Chapter 8 – Epipolar Geometry

Author: Gianmarco Scarano

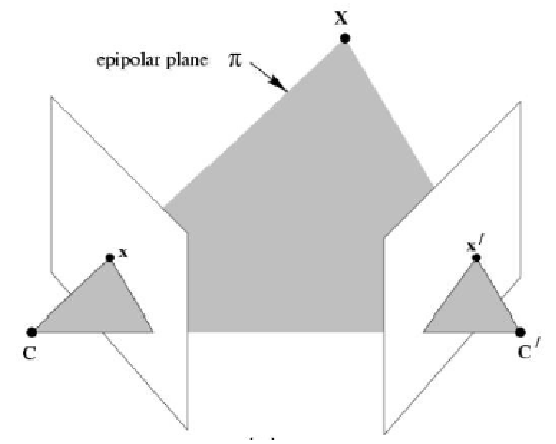
gianmarcoscarano@gmail.com

1. Epipolar Geometry

In Epipolar Geometry, we either have 2 cameras or one moved in order to see how these two cameras correlate with each other.

We use Epipolar Geometry in order to answer 3 questions in general:

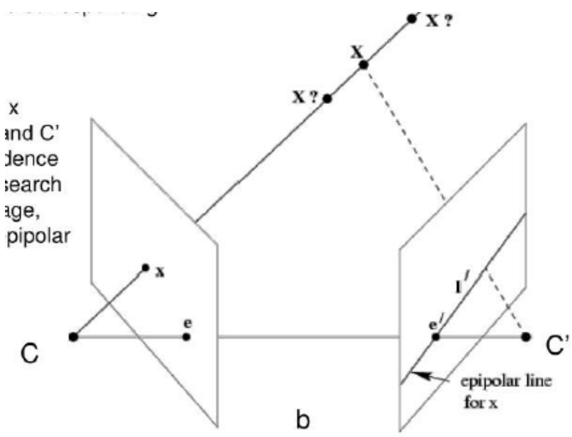
- If we have an image point x in the first image, how does this constrain the position of the corresponding point x' in the second image?
- Given a set of points $\{x_0, x_1 .. x_N\}$ and $\{x'_0, x'_1 .. x'_N\}$, what are the cameras P and P' for the two images?
- Given a set of points $\{x_0, x_1 .. x_N\}$ and $\{x'_0, x'_1 .. x'_N\}$, along with cameras P and P', what is the position of their pre-image X in space?



- C and C' are the camera center for P and P' (which are the two images).
- $\boldsymbol{X}$ is the 3D point.
- x and x' is the corresponding $\boldsymbol{X}$ on P and P' (the two images).
- π is the Epipolar plane.

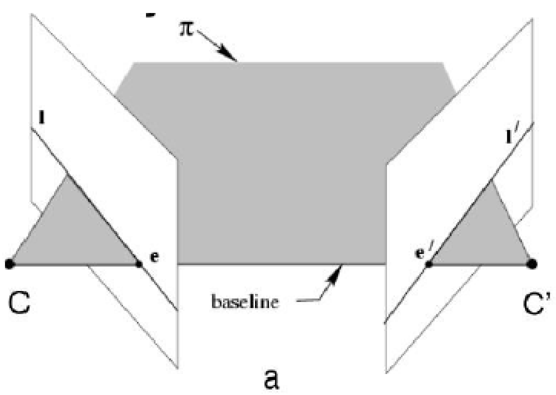
1.1 Point Correspondence Geometry

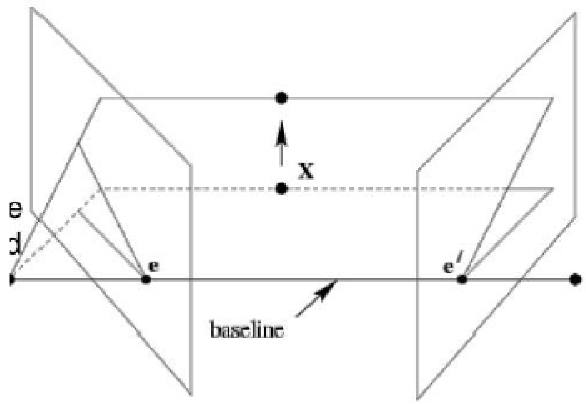
Here, we want to project one camera center in the other hyperplane.



We use the so called Epipolar, which is the point in which the image intersects the camera center and the line that comes from it, it's called Epipolar line, denoted as I'.

So, if we know x, how can we know x'? This is pretty easy, since x' is exactly the point in which the Epipolar line intersects the camera center C'.

