Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

Q1.1

Assuming that the image coordinates are normalized so that the origins coincide with the principal points. If a point in the world $\boldsymbol{w}$ has its image in both the cameras at their origins. Then we can say-

- $\widetilde{\boldsymbol{x_2}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ be the point at which $\boldsymbol{w}$ is projected at in camera 2's image plane.
- $\widetilde{\boldsymbol{x_1}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ be the point at which $\boldsymbol{w}$ is projected at in camera 1's image plane.
- $\boldsymbol{F} = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix}$ be the fundamental matrix which maps the relationship between the 2 views.

Then the below equation holds good-

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

Simplifying we have,

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

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

$$F_{33} = 0$$

Therefore, if the above suppositions regarding the positions of the coordinate axes hold good then the F_{33} element of the fundamental matrix is zero.

Q1.2

Assuming the world coordinate system is centered with the coordinate axes of camera 1. Then we have,

$$\boldsymbol{w} = \boldsymbol{\lambda_1} \widetilde{\boldsymbol{x_1}}$$

Let a rotation matrix $\boldsymbol{\Omega}$ and a translation vector $\boldsymbol{t}$ be the mapping between the world coordinates $\boldsymbol{w}$ and $\widetilde{\boldsymbol{x_2}}$ for camera 2. Then we have,

$$\boldsymbol{\lambda_2} \widetilde{\boldsymbol{x_2}} = \boldsymbol{\Omega w} + \boldsymbol{t}$$

However, it is given that the coordinate axes of camera 2 differs from that of the camera 1, (also the world coordinate axes by our assumption), by a pure translation along the x-axis. Then we have,

$$\boldsymbol{\Omega} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}, \boldsymbol{t} = \begin{bmatrix} \boldsymbol{t_x} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

Therefore, the essential matrix can be calculated as-

$$\boldsymbol{E} = \boldsymbol{t}_\times \boldsymbol{\Omega}$$

Solving we have,

$$\boldsymbol{E} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{t_x} \\ \mathbf{0} & \boldsymbol{t_x} & \mathbf{0} \end{bmatrix}$$

Let Λ_1 and Λ_2 be the intrinsics of camera 1 and 2. Then the fundamental matrix F can be calculated as-

$$\boldsymbol{F} = \boldsymbol{\Lambda_2^{-T} E \Lambda_1^{-1}}$$

Assuming $\boldsymbol{\Lambda_2} = \boldsymbol{\Lambda_1} = \boldsymbol{I}$**,** we have $\boldsymbol{F} = \boldsymbol{E}$
Therefore, the epipolar line due to point $\widetilde{\boldsymbol{x_2}}$ on image 1 can be given as-

$$l_1 = \widetilde{\boldsymbol{x_2}}^{\boldsymbol{T}}\boldsymbol{E}$$

$$l_1 = [x_2 \quad y_2 \quad 1]\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix}$$

$$l_1 = [0 \quad t_x \quad -y_2 t_x]$$

The equation l_1 can be given as-

$$t_x y = y_2 t_x$$

Simplifying we have the equation for the epipolar line in image 1 as,

$$y = y_2$$

The epipolar line due to point $\widetilde{\boldsymbol{x_1}}$ on image 2 can be given as-

$$l_2 = \widetilde{\boldsymbol{x_1}}^{\boldsymbol{T}}\boldsymbol{E}^{\boldsymbol{T}}$$

$$l_2 = [x_1 \quad y_1 \quad 1]\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{bmatrix}$$

The equation l_2 can be given as-

$$l_2 = [0 \quad -t_x y_1 \quad y_1 t_x]$$

Simplifying we have the equation for the epipolar line in image 2 as,

$$y = y_1$$

There we can see that both the epipolar lines derived above are parallel to the x-axis.

