16-720 HW4 3D Reconstruction

Xiaoyu Bai (xiaoyub)

Q1.1

Pick correspondence points X=[0 0 1]^T in C1 and X'=[0 0 1]^T in C2. The fundamental matrix F should satisfy that X'^{T}·F·X=0, which means $[0\ 0\ 1]\cdot\begin{bmatrix}F_{11} & F_{12} & F_{12}\\ F_{21} & F_{22} & F_{23}\\ F_{31} & F_{32} & F_{33}\end{bmatrix}\cdot\begin{bmatrix}0\\0\\1\end{bmatrix}=F_{33}=0$.

Q1.2

Let $P=K[I|0]\ and\ P'=K[I|t]$ where t denotes the translation along x-axis. Therefore $P^+=\begin{pmatrix}K^{-1}\\0^T\end{pmatrix}$ and fundamental matrix $F=[e']_\times P'P^+=[e']_\times KIK^{-1}=[e']_\times$ and $e'=\begin{pmatrix}1\\0\\0\end{pmatrix}$. So $F=\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix}$.

Assume $x=\begin{pmatrix}x\\y\\1\end{pmatrix}$ and $x'=\begin{pmatrix}x'\\y'\\1\end{pmatrix}$ then we have $(x'\quad y'\quad 1)\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix}\begin{pmatrix}x\\y\\1\end{pmatrix}=y'-y=0$.
Therefore epipolar lines are parallel to x-axis.

Q1.3

Under world coordinates,

$$\begin{cases} p_i=\begin{pmatrix}R_i & t_i\\0 & 1\end{pmatrix}\cdot P\\ p_{i+1}=\begin{pmatrix}R_{i+1} & t_{i+1}\\0 & 1\end{pmatrix}\cdot P\end{cases}$$

So $p_{i+1}=\begin{pmatrix}R_{i+1} & t_{i+1}\\0 & 1\end{pmatrix}\begin{pmatrix}R_i & t_i\\0 & 1\end{pmatrix}^{-1}\cdot p_i=\begin{pmatrix}R_{rel} & t_{rel}\\0 & 1\end{pmatrix}\cdot p_i$, which means if we say $M=\begin{pmatrix}R_{i+1} & t_{i+1}\\0 & 1\end{pmatrix}$, *M(1:3,1:3)=$R_{rel}$* and *M(1:3,4)=$t_{rel}$* .

Or if we do it in Cartesian coordinates,

$\begin{cases} p_i=R_i(P+t_i)\\ p_{i+1}=R_{i+1}(P+t_{i+1})\end{cases}$ so we have $p_{i+1}=R_{i+1}\cdot inv(R_i)\cdot(p_{i+1}+(t_{i+1}-t_i)\cdot R_i)$

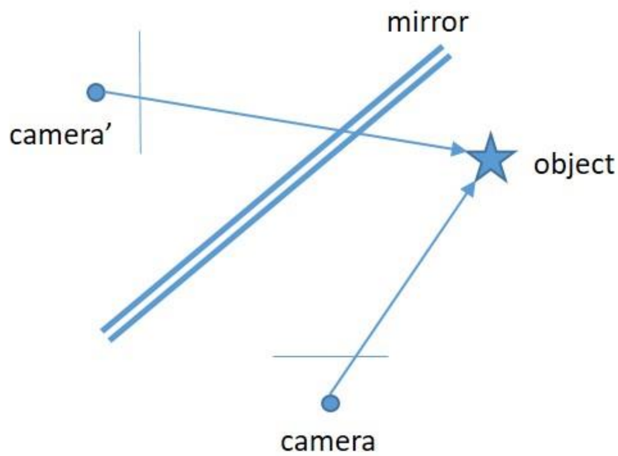
Therefore $R_{rel}=R_{i+1}\cdot inv(R_i),\ t_{rel}=(t_{i+1}-t_i)\cdot R_i$

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

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

Q1.4

Equivalently this scenario is like one object observed by two cameras which are related by mirror reflection, meaning one camera is like the other camera's image. This is shown in the figure below.



Let d be the perpendicular distance from camera to the mirror and $\vec{n}$ be the unit normal vector of the mirror.

