

# Computer Vision, 4th assignment

Mathis Lamarre

## Line fitting with RANSAC

To do line fitting, we apply the RANSAC algorithm. For a given number of iterations, we choose 2 points at random and fit a line through them. We compute the distance between all the points and this line and count the number of inliers, that is to say points whose distance to the line is smaller than a threshold. If this number is bigger than the number of inliers found in all the previous iterations, then we fit the line between all the current inliers and keep the number of inliers as the best.

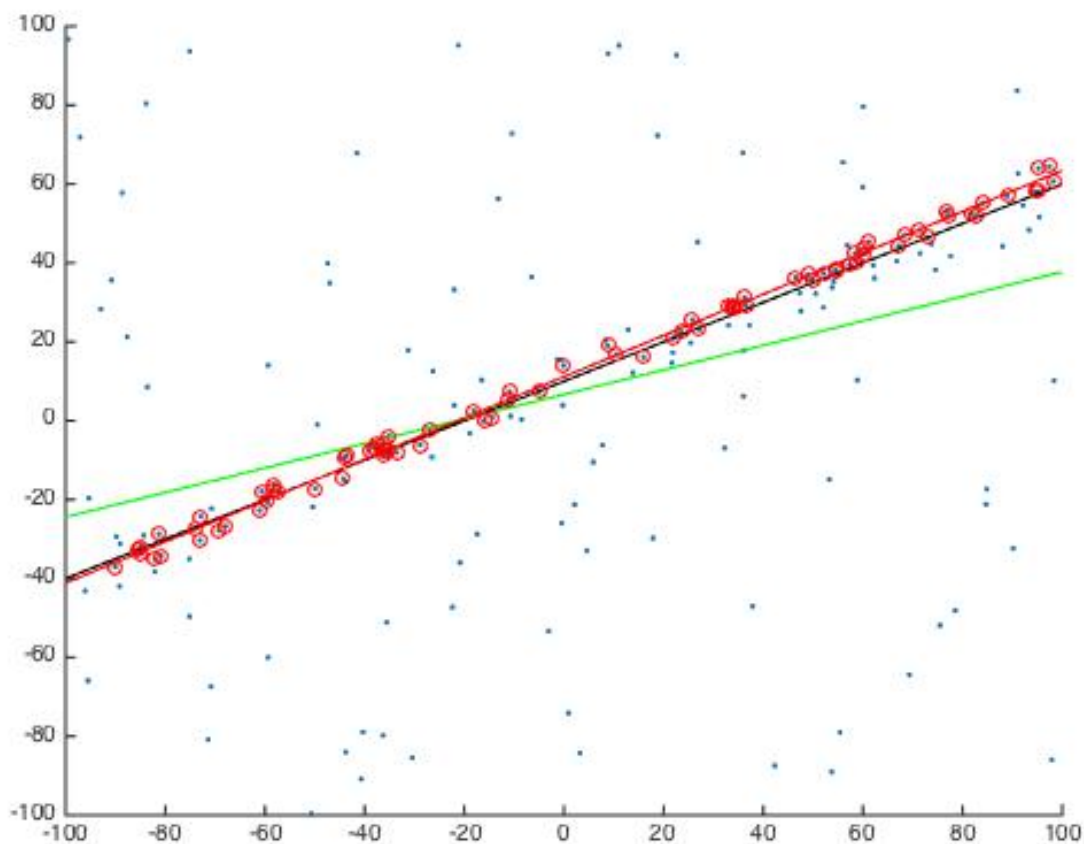


Fig. 1: Line fitting with RANSAC algorithm. The black line is the real model of the line. The green one is the least square fitting. The red one is the one obtained with RANSAC and the circled points are the inliers.

As we can see on Fig.1, the result is very satisfying. The RANSAC line is really close to the real model. As expected it has a bigger error than the least square method, which is designed to have the lowest error but gives too much importance to the outliers.  $err_{real} = 718.9826, err_{ls} = 710.2418, err_{ransac} = 718.3113$

## Fundamental matrix

In order to compute the fundamental matrix, first the points are normalized. Then the 8-point algorithm is applied on all the point correspondences from both images.  $F$  is extracted from the right null vector of the  $A$  matrix, and it is then unnormalized. Finally  $\hat{F}$  by enforcing the singularity constraint on the 3rd singular value of  $F$ .

