16-720 Homework 4: 3D Reconstruction

Austin Windham

November 14, 2023

1 Theory

Q 1.1

There is a scenario where two cameras fixate on point $\mathbf{P}$ such that their principle axes go through the two optical center points $\mathbf{x_1}$ and $\mathbf{x_2}$, so the two points both equal $[0,0,1]^{\mathbf{T}}$. So, with the fundamental matrix $\mathbf{F}$, we then have the equation below

$$\mathbf{x_1}^{\mathbf{T}}\mathbf{F}\mathbf{x_2} = 0 \tag{1}$$

When we expand $\mathbf{F}$ and multiply out the three matrices, we get the following results below showing $\mathbf{F_{33}} = 0$.

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \tag{2}$$

$$F_{33} = 0 \tag{3}$$

Q 1.2

To show that the two epipolar lines in the cameras are parallel to the x-axis when the second camera differs by pure translation parallel to the x-axis, we first need to find $\mathbf{E} = \mathbf{tR}$, where t and R are the translation and rotation matrices respectively. Since there is no rotation and translation is only in the x-direction, the matrices are given in the format below.

$$t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

So, then $\mathbf{E}$ is given below.

$$E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \tag{6}$$

Now, lets say there is $x_1 = \begin{bmatrix} a_1 & b_1 & 1 \end{bmatrix}^T$ and $x_2 = \begin{bmatrix} a_2 & b_2 & 1 \end{bmatrix}^T$, then the epipolar lines $\mathbf{l_1}$ and $\mathbf{l_2}$ can be derived as shown below.

$$\mathbf{l_1^T} = x_2^T E \tag{7}$$

$$= \begin{bmatrix} a_2 & b_2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & t_x & -b_2 t_x \end{bmatrix} \tag{8}$$

$$\mathbf{l_2^T} = x_1^T E^T \tag{9}$$

$$= \begin{bmatrix} a_1 & b_1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -t_x & b_1 t_x \end{bmatrix} \tag{10}$$

So, the equations of the epipolar lines are

$$t_x y - b_2 t_x = 0 \tag{11}$$

$$-t_x y + b_1 t_x = 0 \tag{12}$$

And, since both equations do not have an x component, they are both parallel to the x-axis.

Q 1.3

In order to find the effective rotation and effective translation between two time frames at different time stamps, and then later find the essential and fundamental matrices, lets say that there is a point in the 3D world given by $[u, v, w]$ and that frames at time stamps 1 and 2 are given by $[x_1.y_1]$ and $[x_1.y_1]$. We can then get the following equation below relating the point in frame 1 to the 3D world.

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K(R_1 \begin{bmatrix} u \\ v \\ w \end{bmatrix} + t_1) \tag{13}$$

Then some basic algebraic tactics to get $[u, v, w]$ on the side by itself gives us the results below.

$$\begin{aligned} \begin{bmatrix} u \\ v \\ w \end{bmatrix} &= R_1^{-1}(K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - t_1) && (14) \\ &= R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - R_1^T t_1 && (15) \end{aligned}$$

Now, doing the same for time stamp 2, but substituting in the value we just found for $[u, v, w]$, we obtain the following results:

$$\begin{aligned} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} &= K(R_2 \begin{bmatrix} u \\ v \\ w \end{bmatrix} + t_2) && (16) \\ &= K(R_2(R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - R_1^T t_1) + t_2) && (17) \\ &= KR_2R_1^T K^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - KR_2R_1^T t_1 + Kt_2 && (18) \end{aligned}$$

From the above result, the equations for the effective rotation and translation between two different time stamps can be drawn.

$$\begin{aligned} R_{rel} &= KR_2R_1^T K^{-1} && (19) \\ t_{rel} &= -KR_2R_1^T t_1 + Kt_2 && (20) \end{aligned}$$

With these results, the essential and fundamental matrices were derived below.

