Practice session: Fundamental matrix estimation

1. Goal

The first step for most computer vision algorithms is almost always the computation of the Fundamental matrix, which is of great theoretical and practical importance.
The main goal of this session is to apply different methods for estimating the Fundamental matrix relating two views of a scene and to analyze the performances of these methods.

2. Recall on epipolar geometry

Epipolar geometry is the intrinsic projective geometry between two views. It is independent of scene structure, and only depends on the cameras' internal parameters and relative pose.

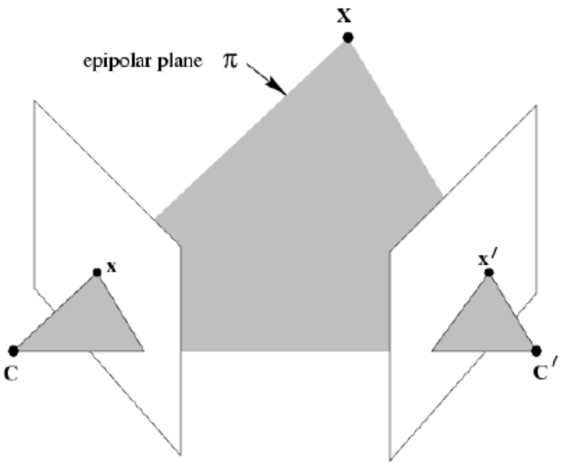

Figure 1: Epipolar geometry

Epipolar geometry gives a constraint between corresponding points, i.e. if a 3D point $\boldsymbol{X}$ of the scene is projected onto $\boldsymbol{x}$ and $\boldsymbol{x}'$ in the two views, then the image points x and x' must satisfy the *epipolar constraint:*

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{F}\,\boldsymbol{x}' = \boldsymbol{0}, \qquad (1)$$

With $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$, $\boldsymbol{x}' = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$ *and* $\boldsymbol{F} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$.

We know that the fundamental matrix $\boldsymbol{F}$ is defined up to a scale factor. Moreover, the fact that the epipole in the second image belongs to all the epipolar lines implies that $\det(F) = 0$. Therefore, $\boldsymbol{F}$ depends on seven parameters only.

3. Computing the Fundamental matrix

Each corresponding couple of points $(\boldsymbol{x}, \boldsymbol{x}')$ yields one equation (1). Therefore, with a sufficient number of correspondences in general position it is possible to determine $\boldsymbol{F}$. No knowledge about the cameras or scene structure is necessary.

The epipolar constraint equation is linear in the entries of $\boldsymbol{F}$ and it can be rewritten as:

$$U^T f = 0,$$

where

$$U = [xx', xy', x, yx', yy', y, x', y', 1]^T,$$

$$f = [f_{11},\ f_{12},\ f_{13}, f_{21},\ f_{22},\ f_{23}, f_{31},\ f_{32},\ f_{33}]^T.$$

With N correspondences, we can form a linear system of the form $Af = 0$ and solve for f.
Using seven points, it is possible to compute F using the rank deficiency constraint $\det(F) = 0$.
However, there is no unique solution. With eight correspondences, there is a unique solution which can be obtained linearly. In practice, we will use more than eight correspondences.

3.1. Linear least square solution with unit constraint on one entry of $\boldsymbol{F}$

This is the simplest method. Since $\boldsymbol{F}$ is defined up to a scale factor, we can take one of its entries to be equal to one: $f_{33} = 1$.
In that case, the linear system to solve becomes $Af = B$, with:

$$A = \begin{bmatrix} x_1x_1' & x_1y_1' & x_1 & y_1x_1' & y_1y_1' & y_1 & x_1' & y_1' \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_Nx_N' & x_Ny_N' & x_N & y_Nx_N' & y_Ny_N' & y_N & x_N' & y_N' \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix}.$$

The LLS solution is thus given by:

$$\hat{f} = (A^T A)^{-1} A^T B$$

3.2. Linear least square solution with unit constraint on the norm of $\boldsymbol{F}$

Another approach is to put a unit constraint on the norm of the matrix $\boldsymbol{F}$, i.e. we seek the LLS solution:

$$\min_f \|Af\| \text{ subject to } \|f\| = 1$$

The solution in this case is given by the right singular vector of A associated with the smallest singular value.

3.3. Enforcing the rank 2 constraint

In practice, the Fundamental matrix obtained by either of these approaches doesn't satisfy the constraint $\det(F) = 0$. So, we have to find the best rank two matrix from $\boldsymbol{F}$. Best in the sense that the obtained matrix $\boldsymbol{F}'$ minimizes the Frobenius norm $\|\boldsymbol{F} - \boldsymbol{F}'\|$.

