## Q1.1

Let's denote the left point as $p_l$ and right point as $p_r$.

Then we can get $p_l = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ and $p_r = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

We also know that $p_l^T F p_r = 0$.

$$\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} F \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

$$\begin{pmatrix} F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

$$F_{33} = 0$$

## Q1.2

Since two cameras view an object such that the second camera differs from the first by a pure translation that is parallel to the x-axis, then $R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $t = \begin{pmatrix} t_x \\ 0 \\ 0 \end{pmatrix}$.

Then the essential matrix $E = t \times R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix}$ Assume a point on the left camera is $p_l = \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$ and the corresponding point on the right camera is $p_r = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$.

Therefore the epipolar line in the left camera is $l_l = Ep_l = \begin{pmatrix} 0 \\ -t_x \\ bt_x \end{pmatrix}$ and the epipolar line in the right camera is $l_r = Ep_r = \begin{pmatrix} 0 \\ -t_x \\ vt_x \end{pmatrix}$.

Since their $x$ component is 0, they are both parellel to the x-axis.

## Q1.3

Assume a point $P$ in the world coordinate and two points $p_l$ and $p_r$, which denote the position of $P$ on the image plane of camera coordinate at frame $i$ and $i+1$.
Then we have $p_l = R_i P + t_i$ and $p_r = R_{i+1} P + t_{i+1}$.
Since $R_i$ is invertable, then we have $P = R_i^{-1}(p_l - t_i)$.
Then $p_r = R_{i+1} R_i^{-1}(p_l - t_i) + t_{i+1} = R_{i+1} R_i^{-1} p_l - R_{i+1} R_i^{-1} t_i + t_{i+1}$.
Therefore, the effective rotation $R_{\text{rel}} = R_{i+1} R_i^{-1}$ and effective translation $t_{\text{rel}} = t_{i+1} - R_{i+1} R_i^{-1} t_i$.
The essential matrix is $E = t_{\text{rel}} \times R_{\text{rel}}$.
The fundamental matrix is $F = K^{-T}(t_{\text{rel}} \times R_{\text{rel}}) K^{-1}$.

## Q1.4

Assume the 3d object is $P$ and its reflection in the mirror is $P'$. The real camera intrinsic is denoted by $K$ and the extrinsic is parameterised by $[\mathbf{I}, \mathbf{0}]$. $M_1$ is the projection matrix for the real camera, where $M_1 = K[\mathbf{I}, \mathbf{0}]$.
Let's assume the plane mirror is $Z = 1$.
Let's say $p$ and $p'$ is the 2d projection of $P$ and $P'$ by the real camera, respectively. Thus we can obtain:

$$p' = MP' \tag{1}$$

Now we consider the transformation matrix from $P'$ to $P$, i.e. $P = QP'$. Since the plane mirror is $Z = 1$, the homogeneous transformation matrix can be written as $Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.
In Equation 1, if we want to keep $p'$ unchanged, we can insert $Q$ and $Q^{-1}$ into it. Then we will get:

$$p' = MQ^{-1}QP' = MQ^{-1}P$$

, where $Q^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = Q$.
Geometrically, it represents the same transformation as reflecting w.r.t. the plane $Z = 1$.
Thus, the projection matrix for camera 2 is

$$M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

,where $R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ and $t = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ To consider the fundamental matrix, we can write $p'^T FP = 0$.
Here, $F = K^{-T}EK^{-1}$, where $E$ is the essential matrix.

$$E = [t]_x R = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Since $E^T = -E$, then we can say that $E$ is a skew-symmetric matrix. Then $F$ is also a skew-symmetric matrix.
Therefore, the situation is equivalent to having two images of the object which are related by a skew-symmetric fundamental matrix.

## Q2.1

The recovered F is shown in Figure 1.

```
F: [[ 9.78833284e-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]]
```

Figure 1: F Matrix.

The epipolar line visualization is shown in Figure 2.

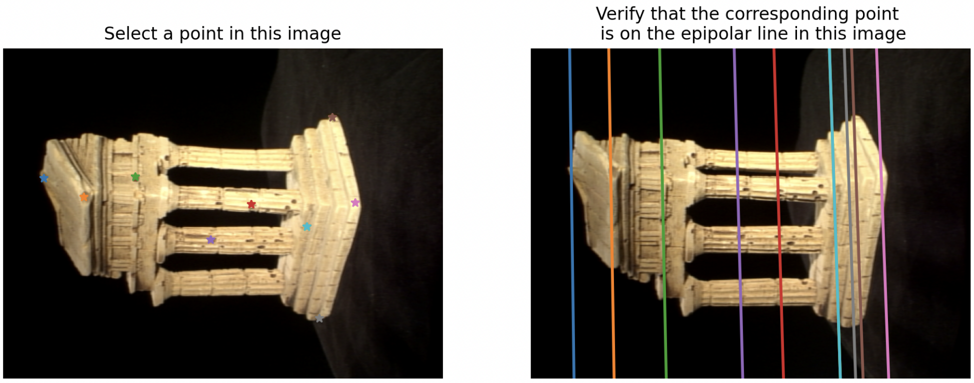


Figure 2: Epipolar Line Visualization.