$$\begin{aligned} &E = t_{rel} \times R_{rel} && (21) \\ &F = \left(K^{-1}\right)^T EK^{-1} = \left(K^{-1}\right)^T (t_{rel} \times R_{rel})\, K^{-1} && (22) \end{aligned}$$

Q 1.4

If a camera views a flat object and views the object's reflection on a planar mirror then this can be modeled as two camera frames where there is just translation and no rotation between the frames. This means that the rotation matrix is an identity matrix and the shift is given by a variable **d** in the three directions, and so we can use the above fundamental matrix equation already derived to solve for F.

$$R_{rel} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{23}$$

$$t_{rel} = \begin{bmatrix} d_x & d_y & d_z \end{bmatrix} \tag{24}$$

Substituting into previous equation,

$$F = \left(K^{-1}\right)^T (t_{rel} \times R_{rel})\, K^{-1} \tag{25}$$

$$= \left(K^{-1}\right)^T \begin{bmatrix} 0 & d_z & -d_y \\ -d_z & 0 & d_x \\ d_y & -d_x & 0 \end{bmatrix} K^{-1} \tag{26}$$

Since a skew-symmetric matrix is given by $F^T = -F$, we can see that F above is a skew-symmetric matrix.

2 Fundamental Matrix Estimation

Q 2.1 Eight Point Algorithm

The calculated fundamental matrix is below along with an image showing points in image 1 and the corresponding epipolar lines in image 2.

$$\mathbf{F} = \begin{bmatrix} 9.78833281e-10 & -1.32135929e-07 & 1.12585666e-03 \\ -5.73843315e-08 & 2.96800276e-09 & -1.17611996e-05 \\ -1.08269003e-03 & 3.04846703e-05 & -4.47032655e-03 \end{bmatrix} \tag{27}$$

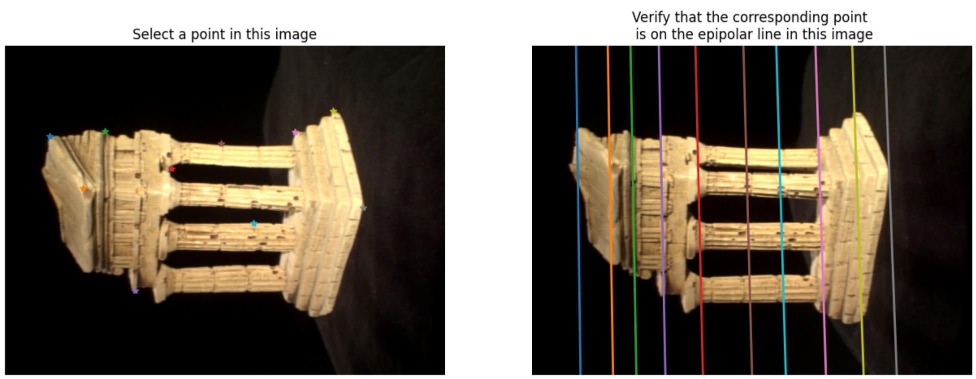

Figure 1: Epipolar Lines from Fundamental Matrix Implementation

3 Metric Reconstruction

Q 3.1

The essential matrix is given below. The report did not say specifically to include the matrix, but I showed it since we did the same for the fundamental matrix.

$$\mathbf{E} = \begin{bmatrix} 2.26268683e-03 & -3.06552495e-01 & 1.66260633e+00 \\ -1.33130407e-01 & 6.91061098e-03 & -4.33003420e-02 \\ -1.66721070e+00 & -1.33210351e-02 & -6.72186431e-04 \end{bmatrix} \tag{28}$$

Q 3.2

In order to solve for the 3D points of the object in the image given by $\mathbf{w_i}$, we need to find matrix $\mathbf{A_i}$ such that $\mathbf{A_i}\mathbf{w_i} = 0$ And, with camera matrices $\mathbf{C_1}$ and $\mathbf{C_2}$ as $4x3$ matrices, $\mathbf{w_i}$ as $4x1$ homogeneous 3D coordinates, and $\mathbf{x_i}$ as $3x1$ 2D homogeneous coordinates, the equations below can be obtained.

