Lecture 2

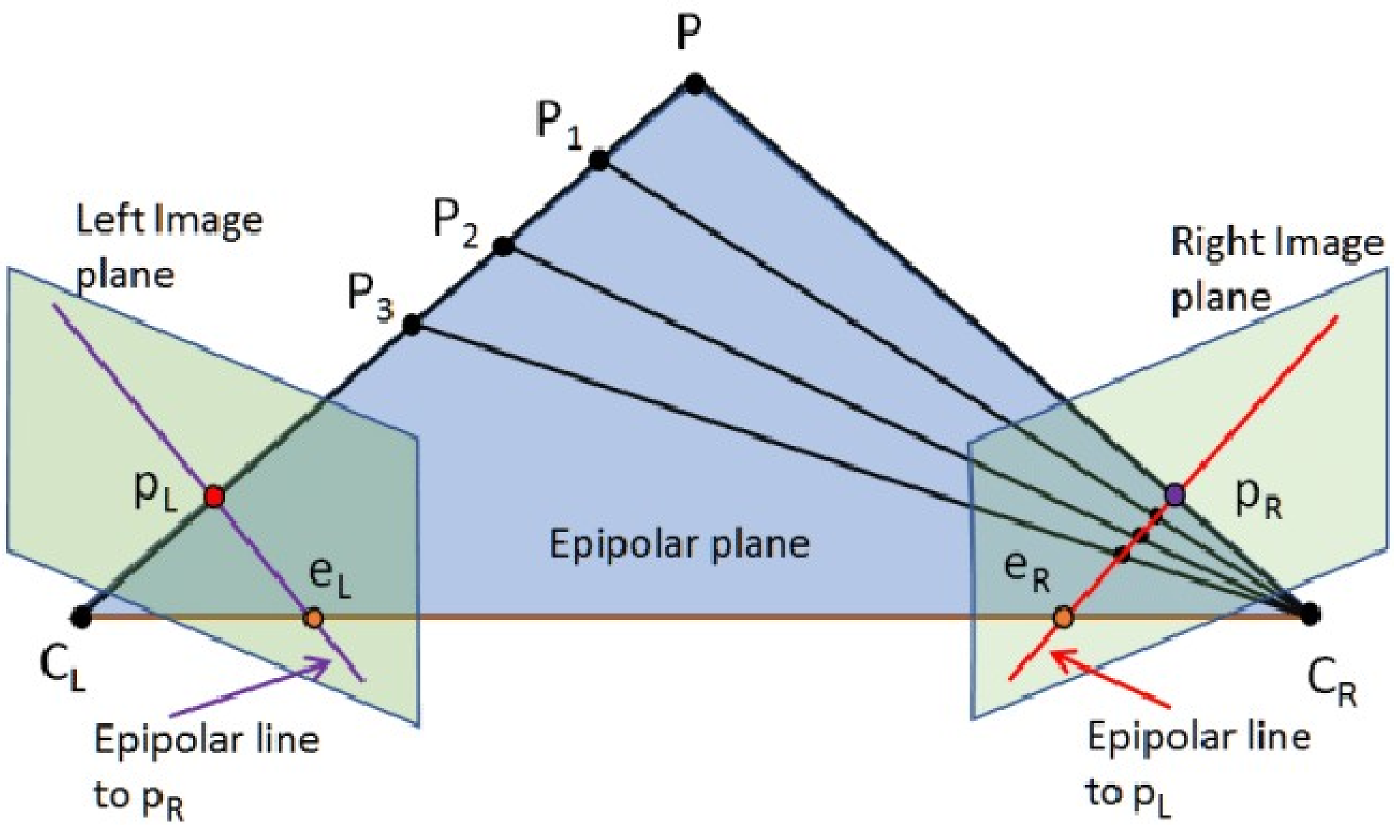
Epi-polar geometry, Essential and Fundamental matrix

Some slides were adapted/taken from various sources, including 3D Computer Vision of Prof. Hee, NUS, Air Lab Summer School, The Robotic Institute, CMU, Computer Vision of Prof. Mubarak Shah, UCF, Computer Vision of Prof. William Hoff, Colorado School of Mines and many more. We thankfully acknowledge them. Students are requested to use this material for their study only and **NOT** to distribute it.