If the SVD of $\boldsymbol{F}$ is $\boldsymbol{F} = \boldsymbol{USV^T}$, with $S = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$, then we obtain $\boldsymbol{F}' = \boldsymbol{US'V^T}$ taking

$$S' = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

3.4. Normalizing image points coordinates

In the linear methods described above, we deal with projective coordinates of pixels. Let take an image of size 512 x 512. A typical point of this image will have projective coordinates whose orders of magnitude are given component wise by the triplet (250, 250, 1). The fact that the first two coordinates are more than one order of magnitude larger than the third one has the effect of causing poor numerical conditioning.

This effect was analyzed by Hartley, who proposed a normalization procedure such that points in both images have their projective coordinates of roughly the same order.

This normalization is achieved by applying two affine transformations to the coordinates systems of the images. We then estimate the Fundamental matrix in the new coordinates systems for greater numerical stability. The result is then expressed in the original coordinate systems by performing the two inverse transformations.

In detail, two affine transformations, represented by two 3 x 3 matrices H and H', are computed in such a way that:

- the centroid of the transformed points is at origin $[0, 0, 1]^T$, and
- the average distance of the transformed points to the origin is $\sqrt{2}$.

Given H and H':

1. transform corresponding points coordinates by applying $\hat{x} = H\,x$ and $\widehat{x'} = H'x'$
2. estimate F' in the transformed coordinates systems by a linear method
3. expressed the result in the original coordinates by $F = H^T F' H'$

4. WHAT DO YOU HAVE TO DO?

You are given a few pair of images for which you have to compute the Fundamental matrices. In this assignment, we will manually select the corresponding points in the images. Use a minimum number of 12 correspondences. A Matlab function is given to select the points.

Question 1:

- Implement the two LLS methods described in Section 3.1 and 3.2 and estimate the Fundamental matrices for the three pairs of images. You, of course, have to make F a rank two matrix.
- Show the epipolar lines for a set of points as in Figure 2.
- For each method and for each pair of images, give an average estimation error which can be computed as the sum of distances between corresponding points and the epipolar lines:

$$\sum_{i=1}^{N} d(x^T F, x')$$

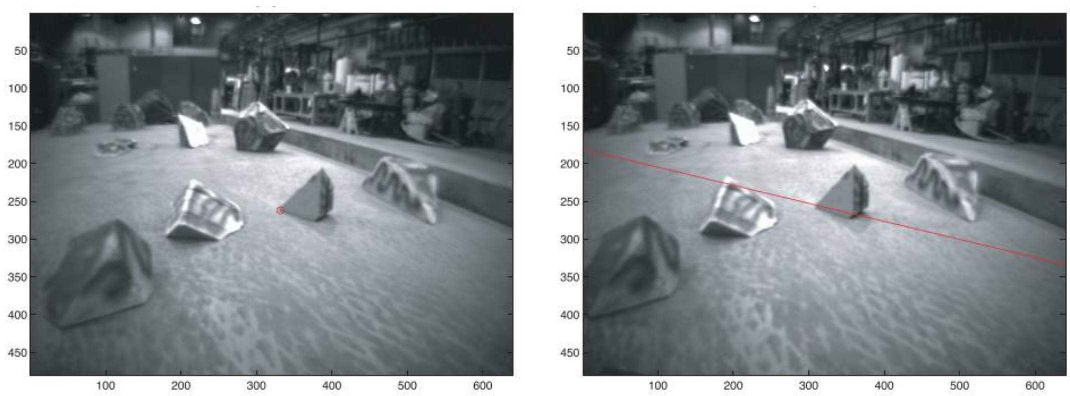

Figure 2: A point selected in the left image and its corresponding epipolar line in the right image.

Question 2:

- Implement the normalization procedure described in Section 3.4 and estimate the Fundamental matrices again.
- Show epipolar lines
- Give for each pair of images, an average estimation error

Question 3:

- We have manually selected the corresponding points. However, in practice they have to be identified automatically (we will later learn how?). Then what kind of new problems will the estimation methods suffer?
- Repeat again the experiments with a few incorrect correspondences (manually selected).

5. WHAT TO SUBMIT?

- Your codes for Fundamental matrix estimation
- A report showing your results (numerical values of $\boldsymbol{F}$, epipolar lines, estimation error, etc) and your comments and conclusions.