

# Computer Vision Exercise 4: Model Fitting

Soomin Lee (leesoo@student.ethz.ch)

October 30, 2020

## 1. Line Fitting

First, two points are sampled in order to get the model  $y = k * x + b$ . Then the distance to the line is calculated using the following equation:

$$d = \frac{|-y + kx + b|}{\sqrt{1 + k^2}}. \quad (1)$$

Then the points that have smaller distances than the threshold are considered as "inliers" and the model was fitted using the inliers. The procedure is repeated for a certain number of iterations and the model with the most inliers is saved as the best model.

The resulting lines are plotted in Figure 1, and the error measured on the true inliers are shown in Table 1. One can observe that the line obtained using RANSAC yields a significantly better result than that of least squares method, and that it is very close to the ground truth model.

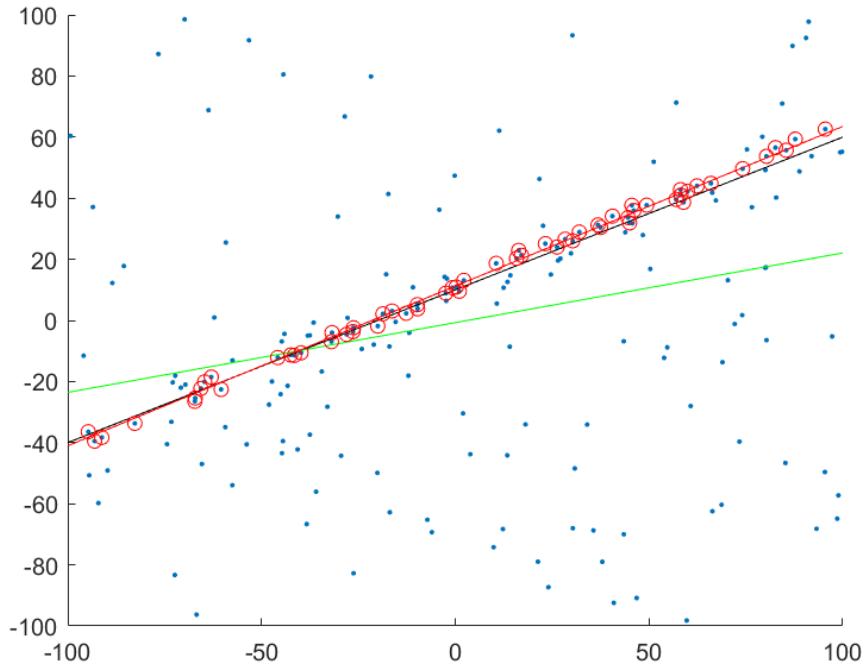


Figure 1: Comparison between the ground truth model (black), Least square (green) and RANSAC line fitting (red) result. Red circles indicate the inliers for the RANSAC model.

err_real	err_ls	err_ransac
44.6503	193.4142	46.0726

Table 1: Error measured on the true inliers for each model.

## 2. Fundamental Matrix Estimation

**Normalization** The points are centralized and normalized for a better numerical conditioning. After converting the representation of the points to  $(x, y, 1)$ , we calculate the mean values of each x and y and subtract it from all the points in order to centralize them. Then using root mean squared distance, we scale the data.

**Calculation of fundamental matrix** To get the fundamental matrix from 8 or more correspondences, we construct an equation in  $Ax = 0$  form. For each point, the constraint can be expressed as in Equation (2) and we can eventually stack up the 8 or more constraints to construct a matrix A.

$$\begin{aligned} x'^T F x &= 0 \\ a * \text{vec}(F) &= [xx' \ xy' \ x \ yx' \ yy' \ y \ x' \ y' \ 1] * \text{vec}(F) = 0 \end{aligned} \quad (2)$$

**Enforcement of singularity constraint** The fundamental matrix we obtained is not guaranteed to satisfy the constraint  $\det(A) = 0$ . Therefore, we can enforce the constraint by performing the singular value decomposition on  $F$  and setting the last element of  $S$  as 0.

