16720: Computer Vision Homework 4

3D Reconstruction

Yuzhang Wu
Andrew ID: yuzhang2
December 30, 2015

Q1.1

As the question describes, we can know that the coordinate origin(0,0) coincides with the principal point. The two point in the plane are the projective point of P, so they are corresponding points and satisfy fundamental matrix.

We can set the coordinates of two points as homogeneous form:

$$p_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad p_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

And fundamental matrix has a general form:

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

And according to fundamental matrix character, we can get:

$$p_1^T \, F \, p_2 = 0$$

So:

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

Simplify it and we can get:

$$F_{33} = 0$$

In conclusion, the F_{33} element of the fundamental matrix is zero.

Q1.2

Firstly we can make a hypothesis that the two camera matrix are

$$P = K[I|0] \quad P' = K[I|t]$$

And also we know the fundamental matrix is

$$F = [e'] \times \mathrm{P}' P^{+}$$

And e' is the epipole of right camera and P^{+} is the pseudo-inverse of P

So we can get:

$$F = [e'] \times \mathrm{P}' P^{+} = [e'] \times K[I|t] \times \begin{bmatrix} K^{-1} \\ 0^T \end{bmatrix} = [e'] \times K' I K^{-1}$$

Since the question describes that the right camera differs from the left by a pure translation that is parallel to the x-axis. So:

$$R = I \quad K = K'$$

$$F = [e']_{\times} K I K^{-1} = [e']_{\times}$$

$$\mathrm{e}' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

So

$$\mathrm{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

We can let the point coordinates in two planes are

$$\mathrm{x} = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \qquad \mathrm{x}' = \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$$

Substitute them into the fundamental matrix formulation:

$$[x_1 \quad y_1 \quad 1] \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = 0$$

Simplify it, we can get:

$$y_1 = y_2$$

Using the epipolar matrix formulation, we can get the coordinates of two different epipolar line in two images.

The first image's epipolar line is:

$$l = F\mathrm{x}' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ y_2 \end{bmatrix}$$

$$\Rightarrow y = y_2$$

The second image's epipolar line is:

$$l = \mathrm{x}^T F = [x_1 \quad y_1 \quad 1] \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = [0 \quad 1 \quad -y_1]$$

$$\Rightarrow y = y_1$$

And the x-axis is

$$y = 0$$

We can find that these three line's coefficient are proportional which means that these three lines are parallel to each other. So the epipolar lines in the two cameras are parallel to the x-axis.

Q1.3

Let us dente X and X' are coordinates in the world of two points. x and x' are separately the projection point in two images of X and X'.

Since we have mirror, so there may be some relations between object coordinate and its reflection coordinate:

$$X' = MX$$

And due to the property of mirror, we have:

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Under the homogeneous form, it is:

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can make a hypothesis that the two camera matrix is:

$$P = K[I|0]$$

Accordingly, we can know the coordinate of the projective matrix:

$$x = K[I|0]\begin{bmatrix} X \\ 1 \end{bmatrix} = KX$$

Similarly,

$$x' = KX'$$

And since

$$X' = MX$$

We can get that:

$$x' = KMX$$

For the fundamental matrix F:

$$x'^T F x = 0$$

$$\Rightarrow X^T M^T K^T F K X = 0$$

We can let:

$$A = M^T K^T F K$$

And

$$F = M^{-T} K^{-T} A^{-1}$$

And for:

$$X^T A X = 0$$

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

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Simplify it and we can get

$$a_{11} = a_{22} = a_{33} = 0$$
$$a_{12} + a_{21} = 0$$
$$a_{13} + a_{31} = 0$$
$$a_{23} + a_{32} = 0$$

So A is a skew-symmetric matrix.

