16-720 Homework 4

Chendi Lin

November 12, 2018

Problem 1.1. Suppose two cameras fixate on a point x in space such that their principal axes intersect at that point. Show that if the image coordinates are normalized so that the coordinate origin (0,0) coincides with the principal point, the F33 element of the fundamental matrix is zero.

Solution 1.1. We are given that for point P in the figure, $\widetilde{x_1} = \widetilde{x_2} = [0 \quad 0 \quad 1]^T$. By fundamental matrix relation, we have

$$\widetilde{x_2}^T F \widetilde{x_1} = 0 \tag{1}$$

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{12} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \tag{2}$$

$$F_{33} = 0 \tag{3}$$

Problem 1.2. 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 1.2. For a pure translation parallel to x-axis,

$$t = \begin{bmatrix} t_1 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

$$E = tR \tag{6}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} t_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & t_1 & 0 \end{bmatrix} \tag{8}$$

Let $\widetilde{x_1^T} = \begin{bmatrix} a_1 & a_2 & 1 \end{bmatrix}$ and $\widetilde{x_2^T} = \begin{bmatrix} b_1 & b_2 & 1 \end{bmatrix}$, we have

$$
\begin{aligned}
l_1^T &= \widetilde{x_2}^T E && (9)\\
&= \begin{bmatrix} b_1 & b_2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & t_1 & 0 \end{bmatrix} && (10)\\
&= \begin{bmatrix} 0 & t_1 & -b_2 t_1 \end{bmatrix} && (11)\\
l_2^T &= \widetilde{x_1}^T E^T && (12)\\
&= \begin{bmatrix} a_1 & a_2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_1 \\ 0 & -t_1 & 0 \end{bmatrix} && (13)\\
&= \begin{bmatrix} 0 & -t_1 & a_2 t_1 \end{bmatrix} && (14)\\
& && (15)
\end{aligned}
$$

Thus, the epipolar line in the first camera is $t_1 y_1 - b_2 t_1 = 0$, and the epipolar line in the second camera is $-t_1 y_2 + a_2 t_1 = 0$. Both of them do not contain x components, so they are both parallel to the x-axis.

Problem 1.3. 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?

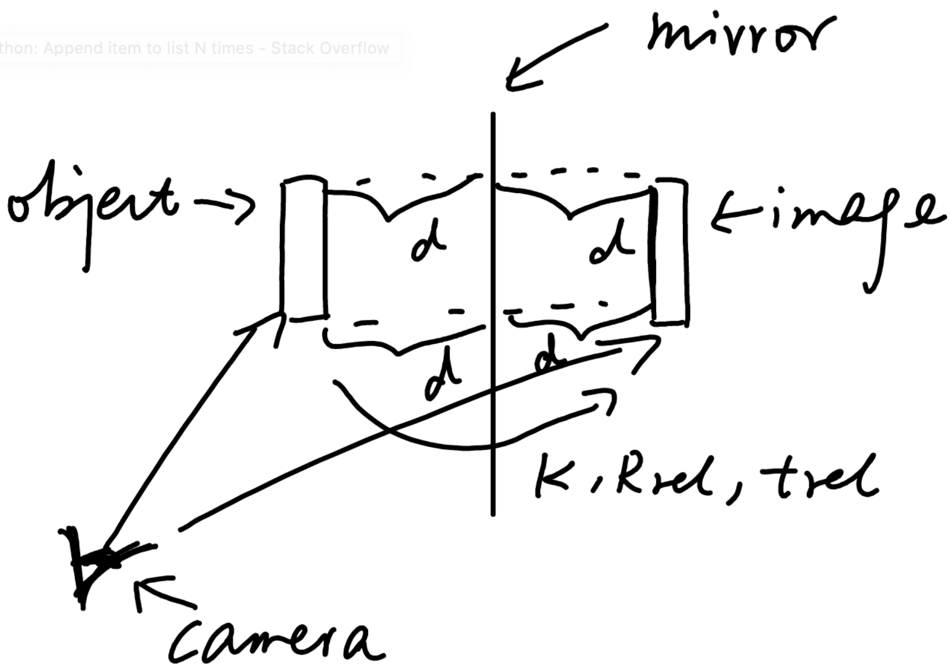
Solution 1.3. Let $\begin{bmatrix} u & v & w \end{bmatrix}^T$ be the coordinate of the object in the 3D world, and let $\begin{bmatrix} x_i & y_i \end{bmatrix}^T$ be the position at time i, so we have