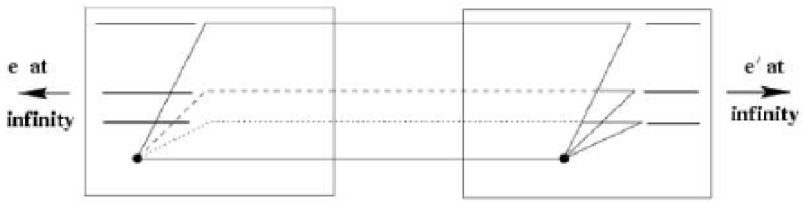




Very important in the Epipolar Geometry, is the 3D point $\boldsymbol{X}$, which basically serves as an angle for the Epipolar plane, thus providing some sort of "rotation" around the baseline.

Also, we can define the epipoles e, e' as the intersection of baseline with image plane or as the projection of the camera center in the other image.

The Epipolar line is the intersection of the Epipolar plane with the image (comes of course in pairs, since we have I' and e').



Note that, if we translate the two images horizontally (parallel translation), we infer that the epipoles are at infinity.

2. Fundamental Matrix

This is a very important concept in the Epipolar Geometry, since we know that a point x in one image is transferred via the plane π to a matching point x' in the second image. The epipolar line through x' is obtained by joining x' to the epipole e'. In this way, we could denote $x' = H_\pi \cdot x$ (which is the Projection Matrix between x and x') and:

- $I' = [e']_x \cdot x'$, which translates into
- $I' = [e']_x \cdot H_\pi \cdot x$, which translates into
- $I' = F \cdot x$

Such that F = $[e']_x \cdot H_\pi$, which is the Fundamental Matrix. F links the x point in the first view to the Epipolar line I' in the 2nd view. It is useful if we don't know I in the first view. So, there should be some sort of connection between x and I' since x and x' point to $\boldsymbol{X}$.

2.1 Geometric derivation

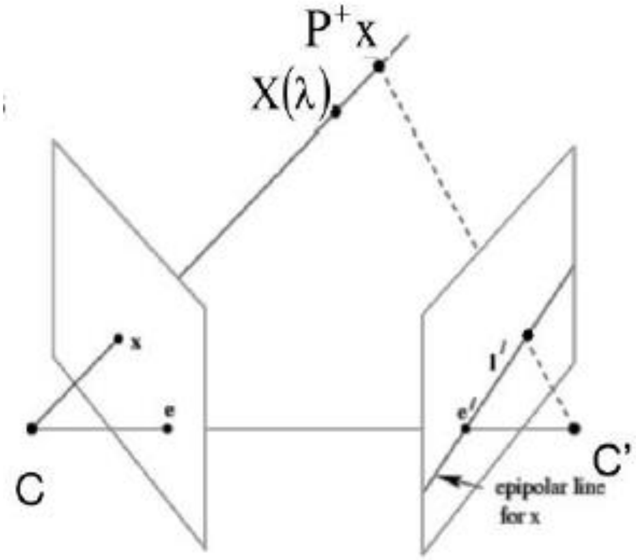
We already recovered the geometric derivation in the point 2. Here follows the skew matrix scheme:

<u>Example</u>:

$a = (a_1, a_2, a_3)^T$

$$[a]_x = \begin{matrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{matrix}$$

2.2 Algebraic derivation



We can recover the epipolar line I' as follows:

$$I' = P'C \times P'P^{+}x$$

Where P^{+} is the pseudo-inverse of P => ($P^{+}P = I$)

$P'C$ is, instead, e' or the epipole on the second view, while $P'P^{+}x$ is the projection $P^{+}x$ on the 2nd view.

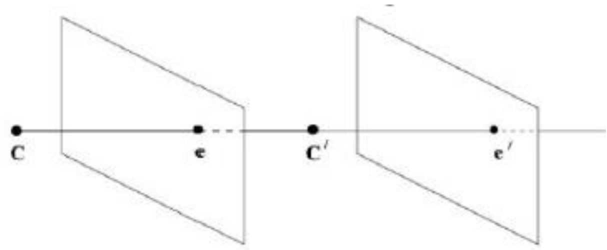
With this being said, very simply, we can rewrite the whole formula as the cross-product between e' and $P'P^{+}x$, such that:

$$F = [e']_{x} P'P^{+}$$

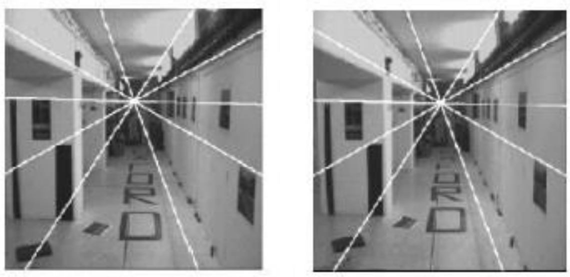
We now have to specify that this is done by taking any value along the X-big axis ($\boldsymbol{X}$), which is denoted as lambda (λ), in fact $\boldsymbol{X}(\lambda)$**.** By changing the λ-value we change the point we're taking along the line (basically we are taking each time a different 3D point), also known as $P^{+}x$. If we multiply this $P^{+}x$ value by P', we have that exact point but on the second view (the second image).

