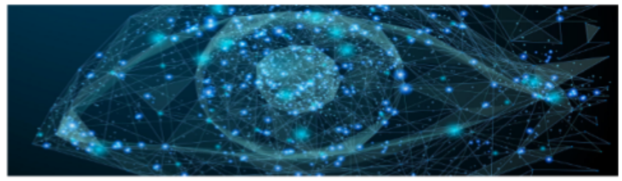
Collaborated with: Dhanvi, Anish, Kshitij and Jash

THE ROBOTICS INSTITUTE Computer Vision 16720-B Fall 2022



16720 (B) 3D Reconstruction - Assignment 5

```
Instructor: Kris                                  TAs: Arka, Jinkun,
Rawal, Rohan, Sheng-Yu
```

Instructions

This section should include the visualizations and answers to specifically highlighted questions from P1 to P4. This section will need to be uploaded to gradescope as a pdf and manually graded (this is a separate submission from the coding notebooks).

1. Students are encouraged to work in groups but each student must submit their own work. Include the names of your collaborators in your write up. Code should Not be shared or copied. Please properly give credits to others by LISTING EVERY COLLABORATOR in the writeup including any code segments that you discussed, Please DO NOT use external code unless permitted. Plagiarism is prohibited and may lead to failure of this course.
2. **Start early!** This homework will take a long time to complete.
3. **Questions:** If you have any question, please look at Piazza first and the FAQ page for this homework.
4. All the theory question and manually graded questions should be included in a single writeup (this notebook exported as pdf or a standalone pdf file) and submitted to gradescope: pdf assignment.
5. **Attempt to verify your implementation as you proceed:** If you don't verify that your implementation is correct on toy examples, you will risk having a huge issue when you put everything together. We provide some simple checks in the notebook cells, but make sure you verify them on more complicated samples before moving forward.
6. **Do not import external functions/packages other than the ones already imported in the files:** The current imported functions and packages are enough for you to complete this assignment. If you need to import other functions, please remember to comment them

out after submission. Our autograder will crash if you import a new function that the gradescope server does not expect.

7. Assignments that do not follow this submission rule will be **penalized up to 10% of the total score**.

Theory Questions (25 pts)

Before implementing our own 3D reconstruction, let's take a look at some simple theory questions that may arise. The answers to the below questions should be relatively short, consisting of a few lines of math and text (maybe a diagram if it helps your understanding).

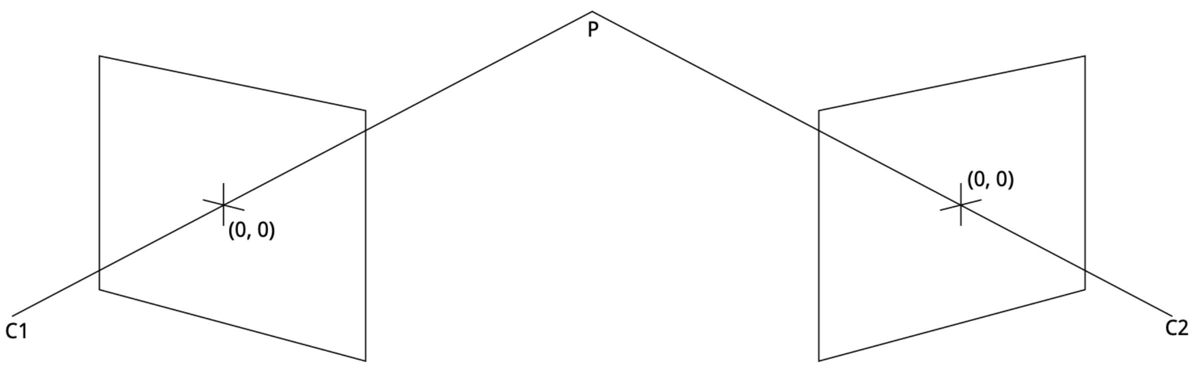


Figure1. Figure for Q1.1. C1 and C2 are the optical centers. The principal axes intersect at point \textbf{w} (P in the figure).

