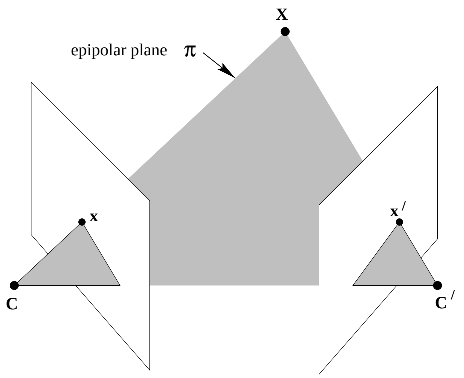
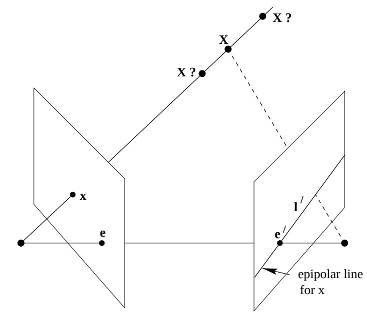
Epipolar geometry and the fundamental matrix

- Let $\mathbf{X}$ be a point in $\mathcal{P}^3$. Let $\mathbf{x}$ and $\mathbf{x}'$ be its mapping in two images through the camera centers $\mathbf{C}$ and $\mathbf{C}'$.
- The point $\mathbf{X}$, the camera centers $\mathbf{C}$ and $\mathbf{C}'$ and the (3D points correspon to) the mapped points $\mathbf{x}$ and $\mathbf{x}'$ will lie in the same plane $\boldsymbol{\pi}$.
- This plane is called the *epipolar plane* for $\mathbf{C}$, $\mathbf{C}'$ and $\mathbf{X}$.



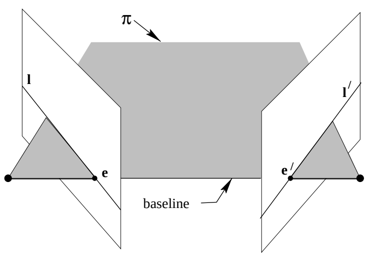
Epipolar lines

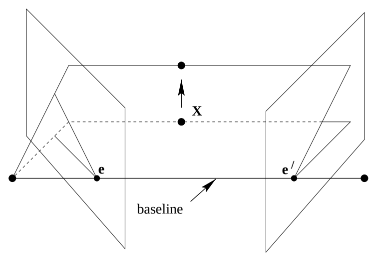
- Given a point $\mathbf{x}$ in image 1, the epipolar plane $\boldsymbol{\pi}$ is defined by the ray through $\mathbf{x}$ and $\mathbf{C}$ and the baseline through $\mathbf{C}$ and $\mathbf{C}'$.
- A corresponding point $\mathbf{x}'$ thus has to lie on the intersecting line $\mathbf{l}'$ between the epipolar plane $\boldsymbol{\pi}$ and image plane 2.
- The line $\mathbf{l}'$ is the projection of the ray through $\mathbf{x}$ and $\mathbf{X}$ in image 2 and is called the *epipolar line* to $\mathbf{x}$.



Epipoles

- The intersection points between the base line and the image planes are called *epipoles*.
- The epipole $\mathbf{e}'$ in image 2 is the mapping of the camera center $\mathbf{C}$.
- The epipole $\mathbf{e}$ in image 1 is the mapping of the camera center $\mathbf{C}'$.
- Since all epipolar planes intersect both camera centers, all epipolar lines will intersect the epipoles.





