Jerry Hsiung
Computer Vision 16-720
Assignment 4
November 13, 2016

Assignment #4

Q1.1 (5 points) Suppose two cameras fixate on a point P (see figure 1) 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 $\mathbf{F}_{33}$ element of the fundamental matrix is zero.

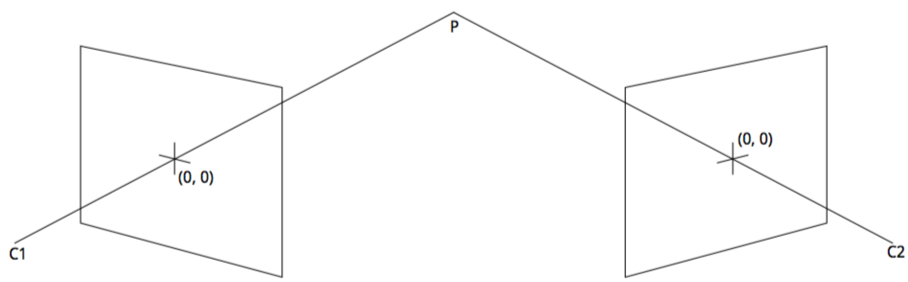


Figure 1: Figure for Q1.1. $C1$ and $C2$ are the optical centers. The principal axes intersect at point P.

Let $\mathbf{x}' = [x', y', 1]^T, \mathbf{x} = [x, y, 1]^T$. According to the definition of fundamental matrix:

$$\mathbf{x}' F \mathbf{x} = 0$$

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

Given that the coordinate are normalized, and $x' = x = (0, 0)$, then to make the equation hold true is when $f_{33} = 0$. ■

Q1.2 (5 points) 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.

Suppose the camera frame 1 is the origin, the camera frame 2 has rotation $R = I$ and translation t with respect to frame 1. The point $\mathbf{x}_2$ in frame 2 can be written as $R\mathbf{x}_1 + t = \mathbf{x}_1 + t$. Also, since camera 1 is at the origin, then frame 2's epipole $\mathbf{e}_2 = t$. Now, denoted the epipolar lines in two camera frames l_1, l_2, then using the definition of epipolar line, we get $l_2 = [\mathbf{e}_2]_\times \mathbf{x}_2$. We can rewrite this definition and get:

$$\begin{aligned} l_2 &= [\mathbf{e}_2]_\times \mathbf{x}_2 \\ &= [t]_\times (R\mathbf{x}_1 + t) \\ &= [t]_\times \mathbf{x}_1 + [t]_\times t && \text{(Since } R = I) \\ &= [t]_\times \mathbf{x}_1 && \text{(Cross product of itself} = 0) \end{aligned}$$

Supposed there is a point t in frame 2, using triple scalar product:

$$\begin{aligned} l_2 &= [t]_\times \mathbf{x}_1 && \text{(Cross product of itself = 0)} \\ t \cdot l_2 &= t \cdot [t]_\times \mathbf{x}_1 \\ &= 0 \end{aligned}$$

This means that the point t and the line l_2 is, in fact, colinear. And since point t can be treated as vector t from the origin in frame 2, the results tells us that the vector t is colinear, and therefore, parallel to l_2. This shows that the epipolar line l_2 is parallel to the translation t, in this case x-axis.

Similar to analyze l_1, we can treated frame 2 as the origin, then the translation t is just the negative direction of x-axis. Then follow the similar analysis, we can see that $l_1 = [-t]_\times \mathbf{x}_1$, and is also parallel to the x-axis. ■

> **Q1.3 (5 points)** Suppose we have an inertial sensor which gives us the accurate positions (R_i and t_i, where R is the rotation matrix and t is corresponding 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}.

The relative rotation will be: $R_{rel} = \frac{R_{i+1}}{R_i}$
and translatoin will be: $t_{rel} = t_{i+1} - t_i$.

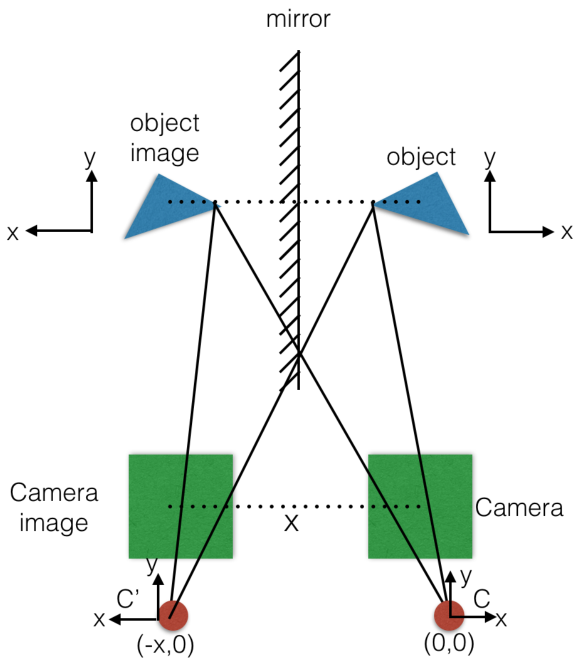
Since the first camera is now the original frame, the camera matrices P, P' are $K[\mathbf{I}|0], K[\mathbf{R}|t]$. Given the definition of fundamental matrix from the text book, we can write F as:

$$\begin{aligned} F &= K^{-T}[\mathbf{t_{rel}}]_\times R_{rel} K^{-1} \\ E &= K^T F K \\ &= K^T [K^{-T}[\mathbf{t_{rel}}]_\times R_{rel} K^{-1}] K \\ &= [\mathbf{t_{rel}}]_\times R_{rel} \end{aligned}$$

■

> **Q1.4 (10 points)** 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. (*Hint:* draw the relevant vectors)

Looking at below picture,



the problem can be seen that the object has flip the axis before and after the reflection of the mirror. Since the vector of the object and its mirror image is always perpendicular to the mirror (parallel to the normal of the mirror), then we can always define the coordinate system such that the x-axis is flip of the object and its mirror image.

Looking at an object that the x-axis is flip, we can instead flip the camera object into the mirror on the x-axis with some translation $t = [x, 0, 0]^T$ since we define the axis perpendicular to the mirror is x. Therefore, assuming the original camera is at $[I|0]$, and the second to be $[R|t]$, so its original rotation:

$$I = R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

, then the second camera rotation can be seen as

$$R' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The definition of the fundamental matrix is

$$\begin{aligned} F &= K'^{-1}[t]_\times R K^{-1} && (K' = K, \text{since it is the same camera}) \\ &= K'^{-T}[t]_\times \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} K^{-1} \end{aligned}$$

Using the definition of skew symmetric, we have:

$$
\begin{aligned}
-F &= K'^{-T} - [t]_\times \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} K^{-1} \\
&= K'^{-T}([t]_\times)^T \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} K^{-1} && \text{(since } [t]_\times \text{ is skew-symmetric)} \\
&= K'^{-T}([t]_\times)^T (\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix})^T K^{-1} \\
&= K'^{-T}(\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} [t]_\times)^T K^{-1} \\
&= (K'^{-T} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} [t]_\times K^{-1})^T \\
&= (K'^{-T}[t]_\times \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} K^{-1})^T && \text{(since } t \text{ only has } x \text{ component)} \\
&= F^T
\end{aligned}
$$

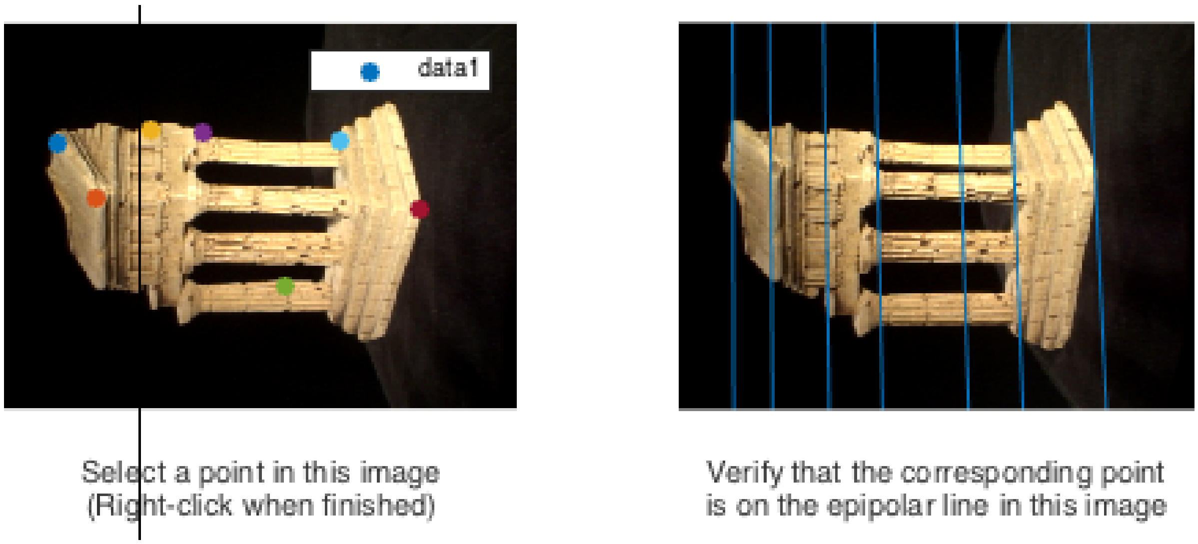
And therefore we have shown that F is skew-symmetric. ■