Since

$$F = M^{-T} K^{-T} A K^{-1}$$

M is a symmetric matrix, so M^{-T} is a symmetric matrix too. $K^{-T}AK^{-1}$ is a skew symmetric matrix due to A is a skew symmetric matrix and $K^{-T}K^{-1}$ cannot change the skew symmetric. And a symmetric matrix multiple a skew symmetric matrix is a skew symmetric matrix, so the fundamental matrix F is a skew-symmetric fundamental matrix.

Q2.1

The 8-point algorithm is the simplest method of computing the fundamental matrix, involving no more than the construction and least-squares solution of a set of linear equations. If care is taken, then it can perform extremely well.

The original algorithm is due to Longuet-Higgins. The key to success with the 8-point algorithm is proper careful normalization of the input data before constructing the equations to solve. In the case of the 8-point algorithm, a simple transformation (translation and scaling) of the points in the image before formulating the linear equations leads to an enormous improvement in the conditioning of the problem and hence in the stability of the result. The added complexity of the algorithm necessary to do this transformation is insignificant.

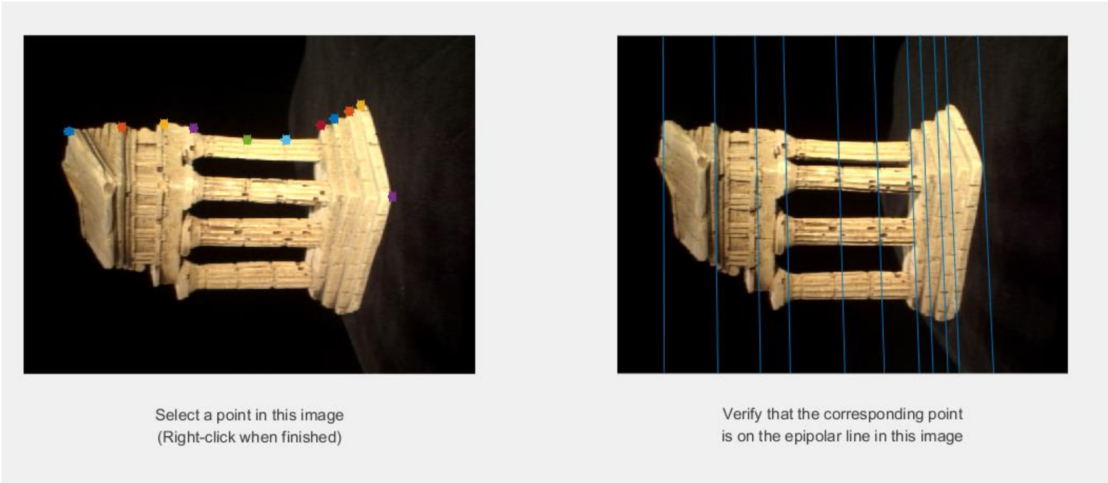
I wrote the eight-point algorithm to calculate the fundamental matrix and the general pseudo-codes are as below:

```
//normalization
norm1 = pts1 / M;
norm2 = pts2 / M;

//compute F
F1 = svd(coefficient matrix);

//constraint enforcement
call refinF function;

//denormalization
F = m' * F * m;
```

And the recovered F is:

$$F = \begin{matrix} -4.6906 & 1.4600 & -0.0019 \\ 1.7891 & -2.5710 & -3.5099 \\ 0.0019 & -2.9982 & 0.0075 \end{matrix}$$

And the output of displayEpipolarF is:



Q2.2

The equation $x_i' F x_i = 0$ gives rise to a set of equations of the form $\mathrm{Af} = 0$. If A has rank 8, then it is possible to solve for f up to scale. In the case where the matrix A has rank seven, it is still possible to solve for the fundamental matrix by making use of the singularity constraint. The most important case is when only 7 point correspondences are known. This leads to a 7 × 9 matrix A, which generally will have rank 7.

