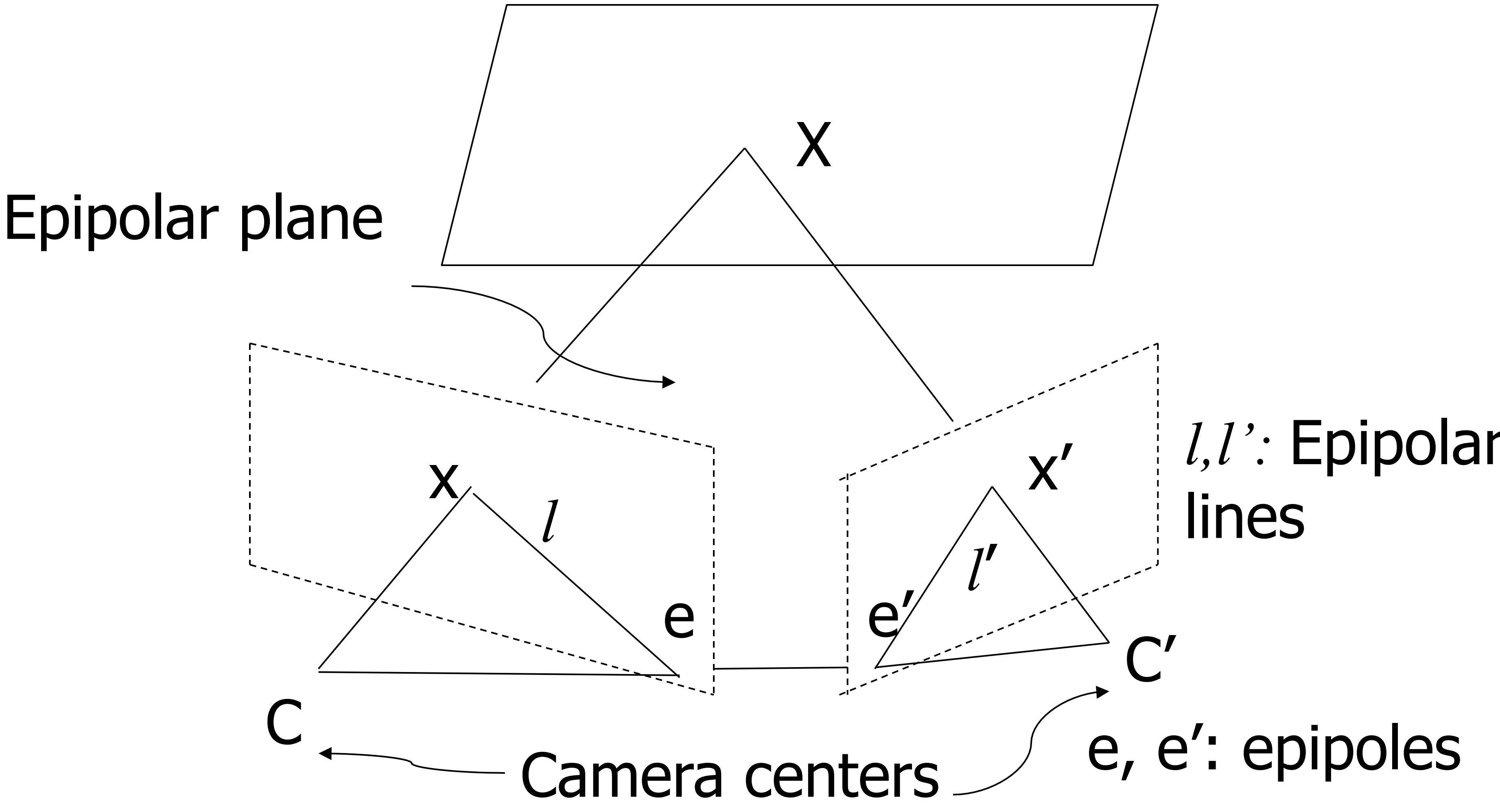
Stereo Geometry
(Week 04-05 Lectures: 16-23)

