Two-view geometry: Epipolar geometry and Fundamental matrix

Nov 21, 2017

EE, KAIST

김창익 (Kim, Changick)

Contents

- Epipolar geometry
- Fundamental matrix
- Essential matrix

They come mostly from the following source.
Marc Pollefeys U. of North Carolina
And
H&Z's book Ch9.

Three questions:

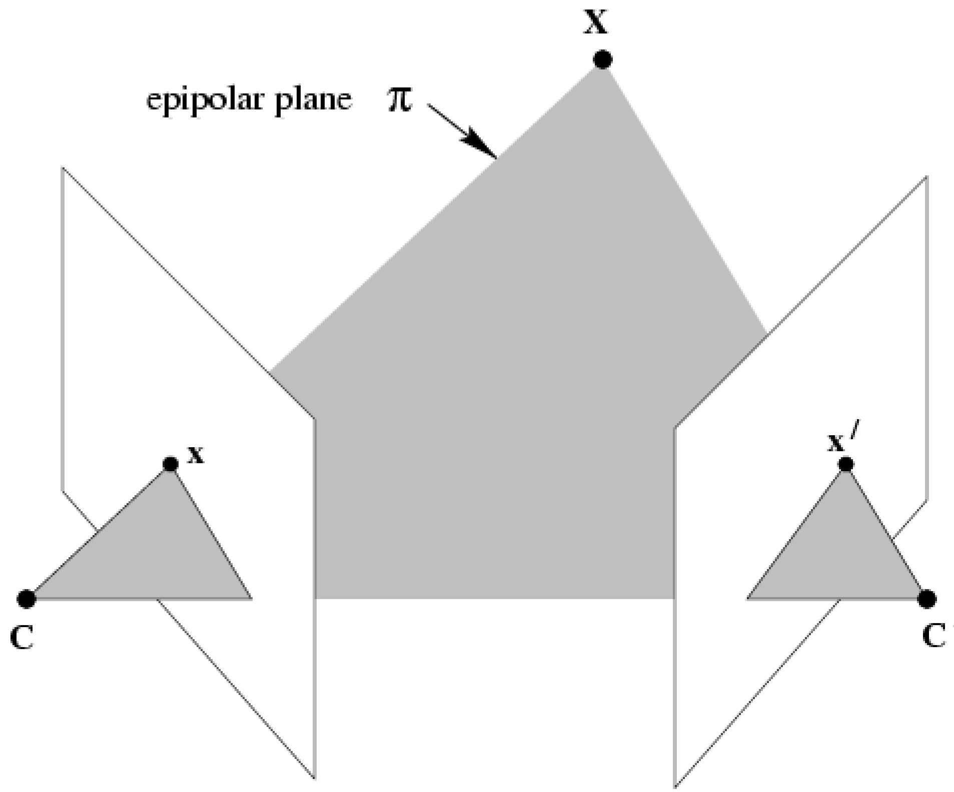
(i) Correspondence geometry: Given an image point x in the first view, how does this constrain the position of the corresponding point x' in the second image?

A) A point in one view defines an epipolar line in the other view.

(ii) Camera geometry (motion): Given a set of corresponding image points $\{x_i \leftrightarrow x'_i\}$, $i=1,\ldots,n$, what are the cameras P and P' for the two views?

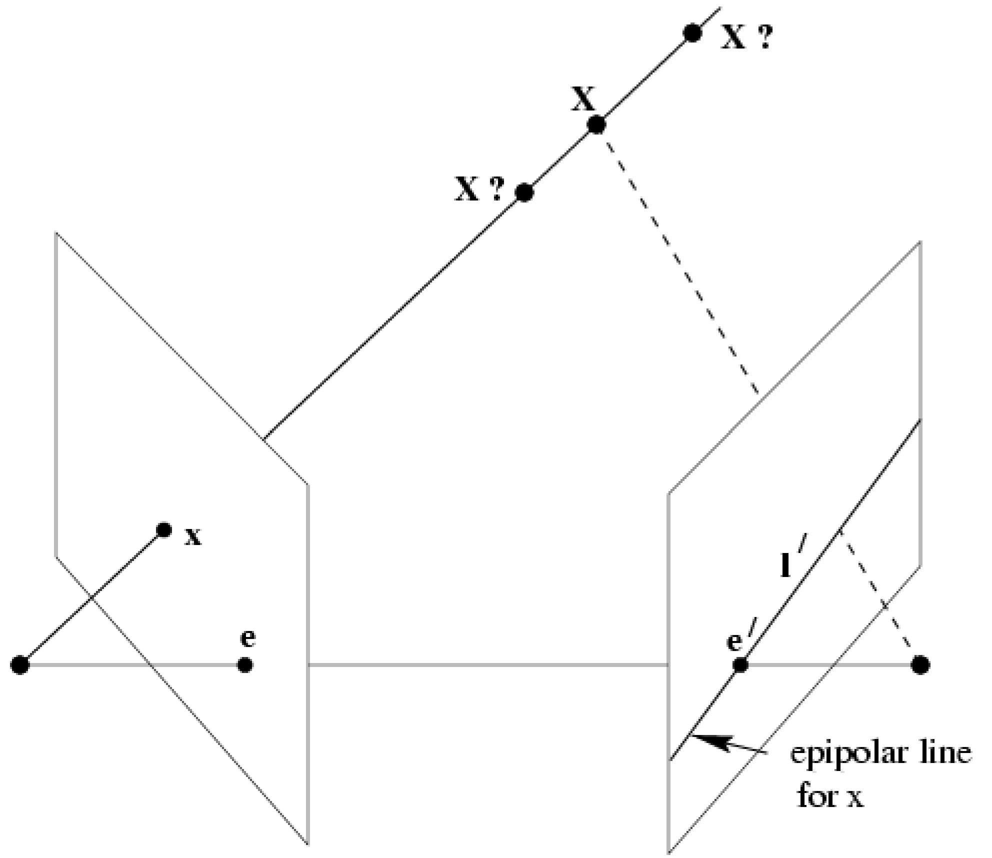
(iii) Scene geometry (structure): Given corresponding image points $x_i \leftrightarrow x'_i$ and cameras P, P', what is the position of (their pre-image) X in space?