Q0.1

Suppose two cameras fixated on a point x (see Figure 1) in space such that their principal axes intersect at the point P. Show that if the image coordinates are normalized so that the coordinate origin $(0,0)$ coincides with the principal point, the $\mathbf{F}_{33}$ element of the fundamental matrix is zero.

Let p_1 and p_2 be the image points for camera 1 and 2 respectively.
Then we know that

$$p_2^T F p_1 = 0$$

And since the image coordinates are normalized so that the coordinate origin coincidess with the principal point:

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

Using this in the above equation:

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

For this condition to be true: f_{33} should be zero

Therefore, f_{33} = 0.

Q0.2

Consider the case of two cameras viewing an object such that the second camera differs from the first by a pure translation that is parallel to the x-axis. Show that the epipolar lines in the two cameras are also parallel to the x-axis. Backup your argument with relevant equations.

The translation matrix for translation parallel to x-axis wiil be:

$$t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix}.$$

Now, for pure rotation and given the above translation matrix; the essential matrix will be:

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

Therefore, the epipolar lines will be:

$$\mathbf{l}_1 = \mathbf{E}^T \tilde{\mathbf{x}}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & t_x \\ 0 & -t_x & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ t_x \\ -t_x y_2 \end{bmatrix}$$

$$\mathbf{l}_2 = \mathbf{E} \tilde{\mathbf{x}}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -t_x \\ t_x y_1 \end{bmatrix}$$

Therefore, the epipolar line for camera 1 will be $t_x y - t_x y_2 = 0$ and for camera 2 will be -$t_x y + t_x y_1 = 0$.

And these lines are parallel to the x-axis.

Q0.3

Suppose we have an inertial sensor which gives us the accurate extrinsics $\mathbf{R}_i$ and $\mathbf{t}_i$ (see Figure 2), the rotation matrix and translation vector of the robot at time i. What will be the effective rotation ($\mathbf{R}_{rel}$) and translation ($\mathbf{t}_{rel}$) between two frames at different time stamps? Suppose the camera intrinsics ($\mathbf{K}$) are known, express the essential matrix ($\mathbf{E}$) and the fundamental matrix ($\mathbf{F}$) in terms of $\mathbf{K}$, $\mathbf{R}_{rel}$ and $\mathbf{t}_{rel}$.

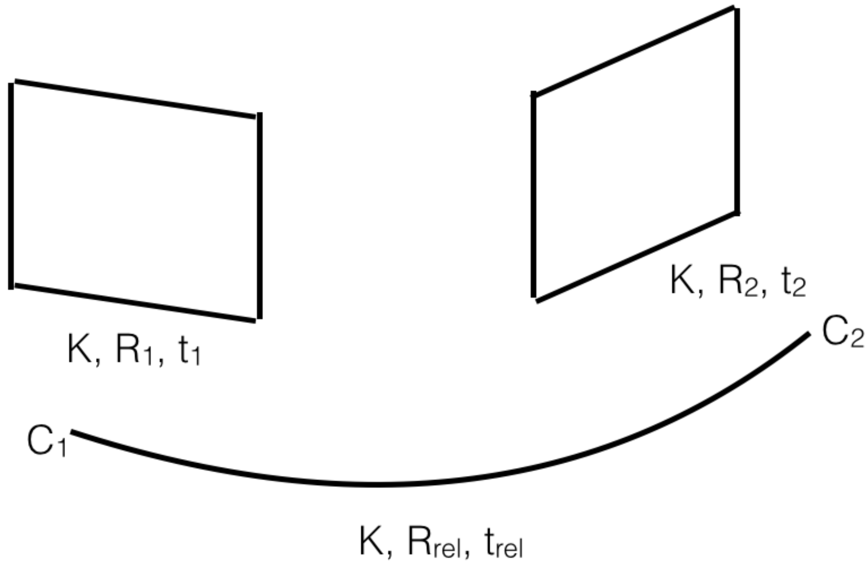


Figure 2. Figure for Q1.3. $C1$ and $C2$ are the optical centers. The rotation and the translation is obtained using inertial sensors. $\mathbf{R}_{rel}$ and $\mathbf{t}_{rel}$ are the relative rotation and translation between two frames.