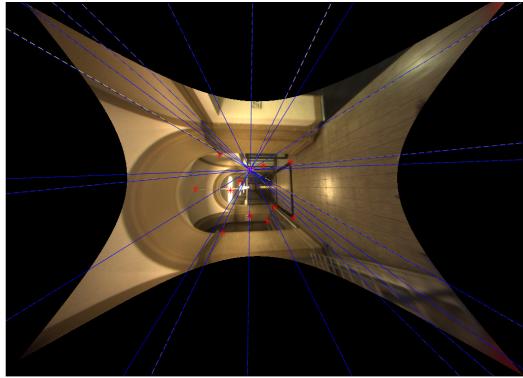
**De-normalization** As we normalized the points in the beginning, we need to de-normalize the points:  $F = T_1^T FT_2$ .

**Result** The results are shown in Figure 2 for 3 different dataset. The epipolar lines for non-singular fundamental matrix and singular fundamental matrix do not differ dramatically and hence they look almost overlapping in the Figure 2. However, if we take a closer look at the first dataset, where both epipoles are present in the images, one can notice the difference. In Figure 3, the yellow dotted lines do not intersect at one point because the singularity constraint is not enforced. However, the blue lines intersect with each other at a single point which is the epipole. The non-singular fundamental matrix, denoted as  $F$ , and the singular fundamental matrix, denoted as  $F_h$ , are calculated as follows for each dataset:

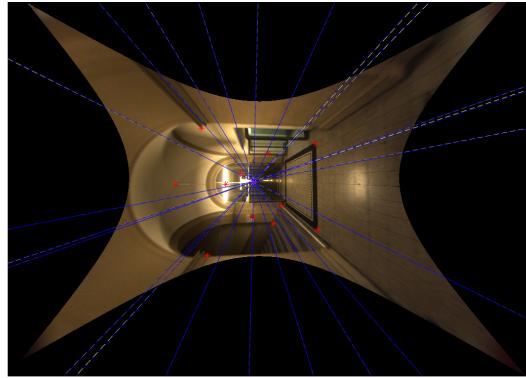
$$\begin{aligned} \text{Dataset 1 : } F &= \begin{bmatrix} 0.0000 & 0.0001 & -0.0528 \\ -0.0001 & 0.0000 & 0.0684 \\ 0.0547 & -0.0719 & -0.5801 \end{bmatrix}, F_h = \begin{bmatrix} 0.0000 & 0.0001 & -0.0528 \\ -0.0001 & 0.0000 & 0.0683 \\ 0.0548 & -0.0719 & -0.5574 \end{bmatrix} \\ \text{Dataset 2 : } F &= \begin{bmatrix} 0.0000 & -0.0000 & -0.0005 \\ -0.0000 & 0.0000 & -0.0010 \\ -0.0007 & 0.0024 & 1.3376 \end{bmatrix}, F_h = \begin{bmatrix} 0.0000 & -0.0000 & -0.0006 \\ -0.0000 & 0.0000 & -0.0010 \\ -0.0007 & 0.0024 & 1.3497 \end{bmatrix} \\ \text{Dataset 3 : } F &= \begin{bmatrix} -0.0000 & 0.0000 & 0.0009 \\ 0.0000 & 0.0000 & -0.0040 \\ -0.0006 & 0.0034 & -0.0968 \end{bmatrix}, F_h = \begin{bmatrix} -0.0000 & 0.0000 & 0.0009 \\ 0.0000 & 0.0000 & -0.0040 \\ -0.0006 & 0.0034 & -0.1010 \end{bmatrix} \end{aligned} \quad (3)$$

## 3. Feature Extraction and Matching

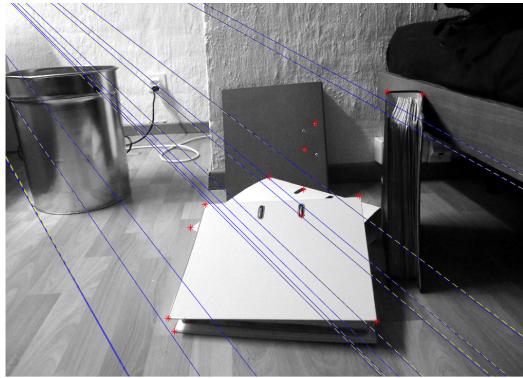
Using VLFeat, SIFT features were extracted from each image pair and matched between the two images. The result is shown in Figure 4 - 6.