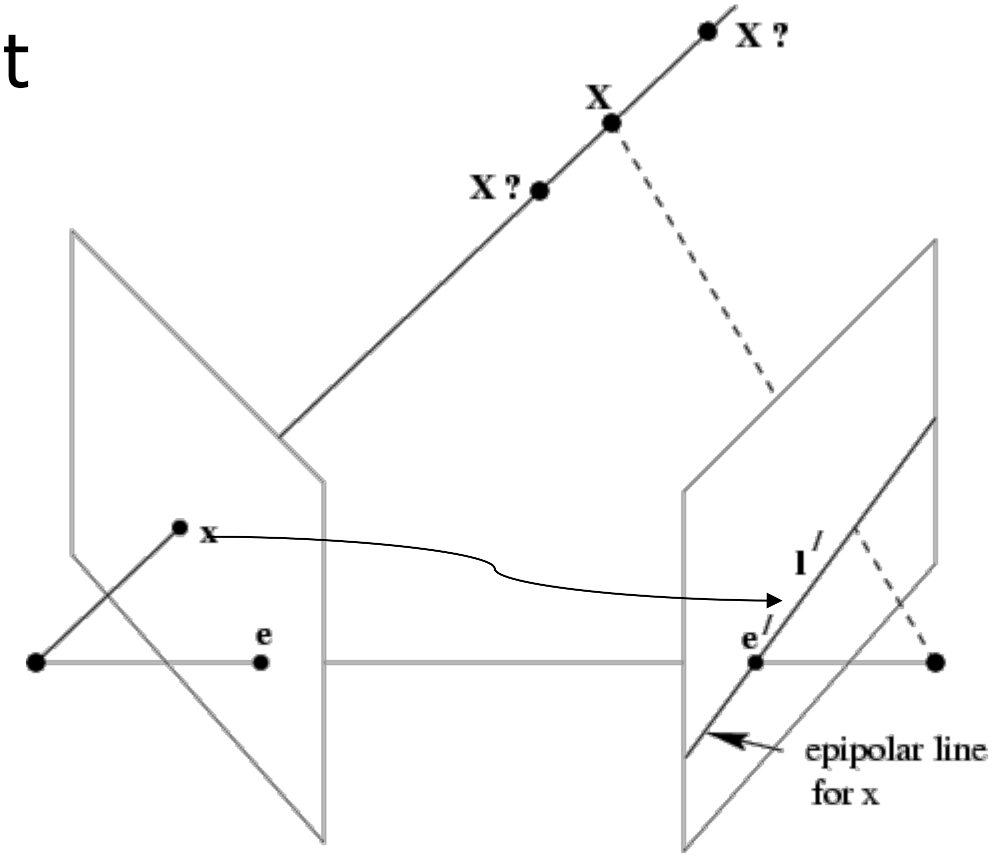
Jayanta Mukhopadhyay
Dept. of Computer Science and Engg.