Q1.3

At time t let the rotational matrix be R_t and translational τ_t and at time t+1 we have R_{t+1} and τ_{t+1}.
Then we can write the following expression-

$$\boldsymbol{R_{t+1}w} + \boldsymbol{\tau_{t+1}} = \boldsymbol{R_{rel}}(\boldsymbol{R_t w} + \boldsymbol{\tau_t}) + \boldsymbol{\tau_{rel}}$$

Simplifying we have,

$$\boldsymbol{R_{t+1}w} + \boldsymbol{\tau_{t+1}} = \boldsymbol{R_{rel}R_t w} + \boldsymbol{R_{rel}\tau_t} + \boldsymbol{\tau_{rel}}$$

Then by comparing coefficients we get the expressions for $\boldsymbol{R_{rel}}$ and $\boldsymbol{\tau_{rel}}$

$$\boldsymbol{R_{t+1}} = \boldsymbol{R_{rel}R_t}$$

$$\boldsymbol{R_{rel}} = \boldsymbol{R_{t+1}R_t^{-1}}$$

Since **R** is an orthonormal matrix, $\mathbf{R}^{-1}=\mathbf{R}^{T}$
Therefore

$$\boldsymbol{R_{rel}} = \boldsymbol{R_{t+1}R_t^T}$$

Also,

$$\boldsymbol{\tau_{rel}} = \boldsymbol{\tau_{t+1}} - \boldsymbol{R_{rel}\tau_t}$$

Therefore,

$$\boldsymbol{\tau_{rel}} = \boldsymbol{\tau_{t+1}} - \boldsymbol{R_{t+1}R_t^T\tau_t}$$

If we say that the world coordinates align with the coordinate system of camera at time t. Then the Rotation matrix and translation vector which represents the motion of camera from time t to time t+1 can be given by $\boldsymbol{R_{rel}}$ and $\boldsymbol{\tau_{rel}}$. Therefore we can give the essential matrix as-

$$\boldsymbol{E} = \boldsymbol{\tau_{rel}} \times \boldsymbol{R_{rel}}$$

$$\boldsymbol{E} = \boldsymbol{\tau_{rel\times}R_{rel}}$$

Where,

$$\boldsymbol{\tau}_{\times} = \begin{bmatrix} 0 & -\tau_z & \tau_y \\ \tau_z & 0 & -\tau_x \\ -\tau_y & \tau_x & 0 \end{bmatrix}$$

The fundamental matrix is given as-

$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$$

Simplifying we have,

$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{\tau}_{rel\times}\boldsymbol{R}_{rel}\boldsymbol{K}^{-1}$$

Q1.4

Let $\boldsymbol{w}$ be a point on the object and $\boldsymbol{w}'$ be the reflection of the point about a surface $\boldsymbol{L}$. Let the image of $\boldsymbol{w}$ be $\boldsymbol{x_1}$ and the image of $\boldsymbol{w}'$ be $\boldsymbol{x}'_{\boldsymbol{1}}$. Then $\boldsymbol{w}'$ is just a translation of $\boldsymbol{w}$ by say $\boldsymbol{d} = (d_x, d_y, d_z)$. Then we can say in homogeneous coordinates-

$$\widetilde{\boldsymbol{w}}' = [\boldsymbol{I}|\boldsymbol{d}]\widetilde{\boldsymbol{w}}$$

It is also given that all points are of the same distance from the mirror and hence their reflections can also be found by translating them by (d_x, d_y, d_z). Hence the reflection of all those points will be similar and can be represented by the properties of a single point $\boldsymbol{w}$ in that set of points.

Let the world coordinate frame and the camera coordinate frames coincide then camera matrix is given as $\boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]$. So, the points $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ can be given as-

$$\boldsymbol{x_1} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]\widetilde{\boldsymbol{w}}$$

$$\boldsymbol{x}'_{\boldsymbol{1}} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]\begin{bmatrix} \boldsymbol{w}+\boldsymbol{d} \\ \boldsymbol{1} \end{bmatrix}$$

Which can be simplified as,

$$\boldsymbol{x}'_{\boldsymbol{1}} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{d}]\widetilde{\boldsymbol{w}}$$

The above equation means that the image of the reflection of $\boldsymbol{w}$ (image of $\boldsymbol{w}'$) can be obtained by translating the camera by the vector $\boldsymbol{d}$ and then viewing $\boldsymbol{w}$.

Therefore, we now have an equivalent case of the problem as the translation of the camera instead. Now we can therefore represent the relation between-

- the 3D points,
- their reflections about a plane mirror,
- and their images on the camera screen

with a fundamental matrix.

