16720 Computer Vision Assignment4

Yu Fang Chang, Robotics Institute

yufangc@andrew.cmu.edu

Octobor 28, 2015

1 Theory

1.1

Since the camera undergoes fixation, two optical axes intersect at a 3D fixation point which projects onto the principal point in both images. Additionally, the image coordinates are normalized so that the coordinate origin (0,0,1) in homogenous coordinate coincides with the principal point. With fundamental matrix derived in class:

$$x^T F x' = 0, \quad F = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \tag{1}$$

Plug $x = x' = (0, 0, 1)^T$ into (1), we get

$$x^T F x' = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = F_{33} = 0 \tag{2}$$

To satisfy the constraint, $F_{33} = 0$.

1.2

Assume two cameras have the same intrinsic parameter matrix, that is $K = K'$. Since two cameras only differ in pure translation, $R = I_3$, we get the fundamental matrix :

$$F = K'^{-T} E K^{-1} = K^{-T} [t]_\times R K^{-1} = [Kt]_\times = [e']_\times = [e]_\times$$

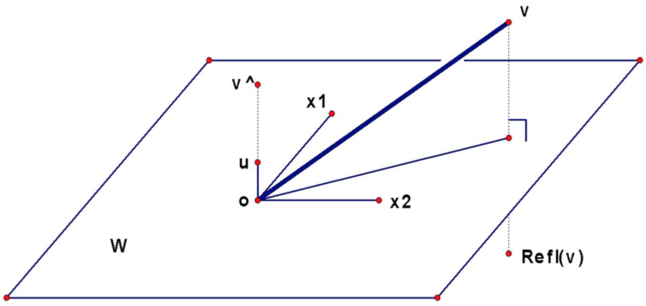
where $[e]_\times$ is the skew-symmetric matrix corresponding to vector e. Since the translation is parallel to the x-axis, $e' = (1, 0, 0)^T$, then

$$F = [e]_\times = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

For corresponding points, $\mathbf{p}, \mathbf{p}'$. Since $\mathbf{p}'^T F \mathbf{p} = 0$,

$$\mathbf{p}'^T F \mathbf{p} = \begin{bmatrix} x'_i & y'_i & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0$$

leads to $y'_i = y_i$. That is, every correspondences in image 2 will have the same y' value as y. Epipolar lines, in this case, can be drawn as rasters parallel to x-axis.



1.3

Assume W a plane mirror passing through the origin in the 3D space with normal unit vector u. We are deriving the point reflecting a vector v across this plane. The projection of v along the line through u is then given by:

$$\hat{v} = Projection_u(v) = u(u^T u)^{-1} u^T v$$

and since u is a unit vector, $u^T u = u \cdot u = 1$, we get:

$$\hat{v} = uu^T v$$

$Ref_W(v)$ is the reflection point of v across W, which lies in the opposite side of W the same distance from W to v and also has the same projection into W as v, then leads to:

$$\begin{aligned} v - Ref_W(v) &= 2\hat{v} = 2uu^T v \\ \Rightarrow Ref_W(v) &= v - 2uu^T v = (I - 2uu^T)v \end{aligned}$$

Suppose $\mathbf{P}_1$ is a 3D point in homogeneous coordinate, we can write its projection in the first image as a 2D points $\mathbf{p}_1 = \mathbf{K}_1[\mathbf{I}_3|0]\mathbf{P}_1$. Given a plane mirror in the 3D space as $\mathbf{n}x + d = 0$ and the above derivation, we get the mirrored 3D points as

$$\mathbf{P}_2 = (\mathbf{I}_3 - 2\mathbf{n}\mathbf{n}^\mathbf{T})\mathbf{P}_1 + \mathbf{n}(-2d)$$

Let the mirrored point be a projection on the second image, we can represent its 2D projection on the image as $\mathbf{p}_2 = \mathbf{K}_2[\mathbf{I}_3|0]\mathbf{P}_2 = \mathbf{K}_2[\mathbf{R}|\mathbf{t}]\mathbf{P}_1$, where the rotation matrix $\mathbf{R} = \mathbf{I}_3 - 2\mathbf{n}\mathbf{n}^\mathbf{T}$, and transformation vector $\mathbf{t} = -2d\mathbf{n}$. We assume $\mathbf{K}_1 = \mathbf{K}_2 = \mathbf{I}_3$, we can write our essential matrix as $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$ and fundamental matrix $\mathbf{F} = \mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$. Since $[\mathbf{t}]_\times$ is a skew-symmetric matrix and $\mathbf{R}$ is a symmetric matrix, we prove that the fundamental matrix $\mathbf{F}$ is skew-symmetric.