2.1 Write down recovered F, and a screen shot of the eight-point algorithm

My recovered F is:

```
F=
  -0.000000387339232  -0.000016376057910   0.230251620507768
  -0.000022365870885   0.000000270614880   0.000712358533553
  -0.221134562248653   0.003244227691014  -0.949034123261248
```

The example output is also shown below:



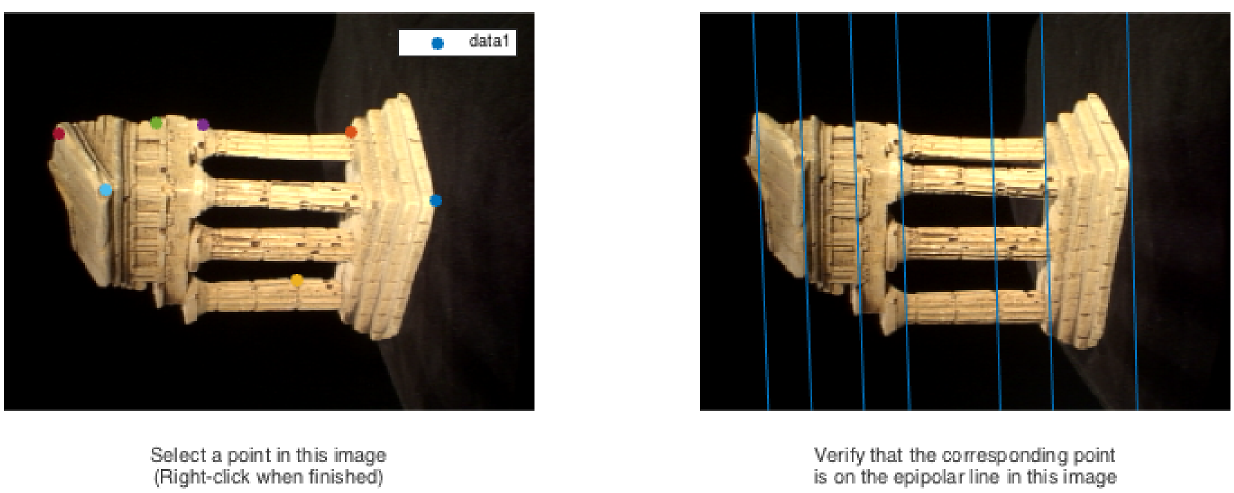
■

2.2 Write down recovered F, and a screen shot of the seven-point algorithm

The corrected recovered F is:

```
F=
   0.000002359180802  -0.000024943546940   0.119260923224233
   0.000003283957910   0.000000763477904  -0.004834507315101
  -0.115202554620373   0.008647145570517  -0.986203369964319
```

The example output is also shown below:



■

2.3 Write down the estimated E using the F from the eight-point algorithm

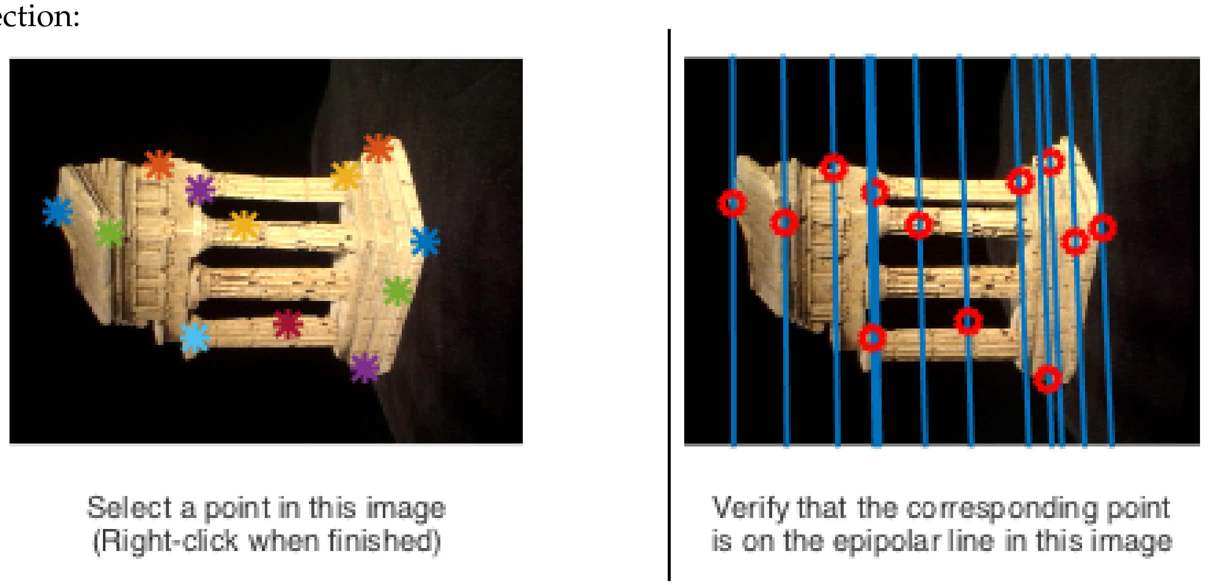
My estimated E is:

```
E=
  1.0e+02 *
  -0.008953796271707  -0.379920999744570   3.437499160013362
  -0.518883364564895   0.006300917861968  -0.091286736435895
  -3.447858595398019  -0.025021333457005  -0.001263668309641
```