$$\mathbf{C}_1\mathbf{w}_i = \mathbf{x}_{1i} \tag{29}$$

$$\mathbf{C}_1 \begin{bmatrix} u_i \\ v_i \\ w_i \\ 1 \end{bmatrix} = \begin{bmatrix} x_{1i} \\ y_{1i} \\ 1 \end{bmatrix} \tag{30}$$

$$\tag{31}$$

$$\mathbf{C}_2\mathbf{w}_i = \mathbf{x}_{2i} \tag{32}$$

$$\mathbf{C}_2 \begin{bmatrix} u_i \\ v_i \\ w_i \\ 1 \end{bmatrix} = \begin{bmatrix} x_{2i} \\ y_{2i} \\ 1 \end{bmatrix} \tag{33}$$

From this, we can see that when the first row, second row, and third row of the camera matrices are multiplied by the 3D coordinates $\mathbf{w_i}$ the corresponding results are the x-coordinate, y-coordinate, and 1. Knowing these relationships, mainly that the 3rd row multiplied by the 3D coordinates equals 1, allows us to create the equation below.

$$\begin{bmatrix} x_{1i} - x_{1i} \\ y_{1i} - y_{1i} \\ x_{2i} - x_{2i} \\ y_{2i} - y_{2i} \end{bmatrix} = \begin{bmatrix} x_{1i}\mathbf{C}_{1,row3}\mathbf{w_i} - \mathbf{C}_{1,row1}\mathbf{w_i} \\ y_{1i}\mathbf{C}_{1,row3}\mathbf{w_i} - \mathbf{C}_{1,row2}\mathbf{w_i} \\ x_{2i}\mathbf{C}_{2,row3}\mathbf{w_i} - \mathbf{C}_{2,row1}\mathbf{w_i} \\ y_{2i}\mathbf{C}_{2,row3}\mathbf{w_i} - \mathbf{C}_{2,row2}\mathbf{w_i} \end{bmatrix} = \begin{bmatrix} x_{1i}\mathbf{C}_{1,row3} - \mathbf{C}_{1,row1} \\ y_{1i}\mathbf{C}_{1,row3} - \mathbf{C}_{1,row2} \\ x_{2i}\mathbf{C}_{2,row3} - \mathbf{C}_{2,row1} \\ y_{2i}\mathbf{C}_{2,row3} - \mathbf{C}_{2,row2} \end{bmatrix} \mathbf{w_i} = 0 \tag{34}$$

Thus our equation for $\mathbf{A_i}$ is given below.

$$\mathbf{A_i} = \begin{bmatrix} x_{1i}\mathbf{C}_{1,row3} - \mathbf{C}_{1,row1} \\ y_{1i}\mathbf{C}_{1,row3} - \mathbf{C}_{1,row2} \\ x_{2i}\mathbf{C}_{2,row3} - \mathbf{C}_{2,row1} \\ y_{2i}\mathbf{C}_{2,row3} - \mathbf{C}_{2,row2} \end{bmatrix} \tag{35}$$

Q 3.3

Code is included in the submission.

4 3D Visualization

Q4.1

I added a save command within the helper function script used to display these points and lines in order to save the points in both images. The save command line is commented out in the submission in the helper script.

As you can see below, the function was able to find the correspondences on the epipiolar lines.