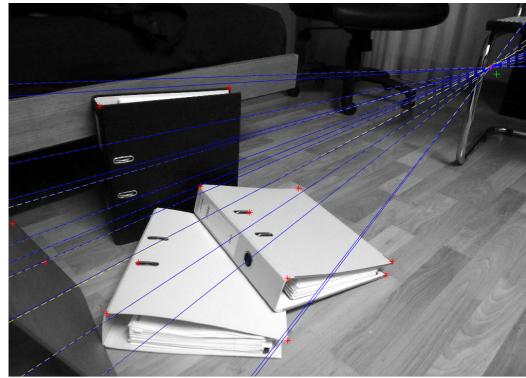
(a) Image 1 from Dataset 1



(b) Image 2 from Dataset 1



(c) Image 1 from Dataset 2



(d) Image 2 from Dataset 2

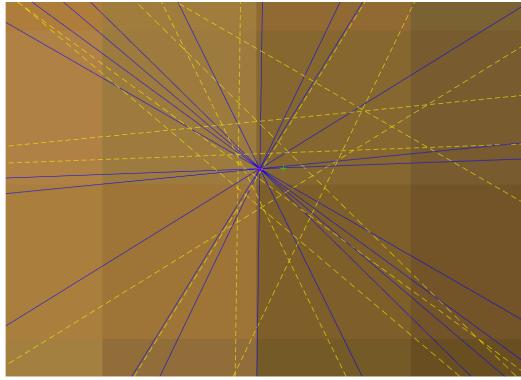


(e) Image 1 from Dataset 3

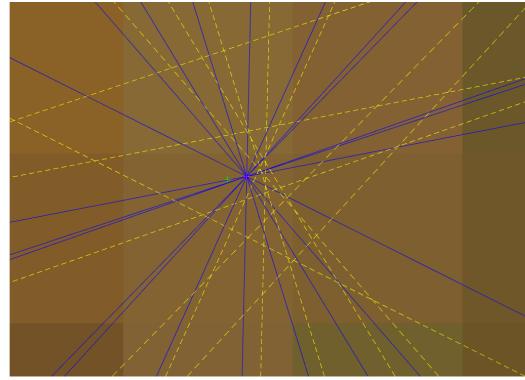


(f) Image 2 from Dataset 3

Figure 2: Red \* points indicate manually clicked points. Yellow dotted lines are epipolar lines for non-singular fundamental matrix  $F$ , and blue lines are epipolar lines for singular fundamental matrix  $\hat{F}$ . Epipoles are marked as a green + point and a magenta + point for non-singular fundamental matrix and singular fundamental matrix, respectively, if they exist within the image.



(a) Close-up image 1 from Dataset 1



(b) Close-up image 2 from Dataset 1

Figure 3: Yellow dotted lines are epipolar lines for non-singular fundamental matrix  $F$ , and blue lines are epipolar lines for singular fundamental matrix  $\hat{F}$ . Epipoles are marked as a green + point and a magenta + point for non-singular fundamental matrix and singular fundamental matrix, respectively.