Let us now represent the problem as a translation of camera position. Then the camera's initial orientation is $[\boldsymbol{I}|\boldsymbol{0}]$ and the camera's second (equivalent) orientation is $[\boldsymbol{I}|\boldsymbol{d}]$. Therefore the $\boldsymbol{R}_{rel} = \boldsymbol{I}$ and $\boldsymbol{\tau}_{rel} = \boldsymbol{d}$. Then The fundamental matrix can be given as-

$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{\tau}_{rel\times}\boldsymbol{R}_{rel}\mathbf{K}^{-1}$$

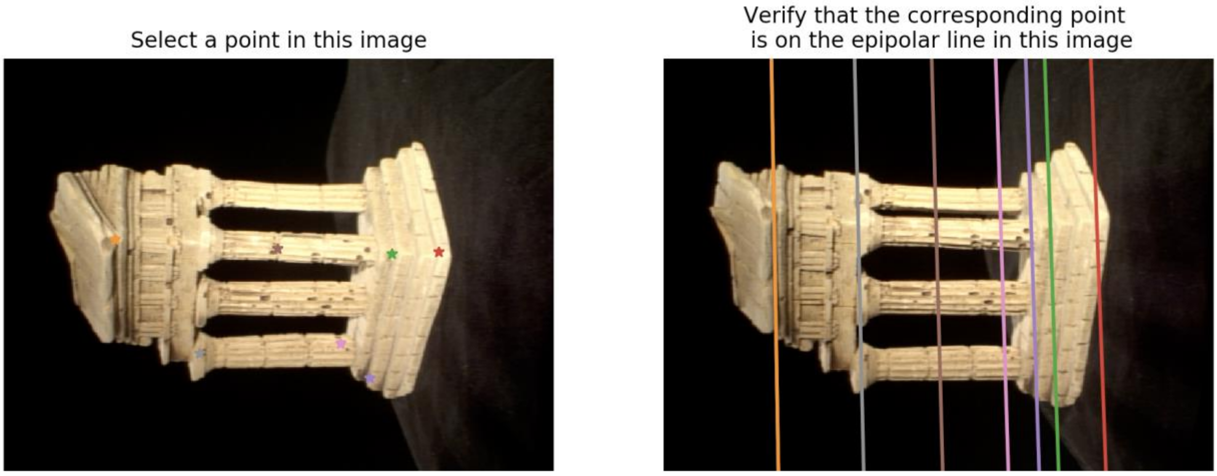
$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{\tau}_{rel\times}\boldsymbol{I}\boldsymbol{K}^{-1} = \boldsymbol{K}^{-T}\boldsymbol{\tau}_{rel\times}\boldsymbol{K}^{-1}$$

$$\boldsymbol{F} = \boldsymbol{K}^{-T}\boldsymbol{\tau}_{rel\times}\boldsymbol{K}^{-1}$$

Therefore since $\boldsymbol{\tau}_{rel\times}$ is skew-symmetric **F** is also skew-symmetric.