2 Fundamental matrix estimation

2.1 The Eight Point Algorithm

I applied all provided correspondense in data/some_corresp.mat to get the fundamental matrix. And the visualization please refer to Figure 1.

$$F = \begin{pmatrix} -1.311 \cdot 10^{-9} & -5.921 \cdot 10^{-8} & -0.00108 \\ -1.308 \cdot 10^{-7} & 3.564 \cdot 10^{-9} & 3.048 \cdot 10^{-5} \\ 0.001125 & -1.65 \cdot 10^{-5} & -0.004166 \end{pmatrix}$$

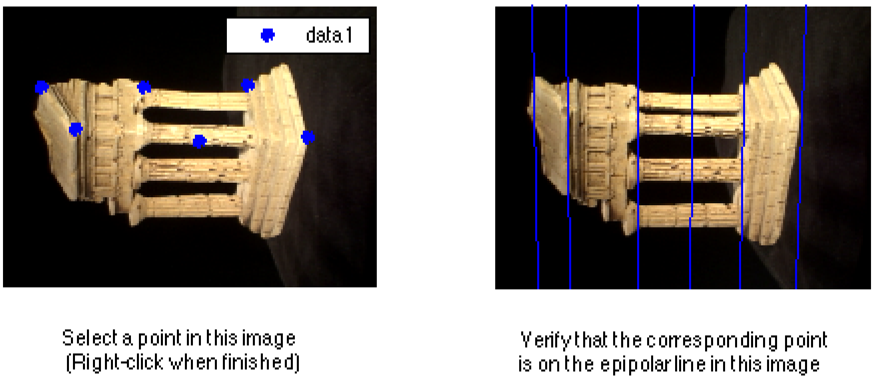


Figure 1: An output of displayEpipolarF for 8-point algorithm.

2.2 The Seven Point Algorithm

I have tried both manually select correspondense with cpselect and randomly pick the correspondense in data/some_corresp.mat. Since I found that it is quite sensitive to the noise, I have to manually select the point with higher accuruacy. Following is the fundamental matrix with randomly select 7 points in data/some_corresp.mat. Please refer to Figure 2 and Figure 3 corresponding to manually selected and randomly selected correspondense.

$$F = \begin{pmatrix} -2.731 \cdot 10^{-22} & 9.541 \cdot 10^{-7} & -0.0001468 \\ -9.541 \cdot 10^{-7} & 3.23 \cdot 10^{-22} & -0.000968 \\ 0.0001468 & 0.000968 & 8.61 \cdot 10^{-19} \end{pmatrix}$$

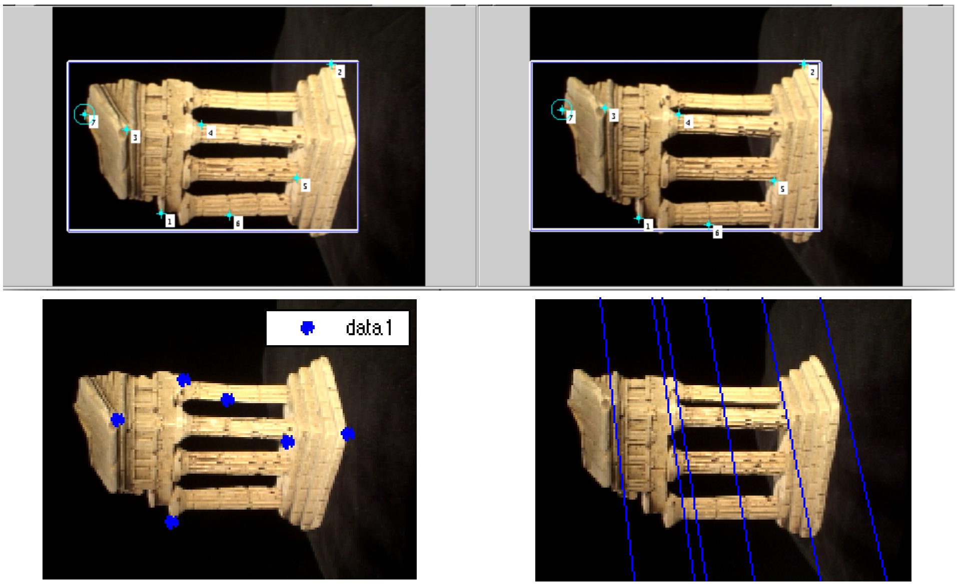


Figure 2: An output of manually select correspondence for 7-point algorithm.

