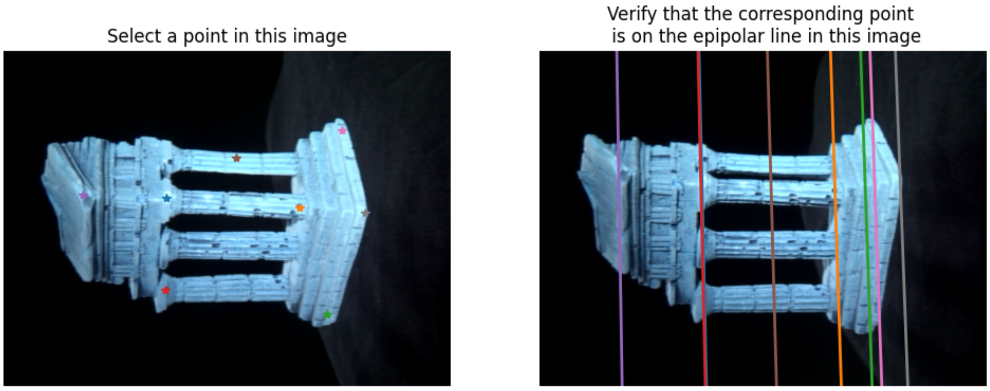
We can now let $F = K^{-T}[t_\times]RK^{-1}$ to get:

$$X_l'^T F X_l = 0$$

and since $[t_\times]$ is skew-symmetric, F will be as well. Since F is skew symmetric, this scenario can be thought of as two cameras taking a picture of X with fundamental matrix F relating the two cameras. Where the object at X appears at the coordinates X_l' in one of the images and X_l in the other image.

Part 2

2.1



$$F = \begin{bmatrix} 9.80213858e-10 & -1.32271663e-07 & 1.12586847e-03 \\ -5.72416248e-08 & 2.97011941e-09 & -1.17899320e-05 \\ -1.08270296e-03 & 3.05098538e-05 & -4.46974798e-03 \end{bmatrix}$$

Part 3

3.1

E matrix:

```
[[ 2.26587820e-03, -3.06867395e-01,  1.66257398e+00],
 [-1.32799331e-01,  6.91553934e-03, -4.32775554e-02],
 [-1.66717617e+00, -1.33444257e-02, -6.72047195e-04]]
```

3.2

Let the camera matrices C_1 and C_2 be denoted as:

$$C_1 = \begin{bmatrix} c^1_{1,1} & \cdots & c^1_{1,4} \\ \vdots & \ddots & \vdots \\ c^1_{3,1} & \cdots & c^1_{3,4} \end{bmatrix} \qquad C_2 = \begin{bmatrix} c^2_{1,1} & \cdots & c^2_{1,4} \\ \vdots & \ddots & \vdots \\ c^2_{3,1} & \cdots & c^2_{3,4} \end{bmatrix}$$

Let: $w_i = [x_i, y_i, z_i, 1]^T$. The points in images 1 and 2 corresponding to w_i are $p_{1i} = [x_{1i}, y_{1i}, 1]^T$ and $p_{2i} = [x_{2i}, y_{2i}, 1]^T$ correspondingly.

Then:

$$C_1 w_i = \lambda_1 p_{1i} \qquad C_2 w_i = \lambda_2 p_{2i}$$

Where λ is some scale factor.

Using the cross product, we can represent $C_1 w_i = \lambda_1 p_{1i}$ as $\lambda_1 p_{1i} \times C_1 w_i = 0$. This allows us to drop the scale factor via division.

Written explicitly $p_{1i} \times C_1 w_i = 0$ becomes:

$$\begin{bmatrix} x_{1i} \\ y_{1i} \\ 1 \end{bmatrix} \times \begin{bmatrix} c^1_{1,1} & \cdots & c^1_{1,4} \\ \vdots & \ddots & \vdots \\ c^1_{3,1} & \cdots & c^1_{3,4} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} x_{1i} \\ y_{1i} \\ 1 \end{bmatrix} \times \begin{bmatrix} c^1_{1,1}x_i + c^1_{1,2}y_i + c^1_{1,3}z_i + c^1_{1,4} \\ c^1_{2,1}x_i + c^1_{2,2}y_i + c^1_{2,3}z_i + c^1_{2,4} \\ c^1_{3,1}x_i + c^1_{3,2}y_i + c^1_{3,3}z_i + c^1_{3,4} \end{bmatrix} = 0$$

$$\begin{bmatrix} (c^1_{3,1}y_{1i} - c^1_{1,1})x_i + (c^1_{3,2}y_{1i} - c^1_{1,2})y_i + (c^1_{3,3}y_{1i} - c^1_{1,3})z_i + c^1_{3,4}y_{1i} - c^1_{1,4} \\ (c^1_{3,1}x_{1i} - c^1_{2,1})x_i + (c^1_{3,2}x_{1i} - c^1_{2,2})y_i + (c^1_{3,3}x_{1i} - c^1_{2,3})z_i + c^1_{3,4}x_{1i} - c^1_{2,4} \\ \cdots \end{bmatrix}^T \begin{bmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{bmatrix} = 0$$

