COMP90086 – Computer Vision – 2022 S2

Assignment 3 – Fundamental Matrix Calculation

Student: Jiahao Chen   ID: 1118749

I. Explanation of each step & Design choices

Finding keypoints and correspondences: SIFT and FLANN are applied, which is referred from workshop 8.

a. It is better to shift and scale the pixel coordinates to achieve numeric stability. In this assignment, the center pixel is shifted to the origin coordinate. Scale is not applied for simplicity since the images provided do not have high resolutions.

b. Forming design matrix (D) is an important step when computing fundamental matrix (F) because DF = 0, which makes it simple to get F.

$$\begin{pmatrix} p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \\ p_1q_1 & p_2q_1 & q_1 & p_1q_2 & p_2q_2 & q_2 & p_1 & p_2 & 1 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

c. Performing SVD on the design matrix gives the null space of D, which is the vector of F values. The SVD function in numpy.linalg is used.

d. Simply reshaping it to 3 * 3 gives the draft fundamental matrix.

e. Make the draft F have determinant = 0 so that the matrix becomes singular. SVD is also applied to help complete this step.

f. Detect inliers by calculating the distance between points and corresponding epipolar lines. The allowed error of the distance is set to 2 pixels. Since scaling is not applied, there is no need to normalise the error. Correspondences which have error less than 2 in both left and right images are considered as inliers and the rest are outliers.

g. Use RANSAC to compute a good F. The number of loops required to run is calculated using probability theories. Assuming that 50% of all correspondences are inliers (I), if the probability of sampling 8 inliers is required to be over 99% (P), the number of iterations is therefore = log(1-P)/log(1-I^8) = 1177 times. 50% is a relatively conservative assumption and the result should be enough to get a good F.

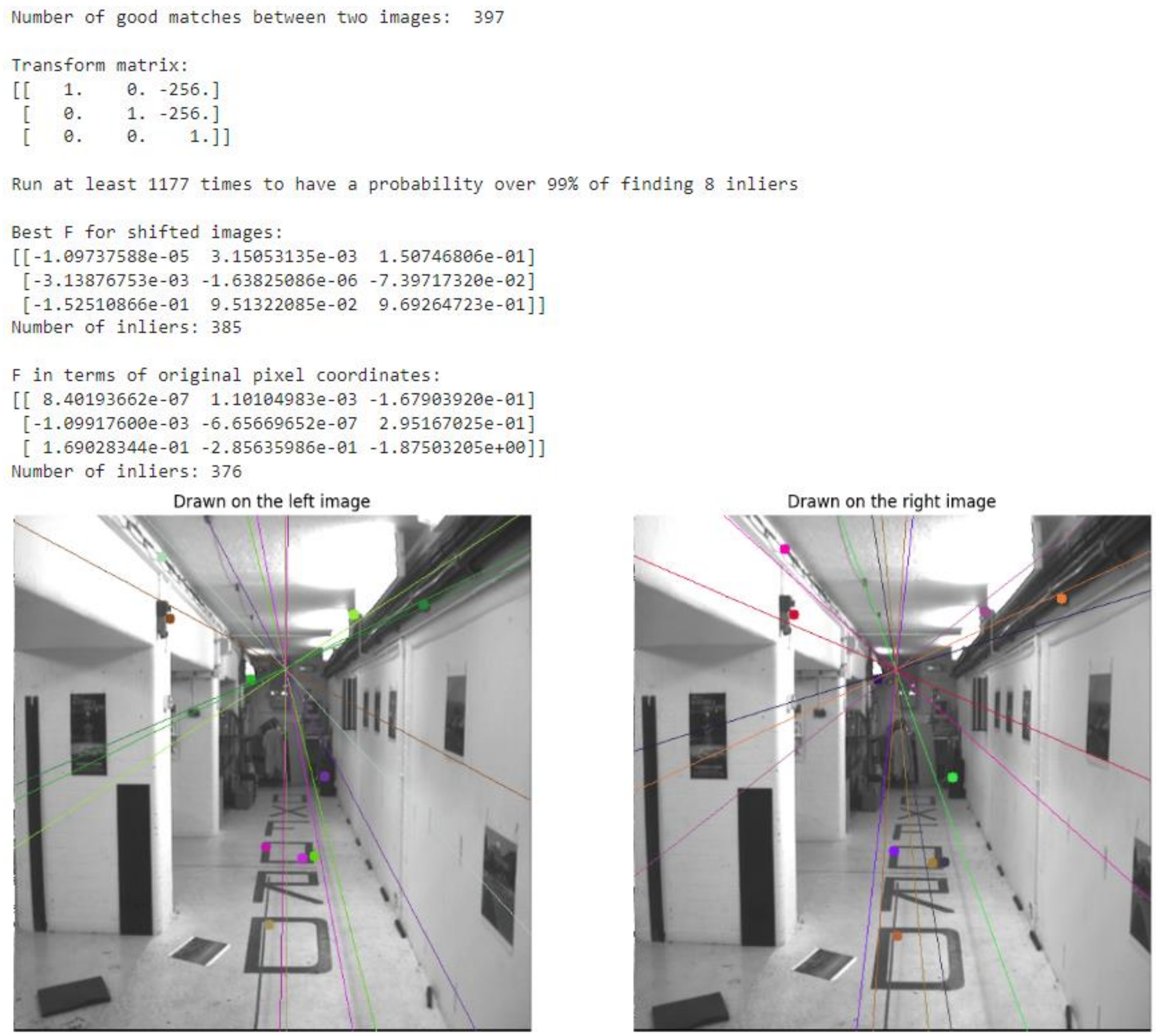
h. Re-estimate F using the inliers computed by RANSAC should make F have better generalisation.