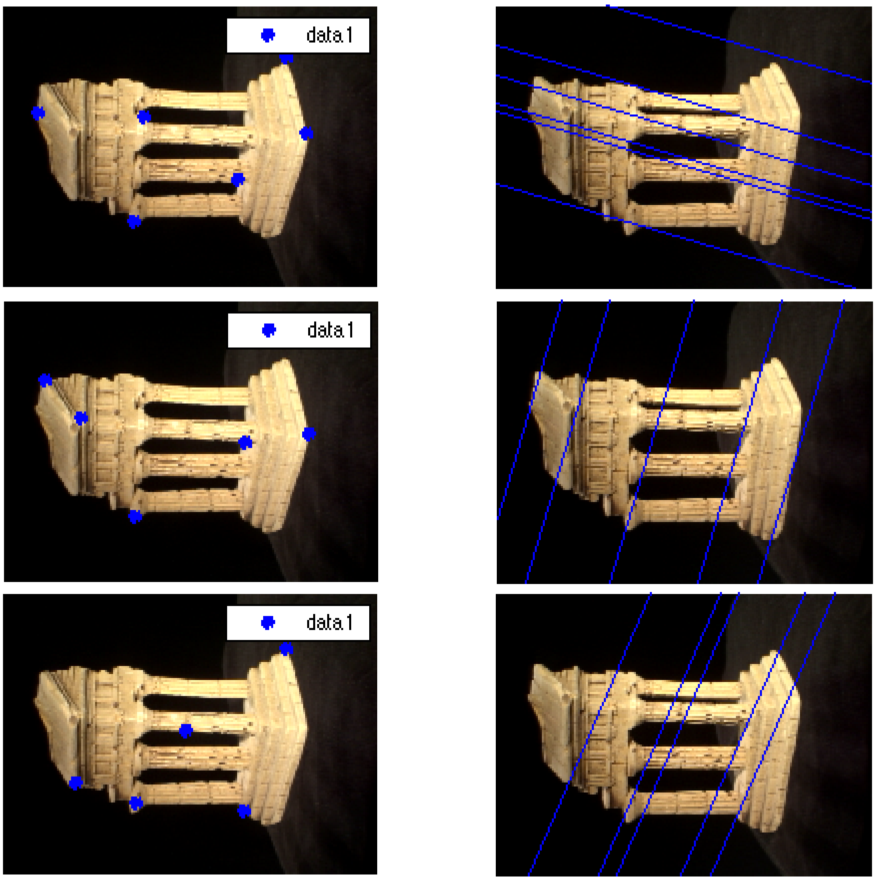


Figure 3: An output of randomly select from displayEpipolarF and get 3 valid F for 7-point algorithm.

2.3

Our estimated E:

$$E = \begin{pmatrix} -0.00303 & -0.1374 & -1.665 \\ -0.3035 & 0.008299 & -0.0125 \\ 1.66 & -0.05115 & -0.001324 \end{pmatrix}$$

2.X We apply RANSAC to remove the outliers in order to get a better result with nosie. Without RANSAC, we get the fundamental matrix:

$$F_{withoutRansac} = \begin{pmatrix} 5.318 \cdot 10^{-7} & 1.279 \cdot 10^{-6} & -0.0004 \\ -8.488 \cdot 10^{-7} & -1.459 \cdot 10^{-6} & 0.0005712 \\ 3.32 \cdot 10^{-5} & 2.97 \cdot 10^{-5} & -0.01918 \end{pmatrix}$$

To implement RANSAC, I first randomly select 7 points to get a temporary fundamental matrix with 7-points algorithm. I take the smallest distance between the epipolar line to the correspondence in image 2, and vice versa, as considering whether they are inliers or not. I finally get the approximately 114 points as inliers, and apply them to 8-points algorithm to get the fundamental matrix:

$$F_{withRansac} = \begin{pmatrix} 1.663 \cdot 10^{-7} & 1.649 \cdot 10^{-6} & -0.0002744 \\ -1.611 \cdot 10^{-6} & -4.53 \cdot 10^{-7} & 0.0002405 \\ 0.0001854 & -4.578 \cdot 10^{-5} & -0.01242 \end{pmatrix}$$

and Figure 4 shows the result of applying eight point algorithm directly from the noise data and the result after implementing RANSAC.