$$
\begin{aligned}
\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} &= K(R_1 \begin{bmatrix} u \\ v \\ w \end{bmatrix} + t_1) && (16)\\
\begin{bmatrix} u \\ v \\ w \end{bmatrix} &= R_1^{-1}(K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - t_1) && (17)\\
&= R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - R_1^T t_1 && (18)\\
\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} &= K(R_2 \begin{bmatrix} u \\ v \\ w \end{bmatrix} + t_2) && (19)\\
&= K(R_2(R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - R_1^T t_1) + t_2) && (20)\\
&= K R_2 R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - K R_2 R_1^T t_1 + K t_2 && (21)
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
R_{rel} &= K R_2 R_1^T K^{-1} && (22)\\
t_{rel} &= -K R_2 R_1^T t_1 + K t_2 && (23)\\
E &= t_{rel} \times R_{rel} && (24)\\
F &= (K^{-1})^T E K^{-1} && (25)\\
&= (K^{-1})^T (t_{rel} \times R_{rel}) K^{-1} && (26)\\
& && (27)
\end{aligned}
$$

Problem 1.4. 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 . From the figure shown here we can see that, when a camera views an object and its reflection in a plane mirror, assuming that all points on the object are equal distance to the mirror, the transformation between the object and its reflection is pure translation. So we know that

$$
\begin{aligned}
R_{rel} &= I && (28)\\
t_{rel} &= [t_x, t_y, t_z] && (29)
\end{aligned}
$$

Also, from Problem 1.3 we know that,

$$
\begin{align}
F &= (K^{-1})^T (t_{rel} \times R_{rel}) K^{-1} \tag{30} \\
&= (K^{-1})^T \begin{bmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{bmatrix} K^{-1} \tag{31} \\
& \tag{32}
\end{align}
$$

and

$$
\begin{align}
F^T &= (K^{-1})^T \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} K^{-1} \tag{33} \\
&= -F \tag{34}
\end{align}
$$

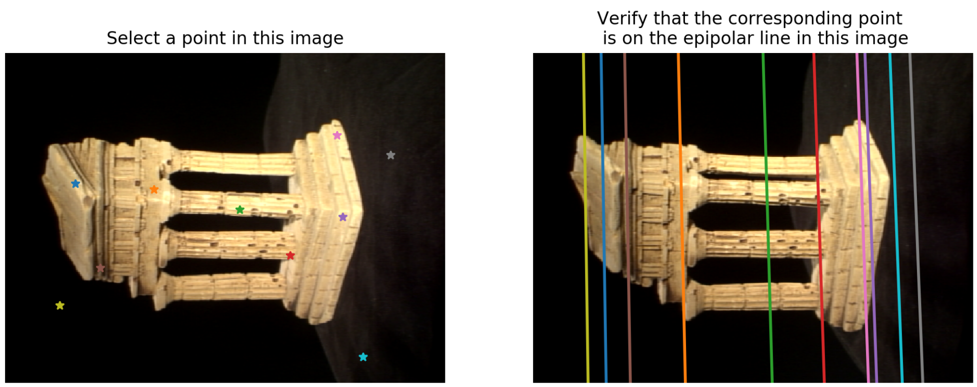
Thus, the fundamental matrix is a skew-symmetric matrix.

Problem 2.1. Eight-Point Algorithm

Solution . The recovered matrix F is

$$F = \begin{bmatrix} 9.80213860e-10 & -1.32271663e-07 & 1.12586847e-03 \\ -5.72416248e-08 & 2.97011941e-09 & -1.17899320e-05 \\ -1.08270296e-03 & 3.05098538e-05 & -4.46974798e-03 \end{bmatrix} \tag{35}$$