Stereo Set-up



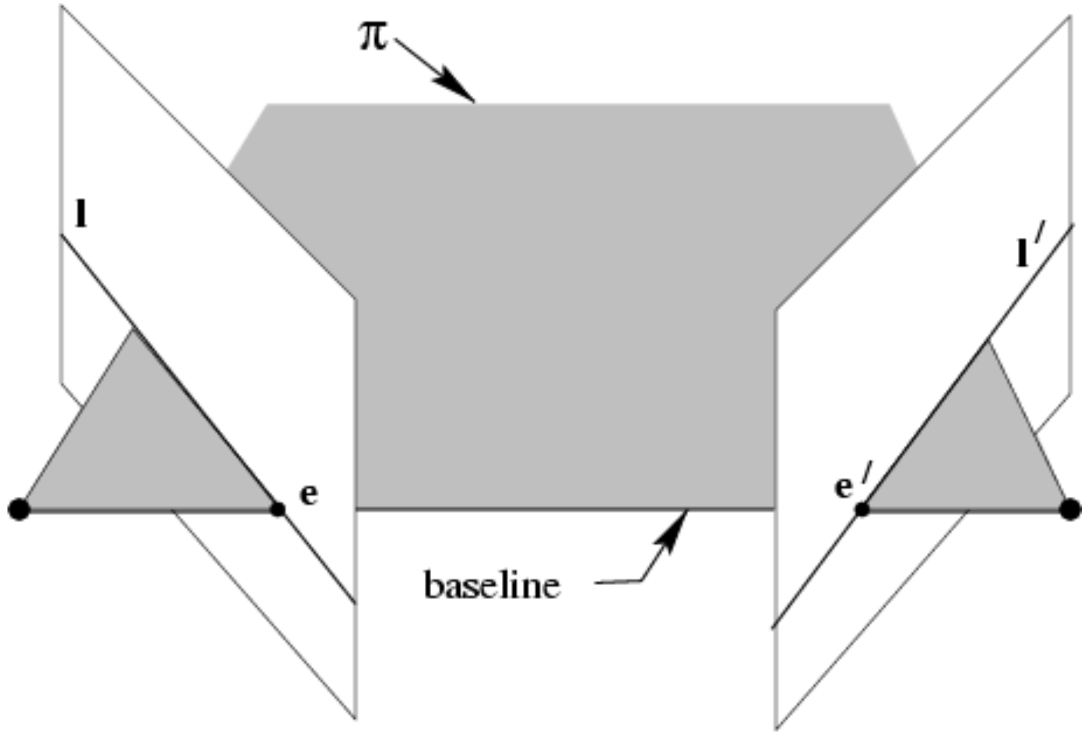
Epipolar geometry

Corresponding point of x in the right image lies on l′.



From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

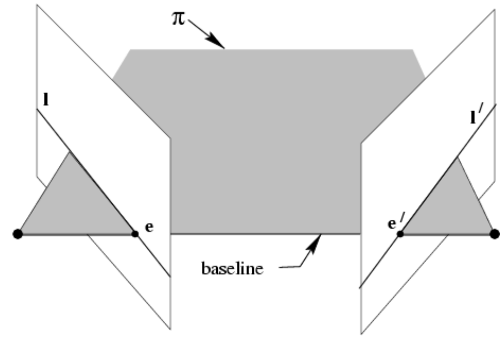
Epipolar geometry



All points on π project on l and l'