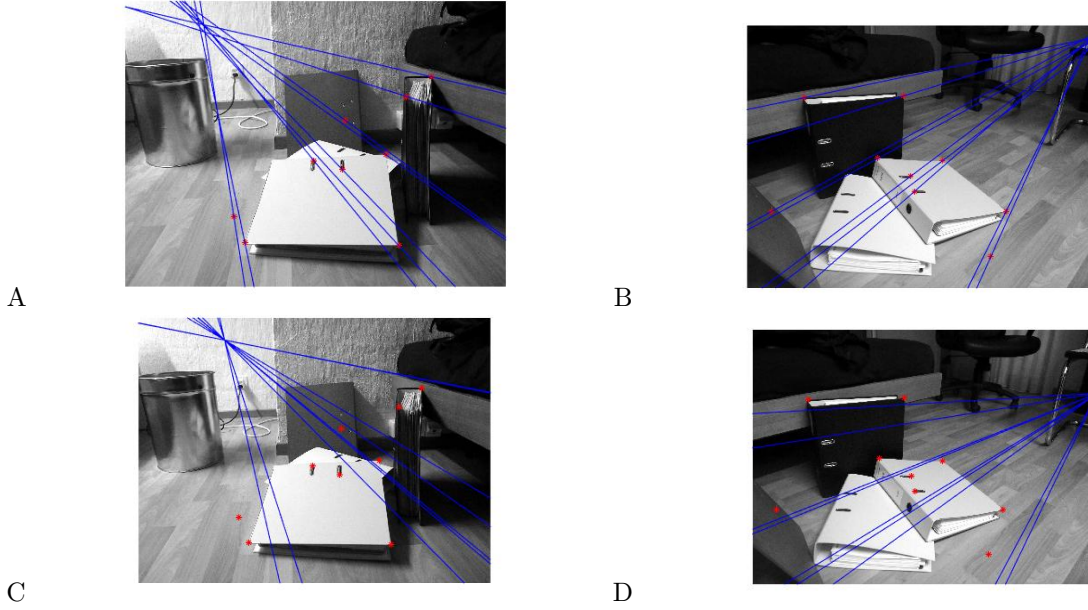


Fig. 2: Epipolar lines from fundamental matrix. A and B show epipolar lines on both images from the  $F$  matrix. C and D show epipolar lines on both images from the  $\hat{F}$  matrix.

The epipolar lines computed with the  $F$  matrix go through all the point correspondences (see Fig.2). However, they do not all intersect in one point. On the other hand, epipolar lines computed with the  $\hat{F}$  matrix do not go through all point correspondences but they do intersect at the epipole. The obtained  $F$  matrix is:

$$F = \begin{bmatrix} 0.0000 & -0.0000 & 0.0001 \\ -0.0000 & -0.0000 & 0.0011 \\ 0.0002 & 0.0038 & -0.5751 \end{bmatrix}$$

The  $\hat{F}$  matrix is very similar, with differences only of an order of magnitude of  $10^{-6}$  or smaller.

## Essential matrix

With the given calibration matrices, we can compute the 2-D normalized coordinates points  $\hat{\mathbf{x}} = \mathbf{K}^{-1}\mathbf{x}$ . Applying the 8-point algorithm with these points, we get the essential matrix. The constraint on the first two singular values being equal to their mean and the third one being zero is enforced singular value decomposition. The fundamental matrix can then be derived as  $\mathbf{F} = \mathbf{K}^{-T}\mathbf{E}\mathbf{K}^{-1}$ . Here are the obtained matrices:

$$E = \begin{bmatrix} 0.1739 & -4.8917 & -1.9139 \\ -5.7836 & -1.8213 & -2.4056 \\ -2.1172 & 2.5780 & 0.1734 \end{bmatrix} \quad \hat{E} = \begin{bmatrix} 0.7226 & -5.0864 & -1.7166 \\ -5.5338 & -1.2340 & -2.1630 \\ -2.4429 & 2.8369 & 0.3334 \end{bmatrix} \quad F = \begin{bmatrix} 0.0000 & -0.0000 & 0.0001 \\ -0.0000 & -0.0000 & 0.0011 \\ 0.0002 & 0.0038 & -0.5750 \end{bmatrix}$$

The  $F$  matrix is almost the same as the one found in the previous section, with a difference of an order of magnitude of  $10^{-3}$  on the last element. As we can see on Fig.3, the epipolar lines resemble the ones previously computed.