The solution to the equations $\mathrm{Af} = 0$ in this case is a 2-dimensional space of the form $\alpha F_1 + (1 - \alpha) F_2$ where α is a scalar variable. The matrices F1 and F2 are obtained as the matrices corresponding to the generators f 1 and f 2 of the right null-space of A. Now, we use the constraint that det F = 0. This may be written as det(αF1 + (1 − α)F2) = 0. Since F1 and F2 are known, this leads to a cubic polynomial equation in α. This polynomial equation may be solved to find the value of α. There will be either one or three real solutions (the complex solutions are discarded. Substituting back in the equation F = αF1 + (1 − α)F2 gives one or three possible solutions for the fundamental matrix.

I wrote the seven-point algorithm to calculate the fundamental matrix and the general pseudo-codes are as below:

```
//normalization
norm1 = pts1 / M;
norm2 = pts2 / M;

//compute F
F1 = the ninth column of svd(coefficient matrix);
F2 = the eighth column of svd(coefficient matrix);

//solve for det(αF1 + (1 − α)F2) = 0
```

```
syms numta
a = solve(det(numta*F1+(1-numta)*F2 == 0));

//denormalization
F = αF1 + (1 − α)F2
F = m' * F * m;
```

After calculation, we got 3 answers F. And then I selected a best one and it is:

$$F = \begin{matrix} -1.05031 & 4.2465 & 0.0028 \\ 9.7877 & -7.2399 & -5.0800 \\ -0.0029 & -1.0919 & 0.0149 \end{matrix}$$

And the output of displayEpipolarF is:

Q2.3

The essential matrix is the specialization of the fundamental matrix to the case of normalized image coordinates (see below). Historically, the essential matrix was introduced (by Longuet-Higgins) before the fundamental matrix, and the fundamental matrix may be thought of as the generalization of the essential matrix in which the (inessential) assumption of calibrated cameras is removed. The essential matrix has fewer degrees of freedom, and additional properties, compared to the fundamental matrix.

The defining equation for the essential matrix is

$$x'^T Ex = 0$$

Substituting for x and x' gives $x'^T K'^{-T} EK^{-}x = 0$. Comparing this with the relation $x'^T Fx = 0$ for the fundamental matrix, it follows that the relationship between the fundamental and essential matrices is

$$\mathrm{E} = K'^{T} F K$$

So in this case, the expression of E is:

$$E = K2' * F * K1$$

And the estimated E is:

$$\mathrm{E} = \begin{matrix} -0.0108 & 0.3387 & -2.8773 \\ 0.4151 & -0.0060 & 0.0762 \\ 2.8856 & 0.0206 & 0.0011 \end{matrix}$$

Q2.4

The linear triangulation method is the direct analogue of the DLT method. In each image we have a measurement x = PX, x' = P 'X, and these equations can be combined into a form AX = 0, which is an equation linear in X.

First the homogeneous scale factor is eliminated by a cross product to give three equations for each image point, of which two are linearly independent. For example for the first image, x × (PX) = 0 and writing this out gives

$$\mathrm{x}(p^{3T}X) - (p^{1T}X) = 0$$
$$\mathrm{y}(p^{3T}X) - (p^{2T}X) = 0$$
$$\mathrm{x}(p^{2T}X) - y(p^{1T}X) = 0$$

where p^{iT} are the rows of P. These equations are linear in the components of X.
An equation of the form AX = 0 can then be composed, with

$$\mathrm{A} = \begin{bmatrix} xp^{3T} - p^{1T} \\ yp^{3T} - p^{2T} \\ x'p'^{3T} - p'^{1T} \\ y'p'^{3T} - p'^{2T} \end{bmatrix}$$

where two equations have been included from each image, giving a total of four equations in four homogeneous unknowns. This is a redundant set of equations, since the solution is determined only up to scale.

I wrote the seven-point algorithm to calculate the fundamental matrix and the general pseudo-codes are as below:

```
//transfer all the coordinates to homogenous form
//calculation part
For i = 1 : size
    T = [p*M1, q*M2];
    D = svd(T);
    P = d/d(4); //normalize
End
```

Q2.5

findM2.m is used to obtain the correct M2 from M2s by testing the four solutions through triangulations. The right M2's Z coordinate must be a positive number, so I use this standard to determine whether the M2 is right or not.

