

1. Fundamental Matrix

No, we are not directly optimizing the residual. The residual represents the average distance between the epipolar line of x_2 and point x_1 , and vice-versa, when viewed from both cameras. The reconstruction error is the average distance between the 3D reconstructed point projected onto one of the image planes. The objective relates to the residual since we are trying to measure an average distance metric, but it isn't exactly the same. Note that the residual is the average of terms $d_{1 \rightarrow 2}$ and $d_{2 \rightarrow 1}$. The squared residual is then an average of terms that look like

$$\frac{(x_2^T F x_1)^2}{x_1^T F^T F x_1}$$

while the SVD is minimizing

$$\frac{(x_2^T F x_1)^2}{\|F\|_F^2}$$

These are clearly different.

In essence, we followed the guidelines of the 8-point algorithm. The first step was to normalize the points. To do so, we constructed matrixes T_1 and T_2 , which were linear transformations and, upon multiplying by the matching points, normalized the points. Our matrix T has the following format, where μ_x, μ_y and σ_x, σ_y are the mean and std of all x and y coordinates for one camera:

$$\begin{bmatrix} x_{norm} \\ y_{norm} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_x} & 0 & -\frac{\mu_x}{\sigma_x} \\ 0 & \frac{1}{\sigma_y} & -\frac{\mu_y}{\sigma_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

After normalizing, we constructed matrix A as said in the Homework PDF and performed an SVD on A , taking the last column of V , since we're trying to find the vector f with norm 1 that lies in the nullspace of A . We reshape the last column of V to the matrix F^* , which is not rank 2.

Upon obtaining F^* , we need to convert the matrix to rank 2, so we use Eckhart Young theorem, which just does SVD on F^* and sets the smallest singular value to 0.

After obtaining rank 2 version of F , we need to denormalize the matrix, $F \leftarrow T_2^T F T_1$.

For the residual error, we followed the correction on Piazza: $d_{1 \rightarrow 2} = \frac{x_2^T F x_1}{\|F x_1\|}$ and $d_{2 \rightarrow 1} = \frac{x_1^T F^T x_2}{\|F^T x_2\|}$.

Results for House:

$$F = \begin{bmatrix} -6.20566675 \times 10^{-8} & 1.45875690 \times 10^{-6} & 1.33031446 \times 10^{-4} \\ 6.45620291 \times 10^{-6} & -5.56578733 \times 10^{-7} & -1.65983285 \times 10^{-2} \\ -1.04085120 \times 10^{-4} & 1.52331793 \times 10^{-2} & -1.01024540 \times 10^{-2} \end{bmatrix}$$

$$\text{Residual} = 2.638036824523416 \times 10^{-6}$$

Results for Library:

$$F = \begin{bmatrix} -4.86210796 \times 10^{-8} & 9.98237781 \times 10^{-7} & -1.47341026 \times 10^{-4} \\ -5.80698246 \times 10^{-6} & -5.65060960 \times 10^{-8} & 1.07254893 \times 10^{-2} \\ 1.38164930 \times 10^{-3} & -9.63586245 \times 10^{-3} & -2.60674702 \times 10^{-1} \end{bmatrix}$$

$$\text{Residual} = 1.0144671451144764 \times 10^{-5}$$

2. Find Rotation Translation

In total, there are four solutions. Let $E = U\Sigma V^T$, where $U = [u_1 u_2 u_3]$. Also, let

$$R_{90^\circ} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then we have

$$R = UR_{90^\circ}V^T, R = UR_{90^\circ}^T V^T$$

and

$$t = \pm u_3$$

This in total makes four possibilities. In our code, if the determinant of R ended up being -1 , we just set R to be negative, since rotation matrices must have determinants of 1.

3. Find 3D Points

We first solve the homogeneous system for $A\mathbf{X} = 0$, where \mathbf{X} is a vector of X, Y, Z, W and A is a 4x4 matrix for x_1, y_1, x_2, y_2 with coefficients coming from the Triangulation equations found in the Homework PDF.

To solve the homogeneous system, we took the SVD and took the last column of V , and normalized the vector by setting $W = 1$. Thus, we received our 3D prediction \mathbf{X} . Then, we projected \mathbf{X} back to both camera image planes by doing $\mathbf{x}_1 = P_1\mathbf{X}$ and $\mathbf{x}_2 = P_2\mathbf{X}$.

To calculate the reconstruction error, we look the L2 norm between our projected points \mathbf{x}_1 and \mathbf{x}_2 and our real points $(x_1, y_1), (x_2, y_2)$. Then we summed it up over all matching points, dividing the reconstruction error by two times the number of matching points.

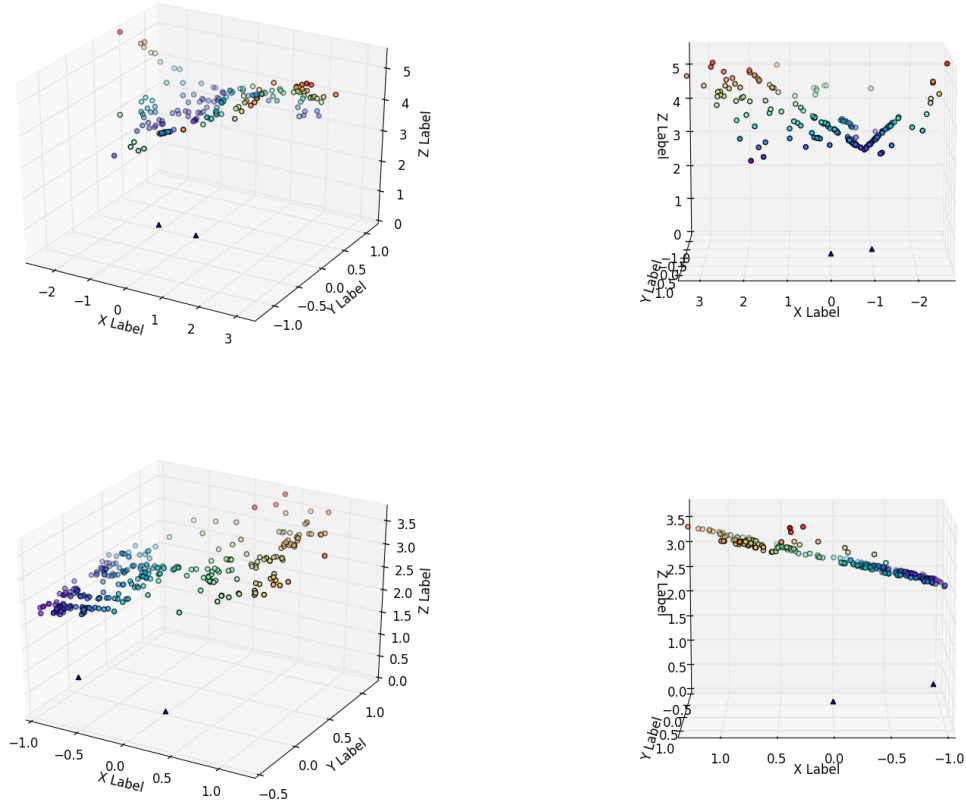
Results for House:

$$\text{Reconstruction error} = 0.2378123445095882$$

Results for Library:

$$\text{Reconstruction error} = 0.31136431177829127$$

4. Plot 3D



These top pair of figures show the 3D point locations for the house from two different views; the corner of the house is clearly visible. The bottom set of figures show the 3D point locations for the library.