

16-720B Computer Vision: Homework 4

3D Reconstruction

Heethesh Vhavle
Andrew ID: hvhavlen

11 November 2018
(Using **TWO LATE DAYS**)

1 Theory

1.1

It is given that the principal axes for both cameras intersect at the point P. Thus, they are corresponding points and we have the following relation when the points are at the origins.

$$\begin{aligned} \mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 &= 0 \\ \begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} &= 0 \\ \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= 0 \end{aligned} \tag{1}$$

On simplifying this, all the terms that have the image point coordinates get cancelled at the origin and therefore, we have the following relation.

$$F_{33} = 0 \tag{2}$$

1.2

For pure translation along the x-axis, we have the following transformation matrices.

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{t} &= \begin{bmatrix} t_1 \\ 0 \\ 0 \end{bmatrix} \end{aligned} \tag{3}$$

We also know the following relation for an essential matrix.

$$\mathbf{E} = \mathbf{t}_\times \mathbf{R} \tag{4}$$

We can write \mathbf{t}_\times as follows and obtain the essential matrix.

$$\begin{aligned}\mathbf{t}_\times &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & t_1 & 0 \end{bmatrix} \\ \mathbf{E} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & t_1 & 0 \end{bmatrix}\end{aligned}\tag{5}$$

Now, for any given point in image 1, we have an epipolar line in image plane 2.

$$\begin{aligned}\mathbf{l}_2 &= \mathbf{E}\mathbf{x}_1 \\ \mathbf{l}_2 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_1 \\ 0 & t_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \\ \mathbf{l}_2 &= \begin{bmatrix} 0 \\ -t_1 \\ y_1 \end{bmatrix}\end{aligned}\tag{6}$$

We can rewrite the above in the line equation form as follows for image 2.

$$y_2 = y_1\tag{7}$$

Therefore, this is a line parallel to the x-axis in image 2 and similarly, it can be shown that the epipolar line in image 1 is also parallel to the x-axis for a point in image 2.

1.3

We can use the concept of composition of homogeneous transformations to solve this problem. We know the following transformations with respect to the inertial frames, measured by the sensors.

$$\begin{aligned}\mathbf{H}_1^0 &= [\mathbf{R}_1 | \mathbf{t}_1] \\ \mathbf{H}_2^0 &= [\mathbf{R}_2 | \mathbf{t}_2]\end{aligned}\tag{8}$$

The relative transformation,

$$\begin{aligned}\mathbf{H}_2^1 &= [\mathbf{R}_{\text{rel}} | \mathbf{t}_{\text{rel}}] \\ \mathbf{H}_2^1 &= (\mathbf{H}_1^0)^{-1} \mathbf{H}_2^0 \\ \mathbf{H}_2^1 &= \begin{bmatrix} \mathbf{R}_1^T & -\mathbf{R}_1^T \mathbf{t}_1 \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_2 & \mathbf{t}_2 \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{H}_2^1 &= \begin{bmatrix} \mathbf{R}_1^T \mathbf{R}_2 & \mathbf{R}_1^T \mathbf{t}_2 - \mathbf{R}_1^T \mathbf{t}_1 \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{R}_{\text{rel}} &= \mathbf{R}_1^T \mathbf{R}_2 \\ \mathbf{t}_{\text{rel}} &= \mathbf{R}_1^T (\mathbf{t}_2 - \mathbf{t}_1)\end{aligned}\tag{9}$$

We can now use the known relations for the essential and fundamental matrices.

$$\begin{aligned}
\mathbf{E} &= \mathbf{t}_{\times \text{rel}} \mathbf{R}_{\text{rel}} \\
\mathbf{F} &= \mathbf{K}^{-T} \mathbf{E} \mathbf{K} \\
\mathbf{F} &= \mathbf{K}^{-T} \mathbf{t}_{\times \text{rel}} \mathbf{R}_{\text{rel}} \mathbf{K}
\end{aligned} \tag{10}$$