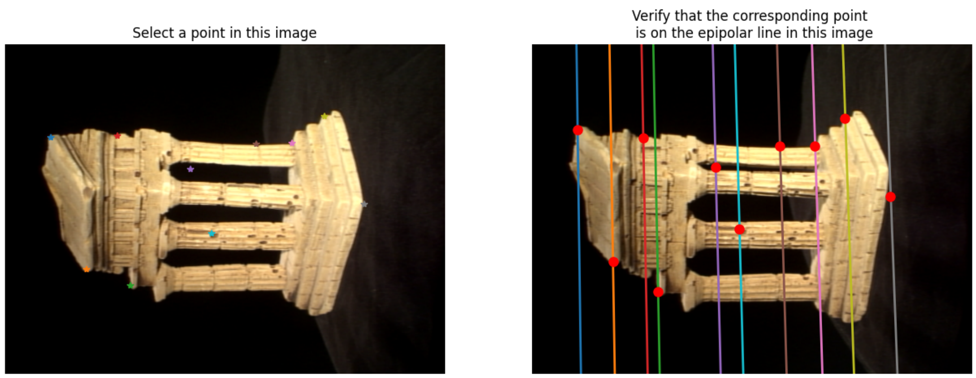

Figure 2: Epipolar Correspondences Plot

Q4.2

The results for the 3D visualization are below. As you can see, the temple image reconstruction was successful.

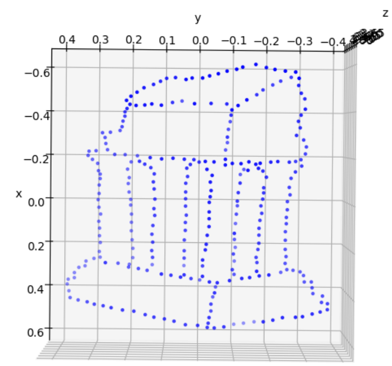

Figure 3: 3D Visualization View 1

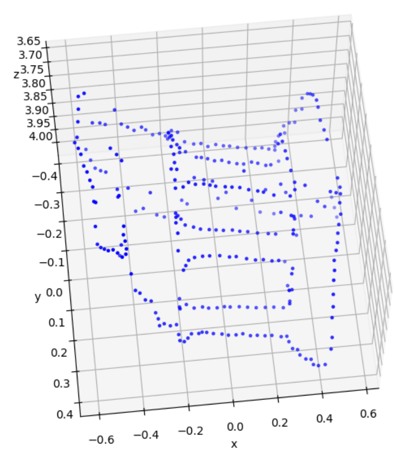

Figure 4: 3D Visualization View 2

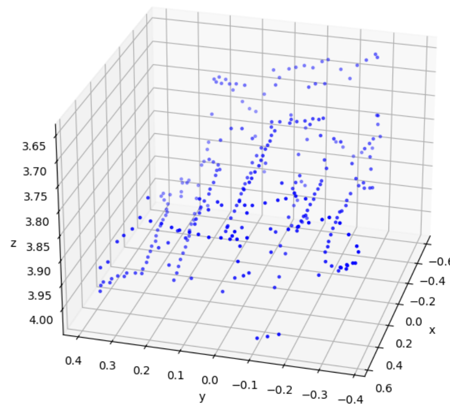

Figure 5: 3D Visualization View 3

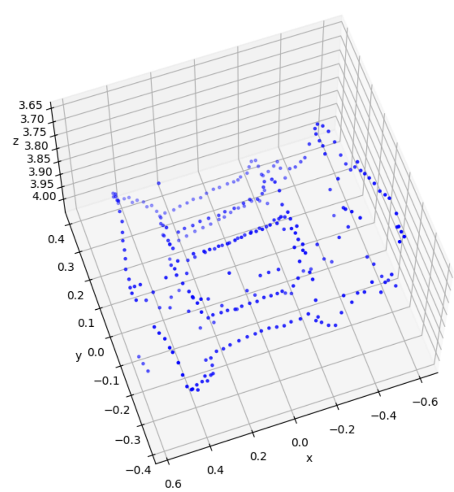

Figure 6: 3D Visualization View 4

5 Bundle Adjustment (Extra Credit)

Q 5.1

Results from using the noisy correspondences are below. As you can see, the RANSAC method results are far better. The no RANSAC method drew epipolar lines that were very incorrect.