From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

Epipolar geometry

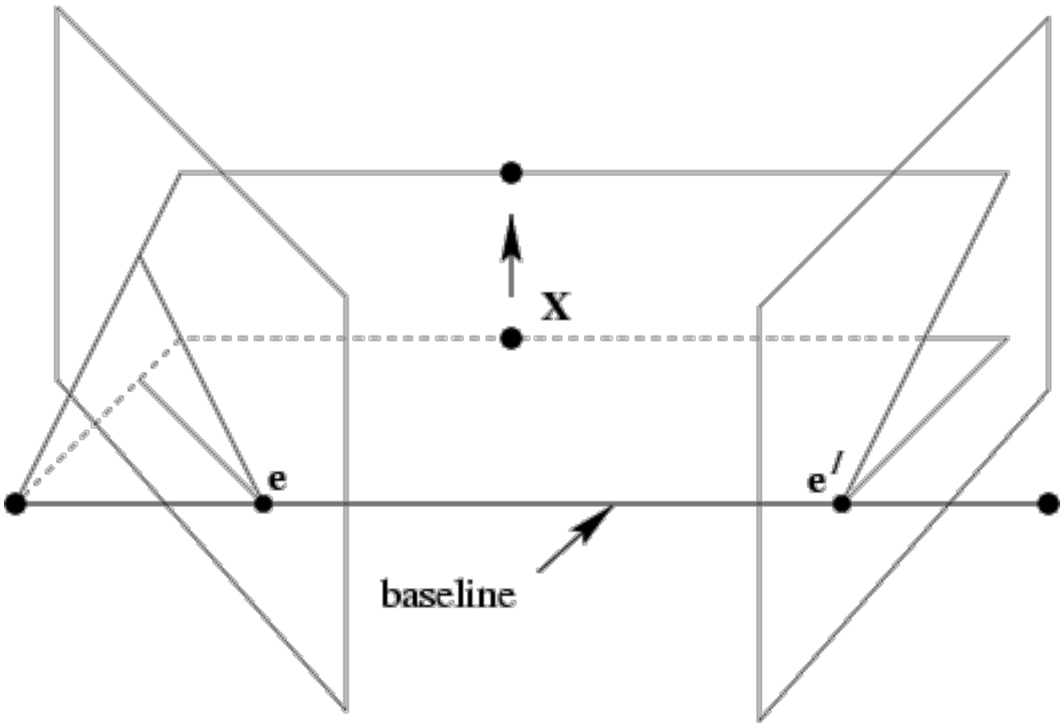


Epipoles e,e'
= intersection of baseline with image plane
= projection of projection center in other image
= vanishing point of camera motion direction

An epipolar plane: plane containing baseline (1-D family)

An epipolar line: intersection of epipolar plane with image (always come in corresponding pairs)