The epipolar geometry



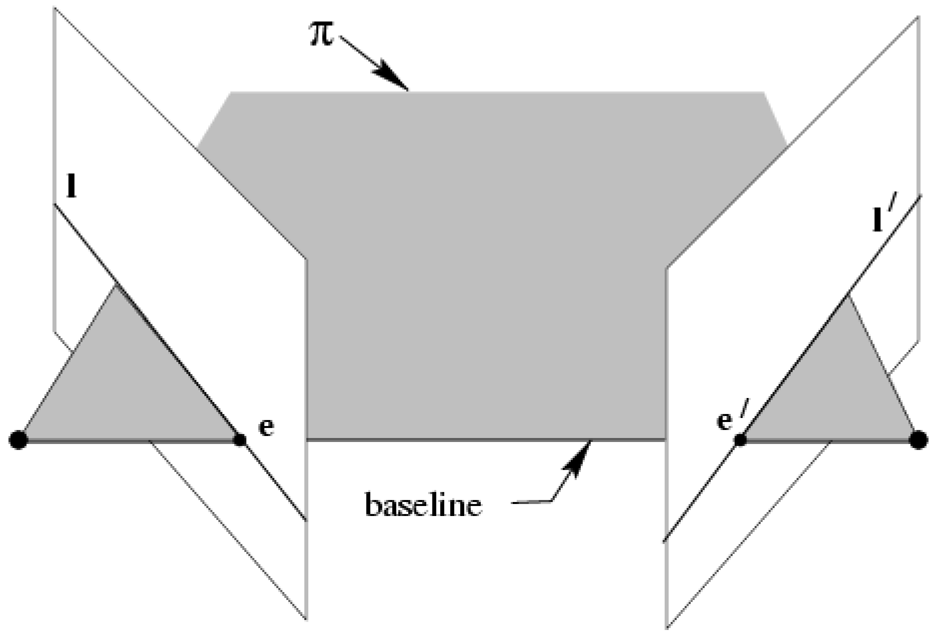
C,C',x,x' and X are coplanar

The epipolar geometry



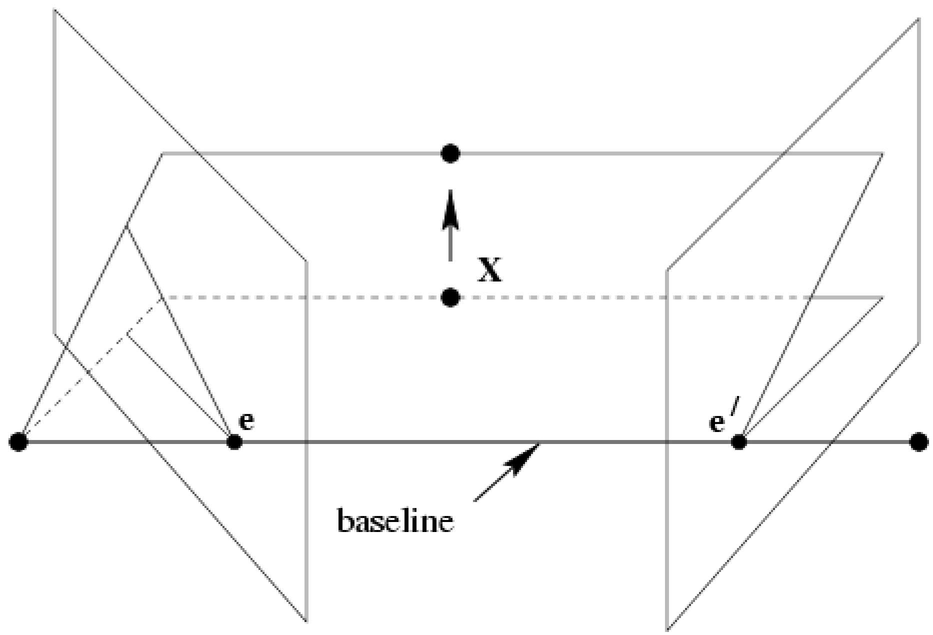
What if only C,C',x are known?

The epipolar geometry



All points on π project on l and l'

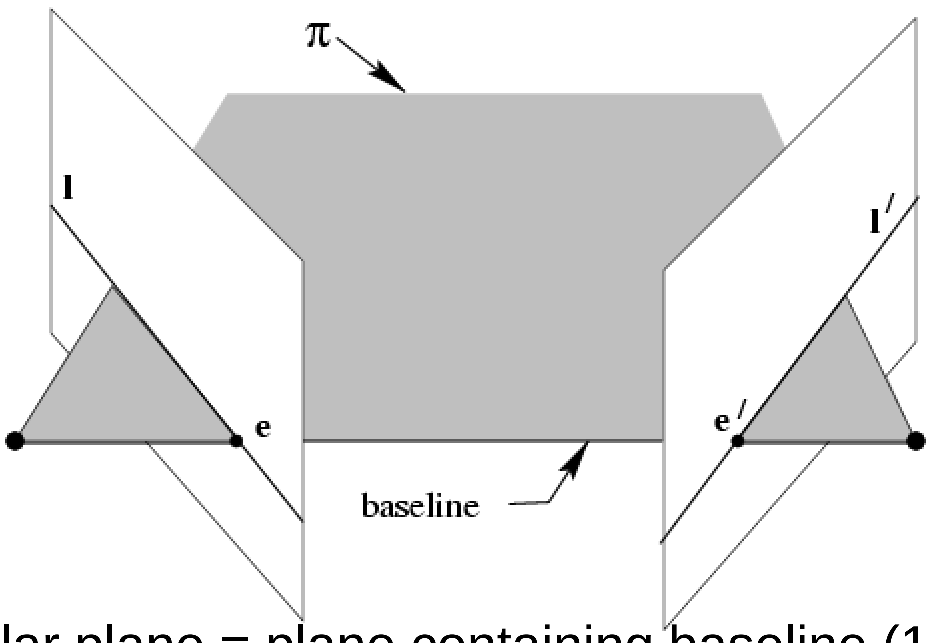
The epipolar geometry



Family of planes π and lines l and l'
Intersection in e and e'

The epipolar geometry

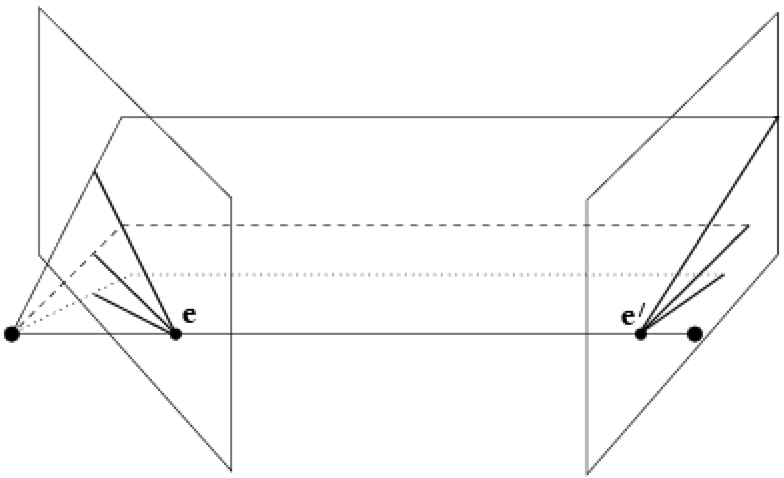
epipoles e,e’
= intersection of baseline with image plane
= projection of projection center in other image
= vanishing point of camera motion direction

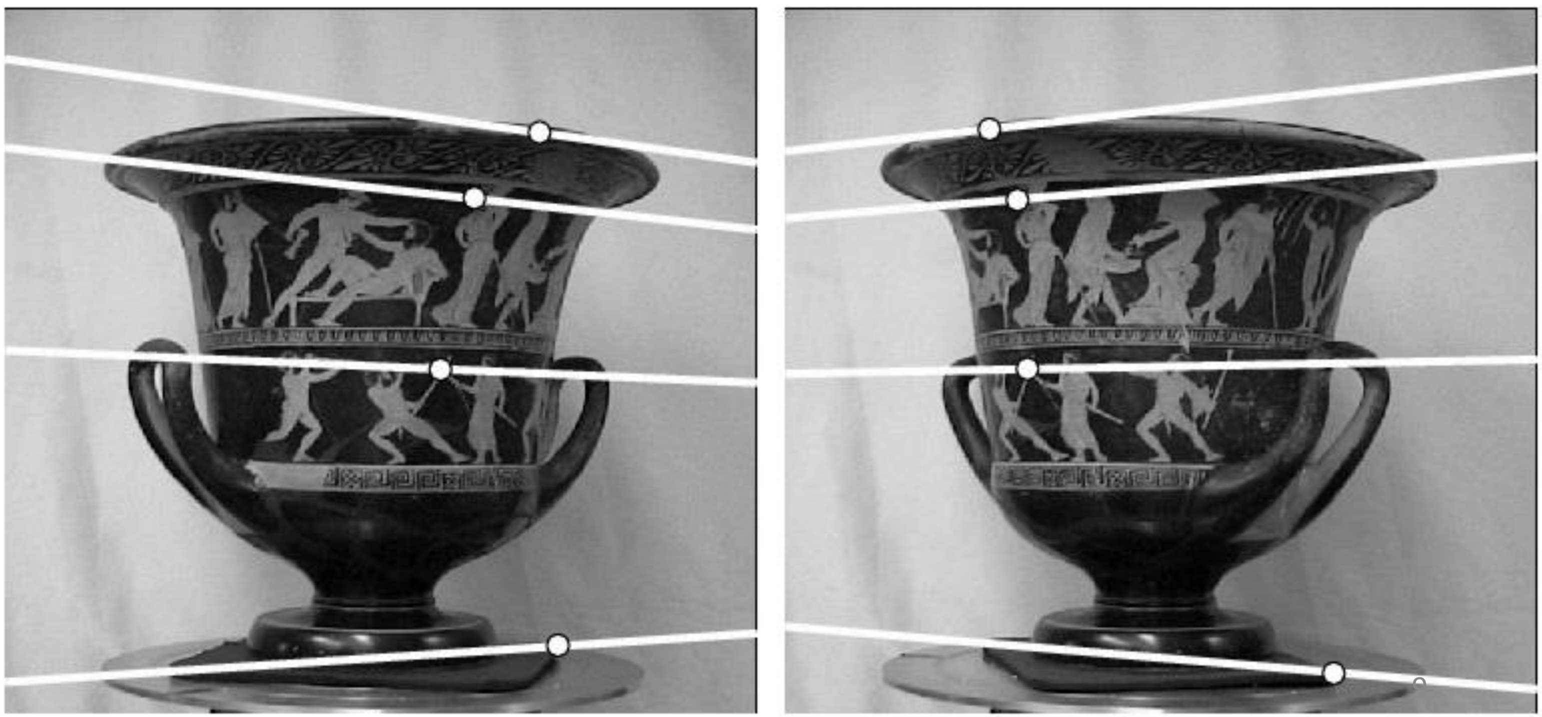


an epipolar plane = plane containing baseline (1-D family)