Assume the 3D coordinates of object in camera to be *X* and *X'* respectively.

$$X' = (I - 2\vec{n}\vec{n}^T)X + 2d\vec{n}$$

Let p and p' be the 2D projection on camera frame.

$$K^{-1}p' = (I - 2\vec{n}\vec{n}^T)K^{-1}p + 2d\vec{n}$$

$$p'^T K^{-T}\left(2d[\vec{n}]_\times(I - 2\vec{n}\vec{n}^T)\right)K^{-1}p = 0$$

Therefore we have essential matrix

$$E = 2d[\vec{n}]_\times(I - 2\vec{n}\vec{n}^T)$$

And fundamental matrix

$$F = K^{-T}\left(2d\vec{n}_\times(I - 2\vec{n}\vec{n}^T)\right)K^{-1}$$

$$F + F^T = 2d(K^{-T}[\vec{n}]_\times K^{-1} - K^{-T}[\vec{n}]_\times K^{-1}) = 0$$

Reference:

Mariottini, Gian Luca, et al. "Planar mirrors for image-based robot localization and 3-D reconstruction." Mechatronics 22.4 (2012): 398-409.

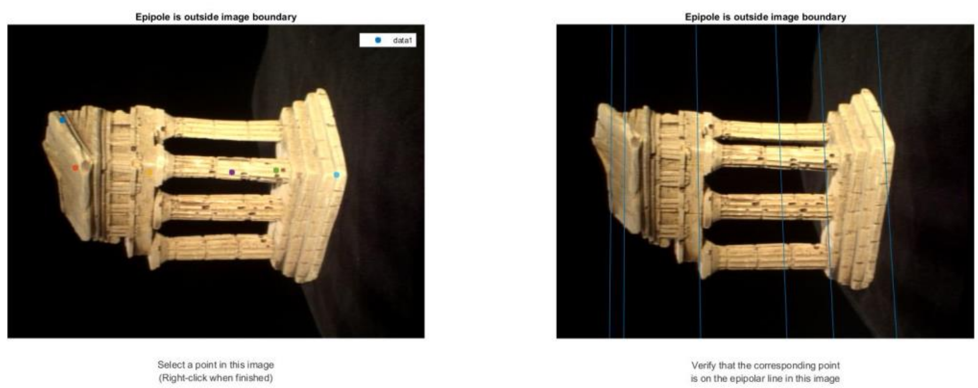
Q2.1

Say $x = \begin{pmatrix} x_1 & x_2 & x_N \\ y_1 & y_2 \cdots & y_N \\ 1 & 1 & 1 \end{pmatrix}$, if we scale to 0~1, then $x_{normalized} = \begin{pmatrix} \frac{1}{width} & 0 & 0 \\ 0 & \frac{1}{height} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot x = T \cdot x$

Therefore during unscaling, $F_{unnormalized} = T^T F T$.

Fundamental matrix F that I calculated:

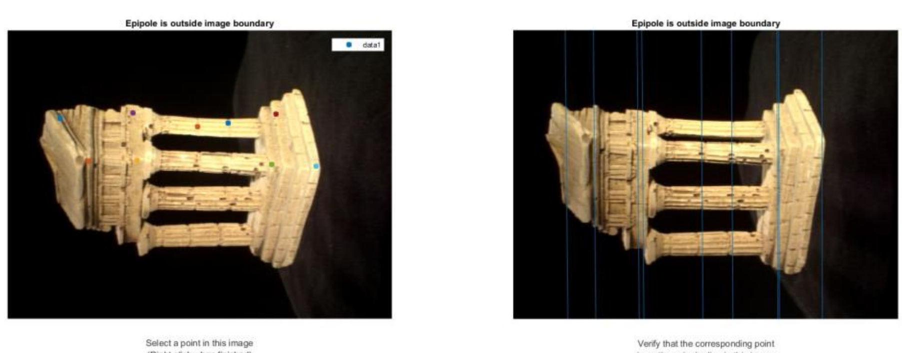
$$F = \begin{pmatrix} -2.57831189150015e-09 & -7.79979958742389e-08 & 0.00102108446795867 \\ -9.36203792670041e-08 & 1.38055399689906e-09 & 1.45416555219677e-06 \\ -0.000980150758986550 & 1.59426453304359e-05 & -0.00425887331720721 \end{pmatrix}$$



Q2.2

The points which are selected by RANSAC are [52 100 39 8 29 82 17]. And the corresponding fundamental matrix F that I calculated:

$$F = \begin{pmatrix} -1.94095238654772e-08 & -2.55282646521840e-07 & 0.00149907103452389 \\ 2.23338121951237e-08 & 2.15355388826295e-09 & -1.58680603446203e-05 \\ -0.00143224964635102 & 3.66100087465998e-05 & -0.00642913108729118 \end{pmatrix}$$



Q2X RANSAC

Once I calculated a fundamental matrix F from randomly seven points, I calculated the epipolar line corresponding to the point in image 1 which is F·p1. Then normalize the line by [l(1) l(2) l(3)]/sqrt(l(1)^2+l(2)^2). Then I calculated the perpendicular distance between this epipolar line and p1's correspondence point p2. If this distance is shorter than a certain tolerance value, which is set to be 0.5 in my program, then this pair of p1 and p2 is considered as in-liner. The maximum iteration number is set to be 500 which yields good results.