i. Compute F in terms of the original pixel coordinates by using the transform matrix applied before for shifting the pixels. Assuming the transform matrix for left image is TL and TR for the right one, original F = Transpose (TR) * F * TL, which gives the final result.

II. Performance on different images

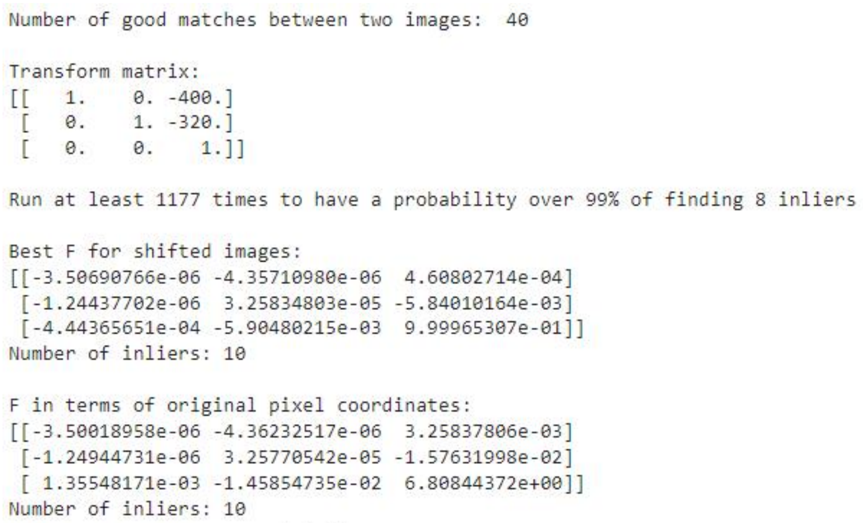
All of the steps above are encapsulated as a single function which is performed on all image pairs from kusvod2. Results are presented at the end of the Jupyter notebook. It is obvious that some of them work much better than others. For example, as shown in figure 1, a huge proportion of matches in the images of corridor are detected as inliers and draws a good result. By contrast, the images presented in figure 2 have a poor performance.

It is mainly because of the difference of the camera positions between 2 images. Apparently, cameras in figure 1 are close to each other and have almost the same direction. While in figure 2, the left camera has a very lateral view and the right one is in front of the scene. It turns that huge difference of camera positions will significantly reduce the performance of the algorithm. Besides, some images have a huge number of pixels with similar color so that there are many noises, which also reduces the performance.