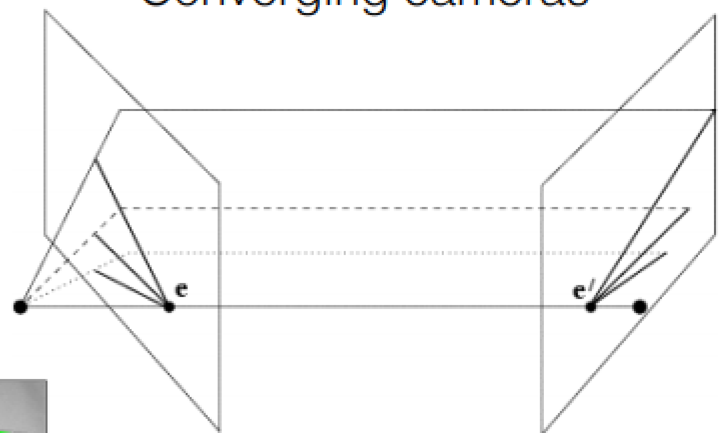
an epipolar line = intersection of epipolar plane with image
(always come in corresponding pairs)

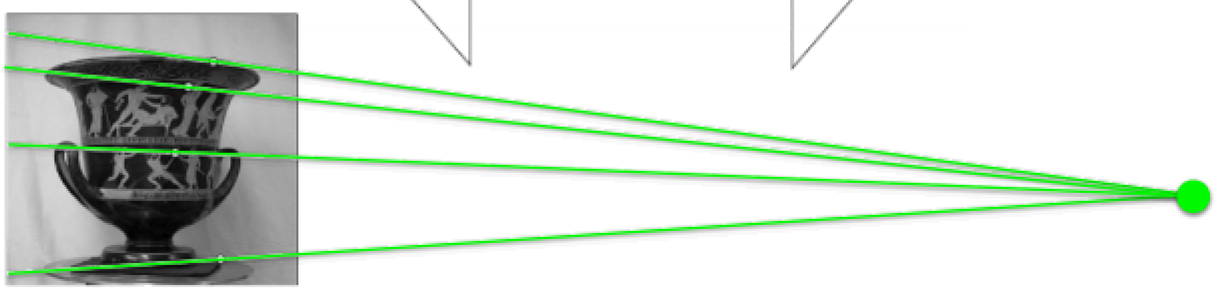
Example: converging cameras





Converging cameras

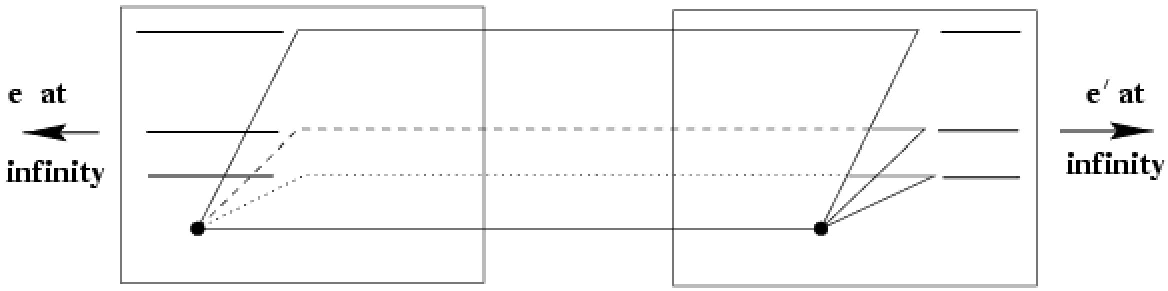




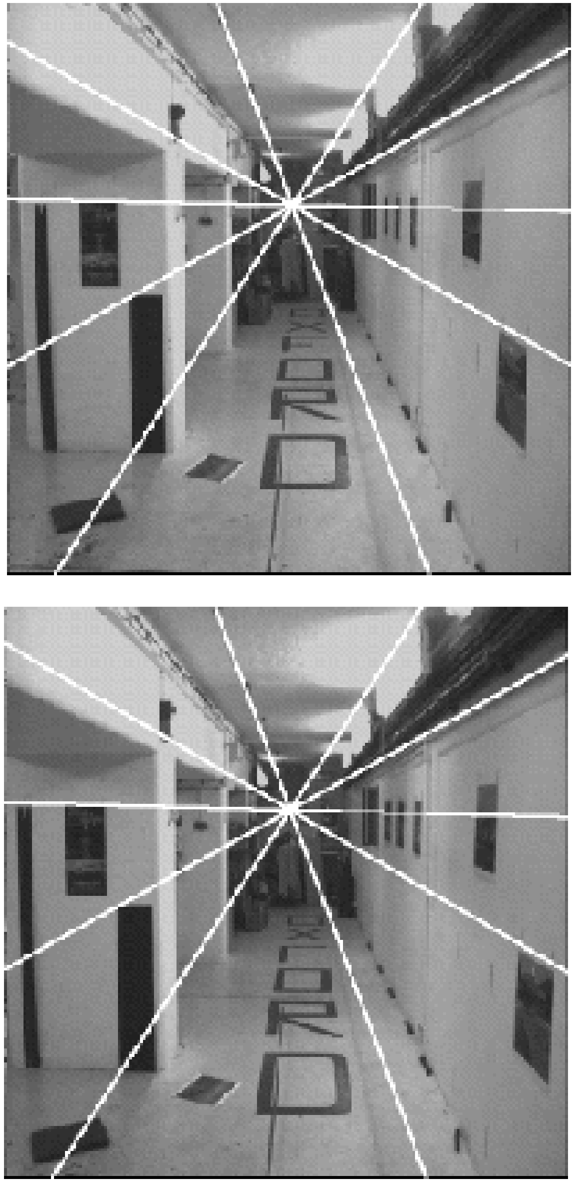
Where is the epipole in this image?

It's not always in the image

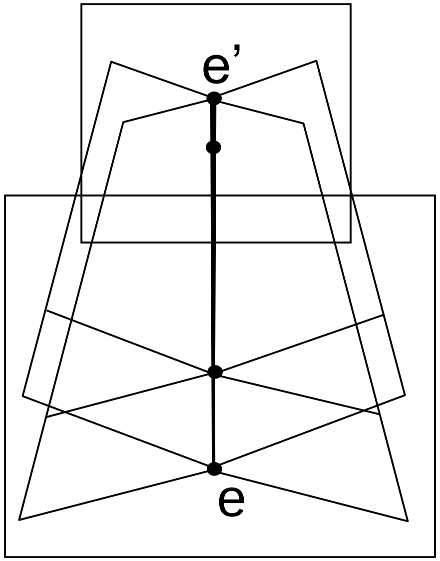
Example: motion parallel with image plane



Example: forward motion



Epipole has same coordinates in both images. Points move along lines radiating from “Focus of expansion”



The epipolar constraint is an important concept for stereo vision

Task: Match point in left image to point in right image

Left image

Right image

Want to avoid search over entire image

(if the images have been rectified)
Epipolar constrain reduces search to a single line