and the visualization is shown below

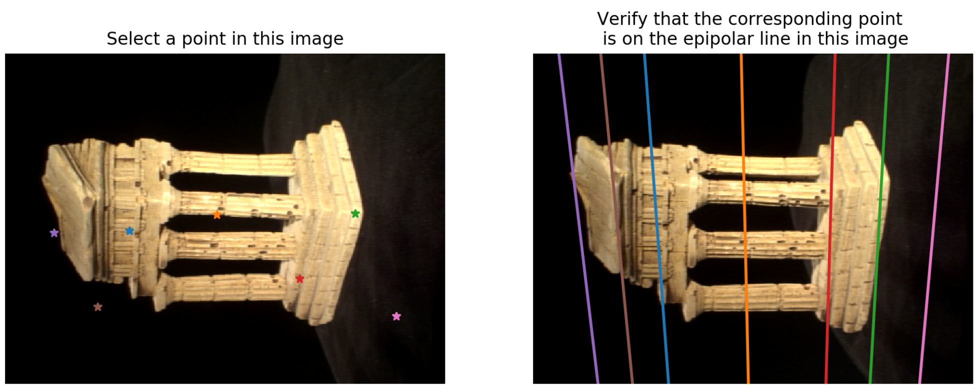


Problem 2.2 Seven-Point Algorithm.

Solution . After several trials, points 70 to 76 were found to perform the best. The seven point algorithm returned three F candidates. After visualization, we manually picked the third F as the best choice. The recovered matrix F is

$$F = \begin{bmatrix} 7.34867779e-08 & -5.07415735e-07 & 1.03710764e-03 \\ 3.71056967e-07 & 1.36989750e-08 & -1.37924363e-04 \\ -1.04235508e-03 & 1.46476677e-04 & 1.51876154e-03 \end{bmatrix} \tag{36}$$

and the visualization is shown below



As we can see, even though we have chosen the seven points carefully, it still did not perform as well as the eight-point algorithm.

Problem 3.1. Write your estimated E using F from the eight-point algorithm.

Solution . E can be calculated as $K_2^T F K_1$. Using eight-point algorithm, the estimated E is

$$E = \begin{bmatrix} 2.26587820e-03 & -3.06867395e-01 & 1.66257398e+00 \\ -1.32799331e-01 & 6.91553934e-03 & -4.32775554e-02 \\ -1.66717617e+00 & -1.33444257e-02 & -6.72047195e-04 \end{bmatrix} \tag{37}$$

Problem 3.2. Triangulate

Solution . Denote C_{1i} as the ith row of C_1 and C_{2i} as the ith row of C_2. If P_i is a 4×1 vector of the 3D coordinates in the homogeneous form, we have

$$C_1 P_i = \widetilde{x_{i1}} \tag{38}$$

$$\begin{bmatrix} C_{11} \\ C_{12} \\ C_{13} \end{bmatrix} \begin{bmatrix} w_i \\ v_i \\ w_i \\ 1 \end{bmatrix} = \begin{bmatrix} x_{i1} \\ y_{i1} \\ 1 \end{bmatrix} \tag{39}$$

$$C_2 P_i = \widetilde{x_{i2}} \tag{40}$$

$$\begin{bmatrix} C_{21} \\ C_{22} \\ C_{23} \end{bmatrix} \begin{bmatrix} w_i \\ v_i \\ w_i \\ 1 \end{bmatrix} = \begin{bmatrix} x_{i2} \\ y_{i2} \\ 1 \end{bmatrix} \tag{41}$$

Therefore,

$$C_{11} P_i = x_{i1} \tag{42}$$
$$C_{12} P_i = y_{i1} \tag{43}$$
$$C_{13} P_i = 1 \tag{44}$$
$$C_{21} P_i = x_{i2} \tag{45}$$
$$C_{22} P_i = y_{i2} \tag{46}$$
$$C_{23} P_i = 1 \tag{47}$$
$$\tag{48}$$