In this case, the final row in the result of the cross product will be a linear combination of the first two rows so it isn't necessary to explicitly write it out.

This gives us two equations:

$$(c^1_{3,1}y_{1i} - c^1_{2,1})x_i + (c^1_{3,2}y_{1i} - c^1_{2,2})y_i + (c^1_{3,3}y_{1i} - c^1_{2,3})z_i + c^1_{3,4}y_{1i} - c^1_{2,4} = 0 \tag{1}$$

$$(c^1_{3,1}x_{1i} - c^1_{1,1})x_i + (c^1_{3,2}x_{1i} - c^1_{1,2})y_i + (c^1_{3,3}x_{1i} - c^1_{1,3})z_i + c^1_{3,4}x_{1i} - c^1_{1,4} = 0 \tag{2}$$

The same process can be done for $C_2 w_i = \lambda_2 p_{2i}$ to get the equations:

$$(c^2_{3,1}y_{1i} - c^2_{2,1})x_i + (c^2_{3,2}y_{1i} - c^2_{2,2})y_i + (c^2_{3,3}y_{1i} - c^2_{2,3})z_i + c^2_{3,4}y_{1i} - c^2_{2,4} = 0 \tag{3}$$

$$(c^2_{3,1}x_{1i} - c^2_{1,1})x_i + (c^2_{3,2}x_{1i} - c^2_{1,2})y_i + (c^2_{3,3}x_{1i} - c^2_{1,3})z_i + c^2_{3,4}x_{1i} - c^2_{1,4} = 0 \tag{4}$$

Combining equations (1) to (4), we can represent them in the form $Aw_i = 0$:

$$\begin{bmatrix} c^1_{1,1} - c^1_{3,1}x_{1i} & c^1_{1,2} - c^1_{3,2}x_{1i} & c^1_{1,3} - c^1_{3,3}x_{1i} & c^1_{1,4} - c^1_{3,4}x_{1i} \\ c^1_{2,1} - c^1_{3,1}y_{1i} & c^1_{2,2} - c^1_{3,2}y_{1i} & c^1_{2,3} - c^1_{3,3}y_{1i} & c^1_{2,4} - c^1_{3,4}y_{1i} \\ c^2_{1,1} - c^2_{3,1}x_{1i} & c^2_{1,2} - c^2_{3,2}x_{1i} & c^2_{1,3} - c^2_{3,3}x_{1i} & c^2_{1,4} - c^2_{3,4}x_{1i} \\ c^2_{2,1} - c^2_{3,1}y_{1i} & c^2_{2,2} - c^2_{3,2}y_{1i} & c^2_{2,3} - c^2_{3,3}y_{1i} & c^2_{2,4} - c^2_{3,4}y_{1i} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = 0$$

Q3.3

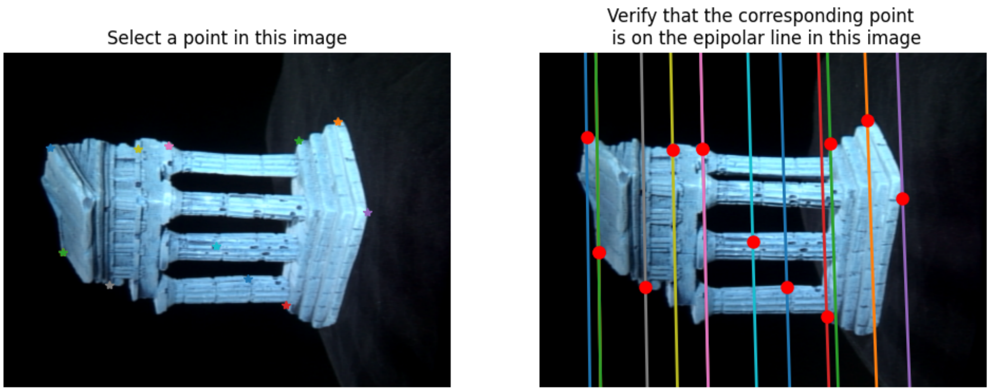
Minimum Error: 351.8979662288687

Best M2 matrix:

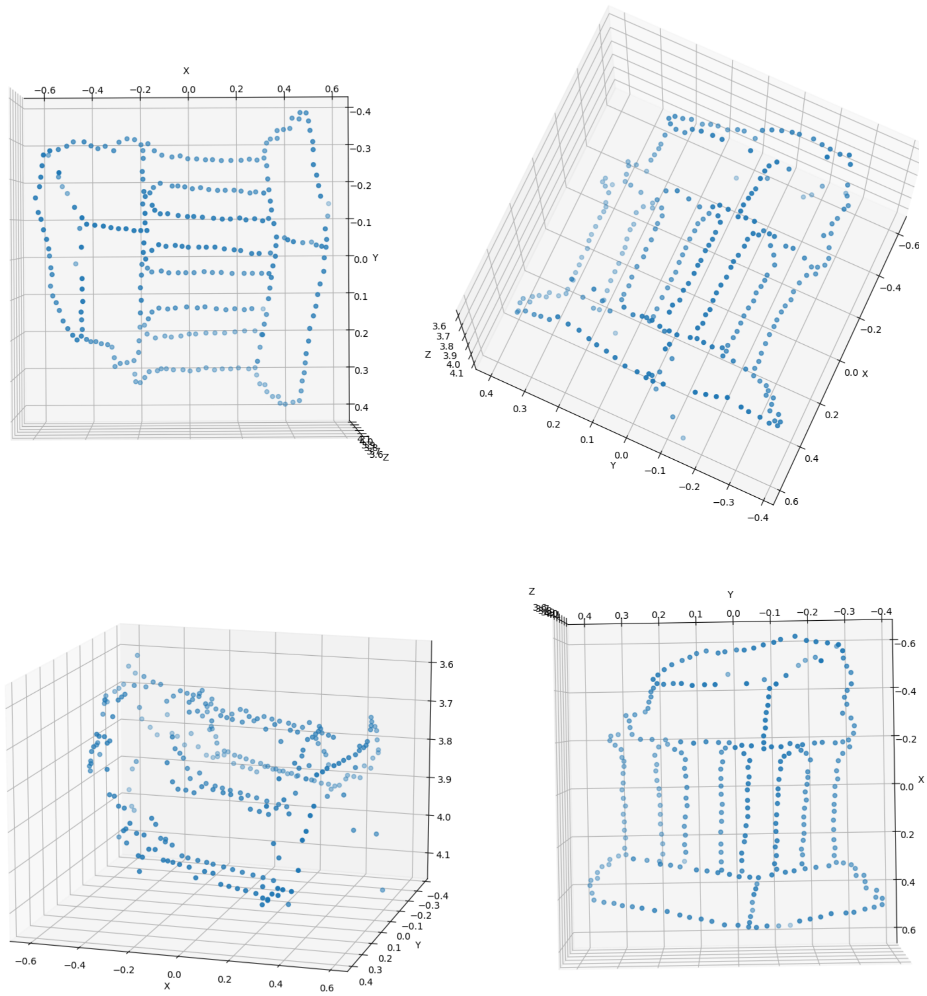
$$\begin{bmatrix} 0.99942697 & 0.03331533 & 0.00598477 & -0.02599827 \\ -0.03372859 & 0.96531605 & 0.25889634 & -1. \\ 0.00284802 & -0.25894984 & 0.96588657 & 0.07961991 \end{bmatrix}$$

Part 4

Q4.1


Select a point in this image
Verify that the corresponding point
is on the epipolar line in this image

Q4.2


X
−0.6 −0.4 −0.2 0.0 0.2 0.4 0.6
Y
−0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Z
Z
3.6 3.7 3.8 3.9 4.0 4.1
X
−0.6 −0.4 −0.2 0.0 0.2 0.4 0.6
Y
0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
3.6 3.7 3.8 3.9 4.0 4.1
X
−0.6 −0.4 −0.2 0.0 0.2 0.4 0.6
Y
−0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Z
Y
0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
X
−0.6 −0.4 −0.2 0.0 0.2 0.4 0.6

Part 5

5.1

nIters	tol	score
50	0.5	71
50	1	88
50	2	93
100	0.5	49
100	1	98
100	2	107
200	0.5	69
200	1	87
200	2	109
400	0.5	72
400	1	98
400	2	110