The fundamental matrix F

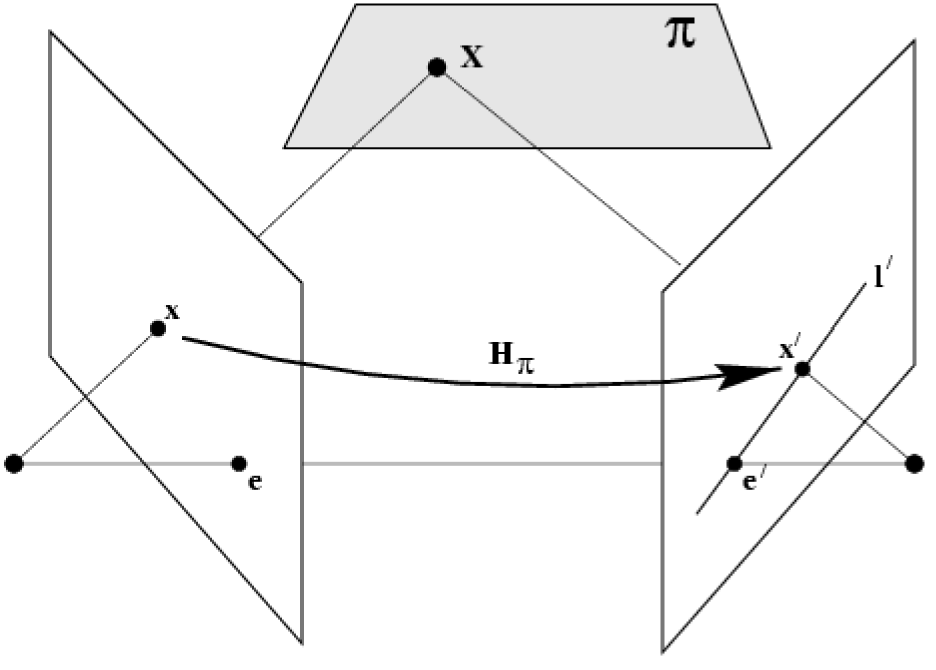
algebraic representation of epipolar geometry

$$x \mapsto l'$$

we will see that mapping is (singular) correlation
(i.e. projective mapping from points to lines)
represented by the fundamental matrix F

The fundamental matrix F

geometric derivation



$$x' = H_\pi x$$

$$l' = e' \times x' = [e']_\times H_\pi x = Fx$$

mapping from 2-D to 1-D family (rank 2)

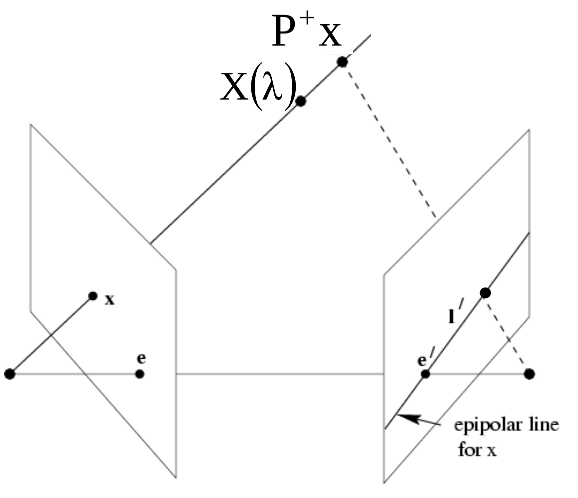
The fundamental matrix F

algebraic derivation

$$X(\lambda) = P^{+}x + \lambda C \qquad \left(P^{+}P = I\right)$$

$$l' = (P'C) \times (P'P^{+}x)$$

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



(note: doesn't work for C=C' $\Rightarrow$ F=0)

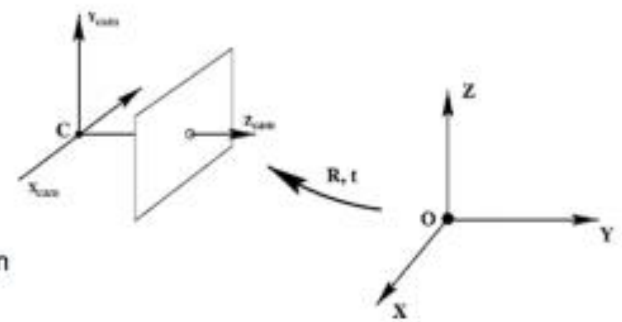
The fundamental matrix F

- choose camera matrices

$$P = K\,[\,R \mid t\,]$$

internal calibration — rotation — translation

from world to camera coordinate frame

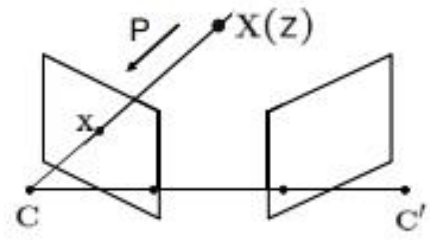


- first camera $\quad P = K\,[\,I \mid 0\,]$

world coordinate frame aligned with first camera

- second camera $\quad P' = K'\,[\,R \mid t\,]$

The fundamental matrix F



Step 1: for a point x in the first image back project a ray with camera $P = K[I \mid 0]$

A point X back projects to a ray

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = ZK^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = ZK^{-1}\mathbf{x}$$

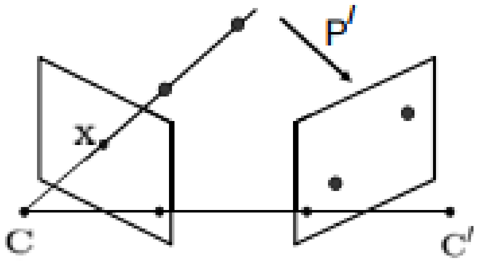
where Z is the point's depth, since

$$X(Z) = \begin{pmatrix} ZK^{-1}\mathbf{x} \\ 1 \end{pmatrix}$$

satisfies

$$PX(Z) = K[I \mid 0]X(Z) = \mathbf{x}$$

The fundamental matrix F



Step 2: choose two points on the ray and project into the second image with camera P′

Consider two points on the ray $X(z) = \begin{pmatrix} zK^{-1}\mathbf{x} \\ 1 \end{pmatrix}$

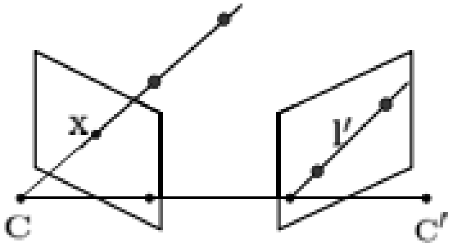
- $\mathbf{Z}$ = 0 is the camera centre $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$
- $\mathbf{Z}$ = ∞ is the point at infinity $\begin{pmatrix} K^{-1}\mathbf{x} \\ 0 \end{pmatrix}$

Project these two points into the second view

$$P' \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = K'[R \mid \mathbf{t}] \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = K'\mathbf{t} \qquad P' \begin{pmatrix} K^{-1}\mathbf{x} \\ 0 \end{pmatrix} = K'[R \mid \mathbf{t}] \begin{pmatrix} K^{-1}\mathbf{x} \\ 0 \end{pmatrix} = K'RK^{-1}\mathbf{x}$$

The fundamental matrix F

Step 3: compute the line through the two image points using the relation $\mathbf{l}' = \mathbf{p} \times \mathbf{q}$



Compute the line through the points $\mathbf{l}' = (\mathtt{K}'\mathbf{t}) \times (\mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x})$ $\longrightarrow$ $\mathrm{F} = [\mathrm{e}']_\times \mathrm{K}'\mathrm{R}\mathrm{K}^{-1}$

Using the identity $(\mathtt{M}\mathbf{a}) \times (\mathtt{M}\mathbf{b}) = \mathtt{M}^{-\top}(\mathbf{a} \times \mathbf{b})$ where $\mathtt{M}^{-\top} = (\mathtt{M}^{-1})^\top = (\mathtt{M}^\top)^{-1}$

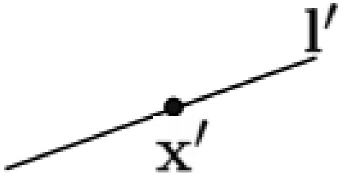
$$\mathbf{l}' = \mathtt{K}'^{-\top}\left(\mathbf{t} \times (\mathtt{R}\mathtt{K}^{-1}\mathbf{x})\right) = \underbrace{\mathtt{K}'^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1}}_{\mathrm{F}}\mathbf{x}$$