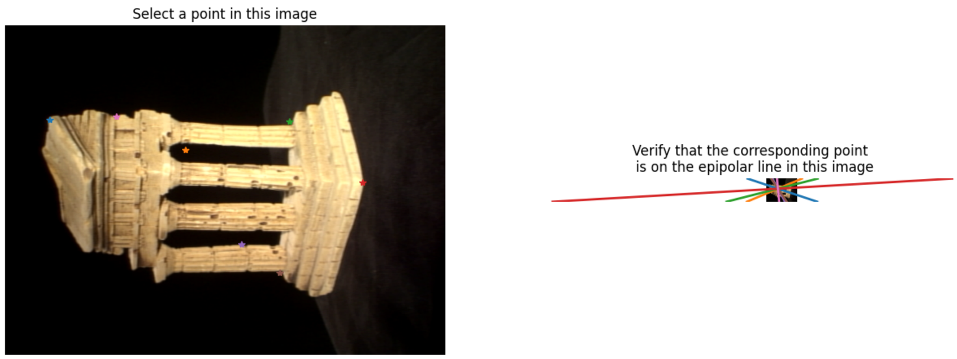

Figure 7: Epipolar Lines without RANSAC

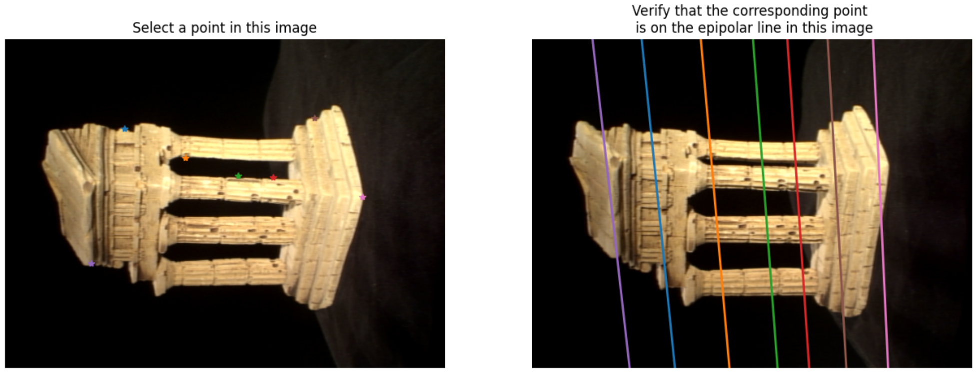

Figure 8: Epipolar Lines with RANSAC

From before, we mentioned the relationship between the fundamental matrix and a pair of matching coordinates between two images that is given below.

$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 = 0 \tag{36}$$

With $\mathbf{\hat{x}_{2i}}$ as our prediction for each set of coordinates in image 2, we can then calculate the error by taking the absolute value of the above equation for each pair of matching points.

$$error = |\hat{\mathbf{x}}_{2i}^T \mathbf{F} \mathbf{x}_{1i}| \tag{37}$$

The inliers are then considered the matches at which the error is less than the tolerance value. I used a tolerance value of 0.8 and 350 iterations for my RANSAC method.

Increasing the tolerance values means that more points are accepted even when they may be inaccurate, so the algorithm may not find the best solution. But, if the tolerance value is too low, then there may not be a fundamental matrix that finds enough inliers to determine if it is even suitable. Increasing the number of iterations means that the algorithm gets more chances to find a better fundamental matrix, so more iterations should theoretically mean a better solution, but increasing the number of iterations also increases the compute time. I largely determined if the fit was suitable enough given the parameters by qualitative inspection of the epipolar lines. Sometimes even when I set the number of iterations to greater than 5000, I did not consistently get near perfect fits. This indicates that an algorithm that could help eliminate some noise would be beneficial for these given points.

Q 5.2

Rodriguez and Inverse Rodriguez functions were created and verified that they perform the correct calculations by comparing the results to those produced from using MATLAB's rotvec2mat3d function in its Computer Vision Toolbox. Code is included in the assignment submission.

Q5.3

I did not complete this extra credit problem.