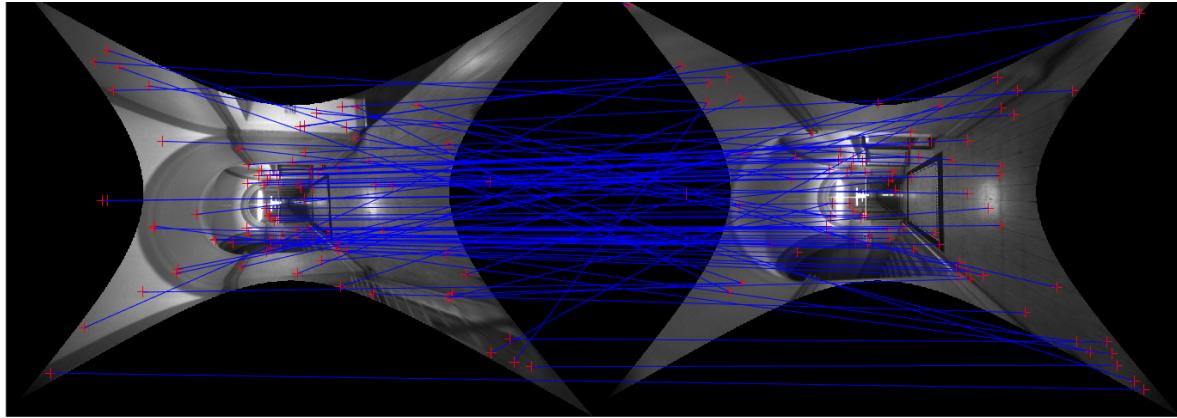


Figure 4: Matching SIFT features for Dataset 1, obtained using VLFeat

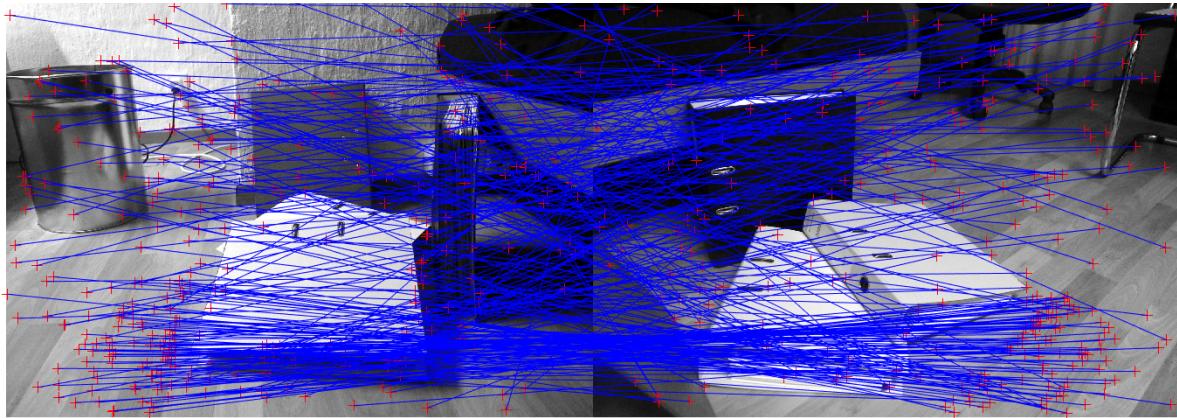


Figure 5: Matching SIFT features for Dataset 2, obtained using VLFeat

#### 4. Eight-Point RANSAC

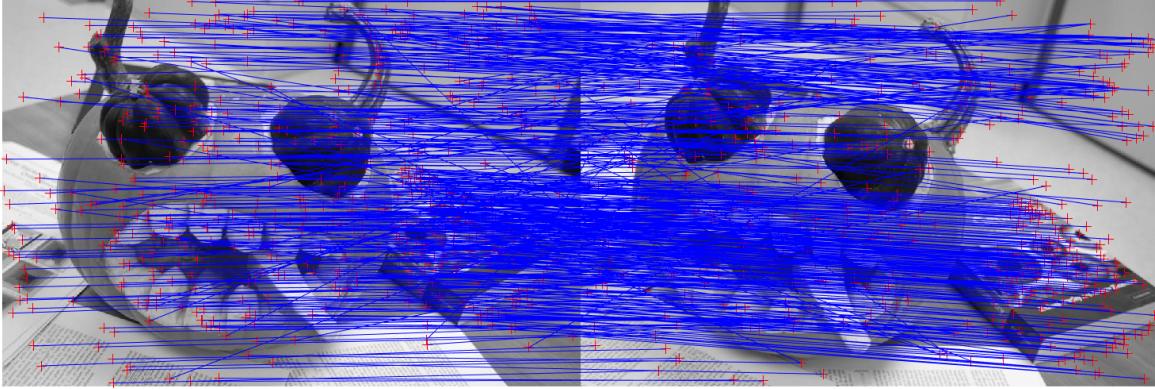


Figure 6: Matching SIFT features for Dataset 3, obtained using VLFeat

**Simple RANSAC** For the simple RANSAC, we fix the number of iteration as 1000. Then in each loop, we randomly select 8 points and estimate the fundamental matrix. Next, calculate the errors with the fundamental matrix, which is defined as the average of the perpendicular distances of the points to the epipolar lines in the two images. The distance between a line with coefficients of  $(A, B, C)$  and a homogeneous point  $(x, y, 1)$  can be obtained using Equation (4). Then we can get the number of inliers of the current fundamental matrix and choose the one with the most inliers as the best model. The result of the number of inliers and the mean error of inliers are shown in Table 2, with different thresholds. The results are visualized in Figure 7 and 8. From both the figures and the numbers in Table 2, we can see that only the good matches are kept as we decrease the error threshold for inliers and thus the number of inliers and the mean error of inliers both decrease.

$$d = \frac{|Ax + By + C|}{\sqrt{A^2 + B^2}} \quad (4)$$

Threshold	Dataset 1		Dataset 2		Dataset 3	
	# of inliers	mean error	# of inliers	mean error	# of inliers	mean error
10	39	4.4739	53	4.4475	319	3.7543
5	30	2.0469	23	2.4566	193	1.9610

Table 2: Number of inliers and mean error for different threshold, using simple RANSAC.

**Adaptive RANSAC** In adaptive RANSAC, we do not fix the number of iterations, but we terminate it after  $M$  trials if we know with a probability  $p$  that at least one of the random samples from these  $M$  trials is free from outliers. Therefore, it was implemented using a `while` loop with the condition  $M > i$ , where  $i$  is the number of iterations done. By comparing Table 2 and 3, one can confirm that in all cases the number of inliers have increased when using the adaptive RANSAC compared to the simple RANSAC. The resulting matches of SIFT features that are considered as the inliers are visualized in Figure 9. We also can see that there are more matches present compared to the Figure 7.