The error metric used for ransacF was distance from the point p_2 to the epipolar line in image 2, Fp_1, where p_2 is the point on the second image, and p_1 is the point on the first image. The formula for this is:

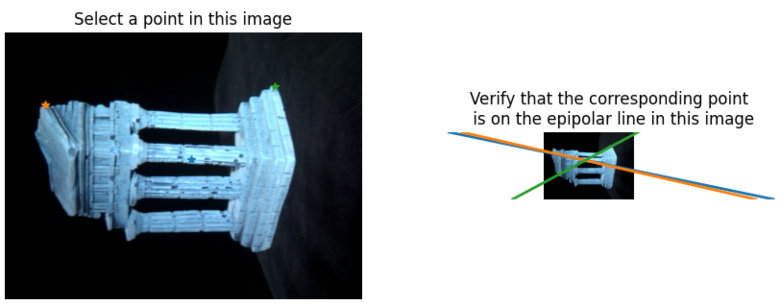
$$\frac{ax_0 + by_0 + c}{\sqrt{a^2 + b^2}}$$

where $ax + by + c = 0$ is the equation of the line described by Fp_l and (x_0, y_0) is the coordinates of p_2.

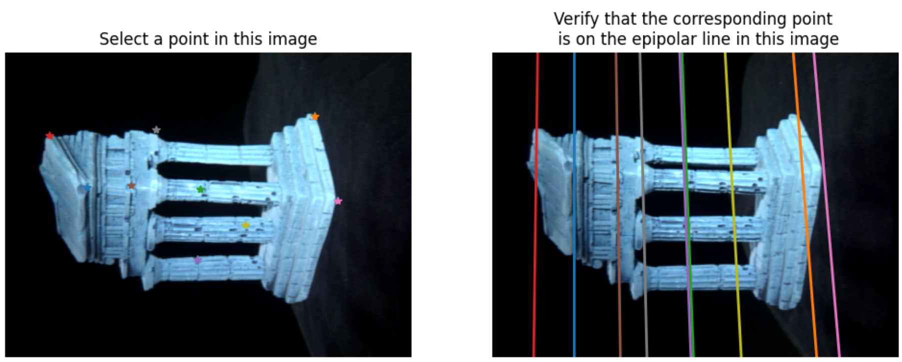
Based on the ablation table changing the number of iterations improves the score (number of inliers) up to a certain point. Increasing tolerance on the other hand more significantly increases score. However a larger score is not necessarily better when increasing tolerance. Since tolerance is used to determine inliers, a larger the tolerance makes it easier to incorrectly identify inliers.

Getting more points allows us to generate a more accurate fundamental matrix so increasing the number of iterations and tolerance (to a certain extent) will improve the accuracy of the fundamental matrix. This is true so long as we don't start adding bad inliers.

Epipolar Lines on noisy data without ransac.



Epipolar Lines on noisy data with ransac.

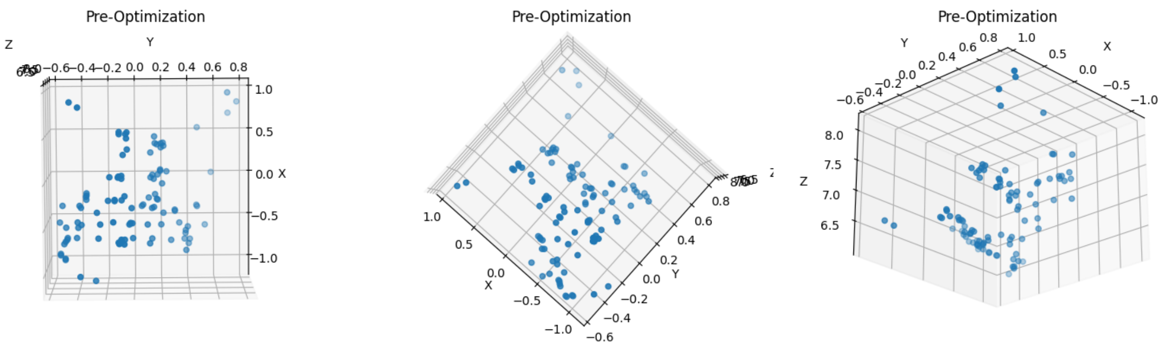


Q5.3

```
Number of Inliers: 105
Initial Error: 3528.7493585341695
Post optimization Error: 9.860943431511027
```

Looking at the scatter plots it looks like the optimized 3D points are shifted, however their relative locations are still similar to the pre-optimized 3D points.

Scatter of inlier points before optimization



Scatter of inlier points after optimization