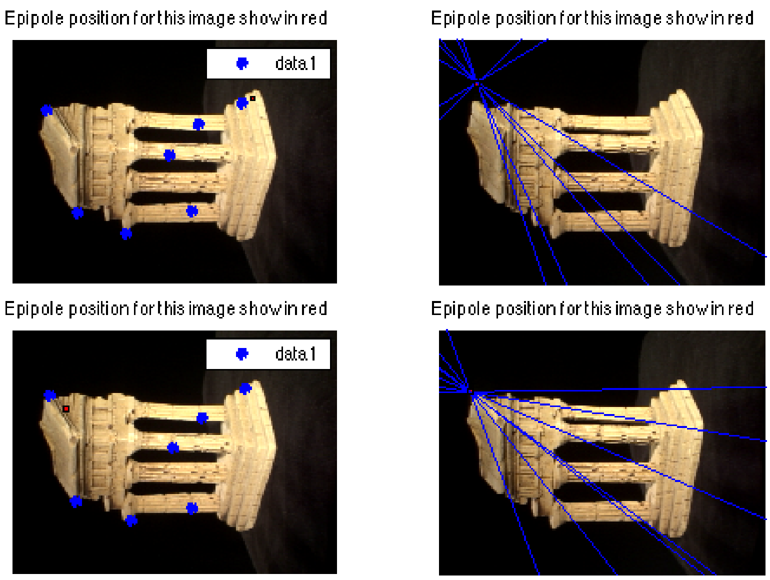


Figure 4: An output of displayEpipolarF without and with RANSAC.

2.4

2.5

Since we will get 4 possible M2s, we would like to find the best M2 among M2s. First we apply each M2 to triangulate the 3D points. To ensure the valid M2, the 3D points should be in front of both cameras and thus we check the Z coordinates whether they are positive. Additionally, since we are finding the best M2, we picked the one which has the least error as we transfer back to 2D. Following is the best M2 I get:

$$M2 = \begin{pmatrix} 0.9993 & -0.03818 & 0.001255 & -0.007655 \\ 0.03719 & 0.9649 & -0.2601 & 1.0 \\ 0.00872 & 0.26 & 0.9656 & 0.1828 \end{pmatrix}$$

2.6

Instead of searching for the matching point at every possible location in im2, we can use F and simply search over the set of pixels that lie along the epipolar line. I tried different window size [5, 11, 21] and also apply gaussian weighting window to focus on the center of the point. Additionally, to get more reasonable index, I transverse the line vertically or horinzontally depending on the slope. Figure 5 shows the result with different window size.

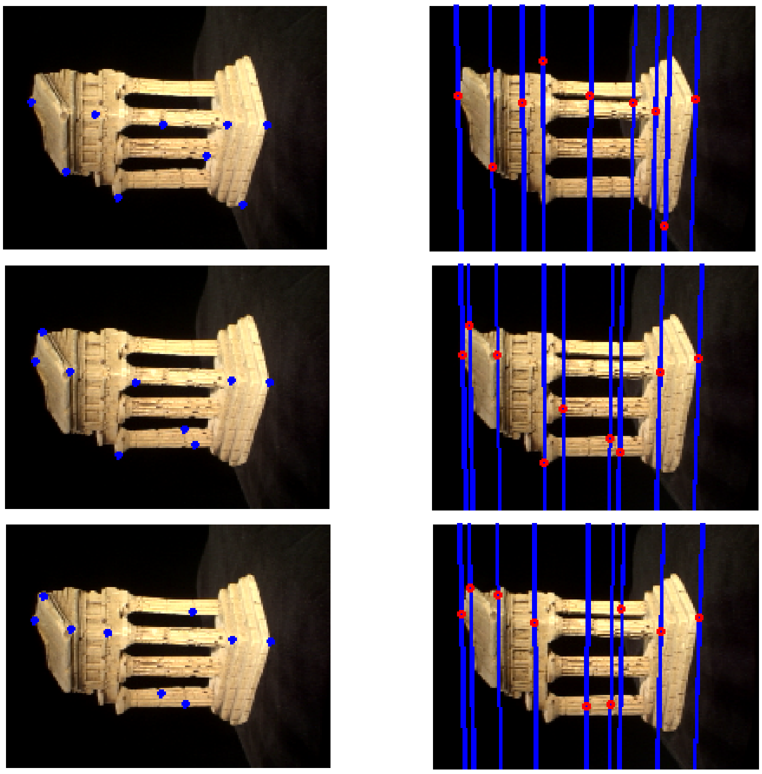


Figure 5: An output of epipolar line with window size [5 11 21].