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

Q2.1

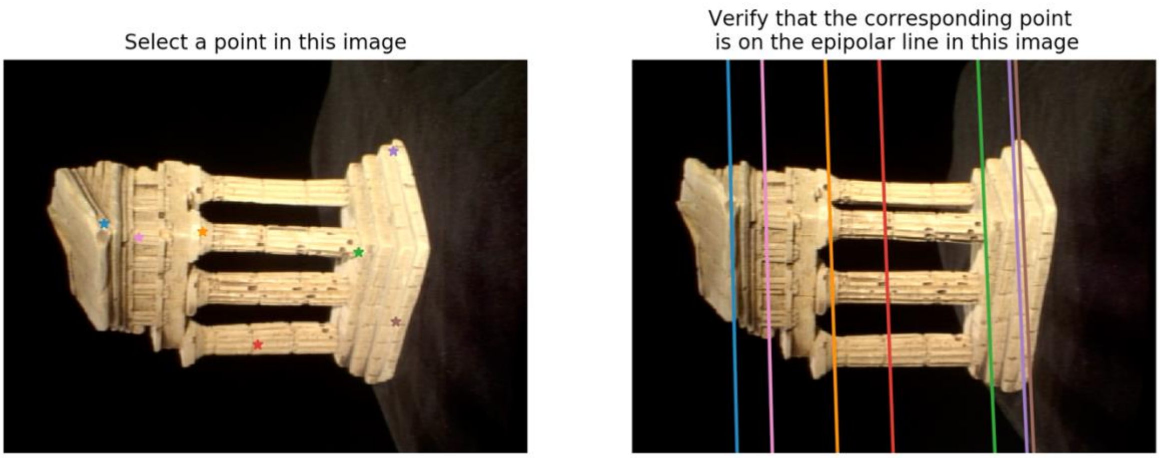


The recovered fundamental matrix is-

$$\begin{bmatrix} 9.80213863e-10 & -1.32271663e-07 & 1.12586847e-03 \\ -5.72416248e-08 & 2.97011941e-09 & -1.17899320e-05 \\ -1.08270296e-03 & 3.05098538e-05 & -4.46974798e-03 \end{bmatrix}$$

Q2.2

The documentation given above the function signature says to expect a N×2 matrix of points but after clarifying in TA office hours went ahead with implementing the function to expect just seven pairs of point correspondences. (i.e. shape of pts1 and pts2 is 7×2)



The F used in the above visualization is-

$$\begin{bmatrix} -4.64397536e-08 & 1.43920153e-07 & -9.51053392e-04 \\ 6.15816958e-08 & -5.19245555e-09 & 1.54187624e-05 \\ 9.28998137e-04 & -4.42375058e-05 & 3.83241747e-03 \end{bmatrix}$$

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

Q3.1

The recovered essential matrix-

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

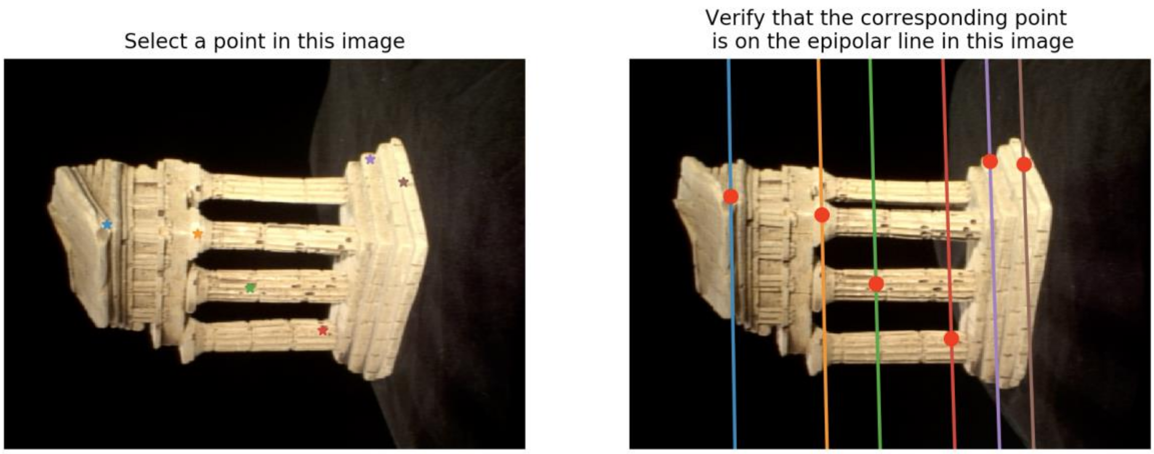
Q3.2

The expression for **A** in the expression $\boldsymbol{A}\widetilde{\boldsymbol{w}} = \boldsymbol{0}$ can be derived by taking advantage of the fact that the points in the two images, $[x, y]$ and $[x', y']$ of the point $\boldsymbol{w_i}$ lie in the same plane. Therefore, the cross product of the vectors forming this triangle of three points is 0. Let the two camera matrices $\boldsymbol{C^1}$ and $\boldsymbol{C^2}$. Also C_{ij}^k is the ith row jth column of the kth camera matrix.

