

Structure From Motion

Chao Ni

November 23, 2020

1 Feature Extraction and initialization with epipolar geometry

To start with feature extraction, we need to first select two images for initialization. We pick the first image and the last image such that they have the biggest baseline and also have a significant overlap. We use SIFT to collect features and apply 8-point RANSAC to compute the fundamental matrix. One trick here is to make the first camera has an identity position, i.e. in the normalized coordinate. We assume two cameras have the same intrinsic matrix, then the essential matrix can be computed by:

$$E = K^T F K \quad (1)$$

When computing the fundamental matrix, We also enforce the singularity constraint. The result of using the first image and the last image can be shown in Fig. 1

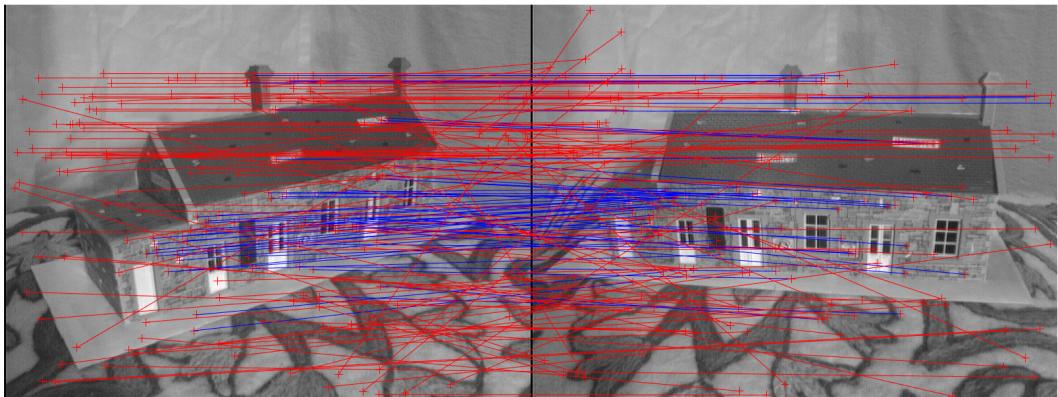


Figure 1: Initialization matches: blue lines are inliers, red lines are outliers. (image 0 and 4)

We also give the epipolar geometry about the initialization as given in Fig. 2

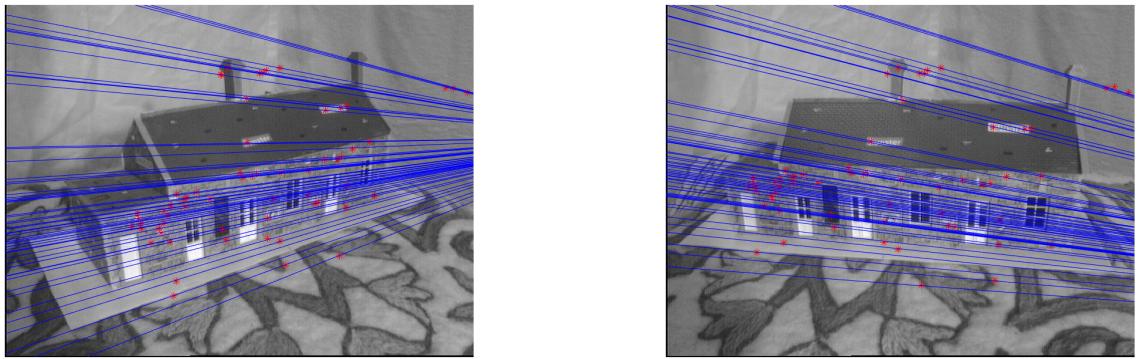
From inliers, we can compute the 3D positions for the matched points after triangulating. Therefore, we will have 2D-3D correspondence and can compute the projection matrix for other cameras.

2 More views

With more images coming into stage, we first extract the SIFT features of the new image, and find those features matching with the validated inlier-labeled matches of the first image. For those new matches, we find the best projection matrix again using 6-point RANSAC algorithm with the indicator being the projection error. After the projection matrix is computed, we need to further require the rotation matrix has positive determinant 1, if the determinant is negative, then we flip signs. The step-by-step results are shown in Fig. 3, 4, 5.