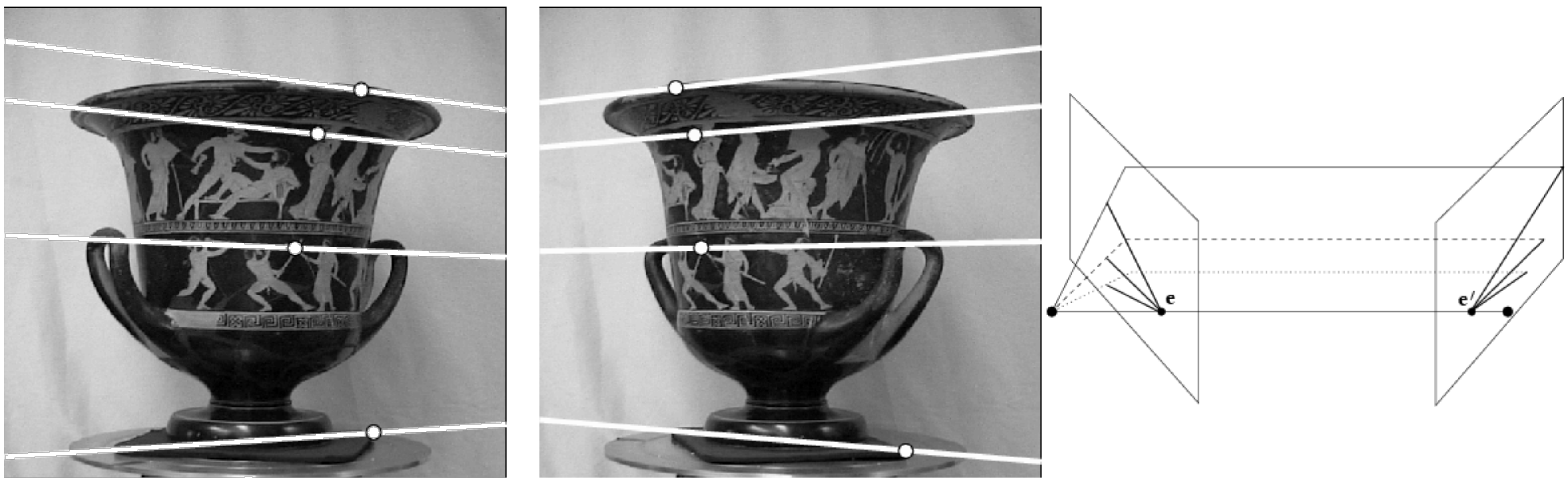
Epipolar geometry



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

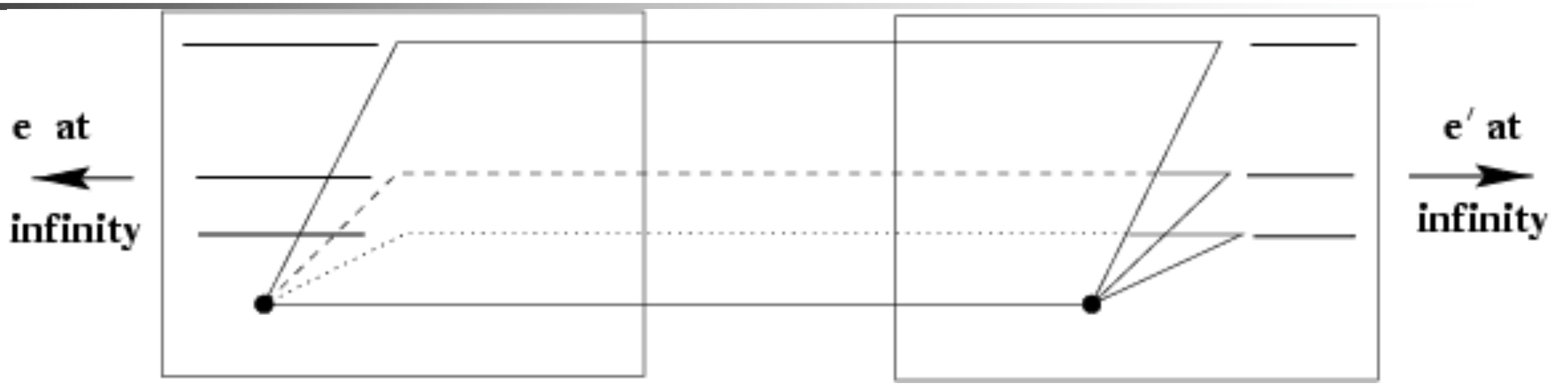
From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

Example: converging cameras



From Hartley and Zisserman, “Multiple view geometry in computer vision”, Cambridge Univ. Press (2000)

Example: motion parallel with image plane



From Hartley and Zisserman, “Multiple view geometry in computer vision”, Cambridge Univ. Press (2000)

Epipolar Geometry

$$[e']_\times = \begin{bmatrix} 0 & -e_z & e_y \\ e_z & 0 & -e_x \\ -e_y & e_x & 0 \end{bmatrix}$$

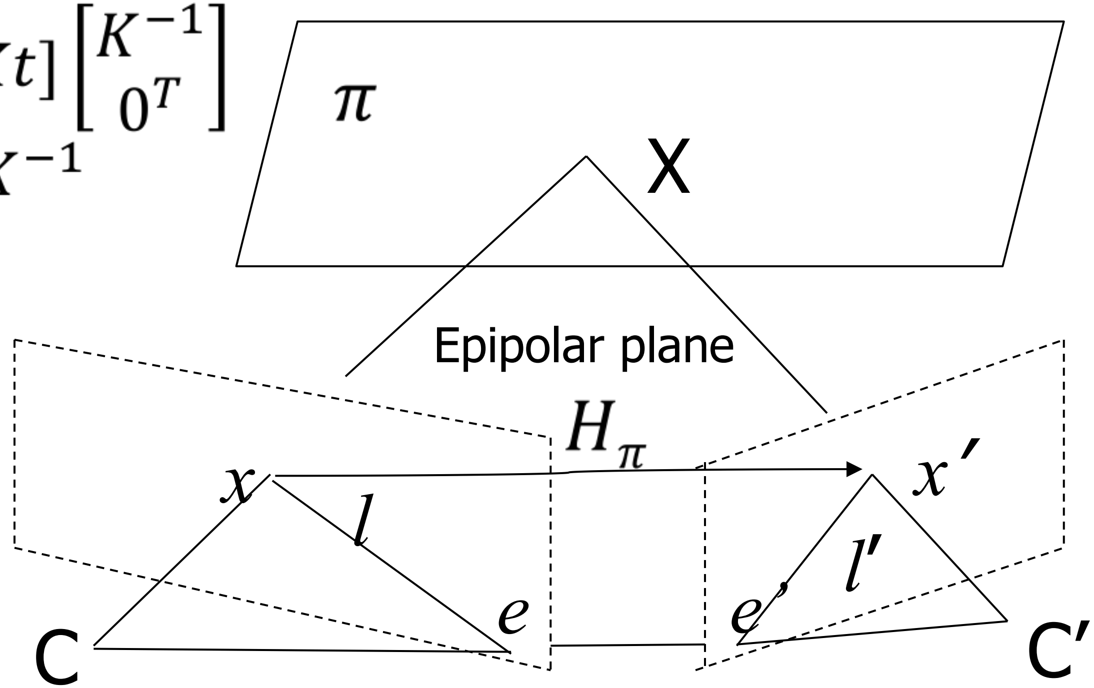
$$H_\pi = [K'R|Kt]\begin{bmatrix} K^{-1} \\ 0^T \end{bmatrix} = K'RK^{-1}$$