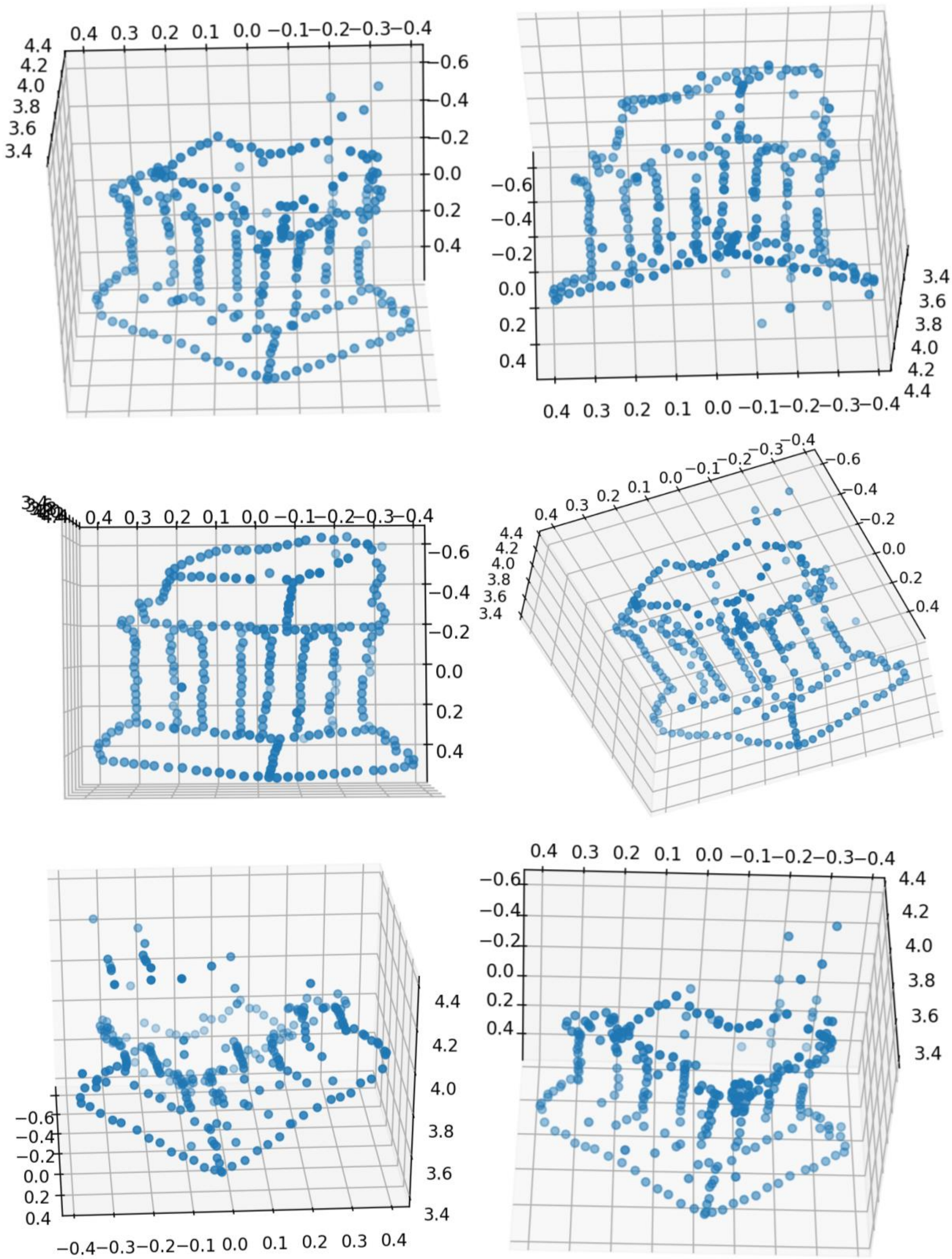
Then solving and simplifying we have-

$$\begin{bmatrix} C_{11}^1 - xC_{31}^1 & C_{12}^1 - xC_{32}^1 & C_{13}^1 - xC_{33}^1 \\ yC_{31}^1 - C_{21}^1 & yC_{32}^1 - C_{22}^1 & yC_{33}^1 - C_{23}^1 \\ C_{11}^2 - x'C_{31}^2 & C_{12}^2 - x'C_{32}^2 & C_{13}^2 - x'C_{33}^2 \\ y'C_{31}^2 - C_{21}^2 & y'C_{31}^2 - C_{21}^2 & y'C_{31}^2 - C_{21}^2 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{bmatrix} = \boldsymbol{0}$$