Rearrange them we have

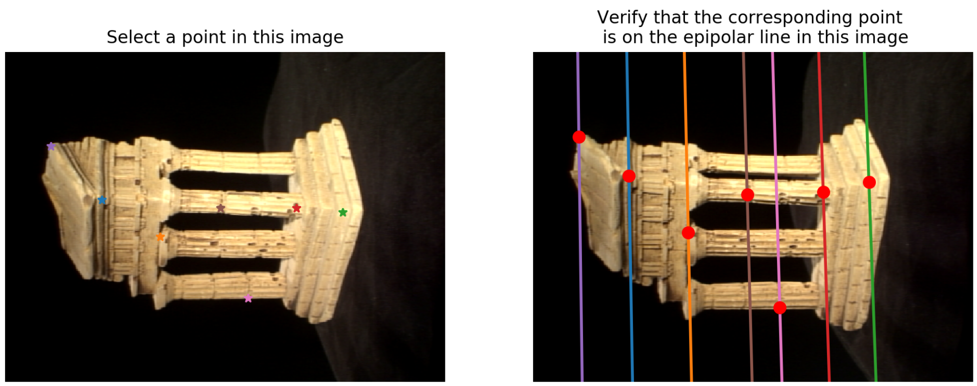
$$(x_{i1} C_{13} - C11) P_i = 0 \tag{49}$$
$$(y_{i1} C_{13} - C12) P_i = 0 \tag{50}$$
$$(x_{i2} C_{23} - C21) P_i = 0 \tag{51}$$
$$(y_{i2} C_{23} - C22) P_i = 0 \tag{52}$$
$$\tag{53}$$

Thus,

$$A_i = \begin{bmatrix} x_{i1} C_{13} - C_{11} \\ y_{i1} C_{13} - C_{12} \\ x_{i2} C_{23} - C_{21} \\ y_{i2} C_{23} - C_{22} \end{bmatrix} \tag{54}$$

Problem 4.1. Epipolar Correspondence

Solution . By several engineering trials, the window size was chosen to be 10*10. The visualization of epipolar correspondence is shown below



The results were pretty good. The corresponding points in different perspectives matched nicely.