F is the fundamental matrix

$$\boxed{\mathbf{l}' = \mathrm{F}\mathbf{x} \qquad \mathrm{F} = \mathrm{K}'^{-\top}[\mathbf{t}]_\times \mathrm{R}\mathrm{K}^{-1}}$$

Points $\mathbf{x}$ and $\mathbf{x}'$ correspond $(\mathbf{x} \leftrightarrow \mathbf{x}')$ then $\mathbf{x}'^\top \mathbf{l}' = 0$

$$\boxed{\mathbf{x}'^\top \mathrm{F}\mathbf{x} = 0}$$



The fundamental matrix F

correspondence condition

The fundamental matrix satisfies the condition that for any pair of corresponding points $x \leftrightarrow x'$ in the two images

$$x'^{T} F x = 0 \qquad \left(x'^{T} l' = 0 \right)$$

The fundamental matrix F

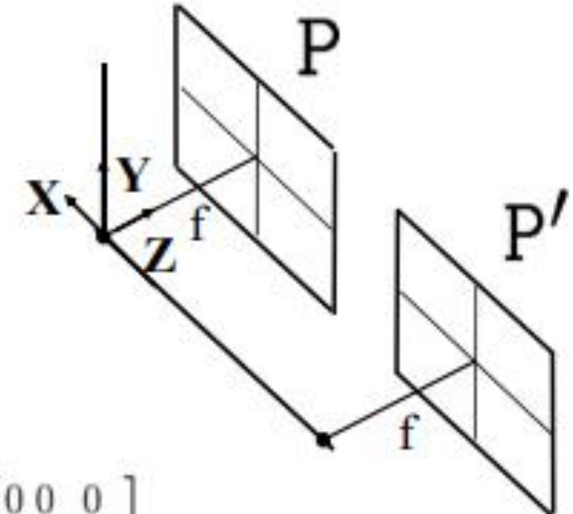
F is the unique 3x3 rank 2 matrix that satisfies $x'^T F x = 0$ for all $x \leftrightarrow x'$

(i) Transpose: if F is fundamental matrix for (P,P'), then F^T is fundamental matrix for (P',P)
(ii) Epipolar lines: $l' = Fx$ & $l = F^T x'$
(iii) Epipoles: on all epipolar lines, thus $e'^T F x = 0, \forall x$ $\Rightarrow e'^T F = 0$, similarly $Fe = 0$
(iv) F has 7 d.o.f. , i.e. 3x3-1(homogeneous)-1(rank2)
(v) F is a correlation, projective mapping from a point x to a line $l' = Fx$ (not a proper correlation, i.e. not invertible)

Example I: compute the fundamental matrix for a parallel camera stereo rig

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

$$\mathtt{K} = \mathtt{K}' = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathtt{R} = \mathtt{I} \quad \mathbf{t} = \begin{pmatrix} t_x \\ 0 \\ 0 \end{pmatrix}$$



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

$$= \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$[\mathbf{v}]_\times = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$$

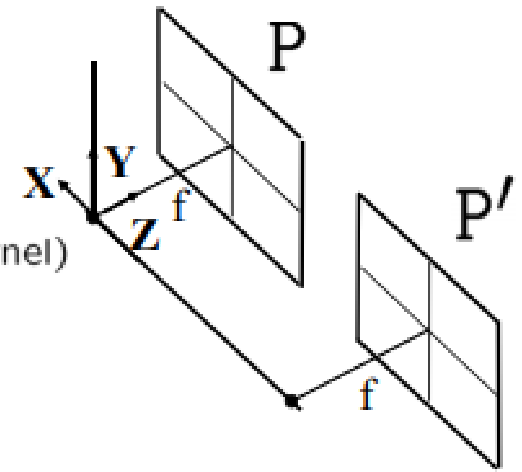
$$\mathbf{x}'^\top \mathtt{F}\mathbf{x} = \begin{pmatrix} x' & y' & 1 \end{pmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

= t_x/f
(but we are in homogeneous space)

- reduces to y = y′ , i.e. raster correspondence (horizontal scan-lines)

F is a rank 2 matrix

The epipole $\mathbf{e}$ is the null-space vector (kernel) of $\mathtt{F}$ i.e. $\mathtt{F}\mathbf{e} = \mathbf{0}$



In this case

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{0}$$

so that

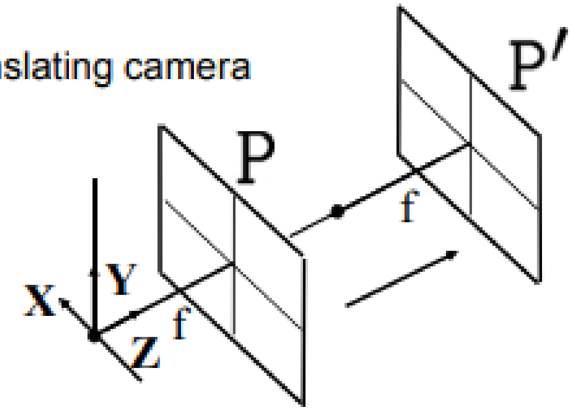
$$\mathbf{e} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Geometric interpretation ?

Example II: compute F for a forward translating camera

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

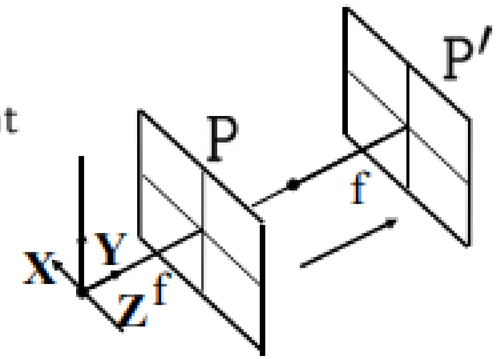
$$K = K' = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad R = I \quad \mathbf{t} = \begin{pmatrix} 0 \\ 0 \\ t_z \end{pmatrix}$$



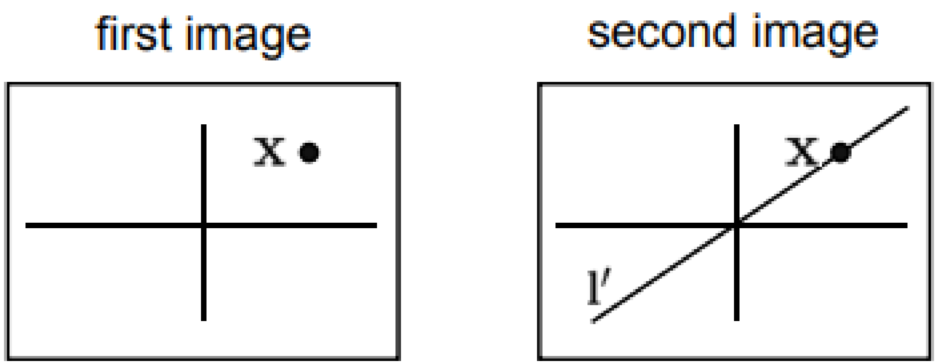
$$\begin{aligned} F &= K'^{-\top}[\mathbf{t}]_{\times} R K^{-1} \\ &= \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \lambda \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

From $\mathbf{l}' = \mathrm{F}\mathbf{x}$ the epipolar line for the point $\mathbf{x} = (x, y, 1)^\top$ is

$$\mathbf{l}' = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$



The points $(x, y, 1)^\top$ and $(0, 0, 1)^\top$ lie on this line



Summary: Properties of the Fundamental matrix

- $\mathtt{F}$ is a rank 2 homogeneous matrix with 7 degrees of freedom.

- Point correspondence:

  if $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$.

- Epipolar lines:
  - $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - $\mathbf{l} = \mathtt{F}^{\mathsf{T}}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.