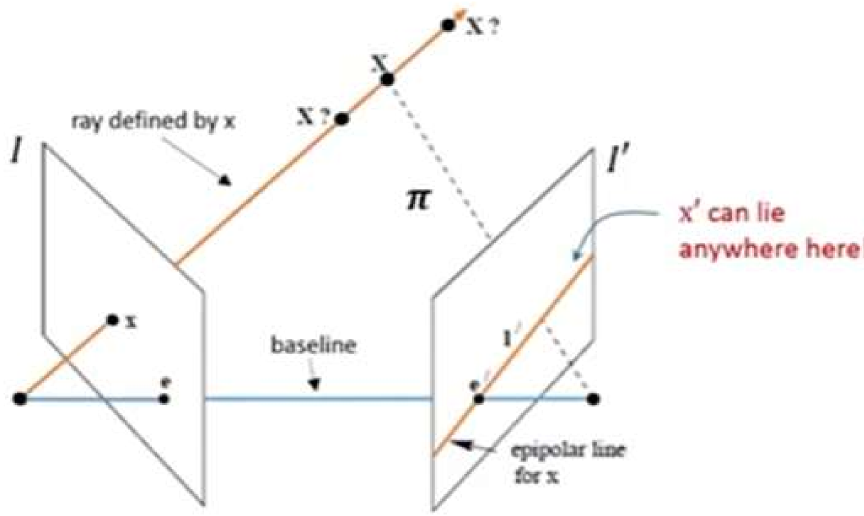
Module I:
3D Computer Vision:

- Pinhole Camera projection model
- Epi-polar geometry, Essential and Fundamental matrix
- RANSAC Algorithm
- Solve camera pose from essential matrix
- Feature detector and descriptor
- Optical Flow: Lucas-Kanade Algorithm
- Camera Pose and depth estimation

The Epipolar Geometry



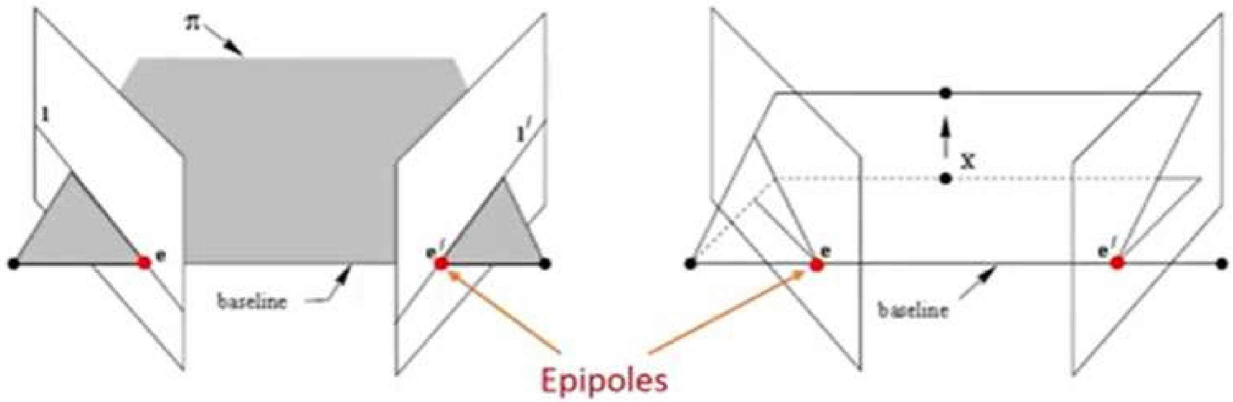
The Epipolar Geometry



- The image point x in I back-projects to a ray, and this ray projects to I' as the epipolar line $\mathbf{l}'$.
- The corresponding point x' can lie anywhere on $\mathbf{l}'$.
- Epipolar plane $\boldsymbol{\pi}$ is determined by the baseline and ray defined by $\mathbf{x}$.

Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

The Epipolar Geometry: Terminology

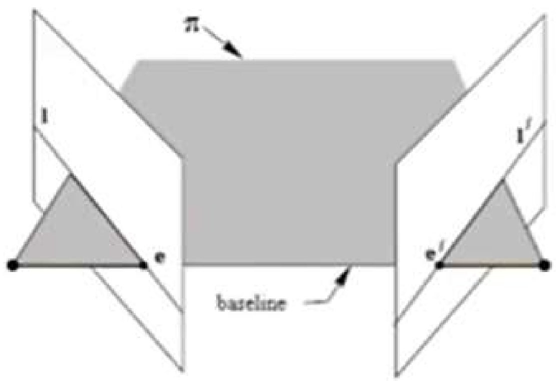


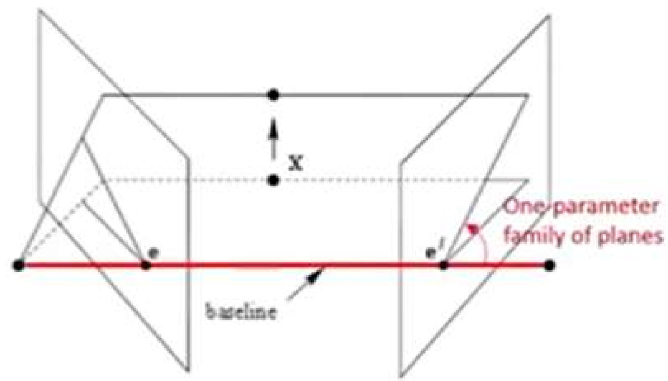
Epipoles (e, e′):

- Point of intersection of the line joining the camera centers (baseline) with the image plane.
- Equivalently, it is the image in one view of the camera center of the other view.
- Also the vanishing point of the baseline (translation) direction.

Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

The Epipolar Geometry: Terminology



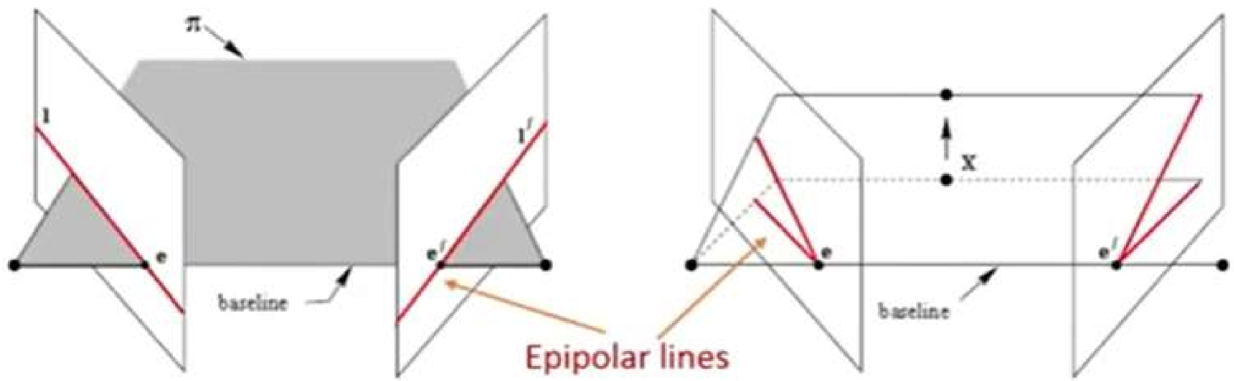


Epipolar plane π :

- A plane containing the baseline.
- There is a one-parameter family (a pencil) of epipolar planes.

Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

The Epipolar Geometry: Terminology



Epipolar lines $(\mathbf{l}, \mathbf{l}')$:

- The intersection of an epipolar plane with the image plane.
- All epipolar lines intersect at the epipole.
- An epipolar plane intersects the left and right image plane in epipolar lines, and defines the correspondences between the lines.

Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

The Fundamental Matrix

- The fundamental matrix is the algebraic representation of epipolar geometry.
- Gives the projective mapping relationship between a point $\mathbf{x}$ on one image to a line $\mathbf{l}'$ on the other.

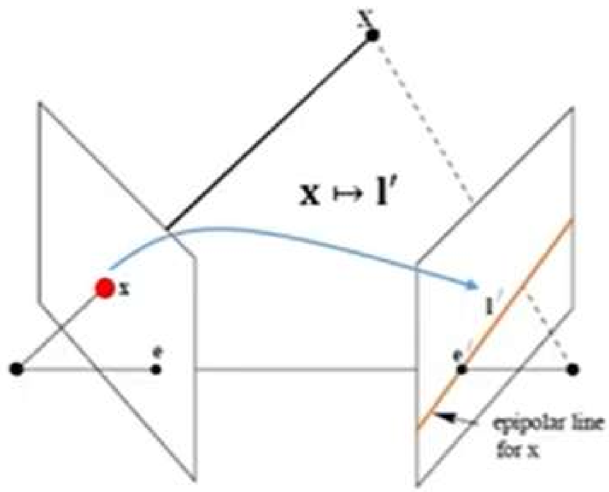


Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

F Matrix: Geometric Derivation

- The mapping $\mathbf{x} \mapsto \mathbf{l}'$ may be decomposed into two steps:

1. The point $\mathbf{x}$ is mapped to some point $\mathbf{x}'$ in the other image lying on the epipolar line $\mathbf{l}$; this point $\mathbf{x}'$ is a potential match for the point $\mathbf{x}$.