I think it is a comprehensive application based on all the previous work, the pseudo-codes are as below:

```
Load all the parameters

//call eightpoint method
F = eightpoint(pts1, pts2, M);

//call essentialMatrix method
E = essentialMatrix(F, K1, K2);

//call camera2 method
M2s = camera2€;

//check the right M2
For i = 1 : 4
    P = triangulate(K1*M1, pts1, K2*M2s(:,:,i),pts2)
    If P(:,3) > 0
        Pcorrect = P;
        Break;
    End
End
```

And at last I found the first M2 is right.

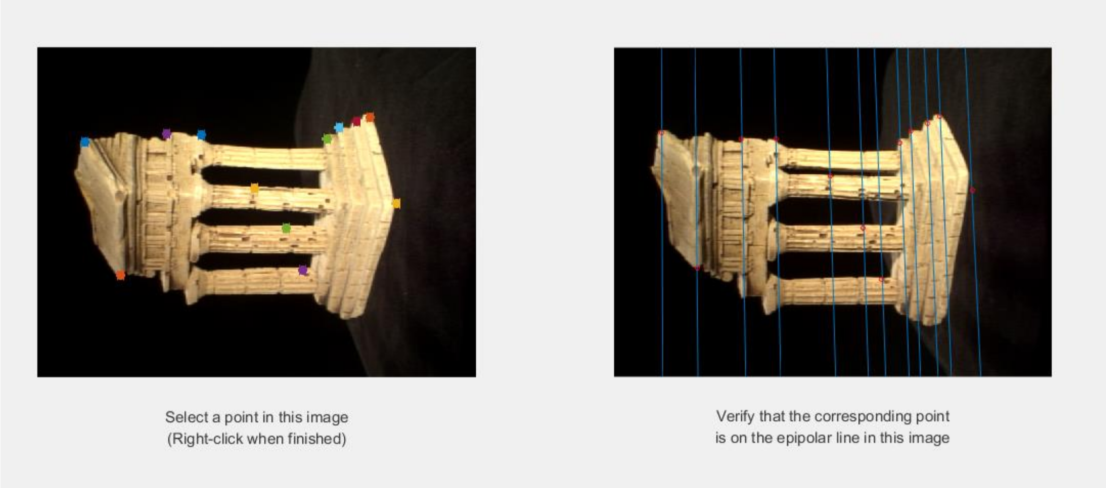
Q2.6

This question takes in the x and y coordinates of a pixel on im1 and fundamental matrix F, and returns the coordinates of the pixel on im2 which correspond to the input point.

Since we have calculated the fundamental matrix F and then we can compute the epipolar line and the corresponding point must be on that line. After that, through the method of comparison the difference in a determined window size. We can estimate location of the most probable point.

1. Compute the epipolar line
2. Estimate an approximate point
3. Let the point move on the estimated line and then compute the difference using Gaussian Filter
4. Find the point in that the errors are least.

I have used several points in the im1 to test the performance of my codes and find that the results are really accurate. The output of epipolarMatchGUI is attached below:



Q2.7

This question requires us to determine the 3D location of these point correspondences using the triangulate function. And then we can plot it.

