16-720 Homework 4 (ONE late day used)

Changsheng Shen (Bobby)
changshs@andrew.cmu.edu

11/13/2018

Q1.1:

Since the both principle points of two image planes coincide with the coordinate origin (0, 0), then the projected point of P in the two image planes are $p_1 = [0, 0, 1]^T$ and $p_2 = [0, 0, 1]^T$.

Denote the fundamental matrix as:

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

Then according to the property of fundamental matrix:

$$p_1^T F p_2 = 0$$

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

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

$$\Longrightarrow F_{33} = 0$$

Q1.2: Since the two cameras differ by a pure translation, with the world frame aligned with the first camera frame, denote the camera projection matrices as:

$$P_1 = K[I|\mathbf{0}]$$

$$P_2 = K[I|\mathbf{t}]$$

Then we can derive:

$$F = [e']_\times P_2 P_1^+ = [e']_\times K K^{-1} = [e']_\times = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The epipolar line l' is:

$$l' = F\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ y \end{bmatrix}$$

and since:

$$x'^T F x = 0$$
$$\implies \begin{bmatrix} x' & y' & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$
$$\implies \begin{bmatrix} x' & y' & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ y \end{bmatrix} = 0$$
$$\implies y' = y$$

Therefore, the epipolar lines in the two cameras are also parallel to the x-axis.

Q1.3:

The homogeneous transformation matrices from the world fixed frame (frame 0) to the frame at timestamp i and j are:

$$H_i = \begin{bmatrix} R_i & t_i \\ 0 & 1 \end{bmatrix}$$

and

$$H_j = \begin{bmatrix} R_j & t_j \\ 0 & 1 \end{bmatrix}$$

Then the relative transformation matrix from frame i to frame j is:

$$H_{rel} = H_j H_i^{-1} = \begin{bmatrix} R_j & t_j \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_i^{-1} & -R_i^{-1} t_i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_j R_i^{-1} & -R_j R_i^{-1} t_i + t_j \\ 0 & 1 \end{bmatrix}$$

Therefore the effective rotation and translation between two frames at different timestamps, specifically from frame i to frame j is:

$$R_{rel} = R_j R_i^{-1}$$

and

$$t_{rel} = -R_j R_i^{-1} t_i + t_j$$

Since the essential matrix can be expressed as:

$$E = [t_{rel}]_\times R_{rel}$$

and also

$$E = K^{-1} F K$$

Then we can express the fundamental matrix as:

$$F = KEK^{-1} = K[t_{rel}]_\times R_{rel} K^{-1}$$

Q1.4:

Denote the camera center by C, the object being viewed as a point X, the mirrored camera center as C' and the mirrored object as a point X'.

Denote the projected points of X in cameras C and C' as x_1 and x_2, respectively. Similarly, denote the projected points of X' as x_1' and x_2'.

Then from the property of fundamental matrix F we have:

$$x_1^T F x_2 = 0$$
$$x_2' F^T x_1' = 0$$

and since x_1, x_1' and x_2, x_2' are symmetric relative to the mirror, and $x_2' F^T x_1' = 0$, we can infer that: $x_1^T F^T x_2 = 0$

Then adding $x_1^T F x_2 = 0$ and $x_1^T F^T x_2 = 0$ together, we get:

$$x_1^T (F + F^T) x_2 = 0$$
$$\implies F + F^T = 0$$
$$\implies F = -F^T$$

Therefore, we can see that F is a skew-symmetric matrix.

Q2.1:

$$F = \begin{bmatrix} 9.80213866e-10 & -1.32271663e-07 & 1.12586847e-03 \\ -5.72416248e-08 & 2.97011941e-09 & -1.17899320e-05 \\ -1.08270296e-03 & 3.05098538e-05 & -4.46974798e-03 \end{bmatrix}$$

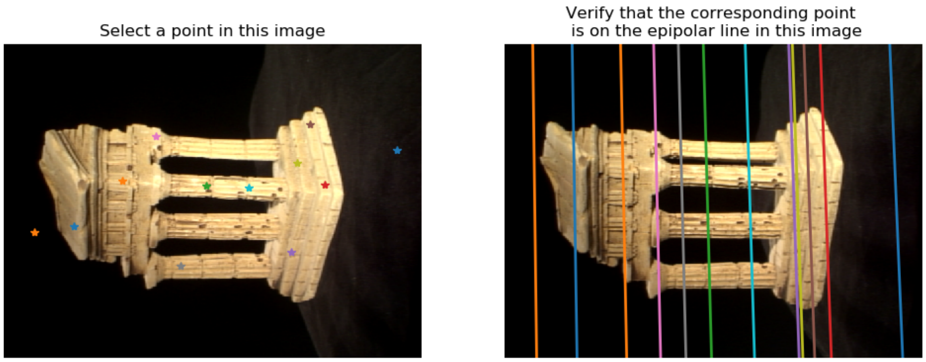

Figure 1: Eight point algorithm

Q2.2:

$$F = \begin{bmatrix} 1.45292584e-07 & -2.62788930e-06 & 6.38443235e-04 \\ 2.80602917e-06 & -1.72122881e-06 & 6.68836902e-04 \\ -6.65138667e-04 & 5.76404051e-06 & -6.28743900e-02 \end{bmatrix}$$