Q4.1



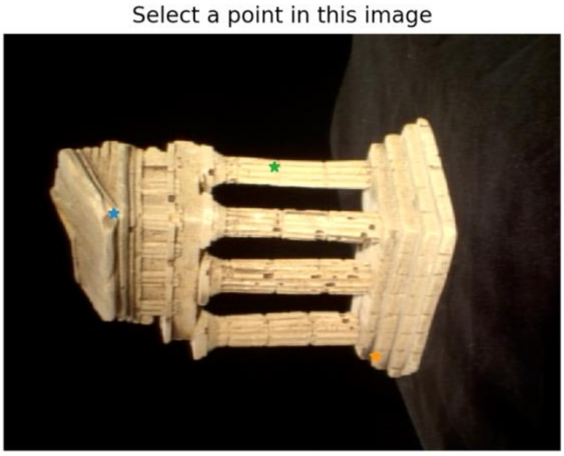
Q4.2

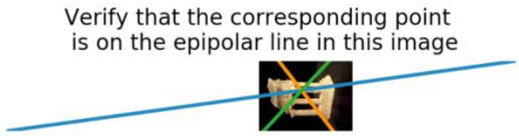

0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
4.4 4.2 4.0 3.8 3.6 3.4
−0.6 −0.4 −0.2 0.0 0.2 0.4
−0.6 −0.4 −0.2 0.0 0.2 0.4
3.4 3.6 3.8 4.0 4.2 4.4
0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
−0.6 −0.4 −0.2 0.0 0.2 0.4
0.4 0.3 0.2 0.1 0.0 −0.1 −0.2 −0.3 −0.4
4.4 4.2 4.0 3.8 3.6 3.4
−0.6 −0.4 −0.2 0.0 0.2 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
4.4 4.2 4.0 3.8 3.6 3.4
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
4.4 4.2 4.0 3.8 3.6 3.4

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

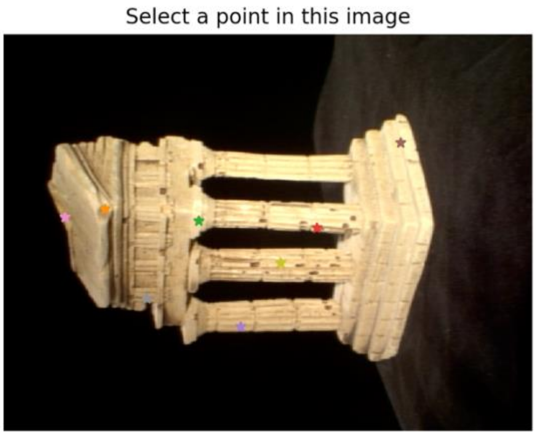
Q5.1

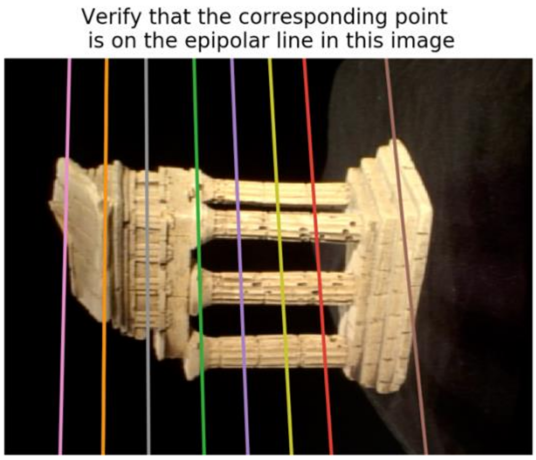
Below is the result obtained using eight-point algorithm with noisy correspondences