Problem 4.2. 3D Visualization

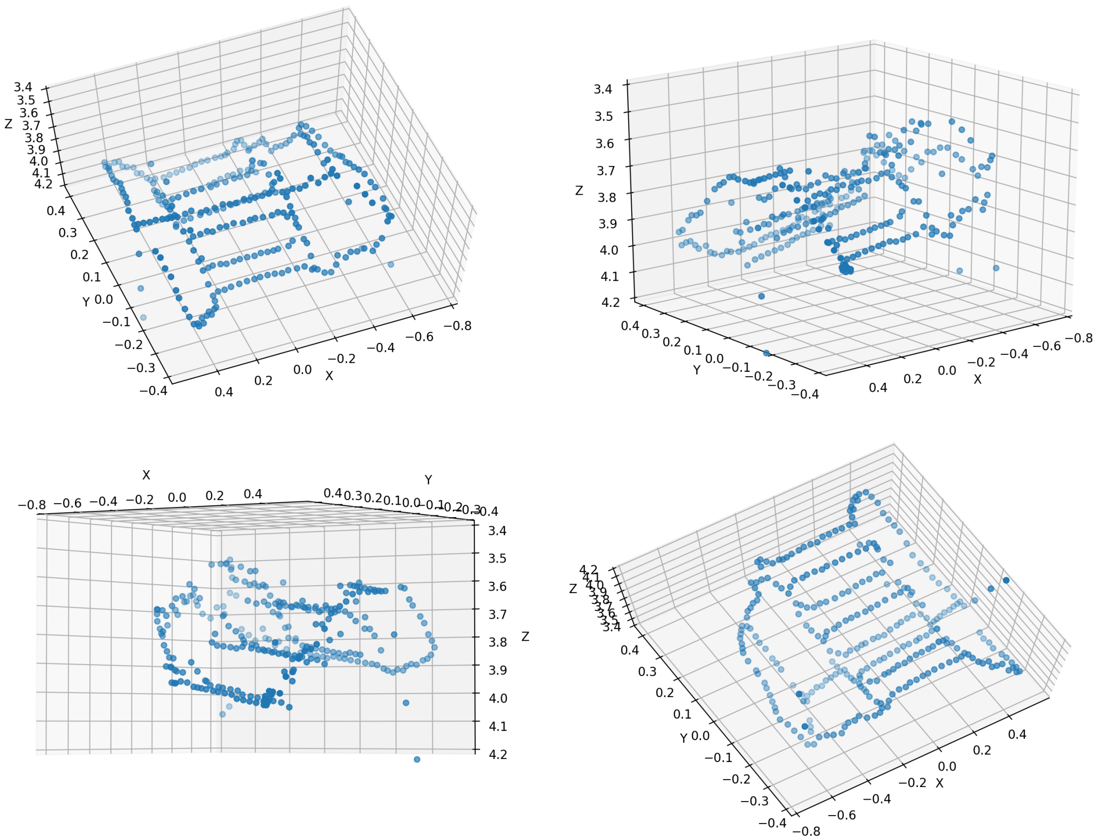


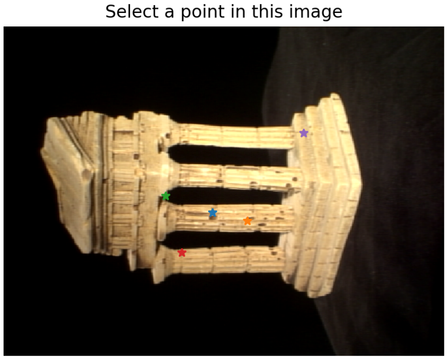
Figure 1: 3D Visualization

The overall reprojection error was 727.8792275245416.

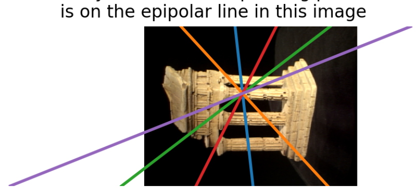
Problem 5.1. RANSAC

Solution . Using "some_corresp_noisy.npz", without RANSAC, the visualization of the epipolar lines looks like

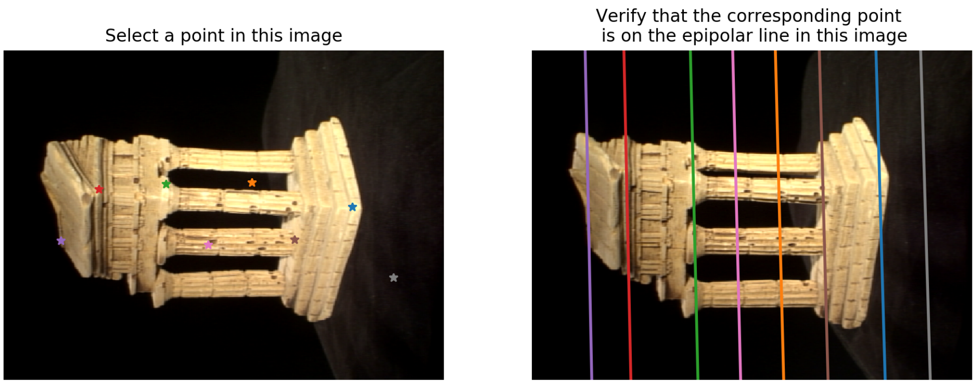
Select a point in this image



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



While with RANSAC, the visualization looks which was much better.



The obtained F from RANSAC was

$$F = \begin{bmatrix} -9.82915852e-10 & 1.32237382e-07 & -1.12586594e-03 \\ 5.72771833e-08 & -2.96995020e-09 & 1.17849254e-05 \\ 1.08270210e-03 & -3.05057289e-05 & 4.46968994e-03 \end{bmatrix} \tag{55}$$

which is the same as the matrix obtained from eight-point algorithm with not-noisy data, except that every entry was negated. It produces the same answer because when negating every entry, the epipolar lines will still be the same.

For fundamental matrix, we know that

$$\widetilde{x_2}^T F \widetilde{x_1} = 0 \tag{56}$$

So the error metrics used to determine if the point i was an inlier was that

$$err = abs(\widetilde{x_{i2}}^T F \widetilde{x_{i1}} - 0) \tag{57}$$

The threshold was set to be $2e-3$ by manual tuning. If the error of a point is less than the threshold, that point was set to be an inlier. After 200 iterations, all the inliers were used in the eight-point algorithm to determine the fundamental matrix F.