3 Camera

After correcting the sign of the projection matrix, we plot the sparse reconstruction of the scene as well as the position of the camera as shown in Fig. 6



(a) Epipolar line of the first image

(b) Epipolar lines of the last image

Figure 2: Epipolar Geometry

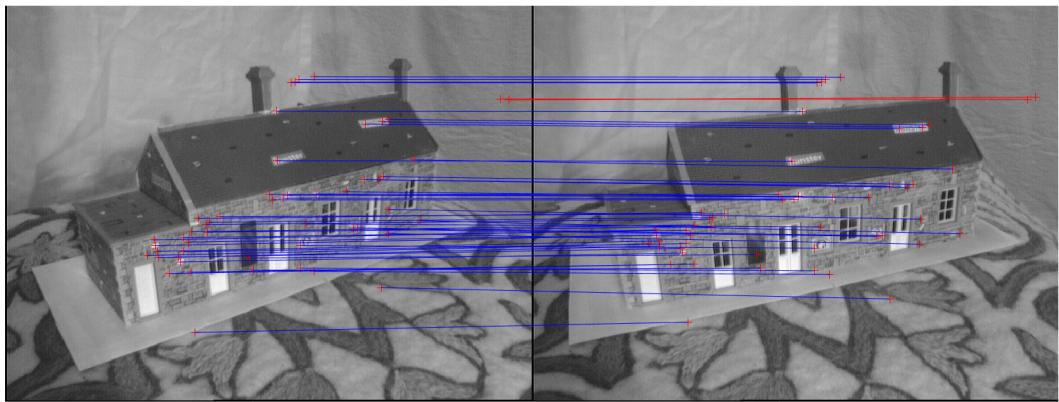


Figure 3: Blue lines are inliers, red lines are outliers. After adding image 1, the number of outliers significantly decreases. (image 0 and 1)

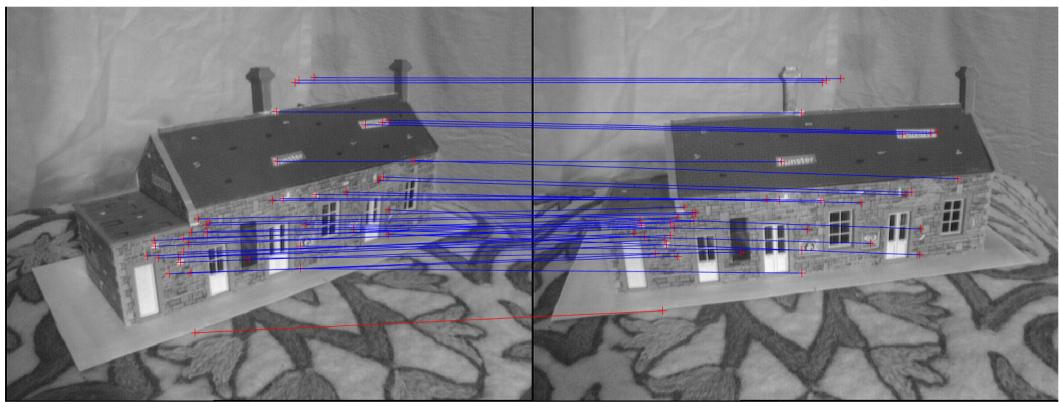


Figure 4: After adding image 2, we can see the number of outliers are further decreased. (image 0 and 2)



Figure 5: After adding image 23, based on our inlier/outlier criteria, there are no outliers in this match. (image 0 and 3)