1.4

Let C be the camera in the real world and C' be the camera in the virtual world and \mathbf{K} be the intrinsic matrix for both the cameras. Let \mathbf{P} and \mathbf{x} be the 3D point in real world and the point in image plane and \mathbf{P}' and \mathbf{x}' be its reflection in the mirror and this point in the image plane. Given the mirror is flat, the transformation between these two points is a pure translation.

$$\begin{aligned}
\mathbf{R} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
\mathbf{t} &= \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \\
\mathbf{P}' &= \mathbf{P} + \mathbf{t} \\
\lambda_1 \mathbf{x} &= K \mathbf{P} \\
\lambda_2 \mathbf{x}' &= K \mathbf{P}'
\end{aligned} \tag{11}$$

Using the above equations, we can now write the relationship between the two points as follows.

$$\lambda_2 K^{-1} \mathbf{x}' = \lambda_1 K^{-1} \mathbf{x} + \mathbf{t} \tag{12}$$

We can simplify the equation and eliminate some terms by taking cross product with \mathbf{t} on both sides, followed by dot product with \mathbf{x}' to get the following.

$$\begin{aligned}
\mathbf{x}'^T K^{-T} \mathbf{t}_{\times} K^{-1} \mathbf{x} &= 0 \\
\mathbf{t}_{\times} &= \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}
\end{aligned} \tag{13}$$

We can see that, \mathbf{t}_{\times} is a skew-symmetric matrix here and we know the relation that $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$. Comparing this with above form, we get the following expression for \mathbf{F} .

$$\mathbf{F} = K^{-T} \mathbf{t}_{\times} K^{-1} \tag{14}$$

Since, \mathbf{t}_{\times} is skew symmetric, it can be shown that \mathbf{F} here will also retain the property of skew-symmetric for a given intrinsic matrix \mathbf{K} .

$$\mathbf{F}^T = -\mathbf{F} \tag{15}$$

Therefore, we can conclude that the two images of the object are related by a skew-symmetric fundamental matrix.

2 Fundamental Matrix Estimation

2.1 The Eight Point Algorithm

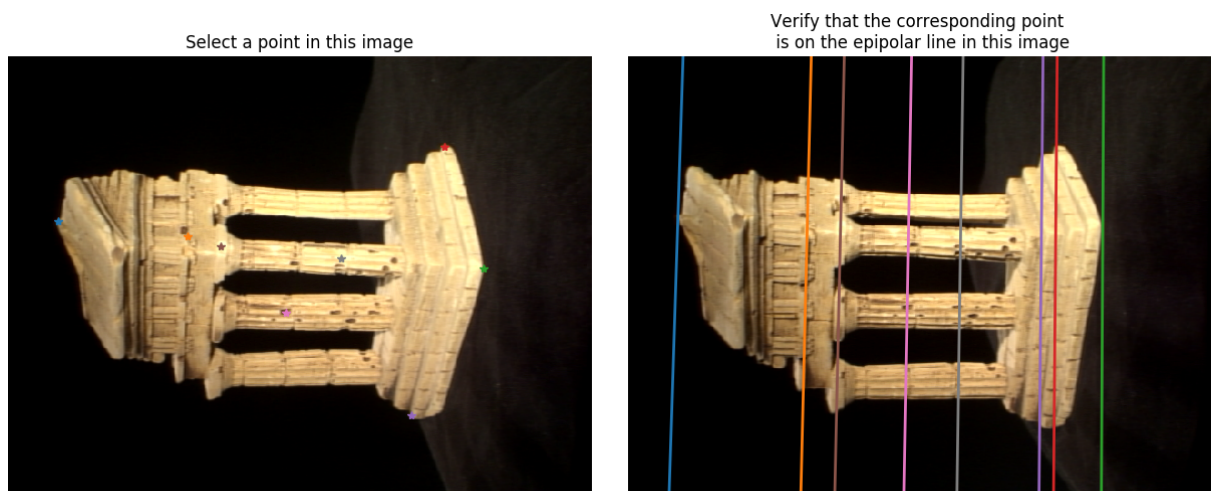


Figure 1: Eight-Point Algorithm Output

```
[[ -8.33149231e-09  1.29538462e-07 -1.17187851e-03]
 [  6.51358337e-08  5.70670060e-09 -4.13435037e-05]
 [  1.13078765e-03  1.91823637e-05  4.16862080e-03]]
```