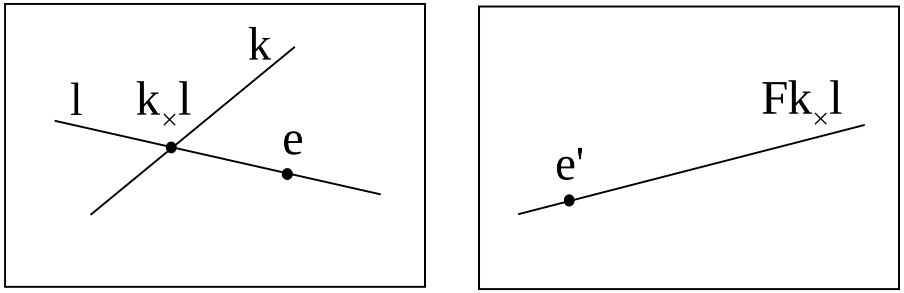
- Epipoles:
  - $\mathtt{F}\mathbf{e} = \mathbf{0}$.
  - $\mathtt{F}^{\mathsf{T}}\mathbf{e}' = \mathbf{0}$.

- Computation from camera matrices $\mathtt{P}, \mathtt{P}'$:

  $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$, $\mathtt{F} = \mathtt{K}'^{-\mathsf{T}}[\mathbf{t}]_{\times}\mathtt{R}\mathtt{K}^{-1}$

The epipolar line geometry

l,l’ epipolar lines, k line not through e
$\Rightarrow$ $l' = F[k]_x l$ and symmetrically $l = F^T[k']_x l'$

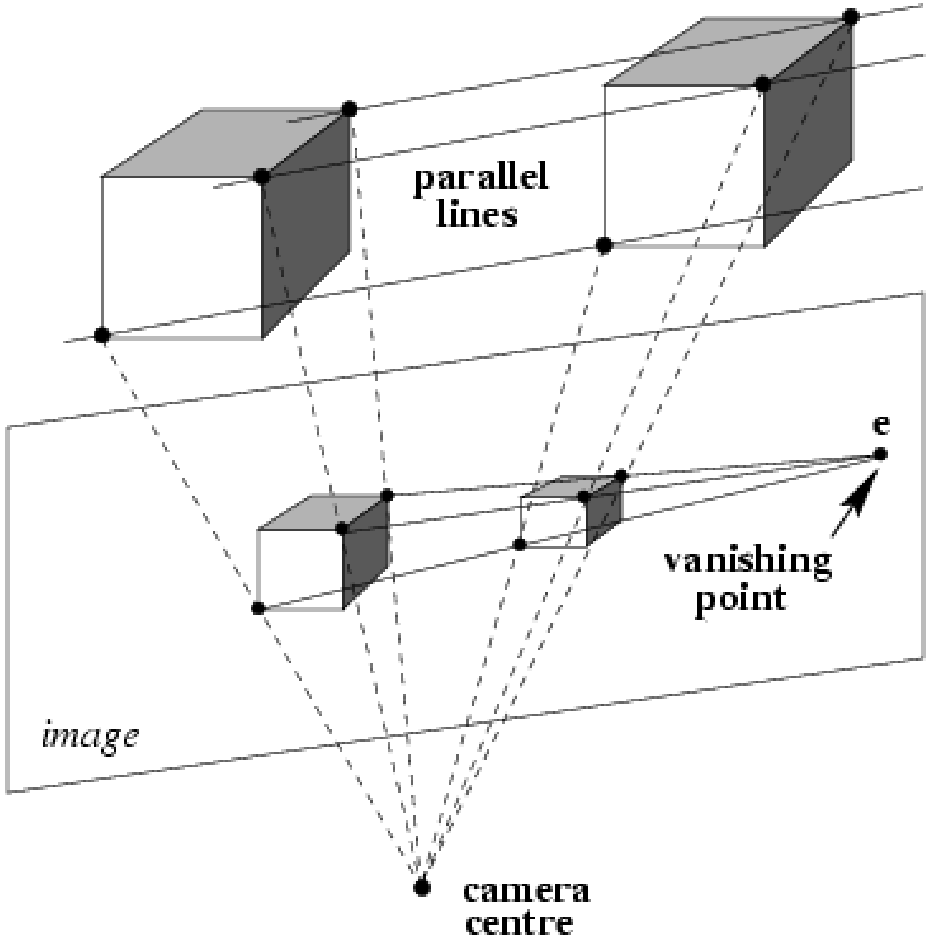


(pick k=e, since $e^T e \neq 0$)

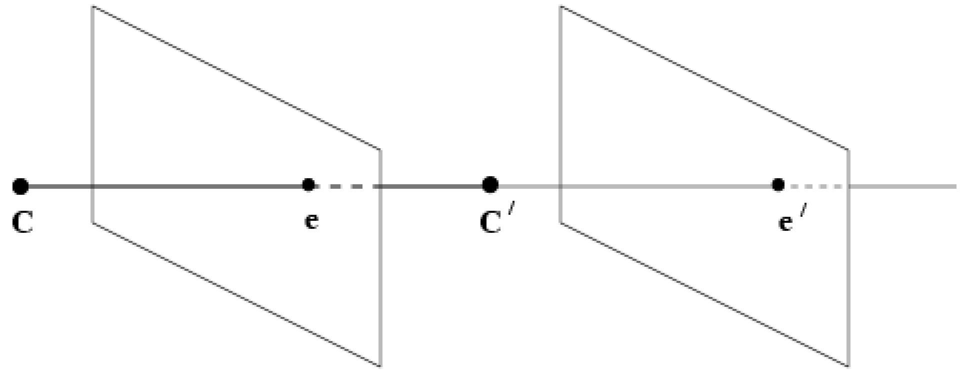
$$l' = F[e]_\times l \qquad l = F^T[e']_\times l'$$

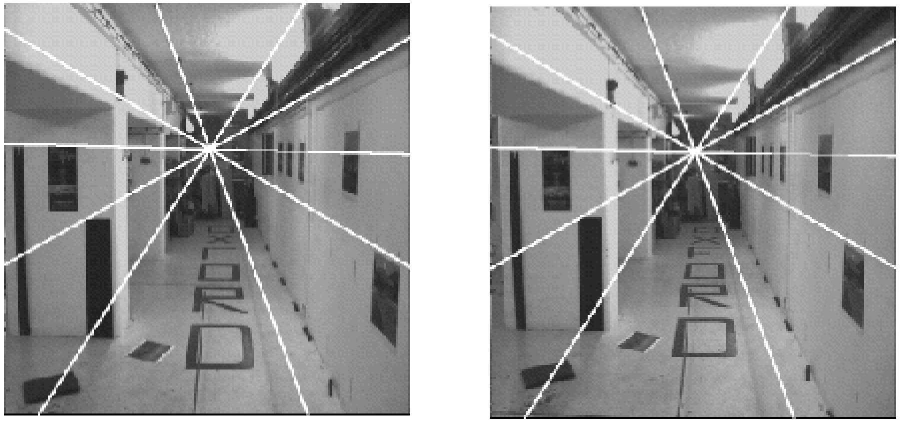
Fundamental matrix for pure translation

Equivalent situation to “Camera is stationary and the world undergoes a translation –**t**.”



Fundamental matrix for pure translation





Fundamental matrix for pure translation

$$F = [e']_\times KK^{-1} = [e']_\times$$

<u>example</u>: when the camera translation is parallel to the x-axis

$$e' = (1,0,0)^T \qquad F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}_\times$$

$$x'^T F x = 0 \Leftrightarrow y = y'$$

$$x = PX = K[I\,|\,0]X \qquad (X,Y,Z)^T = K^{-1}xZ$$

$$x' = P'X = K[I\,|\,t]\begin{bmatrix} K^{-1}x \\ Z^{-1} \end{bmatrix} \qquad x' = x + Kt/Z$$