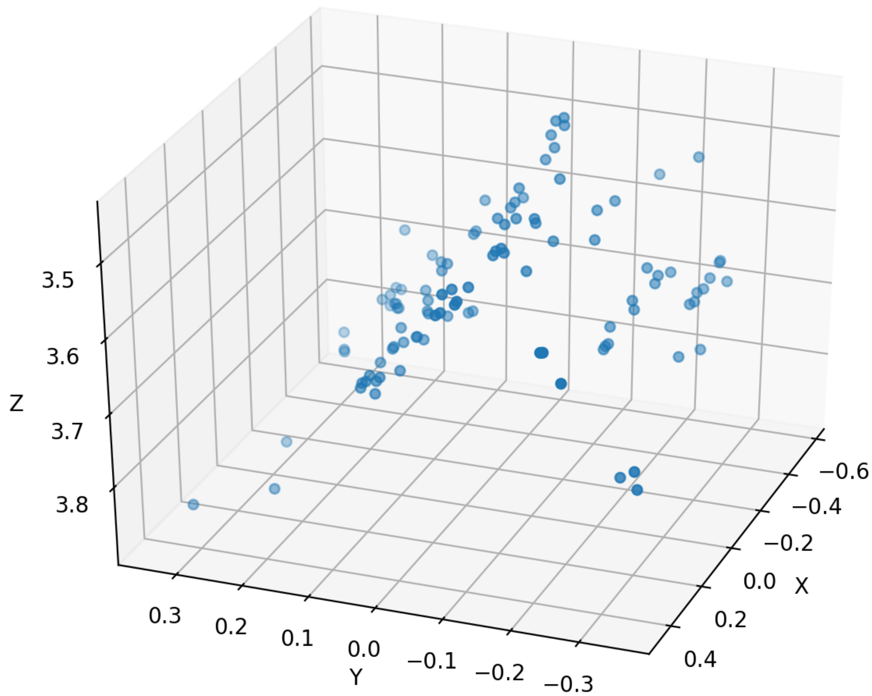


Figure 2: 3D Reconstruction with Noisy Data without Bundle Adjustment

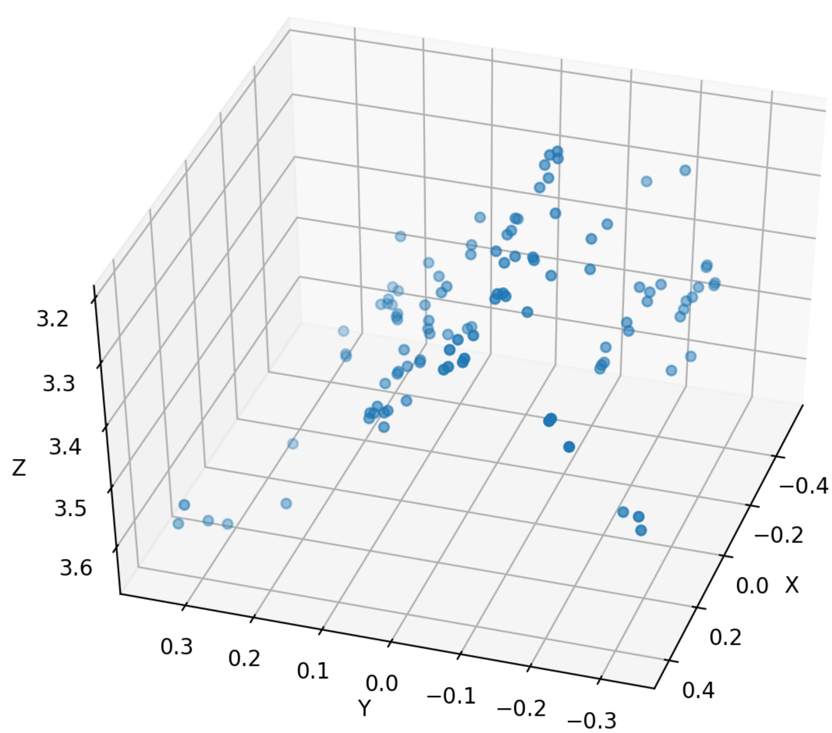


Figure 3: 3D Reconstruction with Noisy Data with Bundle Adjustment

Problem 5.3. Bundle Adjustment

Only the inlier points were plotted above. The numbers of inliers were 106 and 105. Without Bundle Adjustment, the error was 53.67326514786725, while with Bundle Adjustment, the error was 9.512796730969203, which was one-order less.