Figure 2: Fundamental Matrix recovered from Eight-Point Algorithm

2.2 The Seven Point Algorithm

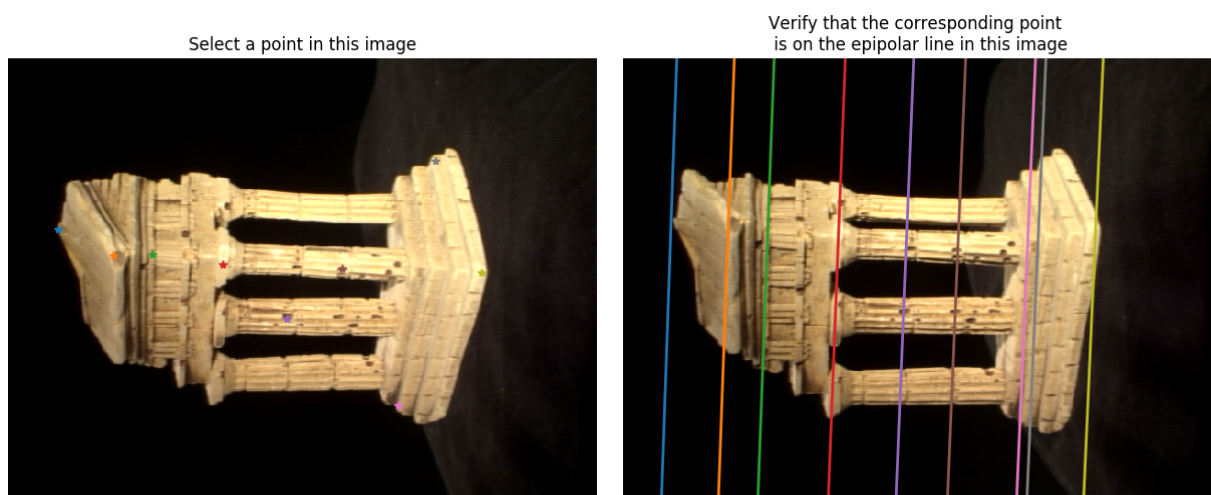


Figure 3: Seven-Point Algorithm Output

The **third fundamental matrix** from the recovered set was found to be the correct one and is used for the visualization.

```

[[-1.29112668e-06,  7.07910920e-06, -2.66910095e-03],
 [-6.12061099e-06,  2.62897062e-06, -5.54238146e-04],
 [ 3.08512019e-03, -8.26588573e-04,  8.43474773e-02]]

[[ 6.90594357e-07, -3.95406822e-06,  1.83830500e-04],
 [ 3.70536589e-06, -1.59312872e-06,  2.89750890e-04],
 [-4.57898759e-04,  4.90784533e-04, -4.62403318e-02]]

[[-6.08872858e-08,  2.29785168e-07, -8.98019857e-04],
 [-2.07091188e-08,  7.91909582e-09, -3.02952995e-05],
 [ 8.85637310e-04, -8.77200649e-06,  3.27942364e-03]]

```

Figure 4: Fundamental Matrices recovered from Seven-Point Algorithm

3 Metric Reconstruction

3.1 Essential Matrix

```

[[-1.92592123e-02  3.00526429e-01 -1.73693252e+00]
 [ 1.51113725e-01  1.32873151e-02 -3.08885272e-02]
 [ 1.73986815e+00  9.11774761e-02  3.90697725e-04]]