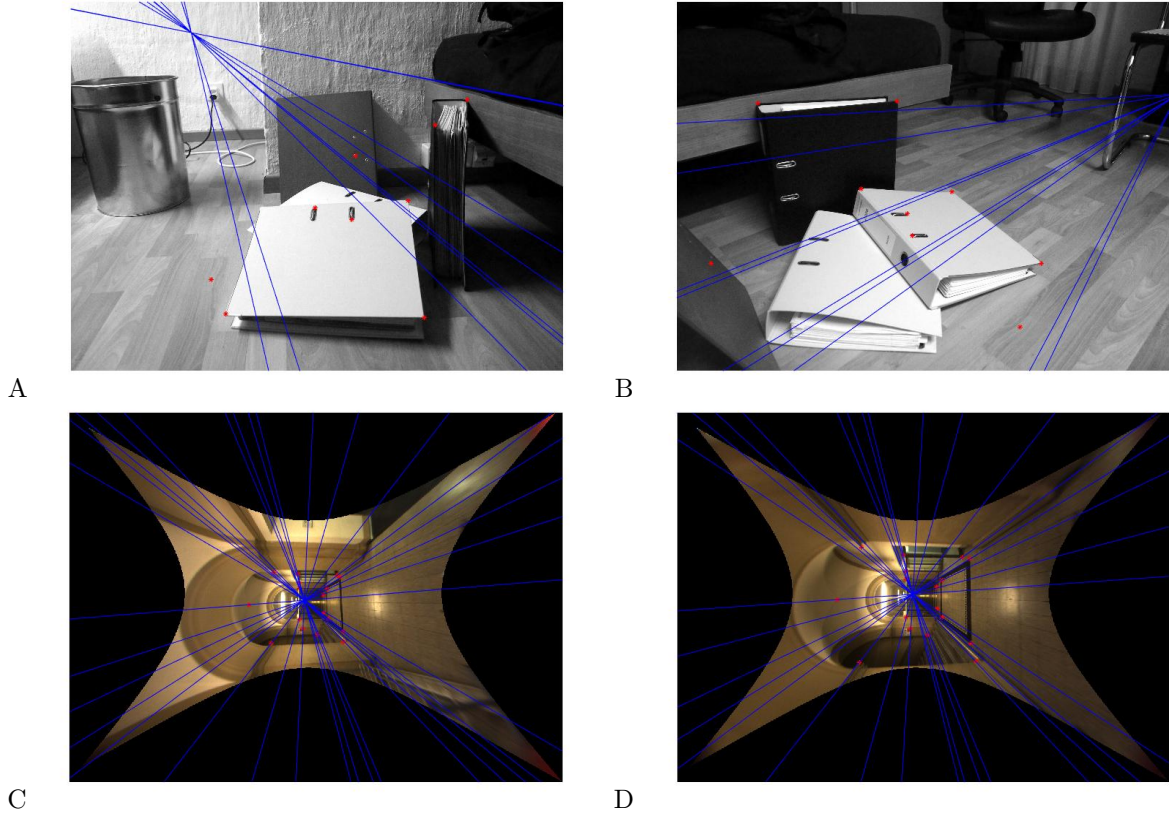


Fig. 3: Epipolar lines with fundamental matrix (enforced singularity constraint) computed for essential matrix. A and B show both views of one dataset. C and D show both views of a different dataset.