Let x_i be the projection of world point 'X' on the image frame at time instant 'i'.
So at time '1',
$x_1 = R_1(X - t_1)$
$\implies R_1^{-1}x_1 + t_1 = X$
and,
$x_2 = R_2(X - t_2)$
$\implies x_2 = R_2(R_1^{-1}x_1 + t_1 - t_2)$
$\implies x_2 = (R_2R_1^{-1})x_1 + R_2(t_1 - t_2)$
$\implies x_2 = (R_{rel})x_1 + (t_{rel})$
where, $R_{rel} = R_2R_1^{-1}$ and $t_{rel} = R_2(t_1 - t_2)$

Now, essential matrix $E = t_{rel}xR_rel$
and fundamental matrix $F = (K^{-1})^TEK^{-1}$

Q0.4

Suppose that a camera views an object and its reflection in a plane mirror. Show that this situation is equivalent to having two images of the object which are related by a skew-symmetric fundamental matrix. You may assume that the object is flat, meaning that all points on the object are of equal distance to the mirror (**Hint:** draw the relevant vectors to understand the relationship between the camera, the object, and its reflected image.)

Let the point be P and its reflection be P'. Let the corresponding points on image 1 and image

2 be x_1 and x_2.

Then we can write:

$\tilde{x_2}^T F \tilde{x_1}^T = 0$

$\implies \tilde{x_1}^T F^T \tilde{x_2} = 0$

$\implies \tilde{x_2}^T F \tilde{x_1}^T + \tilde{x_1}^T F^T \tilde{x_2} = 0$ -----(1)

Now as camera 2 is a reflection of camera 1, we can write:

$M_2 = TM_1$

where, T is a transition matrix between them such that $T^T T = I$

Therefore,

$\tilde{x_1} = KM_1\tilde{P}$, and

$\tilde{x_2} = KTM_1\tilde{P}$

Substituting these values in equation (1):

$\tilde{P}^T M_1^T T^T K^T F K M_1 \tilde{P} + \tilde{P}^T M_1^T K^T F^T K T M_1 \tilde{P} = 0$

$\tilde{P}^T M_1^T (T^T K^T F K + K^T F^T K T) M_1 \tilde{P} = 0$

$\implies T^T K^T F K + K^T F^T K T = 0$

$\implies K^T F K T + T^T K^T F^T K = 0$

$K^T (F + F^T) K = 0$

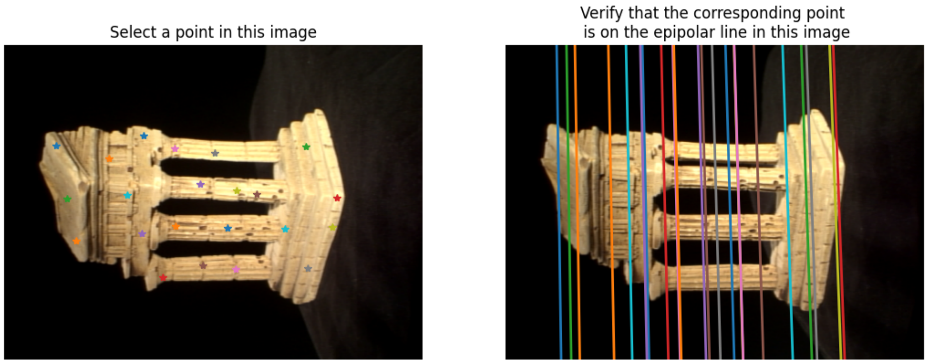
$\implies F = -F^T$

Therefore, the fundamental matrix is a skew-symmetric matrix.