```

Figure 5: Estimated Essential Matrix

3.2 Triangulation

We have the following relationship for given 2D point and the corresponding 3D point for a camera matrix \mathbf{C} .

$$\begin{aligned}
\tilde{\mathbf{x}}_i &= \mathbf{C}\mathbf{w}_i \\
\tilde{\mathbf{x}}_i \times \mathbf{C}\mathbf{w}_i &= \mathbf{0}
\end{aligned}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{0} \quad (16)$$

For the two cameras, we can write it in the following form with \mathbf{A} matrix and solve for the 3D points using SVD.

$$\begin{aligned}
\mathbf{A}_i \mathbf{w}_i &= \mathbf{0} \\
\mathbf{A}_i &= \begin{bmatrix} y_{i1}C_1^{3T} - C_1^{2T} \\ x_{i1}C_1^{3T} - C_1^{1T} \\ y_{i2}C_2^{3T} - C_2^{2T} \\ x_{i2}C_2^{3T} - C_2^{1T} \end{bmatrix} \quad (17)
\end{aligned}$$

3.3 Finding M2

The function has been implemented in `findM2.py`.

4 3D Visualization

4.1 Epipolar Correspondences

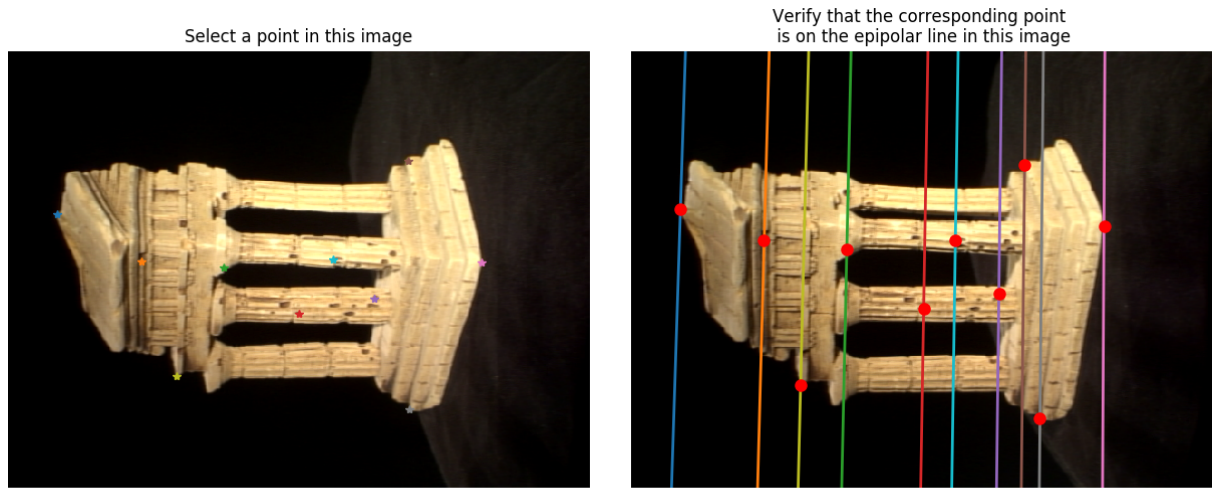


Figure 6: Epipolar Correspondences

4.2 Visualization

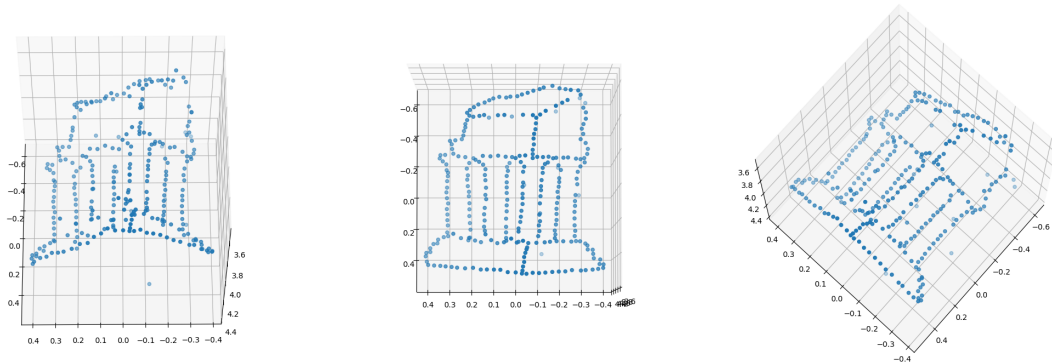


Figure 7: 3D Visualization of the Temple

5 Bundle Adjustment

5.1 The Seven Point Algorithm with RANSAC

The recovered fundamental matrix from eight-point algorithm for all the noisy points did not work well at all with triangulation. In fact, M2 could not be selected such that all points were in the positive z-axis. Therefore, there is no point of reporting a re-projection error for eight-point algorithm, which would be very high. The following re-projection errors obtained using all the noisy points.