## Camera matrix

To compute all four possibilities of the camera matrix, the  $t$  vector is extracted from the last column of the  $U$  matrix of the singular value decomposition of the essential matrix and normalized. The two possible  $R$  matrices are computed with  $U$  and  $V$  matrices of the singular value decomposition as well as a orthogonal  $W$  matrix. If their determinant is -1 we take  $-R$  instead. The 4 possible  $P'$  matrices are:

$$P'_1 = [R_1 | t], P'_2 = [R_1 | -t], P'_3 = [R_2 | t], P'_4 = [R_2 | -t]$$

Then the triangulated 3D points are computed with each possible  $P'$  matrix and we check that they are in front of both cameras. We select the  $P'$  matrix for the which this is the case.

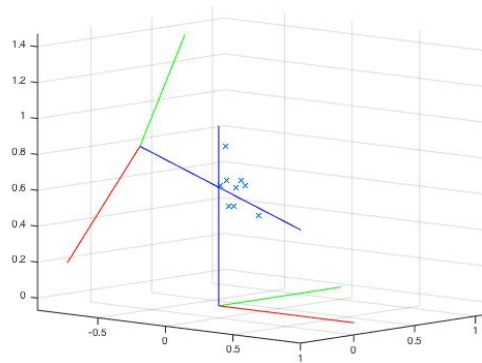


Fig. 4:  $P'$  with chosen  $R$  and  $t$  and corresponding 3D points. The  $x$  axis is in red, the  $y$  axis is in green and the  $z$  axis is in blue.

With the correct  $P'$  matrix, all the 3D points are in front of both cameras (Fig. 4) as they have positive values on the blue axes.

## Feature extraction and matching

To extract the features and match the features, the VLFeat toolbox was used.

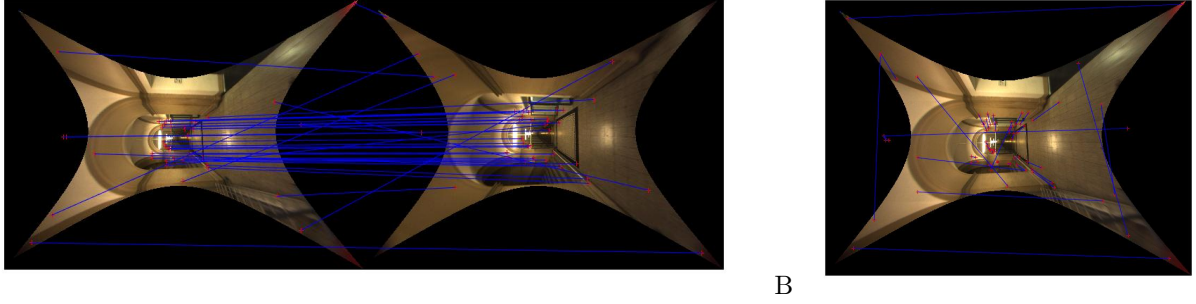


Fig. 5: Features extraction and matching with the VLFeat toolbox. A shows the features and matches on the 2 pictures. B shows the features and matches on the same picture.

On Fig.5-A, we can see that some of the matches are actually wrong as they don't point on the corresponding space on the other image. Fig.5-B shows the motion between the two camera frames.

## 8-point RANSAC

### Normal RANSAC

In order to select the right matches (inliers) and remove the wrong ones (outliers), we use the RANSAC algorithm, with the distance between the points and their corresponding epipolar lines on both images as a threshold criteria.