$$x' = P'X = P'P^+x$$

$$P^+ = \begin{bmatrix} K^{-1} \\ 0^T \end{bmatrix}$$

For scene points lying on plane at infinity.

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



$P=K[I \mid 0]$

$P'=K'[R \mid t]$

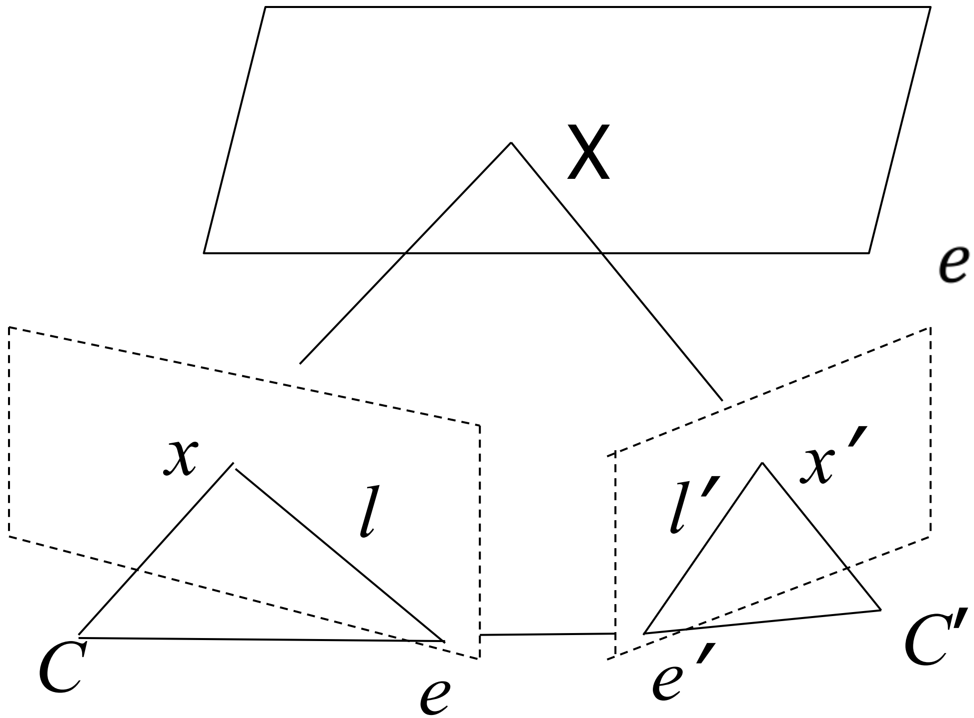
$F=[e']_x K'RK^{-1}$

Coplanar: X, x, x', C, C′, e, e', l, l'