2. The epipolar line $\mathbf{l}'$ is obtained as the line joining $\mathbf{x}'$ to the epipole $\mathbf{e}'$.

F Matrix: Geometric Derivation

Step 1: Point transfer via a plane.

- Consider a plane $\boldsymbol{\pi}$ in space not passing through either of the two camera centres and contains the point $\mathbf{X}$.
- Thus there is a 2D homography $\mathrm{H}_{\boldsymbol{\pi}}$ mapping each $\mathbf{x}_i$ to $\mathbf{x}'_i$.

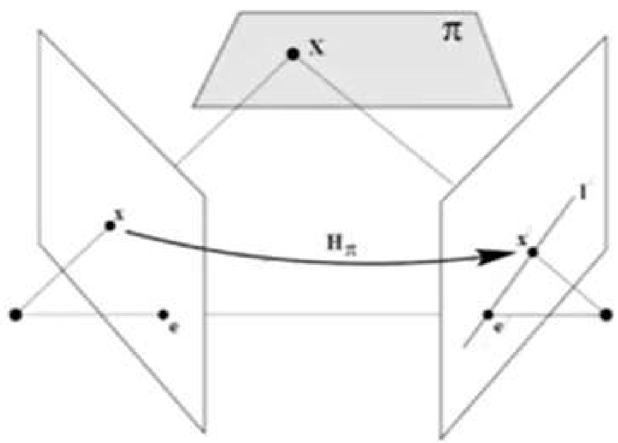


Image Source: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

F Matrix: Geometric Derivation

Step 2: Constructing the epipolar line.

- Given the point $\mathbf{x}'$ the epipolar line $\mathbf{l}'$ passing through $\mathbf{x}'$ and the epipole $\mathbf{e}'$ can be written as $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$.

- Since $\mathbf{x}'$ may be written as $\mathbf{x}' = \mathrm{H}_\pi \mathbf{x}$, we have:

$$\mathbf{l}' = [\mathbf{e}']_\times \mathrm{H}_\pi \mathbf{x} = \mathrm{F}\mathbf{x} \ ,$$

where we define $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_\pi$ as the fundamental matrix.

Cross product as Matrix Multiplication

- Vector cross product can be expressed as the product of a skew-symmetric matrix and a vector:

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times}\mathbf{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

F Matrix: Geometric Derivation

- The fundamental matrix F may be written as:

$$\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_{\boldsymbol{\pi}},$$

- where $\mathrm{H}_{\boldsymbol{\pi}}$ is the transfer mapping from one image to another via any plane.

- Furthermore, since $[\mathbf{e}']_\times$ has rank 2 and $\mathrm{H}_{\boldsymbol{\pi}}$ rank 3, F is a matrix of rank 2.

F Matrix: Geometric Derivation

- Geometrically, F represents a mapping from the 2-dimensional projective plane $\mathbb{P}^2$ of the first image to the pencil of epipolar lines through the epipole $\mathbf{e}'$.

- Thus, it represents a mapping of $\mathbb{P}^2 \mapsto \mathbb{P}^1$, and hence must have rank 2.

- **Note:** The plane is simply used here as a means of defining a point map from one image to another, but not required for F to exist.

Correspondence Condition

- For any pair of corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$ in two images:

$$\mathbf{x}'^T \mathrm{F} \mathbf{x} = 0$$

Proof:

$\mathbf{x}'$ lies on the epipolar line $\mathbf{l}' = \mathrm{F}\mathbf{x}$ corresponding to the point $\mathbf{x}$

$$\Rightarrow 0 = \mathbf{x}'^T \mathbf{l}' = \mathbf{x}'^T \mathrm{F} \mathbf{x} = 0$$

Correspondence Condition

- The importance of the relation $\mathbf{x}'^T \mathrm{F}\mathbf{x} = 0$ is that it gives a way of characterizing the fundamental matrix without reference to the camera matrices.

- That is the relation is only in terms of corresponding image points, and this enables F to be computed from image correspondences alone.

- We will discuss the details later on: how many correspondences are required to compute F from $\mathbf{x}'^T \mathrm{F}\mathbf{x} = 0$?