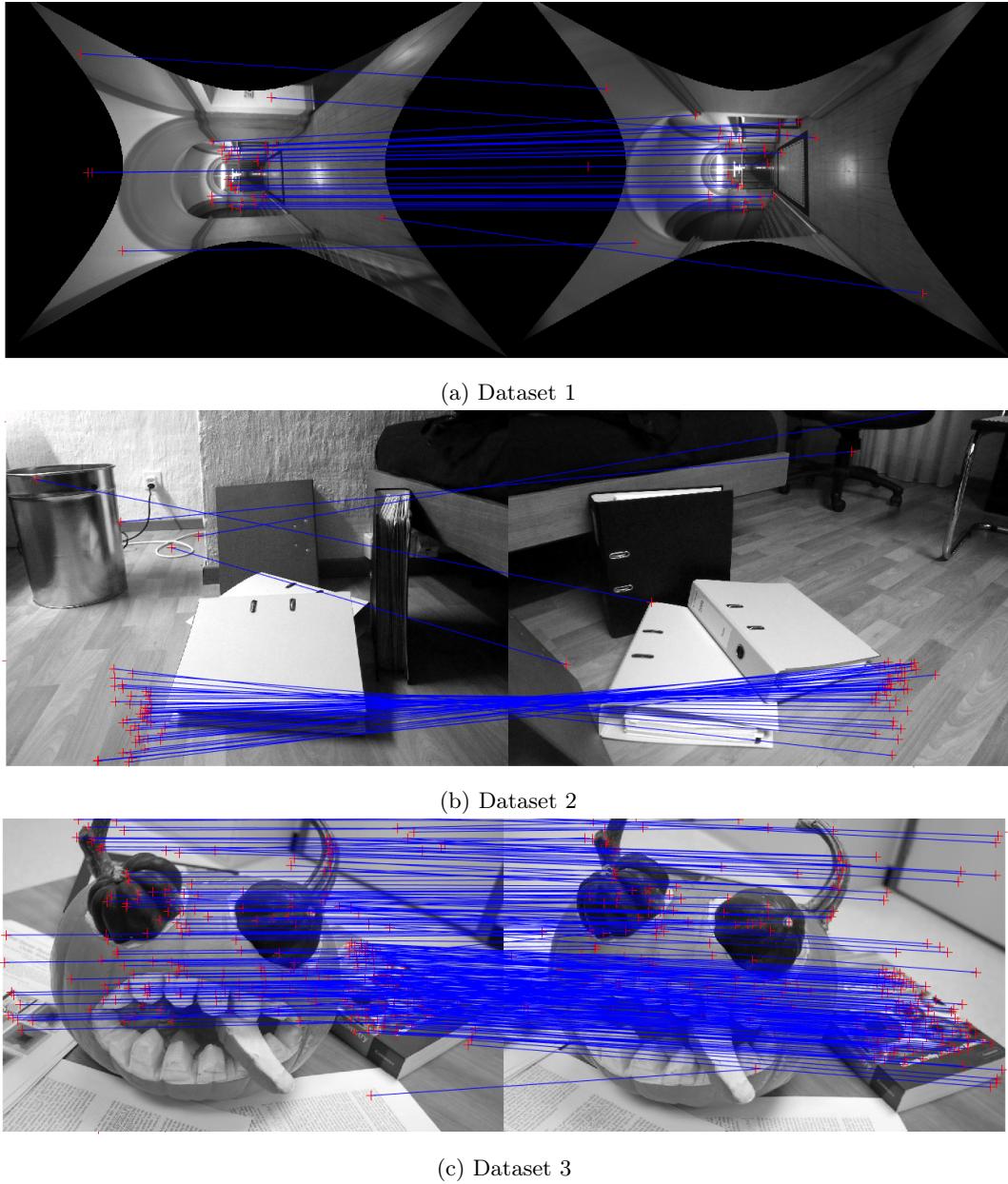
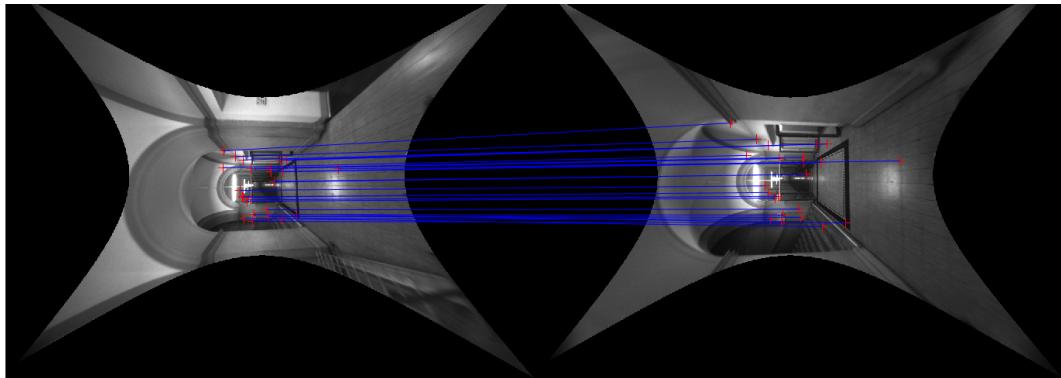