Examples

The fundamental matrix $\mathtt{F}$

- The fundamental matrix $\mathtt{F}$ is the algebraic representation of the epipolar geometry. It describes the mapping $\mathbf{x} \mapsto \mathbf{l}'$ between a point $\mathbf{x}$ in one image and its epipolar line $\mathbf{l}'$ in another image.
- Let $\mathtt{P}$ and $\mathtt{P}'$ be the camera matrices for image 1 and 2. The ray in $\mathcal{P}^3$ that is projected onto the point $\mathbf{x}$ in image 1 is

$$\mathbf{X}(\lambda) = \mathtt{P}^+\mathbf{x} + \lambda\mathbf{C},$$

  where $\mathtt{P}^+$ is the psuedo-inverse to $\mathtt{P}$, i.e. $\mathtt{PP}^+ = \mathtt{I}$, and $\mathtt{P}\mathbf{C} = \mathbf{0}$.
- The line $\mathbf{X}(\lambda)$ intersects the points $\mathtt{P}^+\mathbf{x}$ and $\mathbf{C}$. These points are mapped into the other camera $\mathtt{P}'$ at $\mathtt{P}'\mathtt{P}^+\mathbf{x}$ and $\mathtt{P}'\mathbf{C}$. The epipolar line $\mathbf{l}'$ intersects these projected points, i.e. $\mathbf{l}' = (\mathtt{P}'\mathbf{C}) \times (\mathtt{P}'\mathtt{P}^+\mathbf{x})$.

The fundamental matrix $\mathtt{F}$

- The point $\mathtt{P}'\mathbf{C}$ is the epipole $\mathbf{e}'$, i.e. the projection of the camera center in the other camera. The epipolar line can thus be written as

$$\mathbf{l}' = \mathbf{e}' \times (\mathtt{P}'\mathtt{P}^+\mathbf{x})$$

  or

$$\mathbf{l}' = [\mathbf{e}']_\times (\mathtt{P}'\mathtt{P}^+)\mathbf{x} = \mathtt{F}\mathbf{x},$$

  where

$$\mathtt{F} = [\mathbf{e}']_\times (\mathtt{P}'\mathtt{P}^+).$$

Example

- Assume the camera matrices correspond to a calibrated stereo rig with the world origin in camera center 1.

$$\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}],\ \mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}].$$

- Then $\mathtt{P}^+ = \begin{bmatrix} \mathtt{K}^{-1} \\ \mathbf{0}^\top \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$ and

$$\mathtt{F} = [\mathtt{P}'\mathbf{C}]_\times \mathtt{P}'\mathtt{P}^+ = [\mathtt{K}'\mathbf{t}]_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}[\mathtt{R}^\top\mathbf{t}]_\times \mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^\top[\mathtt{K}\mathtt{R}^\top\mathbf{t}]_\times$$

- Note that the epipoles are

$$\mathbf{e} = \mathtt{P}\begin{bmatrix} -\mathtt{R}^\top\mathbf{t} \\ 1 \end{bmatrix} = \mathtt{K}\mathtt{R}^\top\mathbf{t},\ \mathbf{e}' = \mathtt{P}'\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \mathtt{K}'\mathbf{t}.$$

- We can thus write

$$\begin{aligned} \mathtt{F} = [\mathbf{e}']_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \ldots \quad &= \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^\top[\mathbf{e}]_\times, \\ \mathtt{F}^\top = \quad &= \mathtt{K}^{-\top}\mathtt{R}^\top\mathtt{K}'^\top[\mathbf{e}']_\times. \end{aligned}$$

Correspondence

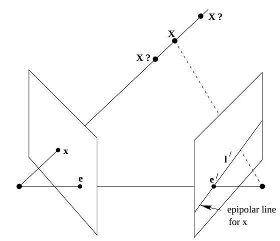
- Given two cameras with different camera centers, the fundamental matrix $\mathtt{F}$ is a 3×3 homogenous matrix with rank 2.
- For each corresponding point pair $\mathbf{x} \leftrightarrow \mathbf{x}'$ it satisfies

$$\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0,$$

  since if $\mathbf{x}$ and $\mathbf{x}'$ are corresponding points, then $\mathbf{x}'$ is on the epipolar line $\mathbf{l}' = \mathtt{F}\mathbf{x}$ corresponding to $\mathbf{x}$, i.e.

$$\mathbf{x}'^\top \mathbf{l}' = 0 = \mathbf{x}'^\top \mathtt{F}\mathbf{x}.$$

- Similarly, $\mathbf{l} = \mathtt{F}^\top\mathbf{x}'$ is the epipolar line in image 1 corresponding to the point $\mathbf{x}'$ in image 2.