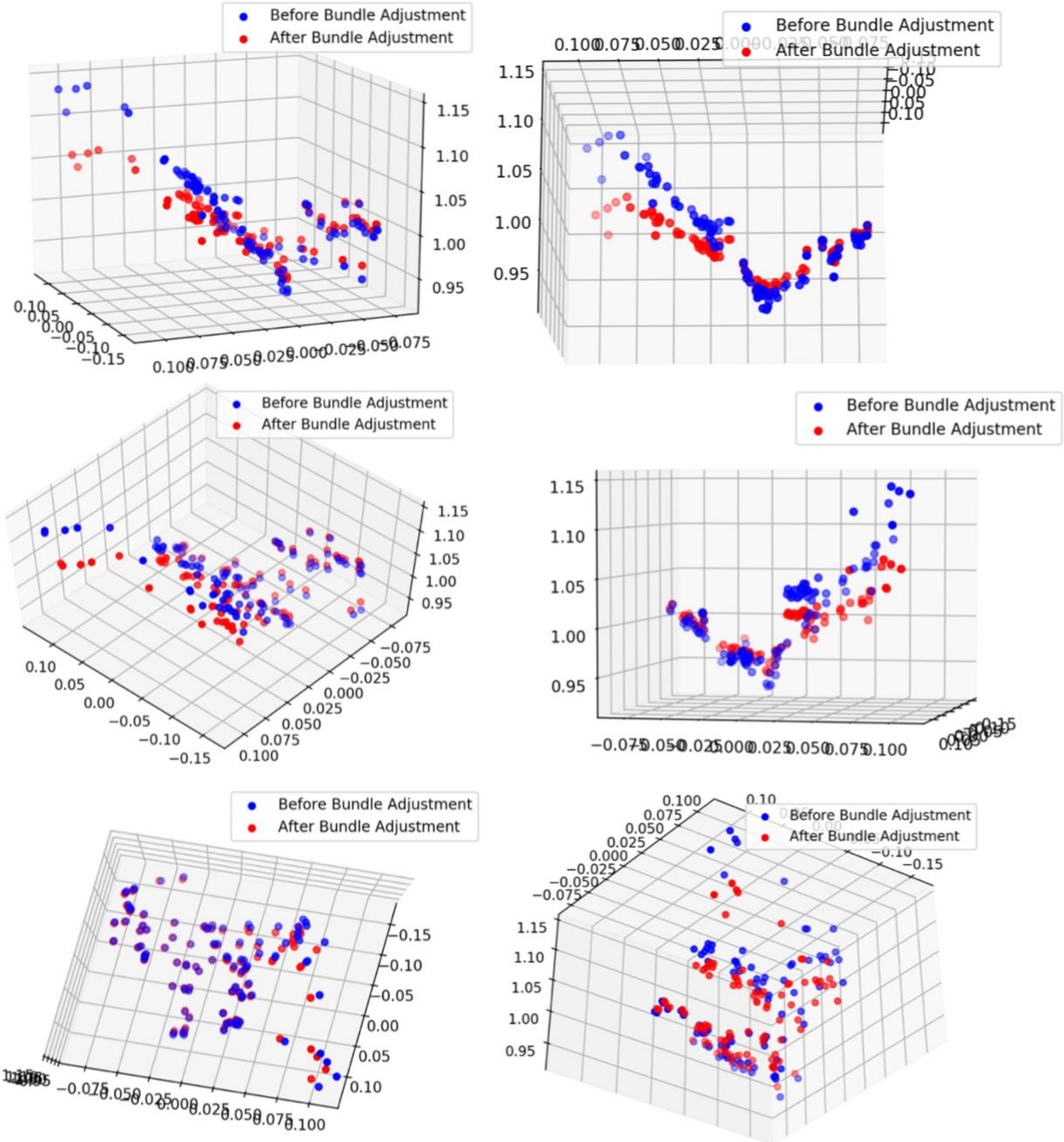
Below is the result obtained using RANSAC and seven-point algorithm with noisy correspondences





Q5.3

Implementation for this visualization is in the file bundle.py



The reprojection error, for this run, before bundle-adjustment is 2946.121888252543.
The reprojection error, for this run, after bundle-adjustment is 8.82792686131785.

We can clearly see that bundle adjustment is bringing far-off points closer. Additionally, when we look at the re-projected points, we get from the optimized world coordinates and camera matrix we can see that coordinates which were previously off by a few decimal values are perfect integers (pixel coordinates) and are the right correspondence values.

Furthermore, we see such a drastic decrease in error because I use an L2-norm and it penalizes the outliers much more and since after bundle adjustment the outliers are optimized the error reduces drastically.