Eight-Point Algorithm: N/A
Seven-Point Algorithm (RANSAC): 476634.5309842234

Figure 8: Re-projection Error for all Noisy Points

The following are the re-projection errors with all the inlier points, that were found using RANSAC.

Eight-Point Algorithm: 588.3440717256182
Seven-Point Algorithm (RANSAC): 1112.0337101361893

Figure 9: Re-projection Error for the Inlier Points

NOTE: *The above comparison is for the fundamental matrix recovered from the algorithm mentioned.*

RANSAC was run for up to 1,000 iterations and the following error metric was used to decide which pair of corresponding points were inlier points and tolerance was set to $1e - 3$.

$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 < tolerance \quad (18)$$

To make RANSAC run faster, the refine method was not called from within seven-point algorithm for every iteration, and was instead called with all the inlier points with the best estimated fundamental matrix, which helped increased the speed of computation.

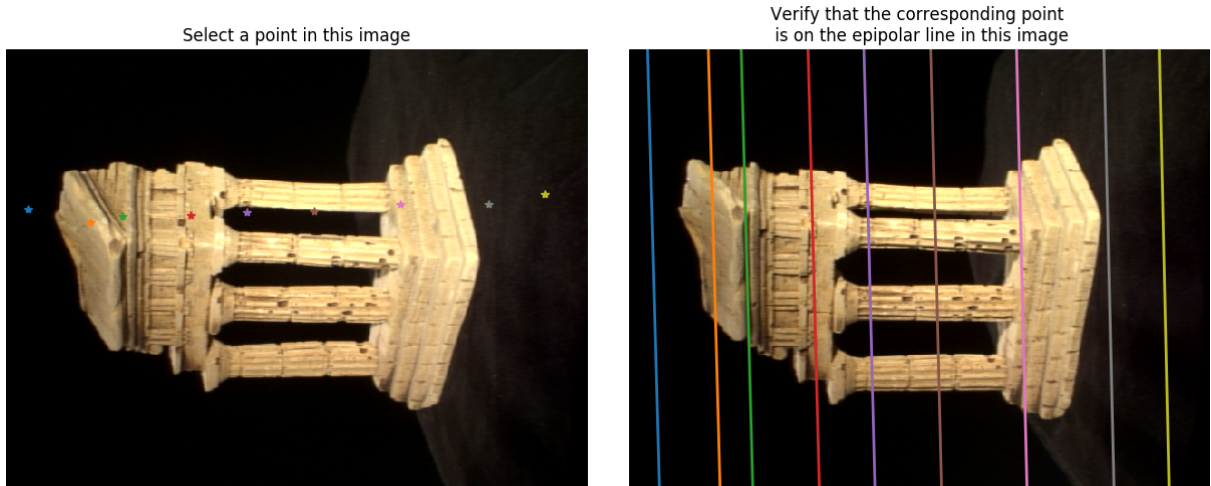


Figure 10: Seven-Point Algorithm with RANSAC Output

5.2 Rodrigues Formula

The formulae were implemented and verified with OpenCV's results as well.