The epipoles

- The epipolar line $\mathbf{l}' = \mathtt{F}\mathbf{x}$ to each point $\mathbf{x}$ (except $\mathbf{e}$) intersects the epipole $\mathbf{e}'$. Thus $\mathbf{e}'$ satisfies $\mathbf{e}'^\top(\mathtt{F}\mathbf{x}) = (\mathbf{e}'^\top\mathtt{F})\mathbf{x} = 0$ for all $\mathbf{x}$.
- This implies that $\mathbf{e}'^\top\mathtt{F} = \mathbf{0}^\top$ or $\mathtt{F}^\top\mathbf{e}' = \mathbf{0}$. The epipole $\mathbf{e}'$ is thus a null vector to $\mathtt{F}^\top$ (in the left null-space of $\mathtt{F}$).
- Similarly, $\mathtt{F}\mathbf{e} = \mathbf{0}$, i.e. $\mathbf{e}$ is a null-vector to $\mathtt{F}$ (in the right null-space of $\mathtt{F}$).

The number of degrees of freedom

- The fundamental matrix $\mathtt{F}$ has 7 degrees of freedom: A 3×3 homogenous matrix has 8 degrees of freedom. The constraint rank$(\mathtt{F}) = 2$ or $\det(\mathtt{F}) = 0$ reduces the number to 7.

Projektive invariance

- The correspondence relation $\mathbf{x}'^\top\mathtt{F}\mathbf{x} = 0$ is invariant under a homography in $\mathcal{P}^2$. If $\hat{\mathbf{x}} = \mathtt{H}\mathbf{x}$ and $\hat{\mathbf{x}}' = \mathtt{H}'\mathbf{x}'$ then

$$\mathbf{x}'^\top\mathtt{F}\mathbf{x} = \hat{\mathbf{x}}'^\top\mathtt{H}'^{-\top}\mathtt{F}\mathtt{H}^{-1}\hat{\mathbf{x}} = \hat{\mathbf{x}}'^\top\hat{\mathtt{F}}\hat{\mathbf{x}},$$

  where $\hat{\mathtt{F}} = \mathtt{H}'^{-\top}\mathtt{F}\mathtt{H}^{-1}$ is the fundamental matrix corresponding to $\hat{\mathbf{x}} \leftrightarrow \hat{\mathbf{x}}'$.
- The fundamental matrix $\mathtt{F}$ is invariant under a homography in $\mathcal{P}^3$. Let $\mathtt{H}$ be a 4×4 matrix corresponding to a projective mapping of $\mathcal{P}^3$. Then the camera pairs $(\mathtt{P}, \mathtt{P}')$ and $(\mathtt{P}\mathtt{H}, \mathtt{P}'\mathtt{H})$ have the same fundamental matrix.
- The points $\mathbf{x} = \mathtt{P}\mathbf{X} = (\mathtt{P}\mathtt{H})(\mathtt{H}^{-1}\mathbf{X})$ and $\mathbf{x}' = \mathtt{P}'\mathbf{X} = (\mathtt{P}'\mathtt{H})(\mathtt{H}^{-1}\mathbf{X})$ are corresponding mappings of $\mathbf{X}$ in the cameras $\mathtt{P}$ and $\mathtt{P}'$ *and* corresponding mappings of $\mathtt{H}^{-1}\mathbf{X}$ in the cameras $\mathtt{P}\mathtt{H}$ and $\mathtt{P}'\mathtt{H}$.
- Thus a homography $\mathtt{H}$ in $\mathcal{P}^3$ does affect the world points $\mathbf{X}$ and cameras $\mathtt{P}, \mathtt{P}'$, but not $\mathtt{F}$.
- This means that the fundamental matrix $\mathtt{F}$ determines the camera matrices $\mathtt{P}, \mathtt{P}'$ up to a right multiplication by a 3D projective transformation.