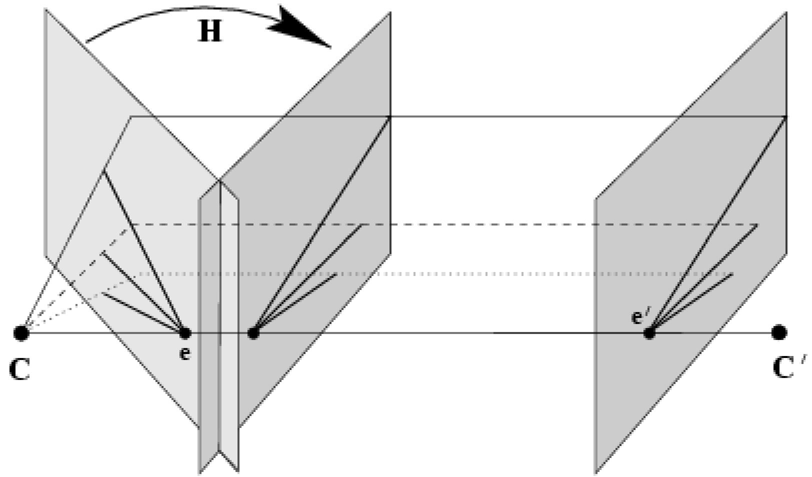
motion starts at x and moves towards e, faster depending on *Z*

pure translation: F only 2 d.o.f., $x^T[e]_x x=0 \Rightarrow$ auto-epipolar

(**x**,**x**’ and **e**=**e**’ are colinear)

General motion

Given two arbitrary cameras, we may rotate the camera used for the first image so that it is aligned with the second camera.
This rotation may be simulated by applying a projective transformation to the first image.



$$x'^{T}[e']_{\times}Hx = 0$$

$$x'^{T}[e']_{\times}\hat{x} = 0$$

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

$$\begin{aligned} x &= K[I \mid 0]X \\ \hat{x} &= K[R \mid 0]X = KRK^{-1}K[I \mid 0]X = KRK^{-1}x \end{aligned}$$

What is Camera Calibration?

- A camera projects 3D world-points onto the 2D image plane
- Calibration: Finding the quantities internal to the camera that affect this imaging process
  - Image center
  - Focal length
  - Lens distortion parameters

Essential Matrix

- The essential matrix is the specialization of the fundamental matrix to the case of normalized image coordinates.
- Historically, the essential matrix was introduced (by Longuet-Higgins) before the fundamental matrix, and the fundamental matrix may be thought of as the generalization of the essential matrix in which the assumption of calibrated cameras is removed.
- The essential matrix has fewer degrees of freedom, and additional properties, compared to the fundamental matrix

Normalized coordinates. Consider a camera matrix decomposed as $\mathrm{P} = \mathrm{K}[\mathrm{R} \mid \mathbf{t}]$, and let $\mathbf{x} = \mathrm{P}\mathbf{X}$ be a point in the image. If the calibration matrix K is known, then we may apply its inverse to the point $\mathbf{x}$ to obtain the point $\hat{\mathbf{x}} = \mathrm{K}^{-1}\mathbf{x}$. Then $\hat{\mathbf{x}} = [\mathrm{R} \mid \mathbf{t}]\mathbf{X}$, where $\hat{\mathbf{x}}$ is the image point expressed in *normalized coordinates*. It may be thought of as the image of the point $\mathbf{X}$ with respect to a camera $[\mathrm{R} \mid \mathbf{t}]$ having the identity matrix I as calibration matrix. The camera matrix $\mathrm{K}^{-1}\mathrm{P} = [\mathrm{R} \mid \mathbf{t}]$ is called a *normalized camera matrix*, the effect of the known calibration matrix having been removed.

Essential Matrix

- Consider a camera matrix decomposed as P = K[R | **t**], and let **x** = P**X** be a point in the image. If the calibration matrix K is known, then we may apply its inverse to the point **x** to obtain the point

$\hat{\mathrm{x}} = \mathrm{K}^{-1}\mathrm{x}$. Then $\hat{\mathrm{x}} = [\mathrm{R} \mid \mathrm{t}]\mathrm{X}$,

- where $\hat{\mathrm{x}}$ is the image point expressed in *normalized coordinates*. It may be thought of as the image of the point **X** with respect to a camera [R | **t**] having the identity matrix I as calibration matrix.
- The camera matrix $\mathbf{K}^{-1}\mathbf{P}$ = [**R** | **t**] is called a *normalized camera matrix*, the effect of the known calibration matrix having been removed.

Essential Matrix

Now, consider a pair of normalized camera matrices $P = [I \mid 0]$ and $P' = [R \mid t]$. The fundamental matrix corresponding to the pair of normalized cameras is customarily called the *essential matrix*, and according to (9.2–*p*244) it has the form

$$E = [t]_\times R = R\,[R^T t]_\times .$$

Definition 9.16. The defining equation for the essential matrix is

$$\hat{x}'^T E \hat{x} = 0 \tag{9.11}$$

in terms of the normalized image coordinates for corresponding points $x \leftrightarrow x'$.

Substituting for $\hat{x}$ and $\hat{x}'$ gives $x'^T K'^{-T} E K^{-1} x = 0$. Comparing this with the relation $x'^T F x = 0$ for the fundamental matrix, it follows that the relationship between the fundamental and essential matrices is

$$E = K'^T F K. \tag{9.12}$$

$$\begin{aligned} F &= [P'C]_\times P'P^+ \\ &= [K't]_\times K'RK^{-1} = K'^{-T}[t]_\times RK^{-1} = K'^{-T}R[R^T t]_\times K^{-1} = K'^{-T}RK^T[KR^T t]_\times \ (9.2) \end{aligned}$$

Essential Matrix

9.6.1 Properties of the essential matrix

The essential matrix, $E = [t]_\times R$, has only five degrees of freedom: both the rotation matrix R and the translation t have three degrees of freedom, but there is an overall scale ambiguity – like the fundamental matrix, the essential matrix is a homogeneous quantity.

The reduced number of degrees of freedom translates into extra constraints that are satisfied by an essential matrix, compared with a fundamental matrix. We investigate what these constraints are.

Result 9.17. *A 3×3 matrix is an essential matrix if and only if two of its singular values are equal, and the third is zero.*

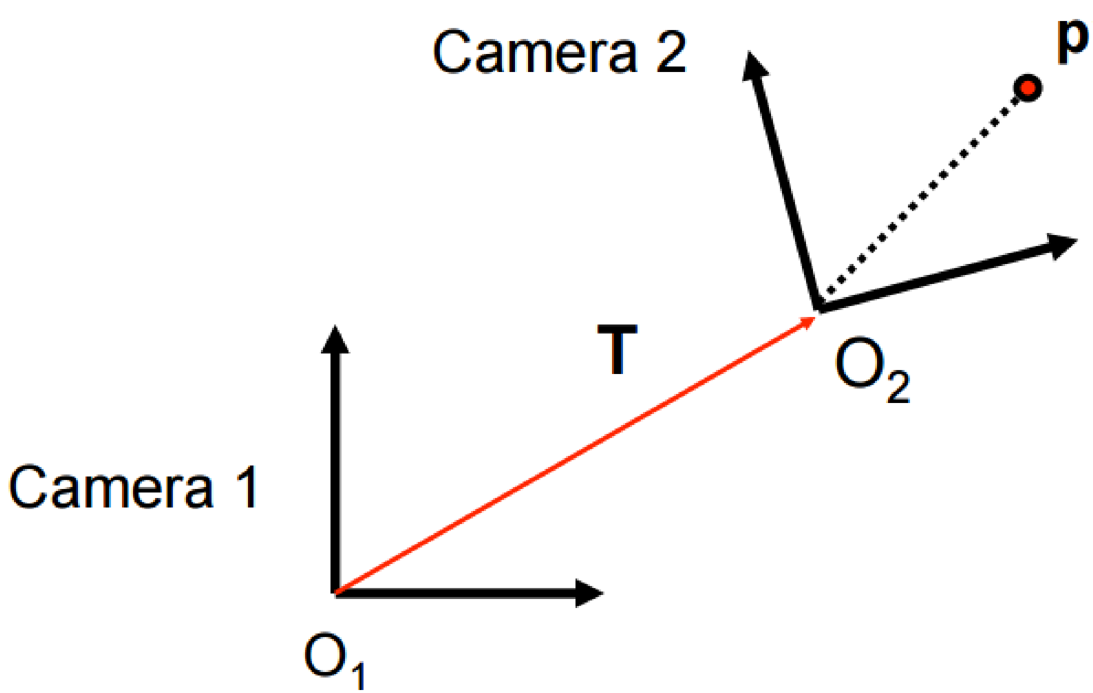
- **Rank of E-matrix (and F-matrix)**

Now, $E = [T_\times]R$ where R is the rotation matrix relating the two camera co-ordinate systems and $[T_\times] = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$. A little bit of manipulation will show that one column of $[T_\times]$ is a linear combination of the other two columns. So $[T_\times]$ has rank 2.