2.7

Figure 6 shows the results of the 3D points we constructed. Our discovered M2 is

$$\begin{pmatrix} 0.9993 & 0.03719 & 0.00872 & -0.03074 \\ -0.03818 & 0.9649 & 0.26 & -1.0 \\ 0.001255 & -0.2601 & 0.9656 & 0.08256 \end{pmatrix}$$

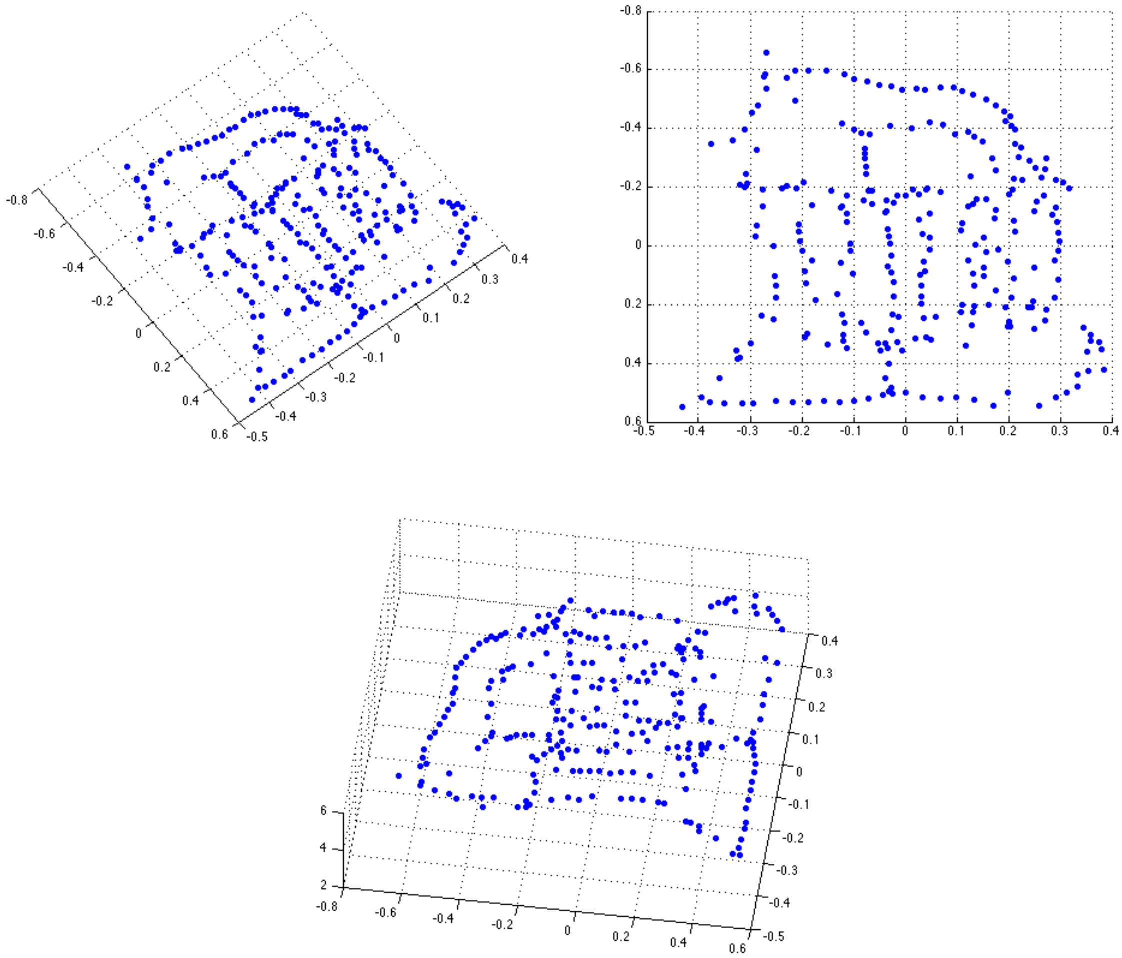


Figure 6: Several views of the 3D visualization we constructed.