## Q3.2

Suppose $C_1 = \begin{pmatrix} C_{11} \\ C_{12} \\ C_{13} \end{pmatrix}$ and there is a 2d point on the left image $\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}$ for a given point $i$.

Then we can have $\begin{pmatrix} C_{11} \\ C_{12} \\ C_{13} \end{pmatrix} \widetilde{w} = \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}$.

$$\begin{cases} C_{11}\widetilde{w} & = x_1 \\ C_{12}\widetilde{w} & = y_1 \\ C_{13}\widetilde{w} & = 1 \end{cases}$$

Then we can obtain $(C_{13}x_1 - C_{11})\widetilde{w} = 0$ and $(C_{13}y_1 - C_{12})\widetilde{w} = 0$.
Similarly, we can obtain $(C_{23}x_2 - C_{21})\widetilde{w} = 0$ and $(C_{23}y_2 - C_{22})\widetilde{w} = 0$ for the second camera.

Therefore, matrix $A_i$ can be written as $\begin{pmatrix} C_{13}x_{i1} - C_{11} \\ C_{13}y_{i1} - C_{12} \\ C_{23}x_{i2} - C_{21} \\ C_{23}y_{i2} - C_{22} \end{pmatrix}$.

## Q4.1

The screenshot of epipolarMatchGUI with some detected correspondences is shown in Figure 3.

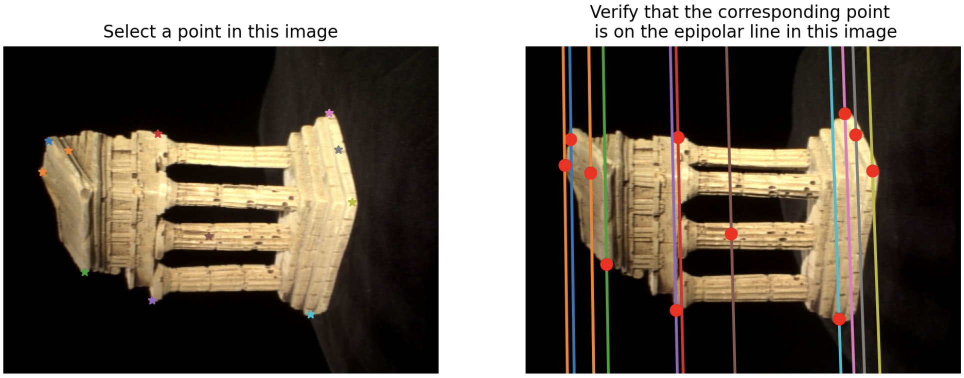


Figure 3: Epipolar Match GUI.

## Q4.2

The visualization results are shown in Figure 4.

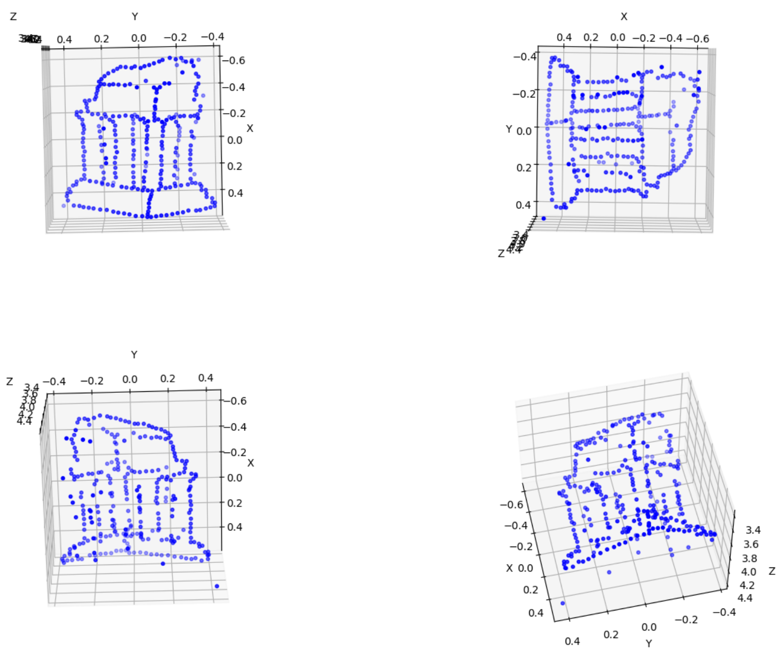

Figure 4: Different views of 3d visualization.