5.3 Optimization

Re-projection Error for all Noisy Points: 474443.6399674097
Re-projection Error for Inlier Points (Before Optimization): 306.2297910673943
Re-projection Error for Inlier Points (After Optimization): 7.742333732219369

Figure 11: Re-projection Errors

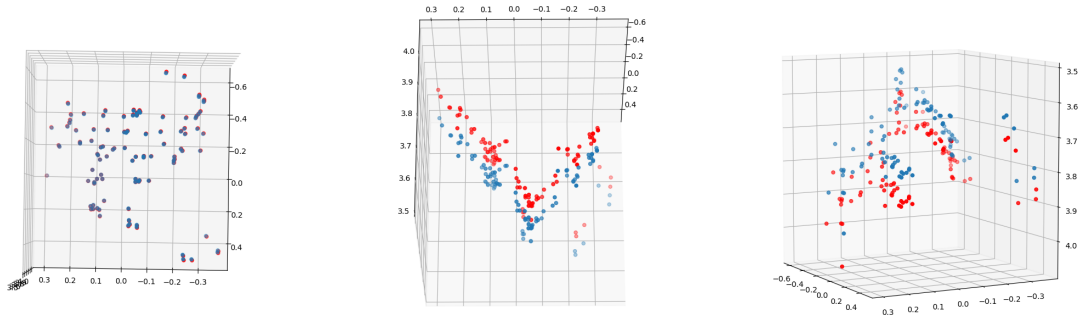


Figure 12: 3D Visualization of the Points (Front, Top and Bottom views) (Red-Before, Blue-After)

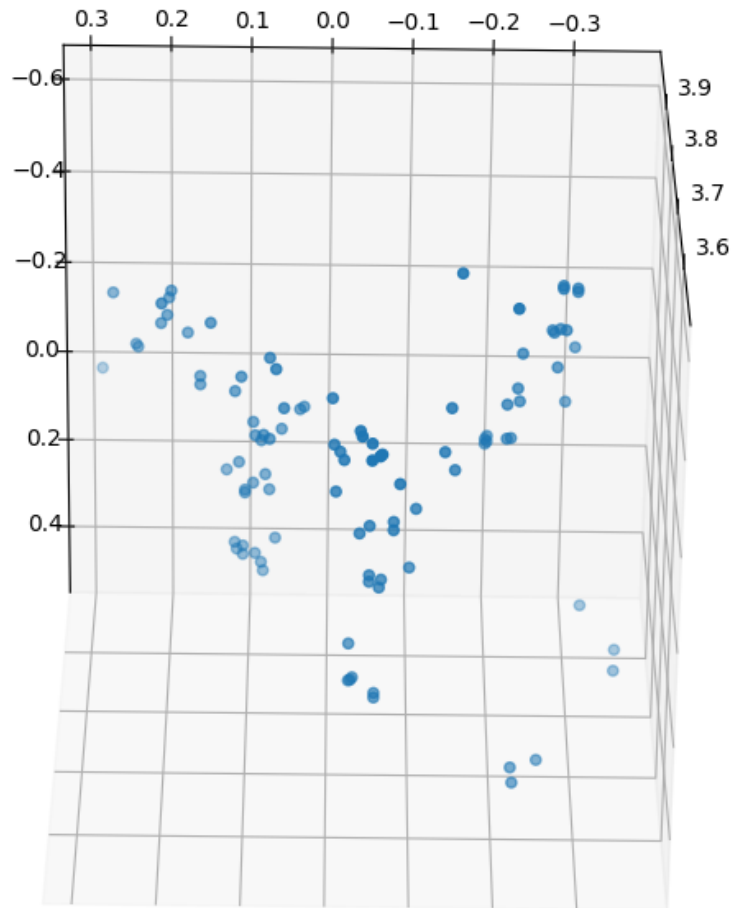


Figure 13: 3D Visualization of the Points after Optimization