CV 2021 HW4 Report

Group 24

0610122 李思賢, 309553049 陳怡安

Introduction

In this assignment, we implemented the Structure from Motion (SfM) to get 3D models from images. SfM technique reconstructs the 3D structure of a scene or an object by utilizing a sequence of 2D images which are created by moving cameras.

Implementation

1. **Find out correspondence across images**
   We have done this part in the previous assignment, so it's basically the same. First, find out the keypoints and corresponding feature descriptions in each image by SIFT. Then calculate the L2-norm between the feature descriptions, and apply the ratio test to avoid ambiguity.

2. **Estimate the fundamental matrix across images (normalized 8 points)**
   We used normalized 8-point algorithm with RANSAC to find out the fundamental matrix. The steps are as follows:
   (1) Normalized the original input images to [-1, 1]x[-1, 1] by finding T and T' matrices.
   $$\widetilde{x} = Tx \quad \widetilde{x'} = T'x'$$
   (2) For every iteration in RANSAC, we randomly picked 8 pairs of matching points to find the matrix $\widetilde{F}$ that satisfies $\widetilde{x'}\widetilde{F}\widetilde{x} = 0$.
   We can rewrite the equation to $Af = 0$, where
   $$A = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_8x_8 & x'_8y_8 & x'_8 & y'_8x_8 & y'_8y_8 & y'_8 & x_8 & y_8 & 1 \end{bmatrix}$$
   $$f = \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{21} & \dots & f_{33} \end{bmatrix}^T$$
   Instead of solving $Af = 0$ directly, we find the f that minimize $||Af||$ using SVD and enforce $det(\widetilde{F}) = 0$.
   (3) Find out the best $\widetilde{F}$ with the most inliners. We take all the good matching point pairs $(\widetilde{x}, \widetilde{x'})$ to calculate the error.
   If $\widetilde{x'}\widetilde{F}\widetilde{x'} < threshold$, then it counts as an inliner.
   (4) Denomarlized the $\widetilde{F}$ by $F = T'^T\widetilde{F}T$.

3. **Draw the interest points on you found in step.1 in one image and the corresponding epipolar lines in another**
   We draw the interest points on the both images, and draw the epipolar lines only on the second images. The epipolar lines can be found by $l = Fx$.

4. **Get 4 possible solutions of essential matrix from fundamental matrix**
   We got the essential matrix E by multiplication of K1, fundamental matrix and K2. According to the formula: E = $[t]_x R$, where $[t]_x$ is a skew-symmetric matrix with 2 equal singular values and another is 0, and the multiplication of the R does not change the singular values.Thus, the singular values of E are also 2 equal values, and the other is 0.
   However, due to noises, we need to force the singular values of E to 2 equal values and one 0, based on the above property. Then, we get the 4 possible solutions by the formula:

$$\begin{aligned} \mathbf{P}_2 &= \left[\mathbf{UWV}^\top \mid +\mathbf{u}_3\right] \\ \mathbf{P}_2 &= \left[\mathbf{UWV}^\top \mid -\mathbf{u}_3\right] \\ \mathbf{P}_2 &= \left[\mathbf{UW}^\top\mathbf{V}^\top \mid +\mathbf{u}_3\right] \\ \mathbf{P}_2 &= \left[\mathbf{UW}^\top\mathbf{V}^\top \mid -\mathbf{u}_3\right] \end{aligned} \qquad \mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   Where U,V^T and u_3 are the results from SVD of E.

5. **Find out the most appropriate solution of essential matrix**
   We utilized the multiplication of intrinsics matrix and each possible solution of essential matrix to get 3D points by triangulation:

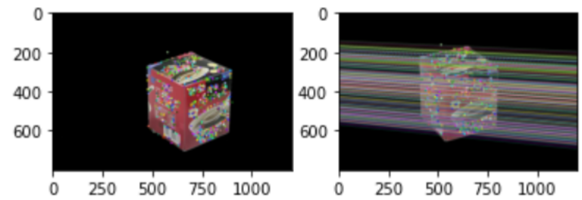
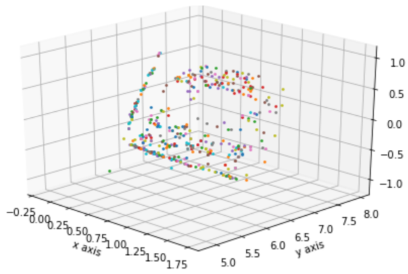
$$x = \mathbf{P}X \quad x' = \mathbf{P}'X$$
$$\Downarrow$$
$$x = w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \mathbf{p}_3^\top \end{bmatrix} \qquad \mathbf{A}X = 0 \quad A = \begin{bmatrix} u\mathbf{p}_3^\top - \mathbf{p}_1^\top \\ v\mathbf{p}_3^\top - \mathbf{p}_2^\top \\ u'\mathbf{p}_3'^\top - \mathbf{p}_1'^\top \\ v'\mathbf{p}_3'^\top - \mathbf{p}_2'^\top \end{bmatrix}$$
$$x' = w\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \quad \mathbf{P}' = \begin{bmatrix} \mathbf{p}_2'^\top \\ \mathbf{p}_2'^\top \\ \mathbf{p}_3'^\top \end{bmatrix}$$