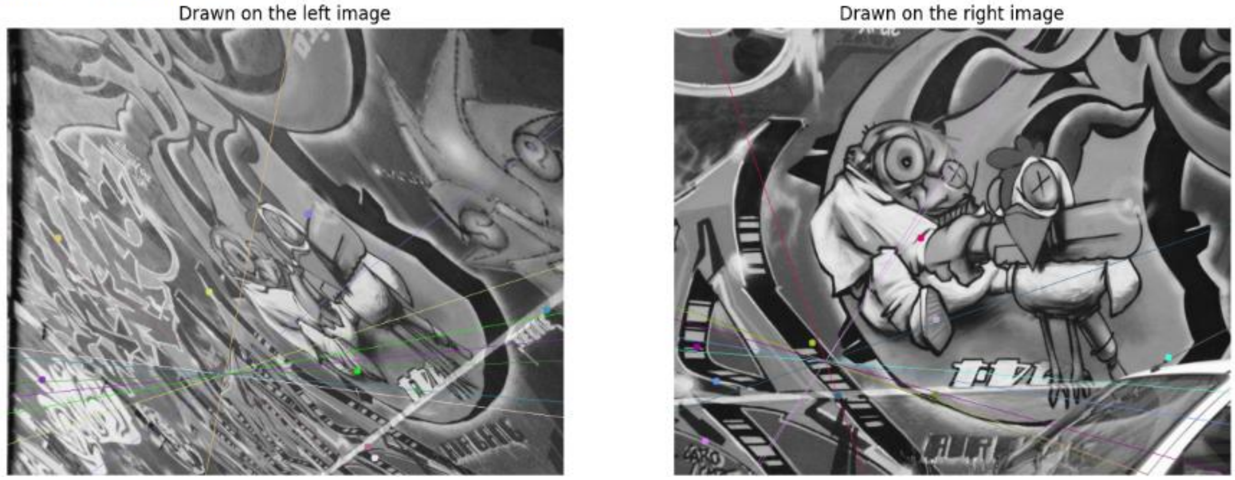


```
Number of good matches between two images:  397

Transform matrix:
[[   1.    0. -256.]
 [   0.    1. -256.]
 [   0.    0.    1.]]

Run at least 1177 times to have a probability over 99% of finding 8 inliers

Best F for shifted images:
[[-1.09737588e-05  3.15053135e-03  1.50746806e-01]
 [-3.13876753e-03 -1.63825086e-06 -7.39717320e-02]
 [-1.52510866e-01  9.51322085e-02  9.69264723e-01]]
Number of inliers: 385

F in terms of original pixel coordinates:
[[ 8.40193662e-07  1.10104983e-03 -1.67903920e-01]
 [-1.09917600e-03 -6.65669652e-07  2.95167025e-01]
 [ 1.69028344e-01 -2.85635986e-01 -1.87503205e+00]]
Number of inliers: 376
```

Figure 1: corrA & corrB



```
Number of good matches between two images:  40

Transform matrix:
[[   1.    0. -400.]
 [   0.    1. -320.]
 [   0.    0.    1.]]

Run at least 1177 times to have a probability over 99% of finding 8 inliers

Best F for shifted images:
[[-3.50690766e-06 -4.35710980e-06  4.60802714e-04]
 [-1.24437702e-06  3.25834803e-05 -5.84010164e-03]
 [-4.44365651e-04 -5.90480215e-03  9.99965307e-01]]
Number of inliers: 10

F in terms of original pixel coordinates:
[[-3.50018958e-06 -4.36232517e-06  3.25837806e-03]
 [-1.24944731e-06  3.25770542e-05 -1.57631998e-02]
 [ 1.35548171e-03 -1.45854735e-02  6.80844372e+00]]
Number of inliers: 10
```



Drawn on the left image

Drawn on the right image

Figure 2: graffA & graffB