Coding Questions (30 pt)

Q1.1: The Eight Point Algorithm

Output: In your write-up: Write your recovered $\mathbf{F}$ and include an image of some example outputs of displayEpipolarF.



The fundamental matrix is:

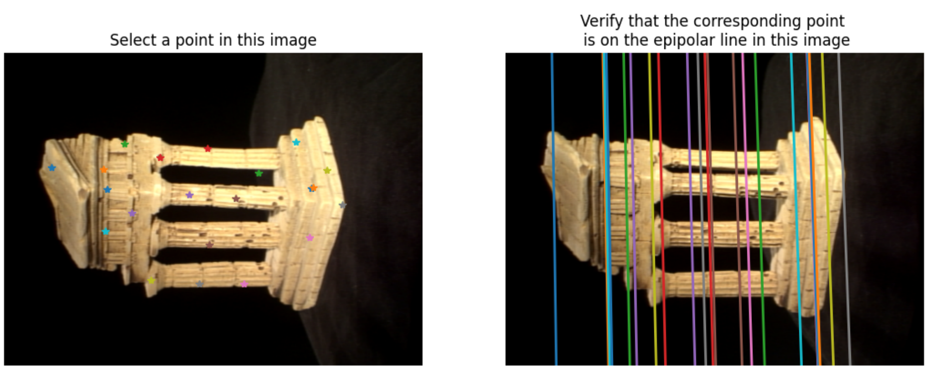
[[-0. 0. -0.2519]

[0. -0. 0.0026]

[0.2422 -0.0068 1.]]

Q1.2: The Seven Point Algorithm

Output: In your write-up: Print your recovered $\mathbf{F}$ and include an image output of `displayEpipolarF`.



The fundamental matrix is:

[[0. , 0. , -0.2011],

[0. , -0. , 0.001],

[0.1923, -0.0045, 1.]]

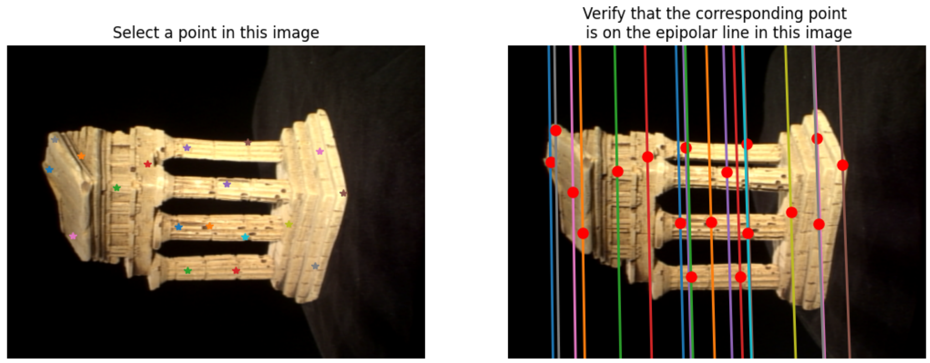
Q2.2 Triangulation and find M2

Output: In your write-up: Write down the expression for the matrix $\mathbf{A}_i$

$$A_i = \begin{bmatrix} y_i C_1[2,:] - C_1[1,:] \\ C_1[0,:] - x_i C_1[2,:] \\ y_i' C_2[2,:] - C_2[1,:] \\ C_2[0,:] - x_i' C_2[2,:] \end{bmatrix}$$