   We solved AX = 0 via SVD to get the 3D points X. After that, we choose the solution with most of the 3D points in front of the cameras according to the conditional formula: $(X - \mathbf{C}) \cdot R(3,:)^\top > 0$
   If the condition is true, it means the 3D point is in front of the camera and otherwise not.

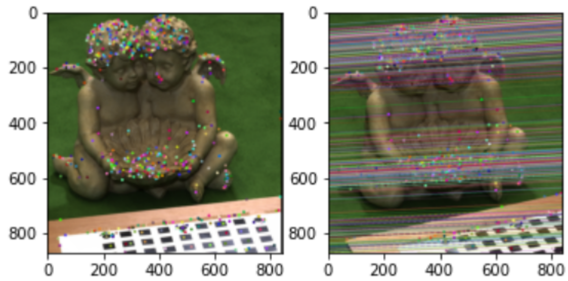
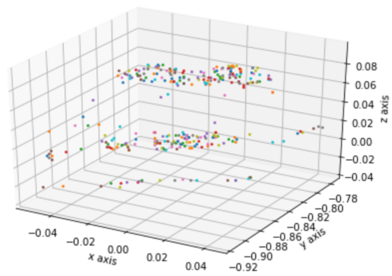
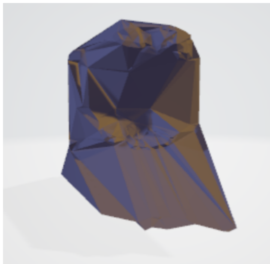
6. **Use texture mapping to get a 3D model**
   In this step, we used the 3D points in the world coordinate and the given matlab code to reconstruct the 3D model.

Experimental results

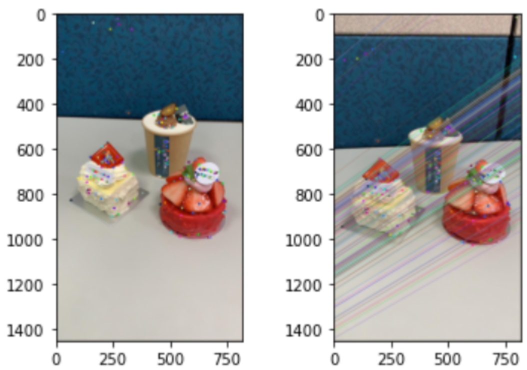
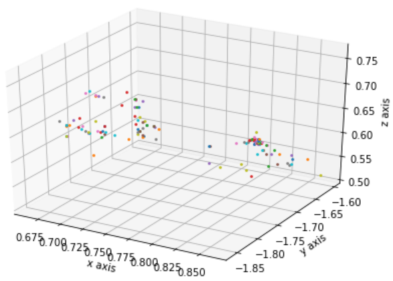
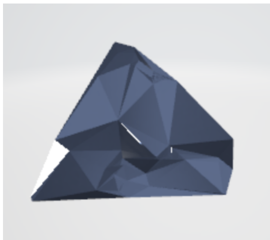
Mesona Box

Interest Points & Epipolar Lines	3D Points	3D Model
		

Statue

Interest Points & Epipolar Lines	3D Points	3D Model
		

Our Data

Interest Points & Epipolar Lines	3D Points	3D Model
		

Discussion

3D models are sensitive to bad feature matches and outliers. It is not easy to choose a good threshold for feature matching and a suitable threshold to distinguish between inliers and outliers. If there are outliers, there will be some spikes in the 3D model, the examples of the mesona box is shown below.

According to our experiment, we got the best results when the feature matching ratio is around 0.5 and the inlier threshold around 0.07.

Conclusion

Using SfM technique, we successfully reconstructed the 3D model from 2D images. The epipolar lines and the 3D models look authentic. Also we found threshold choosing in the ratio test is important and may greatly change the experiment results.

Work assignment plan

code: 李思賢 & 陳怡安
report: 李思賢 & 陳怡安