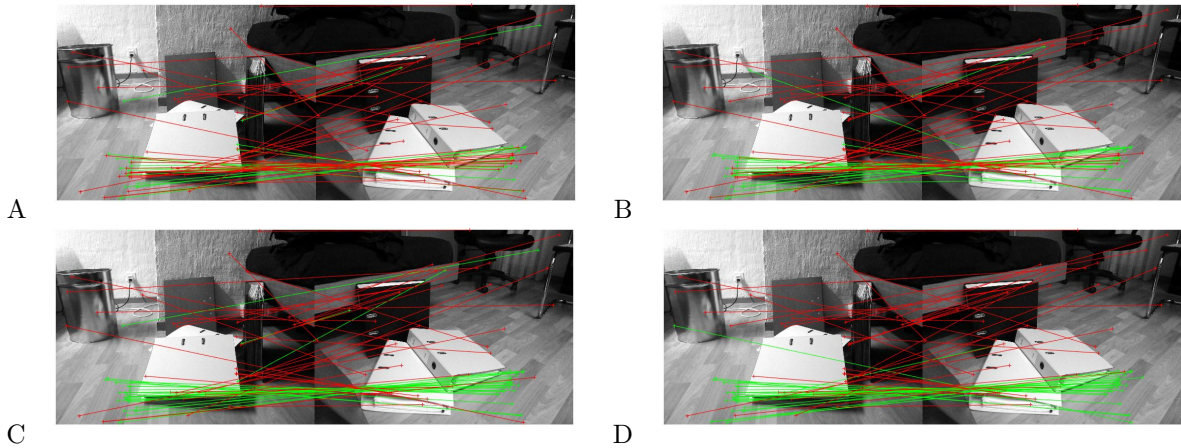


Fig. 6: Selection of inlier matches with 8-point RANSAC. A has a threshold of 1 pixel, B a threshold of 2 pixels, C a threshold of 3 pixels and D a threshold of 4 pixels.

On Fig.6, the inliers ratio is 0.34 on A, 0.43 on B, 0.48 on C and 0.49 on D. As expected, the inliers ratio is bigger when the threshold is bigger because more matches are considered as inliers. Here is the epipolar geometry of the resulting  $F$  matrix:

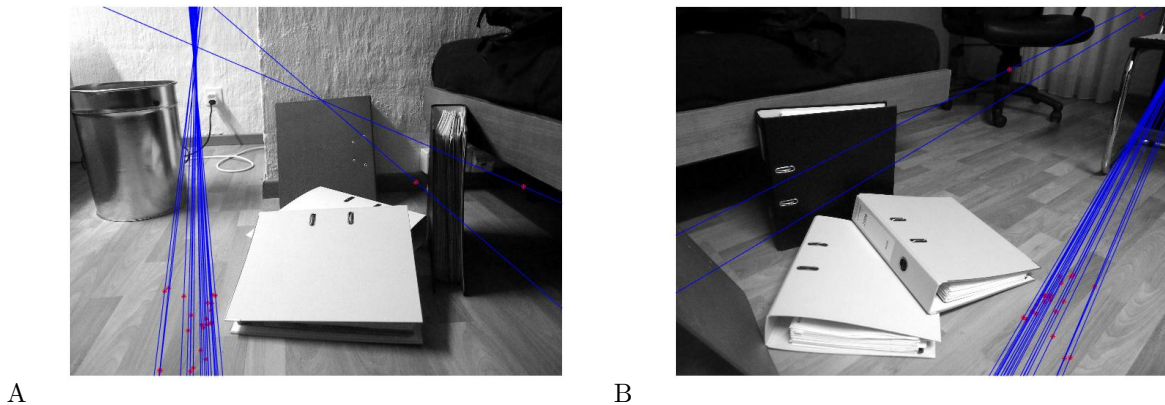


Fig. 7: Epipolar geometry of the  $F$  matrix selected with 8-point RANSAC.

The result resemble the one found on Fig.2. However there are some points that are still considered as inliers even though they should not (in the dark regions of the chair, binder and the bed). The epipole is in the same region as the old one. The imprecision is also due to the fact that most inliers selected are on the same plane of the floor.

### Adaptive RANSAC

As an exhaustive search of all possible set of 8 point correspondences is not computationally feasible, we use an adaptive RANSAC algorithm that will stop after  $M$  trials if the following equation is satisfied:  $p = 1 - (1 - r^N)^M$  with  $p$  the probability that at least one of the random samples of  $N = 8$  points from these  $M$  trials is free from outliers and  $r$  the best inlier ratio found so far.

The number of iterations needed varies with to the threshold and the set of images: for the *ladybug* set, it goes from 1409 with 1 pixel as a threshold to 272 with 3 pixels and for the *rect* set it goes from 12253 with 1 pixel as a threshold to 2190 with 3 pixels.