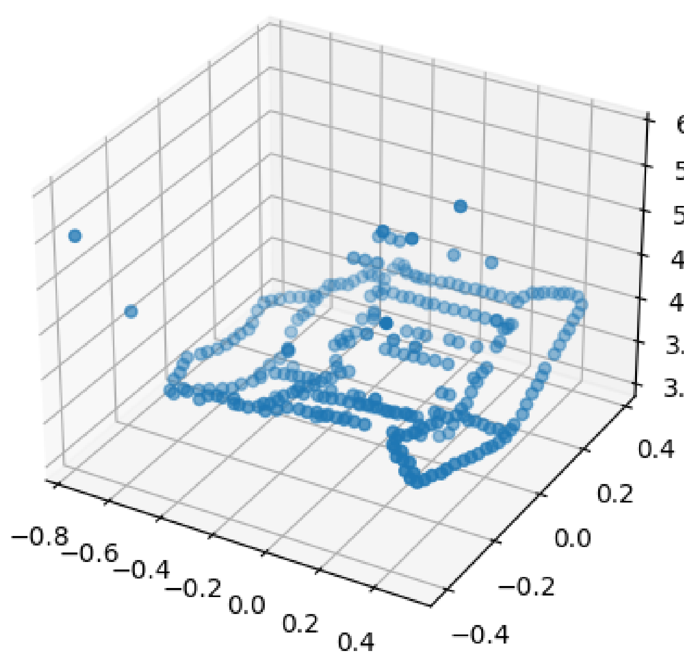
Q2.3 Epipolar Correspondence

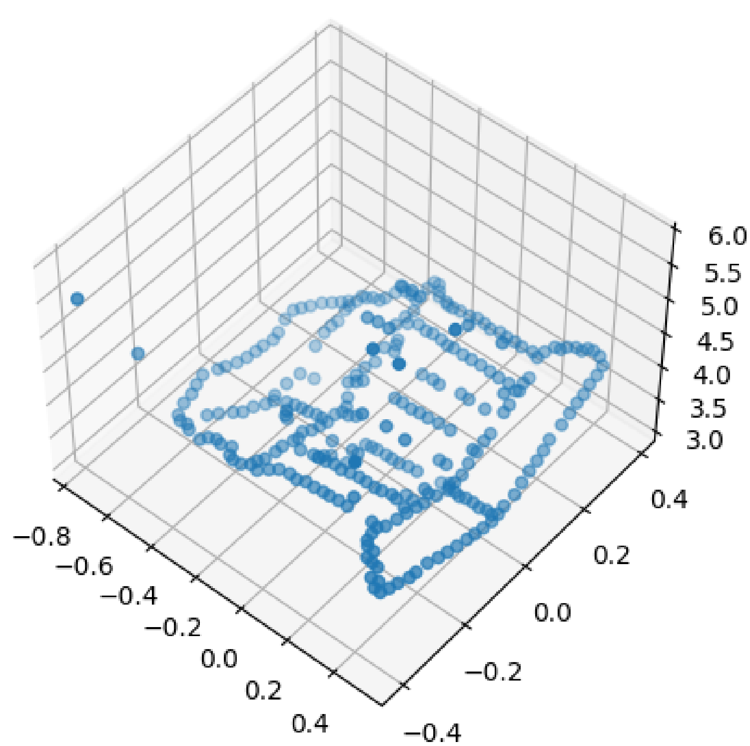
Output: In your write-up, include a screenshot of `epipolarMatchGUI` with some detected correspondences.