■

2.6 Include a screenshot of epipolarMatchGUI with some detected correspondances

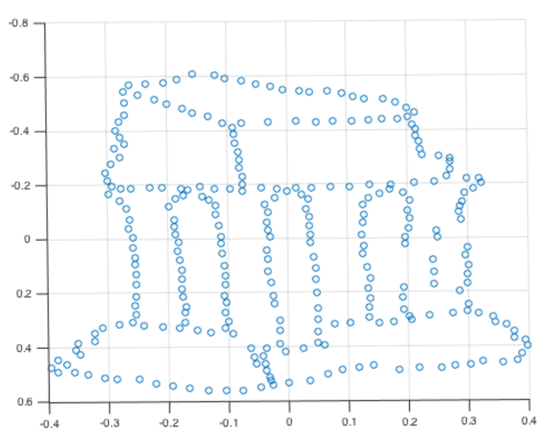
Below is the graph showing the correspondances. Note that ALL of the points are estimated correctly, but most of them are close enough for 3D reconstruction for the next section:

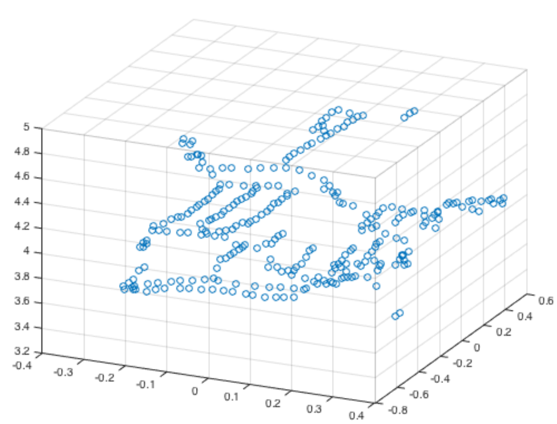


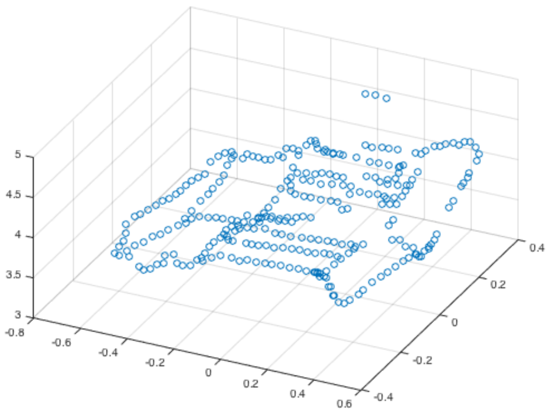
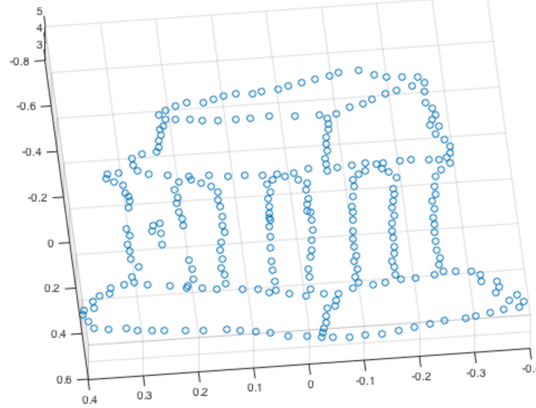
■

2.7 Screenshots of the 3D visualization of the temple.

Since my triangulation method is not optimizing the 3D reprojection, the results for my 3D reconstruction is not perfect. However, the overall shape of the temple and its 3D positions are clearly visible in the below graphs:





■