Canonical form

- Given this ambiguity a *canonical form* for the camera pairs is defined corresponding to a fundamental matrix where the first camera $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$ has center at the origin and world coordinate axes.
- If the second camera is $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$ then the fundamental matrix $\mathtt{F}$ corresponding to the canonical cameras is

$$\mathtt{F} = [\mathbf{m}]_\times \mathtt{M}.$$

- For finite cameras $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}], \mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$ we have

$$\mathtt{F} = [\mathtt{K}'\mathbf{t}]_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}.$$

w symmetry and the fundamental matrix

- A non-zero matrix $\mathtt{F}$ is the fundamental matrix corresponding to the camera pair $\mathtt{P}, \mathtt{P}'$ iff $\mathtt{P}'^\top \mathtt{F}\mathtt{P}$ is skew symmetric.
- The condition that $\mathtt{P}'^\top \mathtt{F}\mathtt{P}$ is skew symmetrical is equivalent to that

$$\mathbf{X}^\top \mathtt{P}'^\top \mathtt{F}\mathtt{P}\mathbf{X} = 0$$

for all $\mathbf{X}$. With $\mathbf{x}' = \mathtt{P}'\mathbf{X}$ and $\mathbf{x} = \mathtt{P}\mathbf{X}$ this becomes

$$\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0,$$

which is the defining equation for the fundamental matrix.

Canonical camera pairs given F

- Let $\mathtt{F}$ be a fundamental matrix and $\mathtt{S}$ an arbitrary skew symmetric matrix. Define the camera matrices

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \text{ and } \mathtt{P}' = [\mathtt{S}\mathtt{F} \mid \mathbf{e}'],$$

where $\mathbf{e}'$ is the left epipole of $\mathtt{F}$, $\mathbf{e}'^\top \mathtt{F} = \mathbf{0}^\top$ and assume that $\mathtt{P}'$ is a valid camera matrix (has rank 3). Then $\mathtt{F}$ is the fundamental matrix corresponding to $(\mathtt{P}, \mathtt{P}')$.
- Check by verifying that

$$\mathtt{P}'^\top \mathtt{F}\mathtt{P} = [\mathtt{S}\mathtt{F} \mid \mathbf{e}']^\top \mathtt{F}[\mathtt{I} \mid \mathbf{0}] = \begin{bmatrix} \mathtt{F}^\top \mathtt{S}^\top \mathtt{F} & \mathbf{0} \\ \mathbf{e}'^\top \mathtt{F} & 0 \end{bmatrix} = \begin{bmatrix} \mathtt{F}^\top \mathtt{S}^\top \mathtt{F} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix}$$

is skew symmetric.

Canonical camera pairs given F

- In order for the matrix $\mathtt{P}'$ to have rank 3 $\mathbf{s}^\top \mathbf{e}'$ has to be non-zero, where $\mathbf{s}$ is a null-vector of $\mathtt{S} = [\mathbf{s}]_\times$. A working choice is $\mathbf{s} = \mathbf{e}'$ leading to the camera pairs

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \text{ and } \mathtt{P}' = [[\mathbf{e}']_\times \mathtt{F} \mid \mathbf{e}'].$$

- The most general formulation for a canonical camera pair is

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}],\ \mathtt{P}' = [[\mathbf{e}']_\times \mathtt{F} + \mathbf{e}'\mathbf{v}^\top \mid \lambda \mathbf{e}'],$$

where $\mathbf{v}$ is an arbitrary 3-vector and λ is a scalar.

Normalized coordinates

- Study a camera matrix $\mathtt{P} = \mathtt{K}[\mathtt{R} \mid \mathbf{t}]$ and let $\mathbf{x} = \mathtt{P}\mathbf{X}$ be an arbitrary point in the image.
- If the camera calibration matrix $\mathtt{K}$ is known we may apply its inverse on the point $\mathbf{x}$ and get $\hat{\mathbf{x}} = \mathtt{K}^{-1}\mathbf{x}$.
- Then $\hat{\mathbf{x}} = [\mathtt{R} \mid \mathbf{t}]\mathbf{X}$ is the projection of $\mathbf{X}$ expressed in *normalized coordinates*.
- The camera matrix

$$\mathtt{K}^{-1}\mathtt{P} = \mathtt{I}[\mathtt{R} \mid \mathbf{t}]$$

is called a *normalized camera matrix* and has camera calibration matrix $\mathtt{K} = \mathtt{I}$.