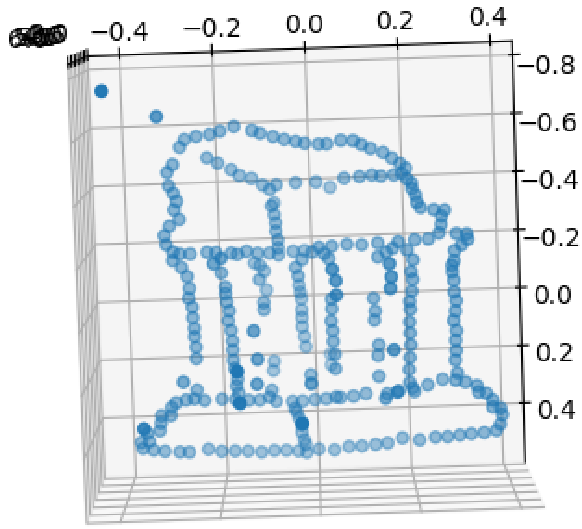


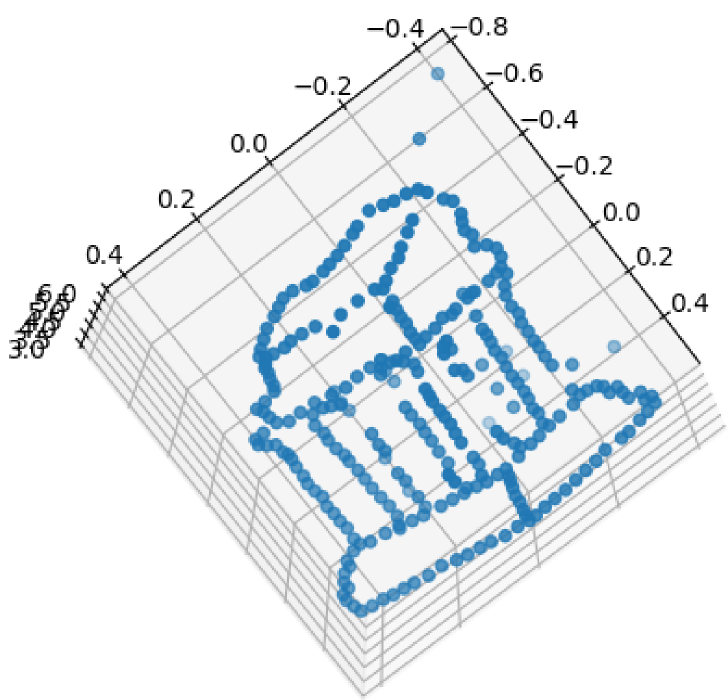
Q2.4 3D Visualization

Output: In your write-up: Take a few screenshots of the 3D visualization so that the outline of the temple is clearly visible.


6.0
5.5
5.0
4.5
4.0
3.5
3.0
0.4
0.2
0.0
−0.2
−0.4
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4


6.0
5.5
5.0
4.5
4.0
3.5
3.0
0.4
0.2
0.0
−0.2
−0.4
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4