Properties of the F Matrix

- **Transpose**:
  - F is the fundamental matrix of the pair of cameras $(\mathrm{P}, \mathrm{P}')$
  - F^T is the fundamental matrix of the pair in the opposite order: $(\mathrm{P}', \mathrm{P})$
- **Epipolar lines**:
  - For any point $\mathbf{x}$ in first image, corresponding epipolar line is $\mathbf{l}' = \mathrm{F}\mathbf{x}$
  - $\mathbf{l} = \mathrm{F}^T\mathbf{x}'$ represents epipolar line corresponding to $\mathbf{x}'$ in second image
- **Epipole**:
  - For any point $\mathbf{x}$ (other than $\mathbf{e}$) the epipolar line $\mathbf{l}' = \mathrm{F}\mathbf{x}$ contains the epipole $\mathbf{e}'$
  - $\mathbf{e}'$ satisfies $\mathbf{e}'^T(\mathrm{F}\mathbf{x}) = (\mathbf{e}'^T\mathrm{F})\mathbf{x} = 0$ for all $\mathbf{x}$
  - $\mathbf{e}'^T\mathrm{F} = \mathbf{0}$, i.e. $\mathbf{e}'$ is the left null-vector of F
  - $\mathrm{F}\mathbf{e} = \mathbf{0}$, i.e. $\mathbf{e}$ is the right null-vector of F

Properties of the F Matrix

- **7 degrees of freedom (9 elements – 2 dof)**:
  - 3 x 3 homogenoueous matrix with 8 independent ratios ⇒ -1 dof
  - $\det(\mathrm{F}) = 0$ ⇒ -1 dof
- **Not a proper correlation (not invertible)**:
  - Projective map taking a point to a line
  - A point in first image $\mathbf{x}$ defines a line in the second $\mathbf{l} = \mathrm{F}\mathbf{x}$, i.e. epipolar line of $\mathbf{x}$
  - If $\mathbf{l}$ and $\mathbf{l}'$ are corresponding epipolar lines then any point $\mathbf{x}$ on $\mathbf{l}$ is mapped to the same line $\mathbf{l}'$
  - This means no inverse mapping, and F is not of full rank

Summary of the F Matrix Properties

- F is a rank 2 homogeneous matrix with 7 degrees of freedom.
- **Point correspondence**: If $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\mathsf{T}}F\mathbf{x} = 0$.
- **Epipolar lines**:
  - ◇ $\mathbf{l}' = F\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - ◇ $\mathbf{l} = F^{\mathsf{T}}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- **Epipoles**:
  - ◇ $F\mathbf{e} = \mathbf{0}$.
  - ◇ $F^{\mathsf{T}}\mathbf{e}' = \mathbf{0}$.
- **Computation from camera matrices** P, P':
  - ◇ General cameras,
    $F = [\mathbf{e}']_{\times} P'P^{+}$, where P^{+} is the pseudo-inverse of P, and $\mathbf{e}' = P'\mathbf{C}$, with $P\mathbf{C} = \mathbf{0}$.
  - ◇ Canonical cameras, $P = [I \mid \mathbf{0}]$, $P' = [M \mid \mathbf{m}]$,
    $F = [\mathbf{e}']_{\times} M = M^{-\mathsf{T}}[\mathbf{e}]_{\times}$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = M^{-1}\mathbf{m}$.
  - ◇ Cameras not at infinity $P = K[I \mid \mathbf{0}]$, $P' = K'[R \mid \mathbf{t}]$,
    $F = K'^{-\mathsf{T}}[\mathbf{t}]_{\times} RK^{-1} = [K'\mathbf{t}]_{\times} K'RK^{-1} = K'^{-\mathsf{T}}RK^{\mathsf{T}}[KR^{\mathsf{T}}\mathbf{t}]_{\times}$.

Source: Page 246, R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

… to continue

Essential Matrix

Essential Matrix

- **Normalized coordinates:** Known calibration matrices K and $\mathrm{K}' \Rightarrow$ we can write $\mathbf{x} \leftrightarrow \mathbf{x}'$ as $\mathrm{K}^{-1}\mathbf{x} \leftrightarrow \mathrm{K}'^{-1}\mathbf{x}'$, i.e. $\hat{\mathbf{x}} \leftrightarrow \hat{\mathbf{x}}'$:

$$\mathbf{x}'^{\mathrm{T}}\boxed{\mathrm{F}}\mathbf{x} = \mathbf{x}'^{\mathrm{T}}\boxed{\mathrm{K}'^{-\mathrm{T}}\mathrm{E}\mathrm{K}^{-1}}\mathbf{x} = 0$$

$$\hat{\mathbf{x}}'^{\mathrm{T}}\boxed{\mathrm{E}}\hat{\mathbf{x}} = 0$$

$$\hat{\mathbf{x}}'^{\mathrm{T}}\boxed{\lfloor\mathbf{t}\rfloor_{\times}\mathrm{R}}\hat{\mathbf{x}} = 0$$