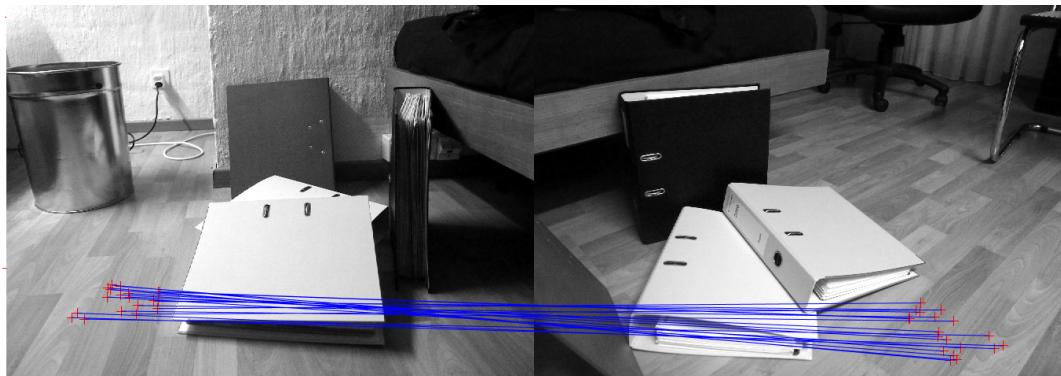
Figure 7: Inliers of the matching SIFT features for each dataset, using simple RANSAC, threshold = 10

Threshold	Dataset 1			Dataset 2			Dataset 3		
	inliers	mean err.	M	inliers	mean err.	M	inliers	mean err.	M
10	47	3.5484	2092	83	4.1375	224810	379	3.9405	267
5	39	1.9949	9316	56	2.5699	5235300	250	2.0740	7488

Table 3: Number of inliers, mean error, and number of iterations for different threshold, using adaptive RANSAC.



(a) Dataset 1

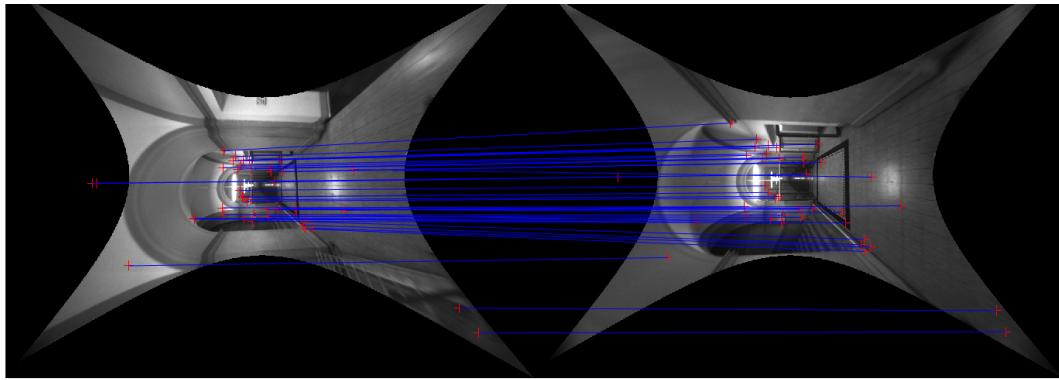


(b) Dataset 2



(c) Dataset 3

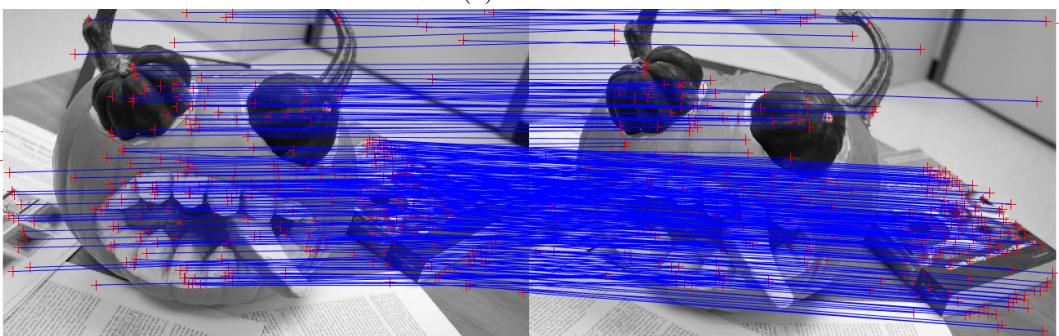
Figure 8: Inliers of the matching SIFT features for each dataset, using simple RANSAC,  $\text{threshold} = 5$



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 9: Inliers of the matching SIFT features for each dataset, using adaptive RANSAC,  $\text{threshold} = 10$