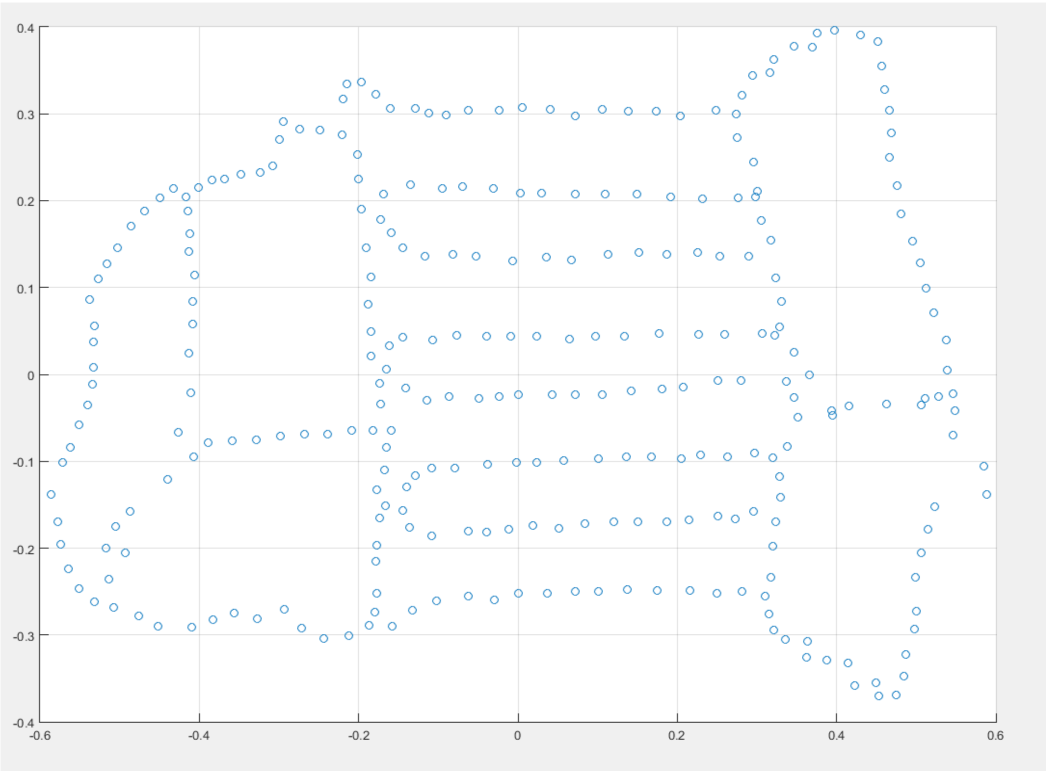
Using a lot of points' coordinates in im1, we can first calculate the location of them in im2. And then, call triangulate function, we can determine the original point coordinates in the 3D world.

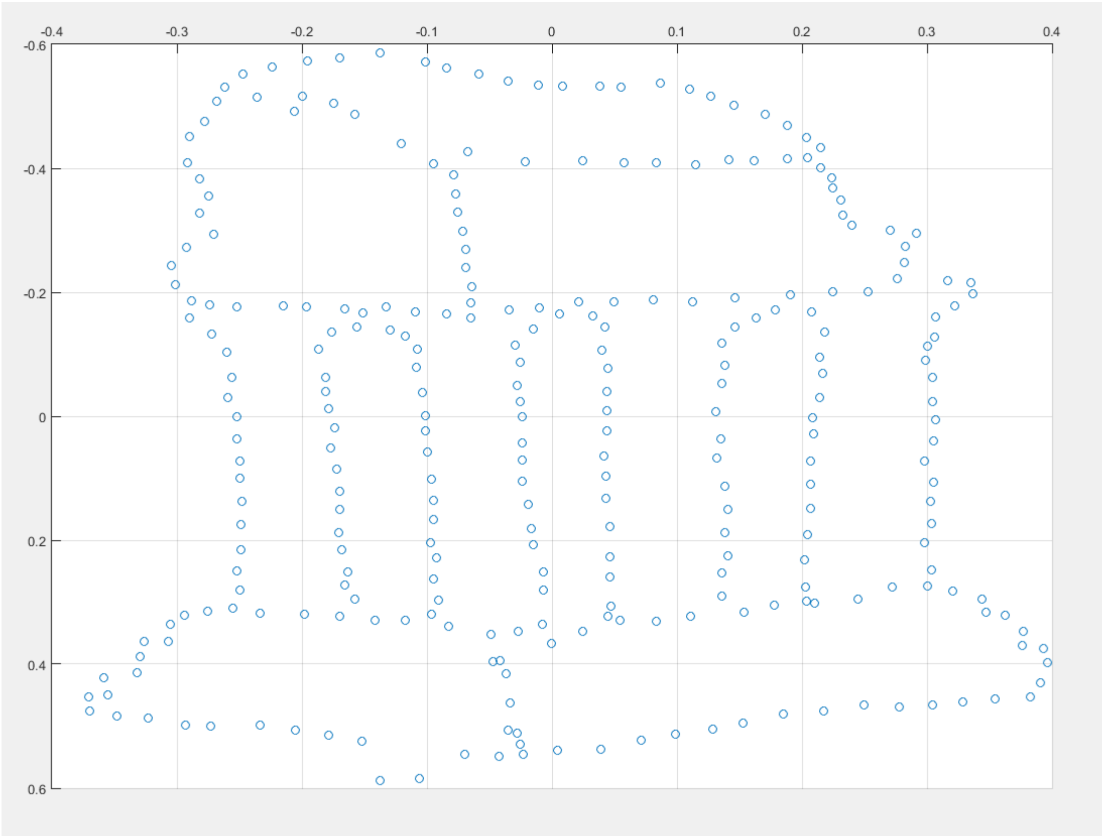
The steps are as follows:

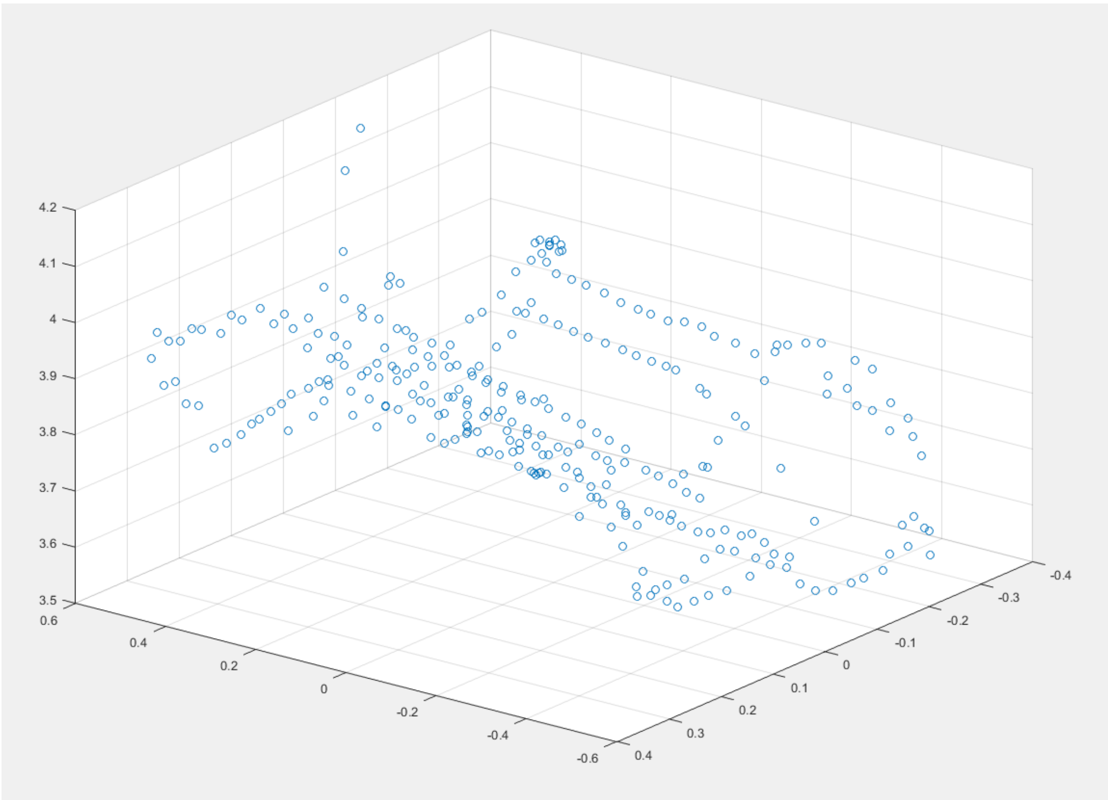
1. Calculate the point coordinate in im2
2. Compute essential matrix
3. Check the right M2 and calculate P using function triangulate