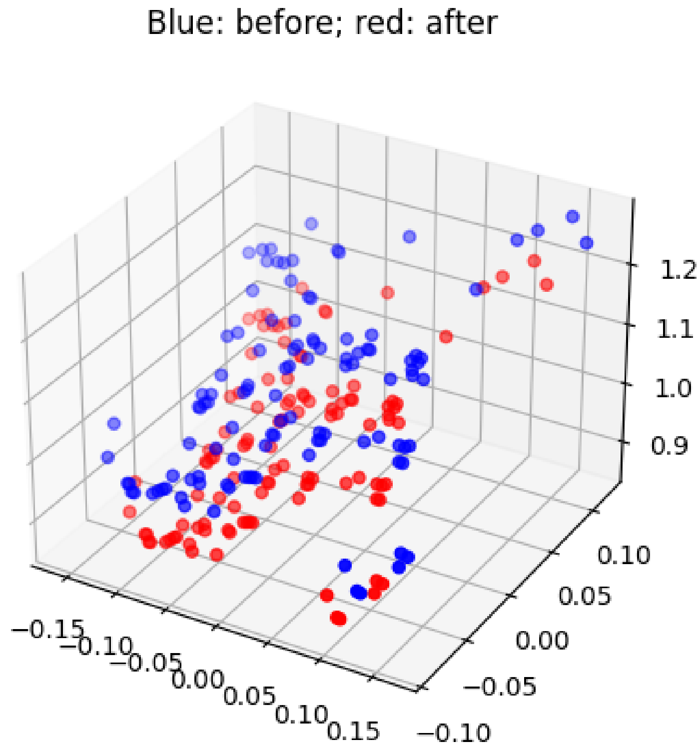


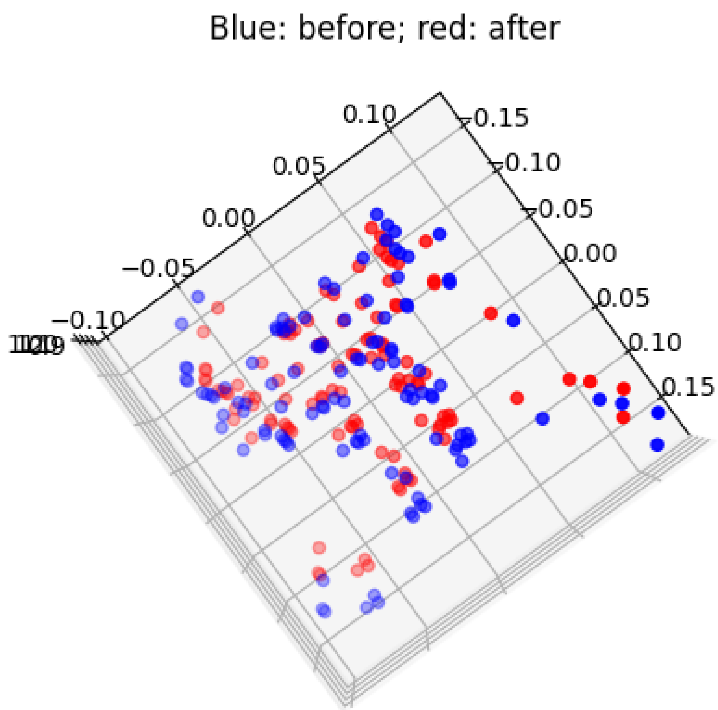


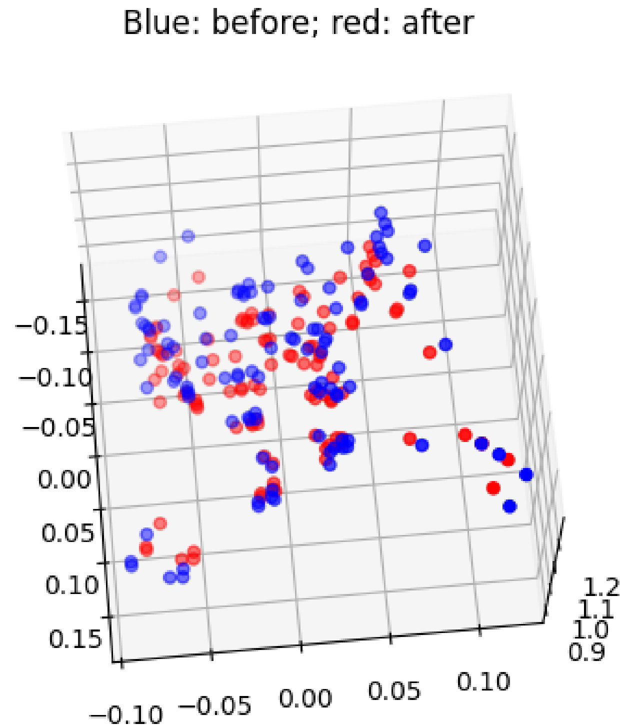
Q3.3 Bundle Adjustment

Output: In your write-up: include an image of output of the `plot_3D_dual` function by passing in the original 3D points and the optimized points. Also include the before and after reprojection error for the `rodriguesResidual` function.

Error before 238.92786884075895, Error after 11.348249171063372


Blue: before; red: after
1.2
1.1
1.0
0.9
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15


Blue: before; red: after
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15


Blue: before; red: after
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
−0.10
−0.05
0.00
0.05
0.10
1.2
1.1
1.0
0.9