## 5.1

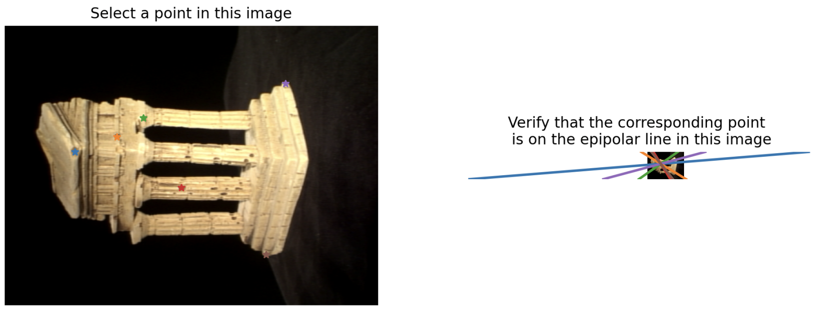


Figure 5: Result of the original eight point algorithm.

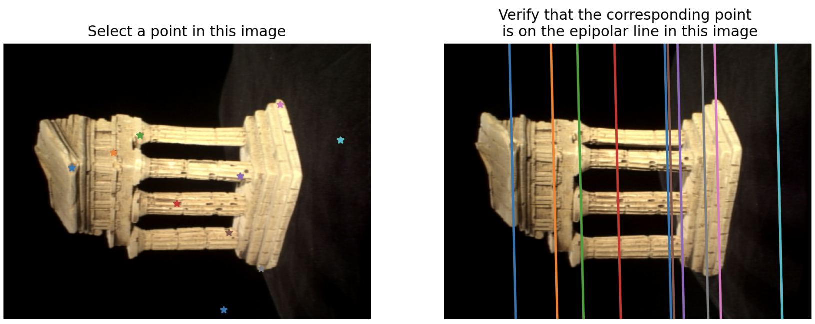


Figure 6: Result of the RANSAC method.

The comparisons between the original eight point algorithm and the RANSAC method is shown in Figure 5 and 6.
The error metric I used is the distance from point 2 to the predicted epipolar line. Let's say the epipolar line is $Ax + By + C = 0$ and point 2 is $[u, v]$. Then it can be written as:

$$d = \frac{|Au + Bv + C|}{\sqrt{A^2 + B^2}}$$

I set the tolerance value to be 1. That is, when $d < 1$, then it will be classified as an inlier. Otherwise, it would be a outlier.
There are two hyper-parameters: number of iterations for RANSAC and the tolerance for classifying the points as inlier/outlier.
- If we increase the number of iterations, the fundamental matrix will be more accurate. It will converge after a certain number of iterations.
- If we use a smaller tolerance value, then fewer points would be considered as inliers, in which case the fundamental matrix will be inaccurate due to overfitting to the few points. On the other hand, if we increase the tolerance value, then more points would be included and it may introduce noise.

## Q5.3

The original 3d points are shown in Figure 7 and the optimized 3d points are shown in Figure 8.
The initial reprojection error is 2841.97 and the reprojection error with optimized matrices is 8.08.

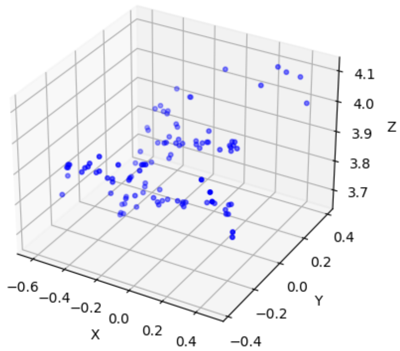


Figure 7: Original 3d points.

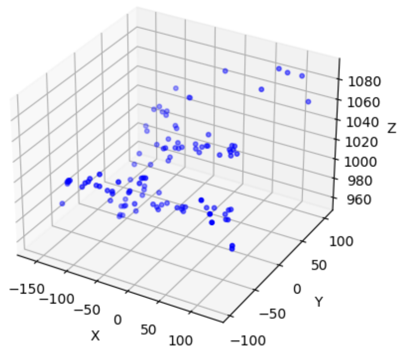


Figure 8: Optimized 3d points.