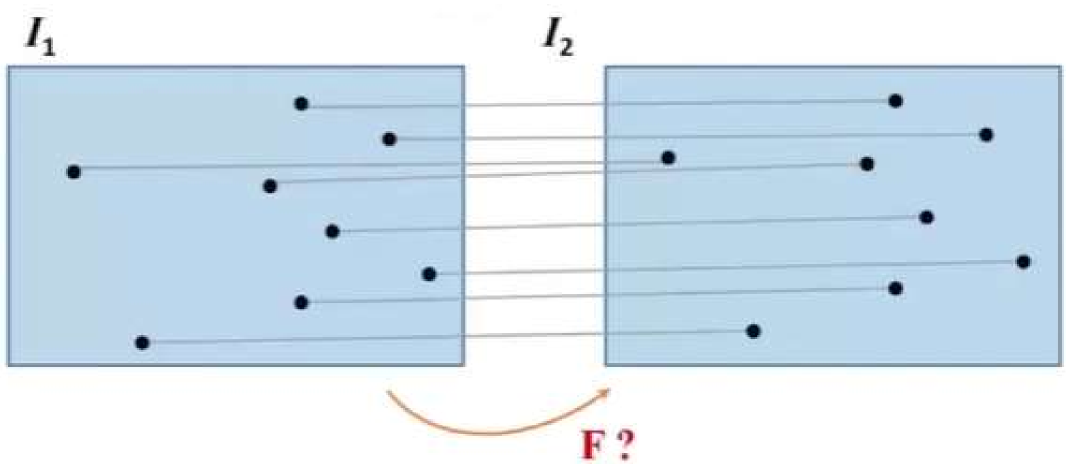
- E is the Essential Matrix which can be expressed in terms of the relative transformation between two image frames.

Properties of Essential Matrix

- Five degree of freedom (3+3-1):
  - R and **t** have 3 degree of freedom each
  - But there is an overall scale ambiguity ⇒ -1 dof
- Singular values:
  - A 3 x 3 matrix is an essential matrix iff two of its singular values are equal, and the third is zero

Linear 8-points Algorithm for F Matrix

- **Given**: A set of points correspondences $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ between two images.
- **Compute**: The Fundamental matrix F.



Linear 8-points Algorithm for F Matrix

- For any pair of matching points $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ in two images, the 3x3 fundamental matrix is defined by the equation:

$$\mathbf{x}'^{\mathsf{T}} \mathrm{F} \mathbf{x} = 0$$

- Let $\mathbf{x} = (x, y, 1)^T$ and $\mathbf{x}' = (x', y', 1)^T$, we rewrite the above equation as:

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

- Let **f** be the 9-vector made up of the entries of F in row-major order, we get:

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1)\,\mathbf{f} = 0$$

Linear 8-points Algorithm for F Matrix

- From a set of *n* point matches, we obtain a set of linear equations of the form:

$$\mathtt{A}\mathbf{f} = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = \mathbf{0}$$

- A is a n x 9 matrix.
- For a non-trivial solution to exist, rank(A)=8 since f is a 9-vector.
- A minimum of 8-point correspondences is needed to solve for f.

Linear 8-points Algorithm for F Matrix

- For noisy data, we obtain the solution of f by finding the least-squares solution.
- Least-squares solution for **f** is the singular vector corresponding to the smallest singular value of A.
- That is the last column of V in the SVD $A = UDV^T$.
- Similar to homography estimation, data normalization is needed.

To continue…

Essential Matrix

Proof:

Previously we seen $F = [\mathbf{e}']_\times P'P^+$, since $P = K[I \mid 0]$ and $P' = K'[R \mid \mathbf{t}]$, we have:

$$P^+ = \begin{bmatrix} K^{-1} \\ 0_{1\times3} \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 0_{3\times1} \\ 1 \end{bmatrix}$$

and

$$F = [\mathbf{e}']_\times P'P^+ = [P'\mathbf{C}]_\times P'P^+$$

$$= [K'\mathbf{t}]_\times K'RK^{-1} = K'^{-T}[\mathbf{t}]_\times RK^{-1}$$