Important properties for the Fundamental Matrix F are that it is a 3x3 singular matrix and that it is transposable for (P, P'), meaning that F^{T} is the Fundamental Matrix for (P', P).

2.3 Pure translation



a



B

c

If we have two images with the same angle but moving in parallel lines, we have a point called vanishing point, since in the end it will end up vanishing.

Pure translation refers to no rotation, no change in internal parameters, camera being stationary and the world follows a translation -t

2.4 Fundamental Matrix for pure translation

Considering the first camera center C being $\mathrm{P} = \mathrm{K}\,[\mathrm{I} \mid 0]$ and the second camera $P' = \mathrm{K}\,[\mathrm{I} \mid \mathrm{t}]$

We have the same internal parameters, since we are in a pure translation. Also given by the fact that 0 becomes t in the second camera.

In this way F reduces only to $F = [e']_x$ since camera translation is parallel to the x-axis.

In the end the position of x' is the same as x, plus a value depending on the translation and divided by Z, which is the depth, the distance of x w.r.t the camera center.

$$x' = x + Kt/Z$$

If Z is too big, it means that the object is very far, thus the quantity Kt/Z will end up being 0.

A clear example is the example of looking out of the train window, since points closer to us appear to move faster than those further away. This highly depends on the value of Z.

Considering a general camera motion, we have a H value, which is a formula that takes into account that there is some sort of rotation between x and x', so that the more general formula becomes:

$$x' = K'RK^{-1}x + K't/Z$$

The first term depends on the image position x, but not on the depth Z and takes into account also the rotation R and the change of some internal parameters of the camera, as explained in H.

Second term, again, depends highly on Z and takes also into account translation t.

3. Projective transformation and invariance

F is invariant to projective transformation, meaning that we can't exactly recover P and P' starting from F.

$$(P, P') \rightarrow F \qquad \text{UNIQUE}$$
$$F \rightarrow (P, P') \qquad \text{NOT UNIQUE}$$

For this reason, we can simply put the first camera $P = [I \mid 0]$ (canonical), such that starting from this canonical form, we can find the uniqueness of P', since we have set P earlier. P', then, is equal to $P' = [M \mid m]$ and therefore $[m]_x M$.

3.1 Projective ambiguity given F

If F is a Fundamental Matrix for both P and P', then we can infer that there exists a non-singular 4x4 matrix H such that $\tilde{P} = PH$ and $\tilde{P}' = P'H$

If we have maximum ambiguity, then we have projective ambiguity. But, if we apply constraints on these ambiguities (such as external information), we end up in affine ambiguity.

3.2 Canonical cameras given F

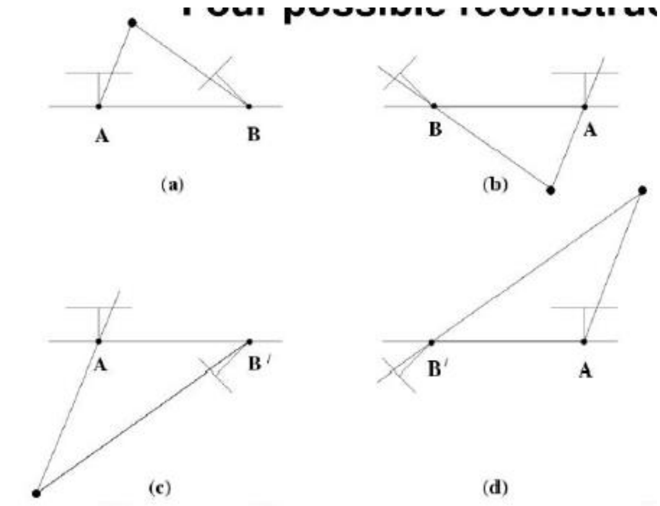
In this subchapter, starting from F and given any S possible, we can recover both P and P'!

F is a Fundamental Matrix and S is any skew-symmetric matrix, then:

$$P = [I \mid 0] \; and \; P' = [SF \mid e']$$

Where SF is $[e']_x F$, basically the translation.

3.3 Essential Matrix



We use the Essential Matrix for calibrated cameras, removing K from x. A 3x3 matrix is an essential matrix if and only if, two of its singular values are equal and the third is zero.

Starting from the Essential Matrix (E), we can recover 4 possible solutions for P and P'. We can test these 4 solutions, but only in 1, the point is in front of the camera, which is the one suitable for us, since it's physically visible and in other views this point is in the back of the camera.

<u>Only after testing we can know which one of these 4, is the real solution.</u>