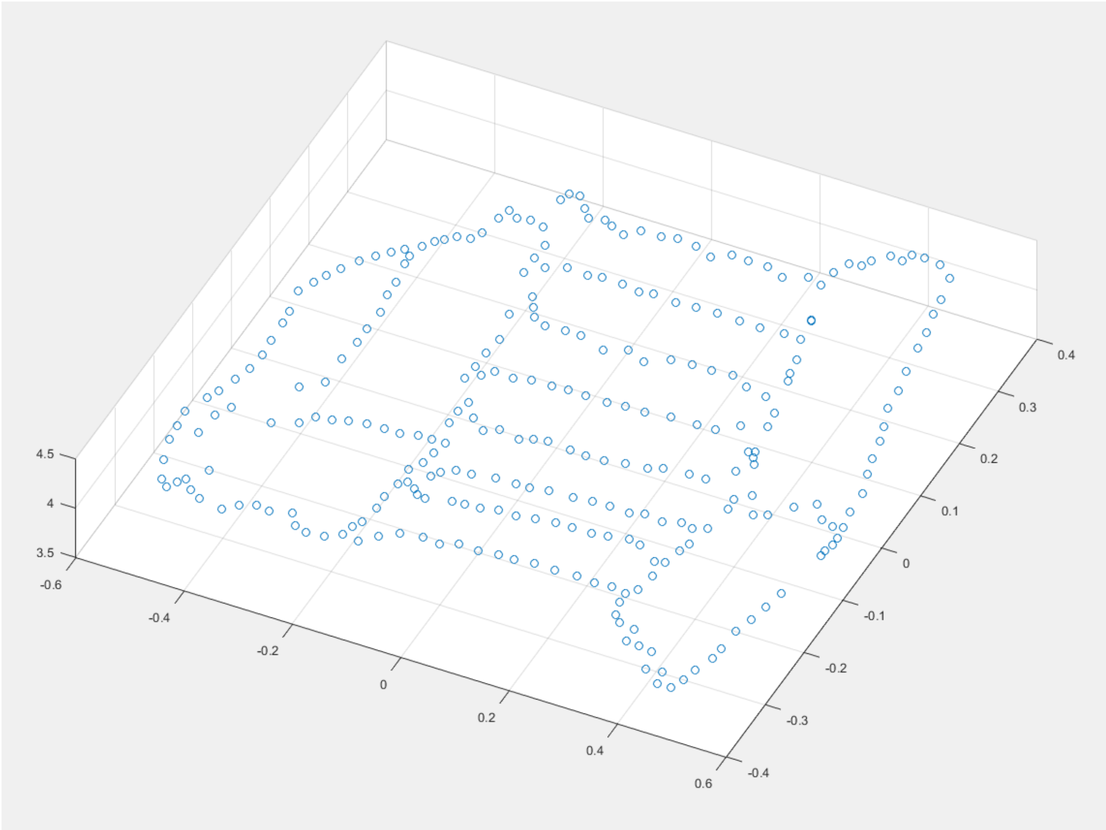
Here are some screenshots of the 3D reconstruction figure in different angles.

We can see that the results is really well and can reflect the original object.




-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
-0.6
-0.4
-0.2
0
0.2
0.4
0.6


4.2
4.1
4
3.9
3.8
3.7
3.6
3.5
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4


4.5
4
3.5
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4