The essential matrix E

- Study a normalized camera pair $P = [I \mid \mathbf{0}]$, $P' = [R \mid \mathbf{t}]$. The fundamental matrix corresponding to normalized camera pairs is called the *essential matrix* and is on the form
$$E = [\mathbf{t}]_\times R = R[R^\top \mathbf{t}]_\times .$$
- The defining equation for the essential matrix is
$$\hat{\mathbf{x}}'^\top E \hat{\mathbf{x}} = 0,$$
expressed in normalized image coordinates for the corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$.
- Substitution with $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ gives
$$\mathbf{x}'^\top K'^{-\top} E K^{-1} \mathbf{x} = 0$$
leading to
$$F = K'^{-\top} E K^{-1} \text{ or } E = K'^\top F K.$$

The number of degrees of freedom for

- The essential matrix $E = [\mathbf{t}]_\times R$ has 5 degrees of freedon; 3 rotation angles in R, 3 elements in $\mathbf{t}$, but arbitrary scale.
- The fewer degrees of freedom correspond to one additional constraint; a 3×3 matrix is an essential matrix if two of its singular values are equal and the last is zero.

The number of degrees of freedom for E

- Study the factorization $E = [\mathbf{t}]_\times R = SR$, where S is skew symmetric. We will use the matrices
$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
that are orthogonal and skew symmetric, respectively. Note that $Z = -\operatorname{diag}(1,1,0)W$.
- A skew symmetric matrix S can be written as $S = kUZU^\top$, where U is orthogonal. Thus
$$S = U \operatorname{diag}(1,1,0) W U^\top$$
up to scale and
$$E = SR = U \operatorname{diag}(1,1,0)(WU^\top R),$$
which is a singular value decomposition of E with two singular values equal and the third equal to zero.

Calculation of the camera matrices fron

- Let the first camera matrix be $P = [I \mid \mathbf{0}]$. In order to calculate the other camera matrix P' it is necessary to factorize E into a product SR by a skew symmetric and a rotation matrix. Given $S = [\mathbf{t}]_\times$ and R, P' is given by $P' = [R \mid \mathbf{t}]$.
- Let E has the singular value decomposition $E = U \operatorname{diag}(1,1,0) V^\top$. Ignoring sign, there are two possible factorizations $E = SR$:
$$S = UZU^\top, \; R = UWV^\top \text{ or } R = UW^\top V^\top$$
- The factorization gives the $\mathbf{t}$ part of the camera matrix P' up to scale from $S = [\mathbf{t}]_\times$. If we choose $\|\mathbf{t}\| = 1$ we get a unit baseline. Furthermore $S\mathbf{t} = 0$
$$S\mathbf{t} = UZU^\top \mathbf{t} = U\left(UZ^\top\right)^\top \mathbf{t} = U \begin{bmatrix} \mathbf{u}_2 & -\mathbf{u}_1 & \mathbf{0} \end{bmatrix}^\top \mathbf{t} = 0$$
gives that $\mathbf{t} = \mathbf{u}_3$, where $\mathbf{u}_i$ is the i:th column of U. The sign of E and hence $\mathbf{t}$ can however not be determined which leads to 4 different possibilities for the second camera P'.

The 4 camera factorizations of $\mathtt{E}$

- Given a singular value decomposition of $\mathtt{E} = \mathtt{U}\,\mathrm{diag}(1,1,0)\mathtt{V}^\top$ and a canonical camera 1 $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, there are 4 alternative camera pairs:

$$\mathtt{P}' = [\mathtt{UWV}^\top \mid +\mathbf{u}_3], \qquad \mathtt{P}' = [\mathtt{UWV}^\top \mid -\mathbf{u}_3],$$
$$\mathtt{P}' = [\mathtt{UW}^\top\mathtt{V}^\top \mid +\mathbf{u}_3], \qquad \mathtt{P}' = [\mathtt{UW}^\top\mathtt{V}^\top \mid -\mathbf{u}_3].$$

- These 4 options have geometric interpretations; baseline reversal and rotation by the second camera $180°$ around the baseline.