Hence any matrix that you construct by multiplying other matrices with $[T_\times]$ (such as E and F) will also have rank 2.

Two-view geometry

- Two cameras are related by R and T



T = $\mathbf{O_2}$ in the camera 1 reference system

R is the rotation matrix such that a vector **p'** in the camera 2 is equal to **R p'** in camera 1.

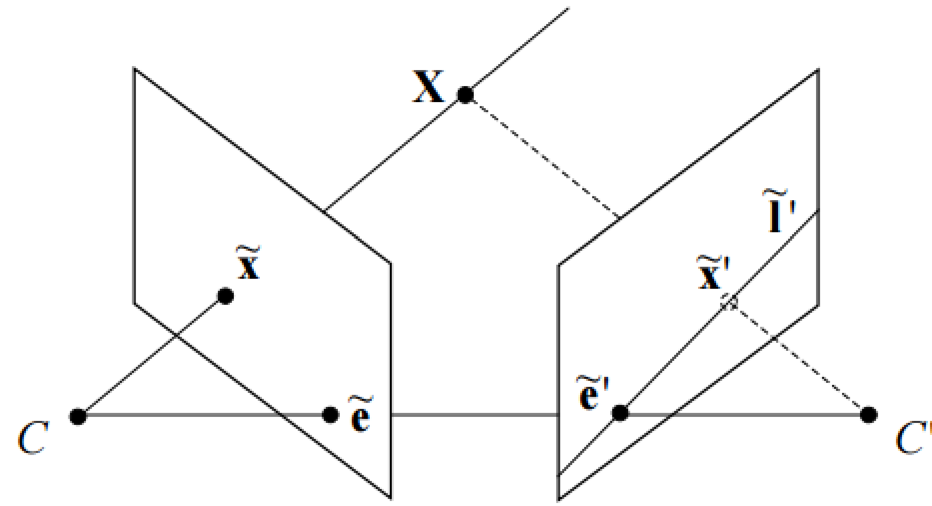
Two-view geometry

- The essential matrix

Note $\tilde{\boldsymbol{x}}, \boldsymbol{T}, \boldsymbol{R}\tilde{\boldsymbol{x}}'$ are coplanar,

and $\tilde{\boldsymbol{x}}'$ is $\boldsymbol{R}\tilde{\boldsymbol{x}}' + \boldsymbol{T}$

in the first camera reference system.



$$\tilde{\boldsymbol{x}} \cdot [\boldsymbol{T} \times ((\boldsymbol{R}\tilde{\boldsymbol{x}}') + \boldsymbol{T})] =$$

$$\tilde{\boldsymbol{x}} \cdot [\boldsymbol{T} \times (\boldsymbol{R}\tilde{\boldsymbol{x}}')] = 0 \Rightarrow \tilde{\mathbf{x}}^\top \mathbf{E} \tilde{\mathbf{x}}' = 0 \Rightarrow \boldsymbol{E} = \boldsymbol{T} \times \boldsymbol{R} = [\boldsymbol{T}_\times]\boldsymbol{R}$$

$$\Downarrow$$

Essential Matrix
(Longuet-Higgins, 1981)

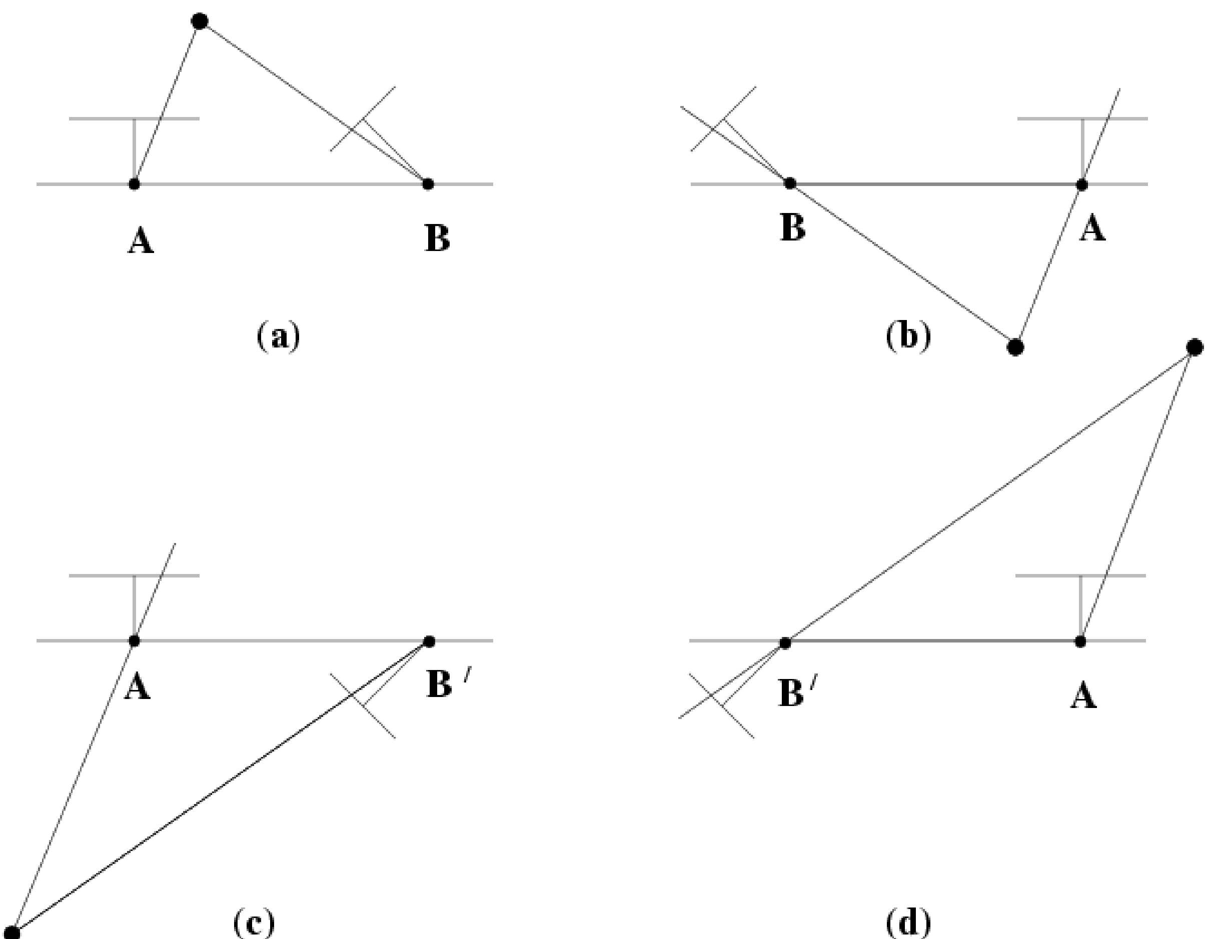
For $\mathbf{T} = [t_x \quad t_y \quad t_z]^\top$, $[\mathbf{T}]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$.

Four possible reconstructions from E

Result 9.19. *For a given essential matrix* $\mathtt{E} = \mathtt{U}\,\mathrm{diag}(1,1,0)\mathtt{V}^\mathsf{T}$, *and first camera matrix* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, *there are four possible choices for the second camera matrix* $\mathtt{P}'$, *namely*

$$\mathtt{P}' = [\mathtt{UWV}^\mathsf{T} \mid +\mathbf{u}_3] \text{ or } [\mathtt{UWV}^\mathsf{T} \mid -\mathbf{u}_3] \text{ or } [\mathtt{UW}^\mathsf{T}\mathtt{V}^\mathsf{T} \mid +\mathbf{u}_3] \text{ or } [\mathtt{UW}^\mathsf{T}\mathtt{V}^\mathsf{T} \mid -\mathbf{u}_3].$$

Thus, testing with a single point to determine if it is in front of both cameras is sufficient to decide between the four different solutions for the camera matrix **P**′.



(only one solution where points is in front of both cameras)