Computervision Lab 4
Multi-view geometry

David Van Hamme

1 Epipoles and the fundamental matrix

The basic mathematics of multi-view geometry are based on the concept of *epipoles*. The goal of multi-view geometry is to recover the three-dimensional aspects of the scene which have been lost by the 3D-to-2D perspective projection described by the pinhole camera model. Points in a camera image corresponds to rays of light in 3D space, but the exact origin of the light on this ray cannot be reconstructed from one image. However, adding a second view of the same point, removes this degree of freedom: the second camera also casts a ray, and the intersection point of the two rays defines a single point in 3D space. This is illustrated in Figure 1.

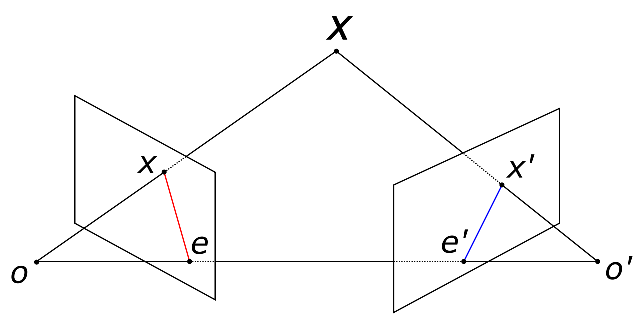

Figure 1: Basic concepts of epipolar geometry.

The projection of the ray through $\mathbf{x}$ (the line $\mathbf{ox}$ in the figure) onto the image plane of the second view is (generally) also a line ($\mathbf{e}'\mathbf{x}'$ in the figure). These lines are called *epipolar lines*, because the projections of different rays $\mathbf{ox}$ from the first camera all go through a single point $\mathbf{e}'$ in the second view's image plane, called the *epipole*. This intersection point is the projection of the first camera's focal point onto the second camera's image plane; it is the only point that all rays share. Similarly, the projection of the second camera's focal point onto the first camera's image plane defines an epipole $\mathbf{e}$; the reasoning is entirely symmetrical.

Usually the exact position of the camera views is not known, so finding the epipoles is part of the problem. Instead, we have to define the problem in terms of the epipolar lines themselves. Note that in homogeneous coordinates, the standard (Cartesian) line equation

$$ax + by + c = 0 \tag{1}$$

can be written as

$$\mathbf{lx} = 0 \tag{2}$$

or as

$$\mathbf{x}^T\mathbf{l} = 0 \tag{3}$$

in which $\mathbf{l}$ is a 3-vector representing the line and $\mathbf{x} = (x, y, 1)$. Therefore the fact that the point $\mathbf{x}'$ lies on the epipolar line $\mathbf{l}'$ can be expressed as

$$\mathbf{x}'^T\mathbf{l}' = 0. \tag{4}$$

We are looking for a mapping $\mathbf{F}$ that expresses the epipolar line as a function of the point $\mathbf{x}$. We will represent the epipolar line $\mathbf{l}'$ as

$$\mathbf{l}' = \mathbf{F}\mathbf{x} \tag{5}$$

in which $\mathbf{F}$ is a 3x3 matrix. This makes sense mathematically: the result of the matrix multiplication is a 3-vector which represents the line. Combining Equations 4 and 5 yields

$$\mathbf{x}'^T \mathbf{F}\mathbf{x} = 0. \tag{6}$$

This relationship means that we can compute $\mathbf{F}$ from point correspondences, e.g., by matching keypoint descriptors across the two images. $\mathbf{F}$ is called the *fundamental matrix*. It can be calculated from at least 8 pairs of corresponding point coordinates using singular value decomposition; usually many more are used in a least-squares solution. From the fundamental matrix, the epipolar lines and hence the epipoles can be computed, and combined with the projection matrices of the cameras, the rotation and translation between the camera pose in the first and second picture can be computed.

Exercise 1

Write a program that computes the fundamental matrix between *im00.png* and *im01.png* and shows the epipolar lines.

Figure 2: Two shots of the same scene, with the same camera.

Assignment 1 Compute SIFT keypoints (or other features of your choice) and descriptors in both images and match them. Use *k nearest neighbor matching* and reject ambiguous matches based on Lowe's ratio test. Show the matches.

Lowe's ratio test is a way to check if a point is sufficiently unique in the picture: if the best match and the second best match have similar matching distances, it means there are at least two similar looking points in the second image and there is a high risk of false correspondence as a result.

Assignment 2 Estimate the fundamental matrix using **findFundamentalMat**, compute the epipolar lines for the feature points you used (you can use **computeCorrespondEpilines**) and draw these onto both images.

Tip: visually try to assess the change in perspective between the two pictures. The epipolar lines should reflect this!

2 The essential matrix

In many cases, we will be trying to estimate the pose change of a camera between two subsequent frames (like the case in Exercise 1). Note that the fundamental matrix combines three transformations: the projective (pinhole) transformation of the first camera, the 3D pose change, and the projective transformation of the second camera. When the first and second camera are the same, we can simplify the problem Using camera calibration (as in the previous lab), we can decouple the three transforms. To do this, we first apply the inverse camera matrix to the image coordinates of each camera, resulting in what we call *normalized coordinates*, they are the coordinates as seen by a camera with the identity matrix as calibration. The term normalized stems from the fact that the units are now the same between the cameras and the origin of each camera coordinate system is in its principal point.

We now compute the fundamental matrix between the point correspondences in normalized coordinates. This matrix will represent purely the pose change, and it is called the *essential matrix*, denoted as $\mathbf{E}$. It can easily be shown that the relationship between the essential and fundamental matrix is given by

$$\mathbf{E} = \mathbf{C}'^T\mathbf{F}\mathbf{C} \tag{7}$$

in which $\mathbf{C}'$ and $\mathbf{C}$ are the calibration matrices of the second and first camera respectively.
The essential matrix describes only the pose change. The rotation and translation component of the pose change can be extracted from this, but the mathematics are rather complicated and will not be described here.

Exercise 2

Compute the essential matrix, given that the camera used for both images has the following calibration matrix:

$$\begin{bmatrix} 792 & 0. & 505 \\ 0. & 791 & 376 \\ 0. & 0. & 1 \end{bmatrix}.$$

Assignment 3

- Compute the essential matrix from the fundamental matrix you estimated in Assignment 2, using Equation 7 and the matrix above.
- Compute the translation and rotation from the essential matrix using **decomposeEssentialMat**. Print the rotation and translation matrices in your report.

Question 1 You get two possible rotation matrices, which one do you think is correct? *Hint: read Question 3 below.*

Assignment 4

- Convert the keypoint coordinates you used in Assignment 2 to normalized coordinates (for both images) by multiplying the inverse camera matrix with them.
- Compute the essential matrix by using **findFundamentalMat** on these normalized points. The resulting matrix is immediately the essential matrix!
- Compute the translation and rotation matrices once again with **decomposeEssentialMat** and print them in your report. They should be close to results from Assignment 3, if not, you have a problem!

Question 2 The translation is only up to a scale factor. Does it correspond to your visual assessment of the perspective difference between the shots? Explain.

Question 3 Assuming that this rotation matrix would be a pure rotation along the Y axis (vertical), what would be the approximate angle of the rotation?