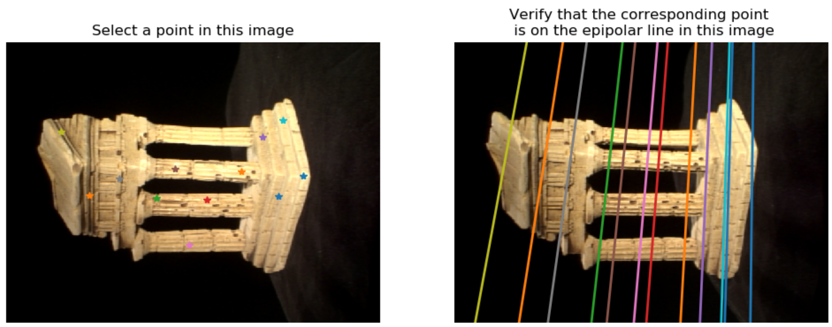

Figure 2: Seven point algorithm

Q3.1:

$$E = \begin{bmatrix} 0.00226588 & -0.3068674 & 1.66257398 \\ -0.13279933 & 0.00691554 & -0.04327756 \\ -1.66717617 & -0.0133444257 & -0.000672047195 \end{bmatrix}$$

Q3.2: Denote the camera matrices as C_1 and C_2, intrinsics as K_1 and K_2, and extrinsics as M_1 and M_2, then we have:

$$C_1 = K_1 M_1$$

$$C_2 = K_2 M_2$$

Then denote C_1 and C_2 as:

$$C_1 = \begin{bmatrix} -\boldsymbol{c_{11}}- \\ -\boldsymbol{c_{12}}- \\ -\boldsymbol{c_{13}}- \end{bmatrix}$$

$$C_2 = \begin{bmatrix} -\boldsymbol{c_{21}}- \\ -\boldsymbol{c_{22}}- \\ -\boldsymbol{c_{23}}- \end{bmatrix}$$

where c_{ij} is the j-th row of C_i

Denote a pair of corresponding points on two images as $p_{i1} = (x_{i1}, y_{i1})$ and $p_{i2} = (x_{i2}, y_{i2})$, respectively, then we can write out the expression of A_i as:

$$A_i = \begin{bmatrix} x_{i1} \cdot \boldsymbol{c_{13}} - \boldsymbol{c_{11}} \\ y_{i1} \cdot \boldsymbol{c_{13}} - \boldsymbol{c_{12}} \\ x_{i2} \cdot \boldsymbol{c_{23}} - \boldsymbol{c_{21}} \\ y_{i2} \cdot \boldsymbol{c_{23}} - \boldsymbol{c_{22}} \end{bmatrix}$$

Q4.1:

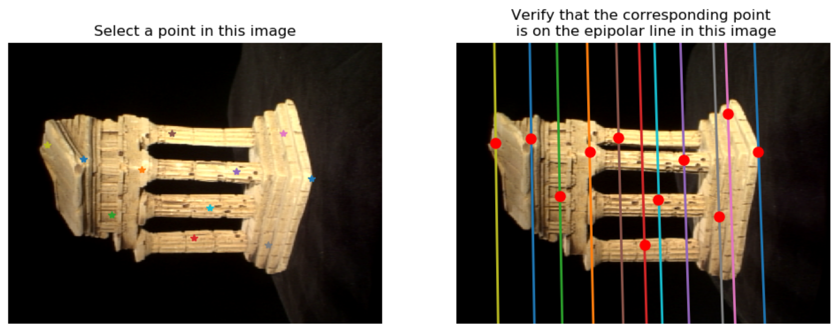

Figure 3: Screenshot of epipolarMatchGUI with detected correspondence

Q4.2:

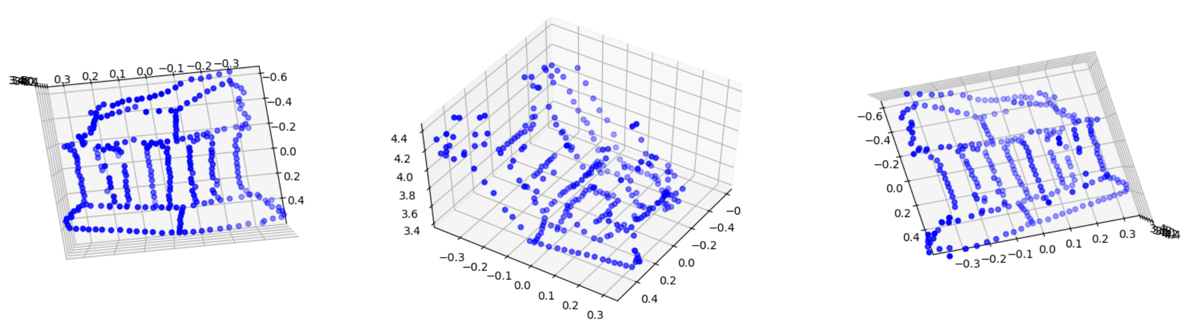

Figure 4: 3D visualization of the reconstructed temple from different angles

Q5.1:

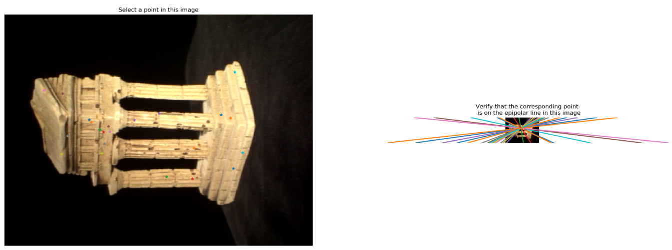

Figure 5: Original result of eightpoint algorithm when running on noisy datasets

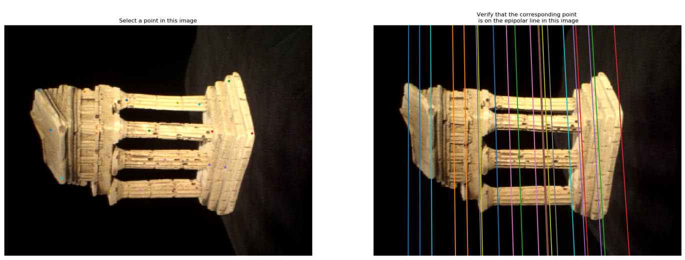

Figure 6: Result of RANSAC

For the computed F_i in the i-th iteration of RANSAC, denote a pair of corresponding points as X_{i1} and X_{i2}, the error is calculated as:

$$error_i = X_{i2}^T \cdot F_i \cdot X_{i1}$$

and if $error_i$ is smaller than a threshold (i.e. 0.001), we consider this pair of points as an inlier.

Q5.3:

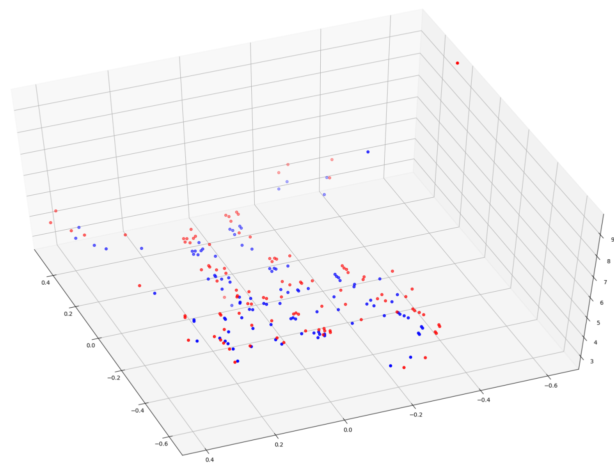

Figure 7: Blue points are the original 3D points, and red points are the optimized points.

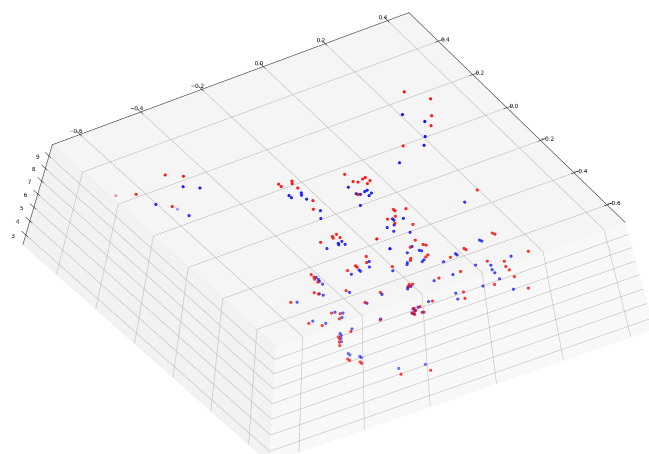

Figure 8: Visualization in another perspective

Reprojection error before bundleAdjustment: 7788.491314034161

Reprojection error after bundleAdjustment: 619.0129714597791