Q2.3

Essential matrix is calculated as $E=K2^T \cdot F \cdot K1$.

Q2.4

We solve for P under the constraint that $\begin{cases} P_1 = M_1 \cdot P \\ P_2 = M_2 \cdot P \end{cases}$

I use the the method in Szeliski Chapter 7 to find the 3D point P that lies closest to all of the 3D rays corresponding to the 2D matching feature locations. But first we need to find out the camera intrinsic K and rotation and translation matrix $[R|t]$.

Camera intrinsic matrix K is a 3x3 upper triangular matrix in the form of $\begin{bmatrix} f_x & S & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$ where f_x and f_y are focal lengths in pixels and x_0 and y_0 are coordinates of image center in pixels. And S is the skew parameter. K has 5 degrees of freedom.

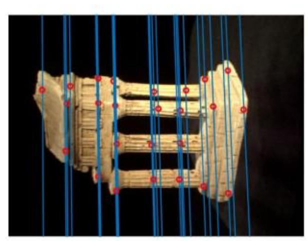
$$x = K[R|t]X = K[R| - R\tilde{C}]X$$

It is straightforward to get $\tilde{C}$ since we left-multiply the last column of M by $-(KR)^{-1}$. And KR can be achieved by taking away the last column of M. Since we have KR and $\tilde{C}$, we can calculate the 3D points according to the method in the book.

Q2.6

Here is a screenshot of some correspondence in image2 from some selected points in image1 according to my function.

Select a point in this image
(Right-click when finished)

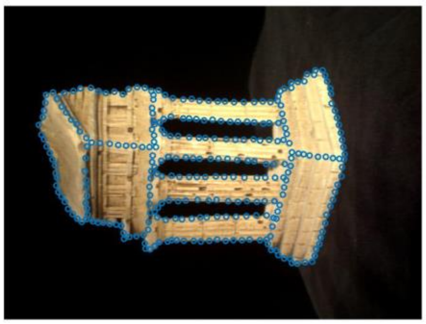


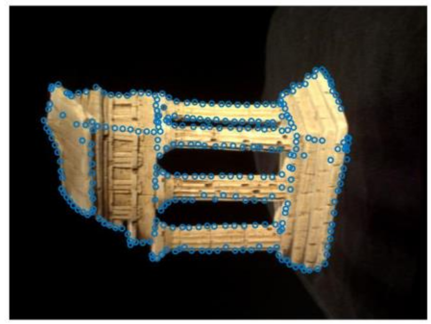
Verify that the corresponding point
is on the epipolar line in this image

In detail, I use a Gaussian mask which has kernel size of 15 pixels and sigma of 2.

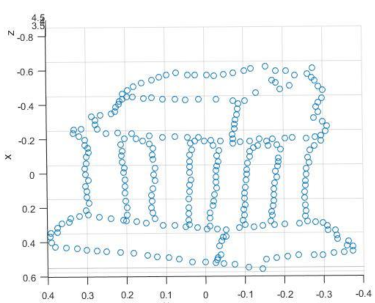
Q2.7

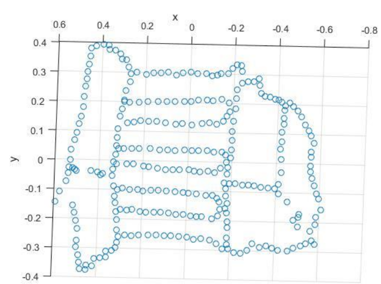
Below are the correspondences in image2 compared with the pre-selected points in image1. Roughly the matching should be correct.

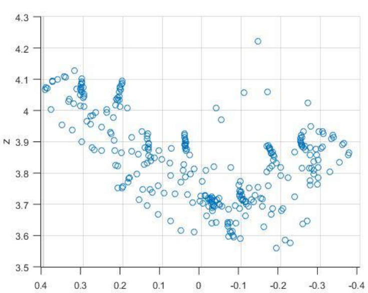




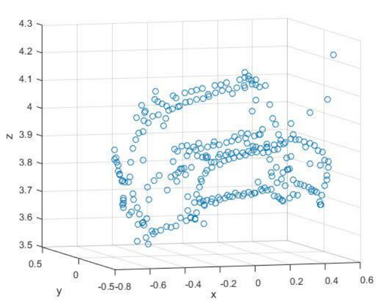
Here are some views of the reconstructed 3-D model:







The 3rd figure is the top view.



The 4th figure is side view.