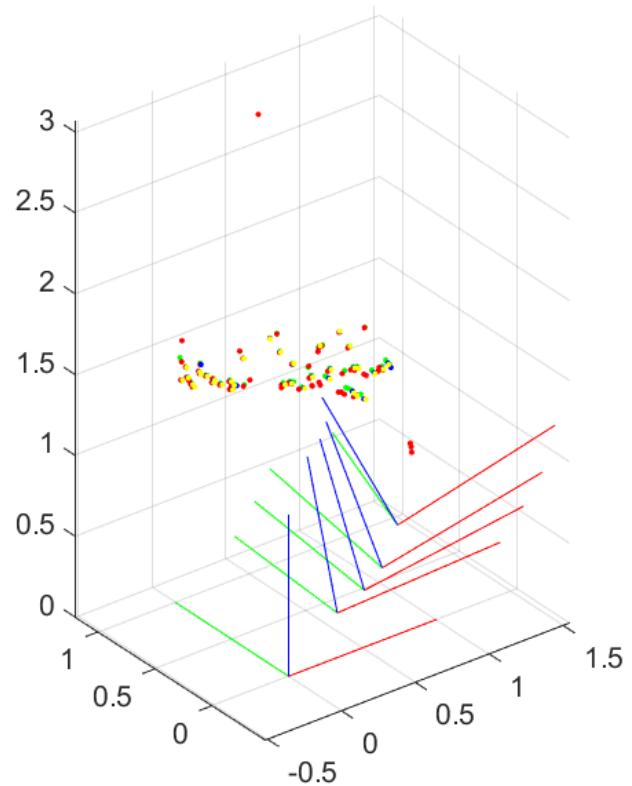


Figure 6: Triangulated points and camera poses plotted in 3D. Red dots are the triangulated points from the initialization, green dots are the triangulated points from the first additional image, blue dots are from the second additional image and yellow dots are from the last additional image.

4 Dense Reconstruction

To get a dense reconstruction, we first rectify the first image and the second image as shown in Fig. 7. Using the code from the stereo matching, we compute the disparity map. Because we don't have RGB picture here, we adjust the code to make it also work for gray images. Fig. 8 shows us the results.

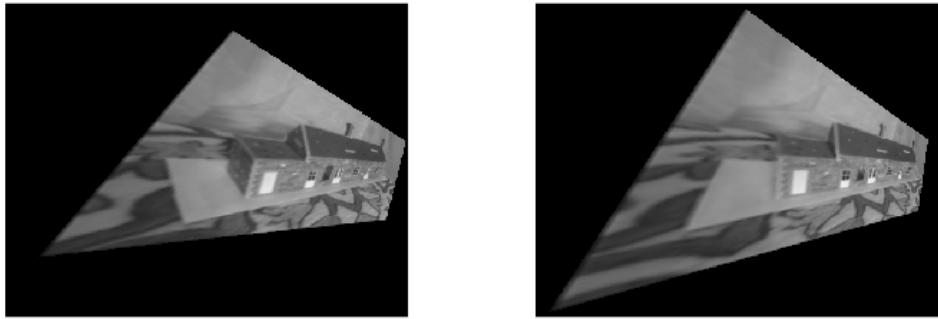
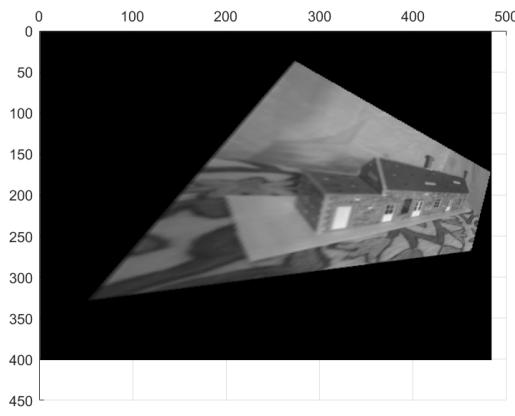
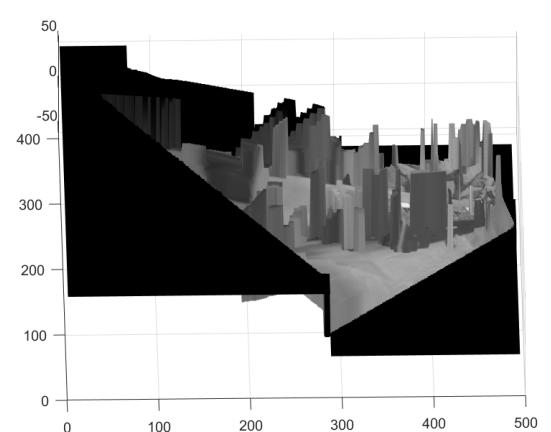


Figure 7: Rectified picture with image 0 and 1.



(a) Dense reconstruction: sky down view.



(b) Dense reconstruction: 3D model

Figure 8: Dense reconstruction