Epipolar Geometry

Fundamental Matrix: F

$$l' = Fx$$



$$e = PC'$$

$$e' = P'C$$

$$e = -KR^T t \equiv KR^T t$$

$$e' = K't$$

$$l' = Fx$$

$$\Rightarrow x'^T l' = 0$$

$$\Rightarrow x'^T Fx = 0$$

$$\Rightarrow x^T F^T x' = 0$$

$$C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$P = K[\, I \mid 0 \,]$

$P' = K'[\, R \mid t \,]$

$$C' = \begin{bmatrix} -R^T t \\ 1 \end{bmatrix}$$

Fundamental and Essential Matrices

$$P = K[I \mid 0]$$
$$P' = K'[R \mid t]$$

$$F = [e']_\times P' P^+$$
$$= [K't]_\times K' R K^{-1}$$

Say, $P=[I \mid 0]$, i.e. $K=I$

$$F = [m]_\times M$$

$$P' = K'[R \mid t] = [K'R \mid K't] = [M \mid m]$$

$$for\ K = I \quad and\ K' = I, \qquad F = [t]_\times R$$

Essential Matrix (E)

Ex. 1

- Consider the following stereo imaging matrices given by $P = [I|0]$ (ref. camera) and P' as follows.

$$P' = \begin{bmatrix} 3 & 4 & 6 & 4 \\ 8 & 7 & 2 & 8 \\ 1 & 5 & 2 & 1 \end{bmatrix}$$

Compute the fundamental matrix of the system.

Ans. 1

$$P' = \begin{bmatrix} 3 & 4 & 6 & 4 \\ 8 & 7 & 2 & 8 \\ 1 & 5 & 2 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} 3 & 4 & 6 \\ 8 & 7 & 2 \\ 1 & 5 & 2 \end{bmatrix} \qquad m = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} \qquad [m]_\times = \begin{bmatrix} 0 & -1 & 8 \\ 1 & 0 & -4 \\ -8 & 4 & 0 \end{bmatrix}$$

$$F = [m]_\times M$$

$$= \begin{bmatrix} 0 & 33 & 14 \\ -1 & -16 & -2 \\ 8 & -4 & -40 \end{bmatrix}$$

Ex. 2

- Consider the following stereo imaging matrices given by P (ref. camera) and P'.

$$P = \begin{bmatrix} 3 & 2 & 4 & -2 \\ 8 & 6 & 0 & 4 \\ 9 & 5 & 7 & 3 \end{bmatrix} \qquad P' = \begin{bmatrix} 3 & 8 & 5 & 2 \\ 2 & 7 & 6 & -3 \\ 6 & 4 & 9 & 8 \end{bmatrix}$$

(a) Compute the fundamental matrix of the system.

(b) Given an image point (15,20) of the reference camera (P), compute the epipolar line and its two end image points of P'.

Ans. 2 (a)

$$P = \begin{bmatrix} 3 & 2 & 4 & -2 \\ 8 & 6 & 0 & 4 \\ 9 & 5 & 7 & 3 \end{bmatrix} \qquad P' = \begin{bmatrix} 3 & 8 & 5 & 2 \\ 2 & 7 & 6 & -3 \\ 6 & 4 & 9 & 8 \end{bmatrix}$$

$$P = [\underbrace{\;\;}_{M} \mid \underbrace{\;}_{p_4}] \qquad P' = [\underbrace{\;\;}_{M'} \mid \cdot\,]$$

$$\tilde{C} = -M^{-1}p_4 \qquad M^{-1} = -\frac{1}{42}\begin{bmatrix} 42 & 6 & -24 \\ -56 & -15 & 32 \\ -14 & 3 & 2 \end{bmatrix}$$

$$C = \begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix} = \frac{1}{42}\begin{bmatrix} -132 \\ 148 \\ 46 \\ 1 \end{bmatrix} \qquad e' = P'C = \begin{bmatrix} 26.23 \\ 21.94 \\ 13.09 \end{bmatrix} = 13.09\begin{bmatrix} 2 \\ 1.67 \\ 1 \end{bmatrix}$$

$$H_\propto = M'M^{-1} \Rightarrow H_\propto = \begin{bmatrix} 9.33 & 2.07 & -4.62 \\ 9.33 & 1.78 & -4.48 \\ 2.33 & -0.07 & -0.05 \end{bmatrix}$$

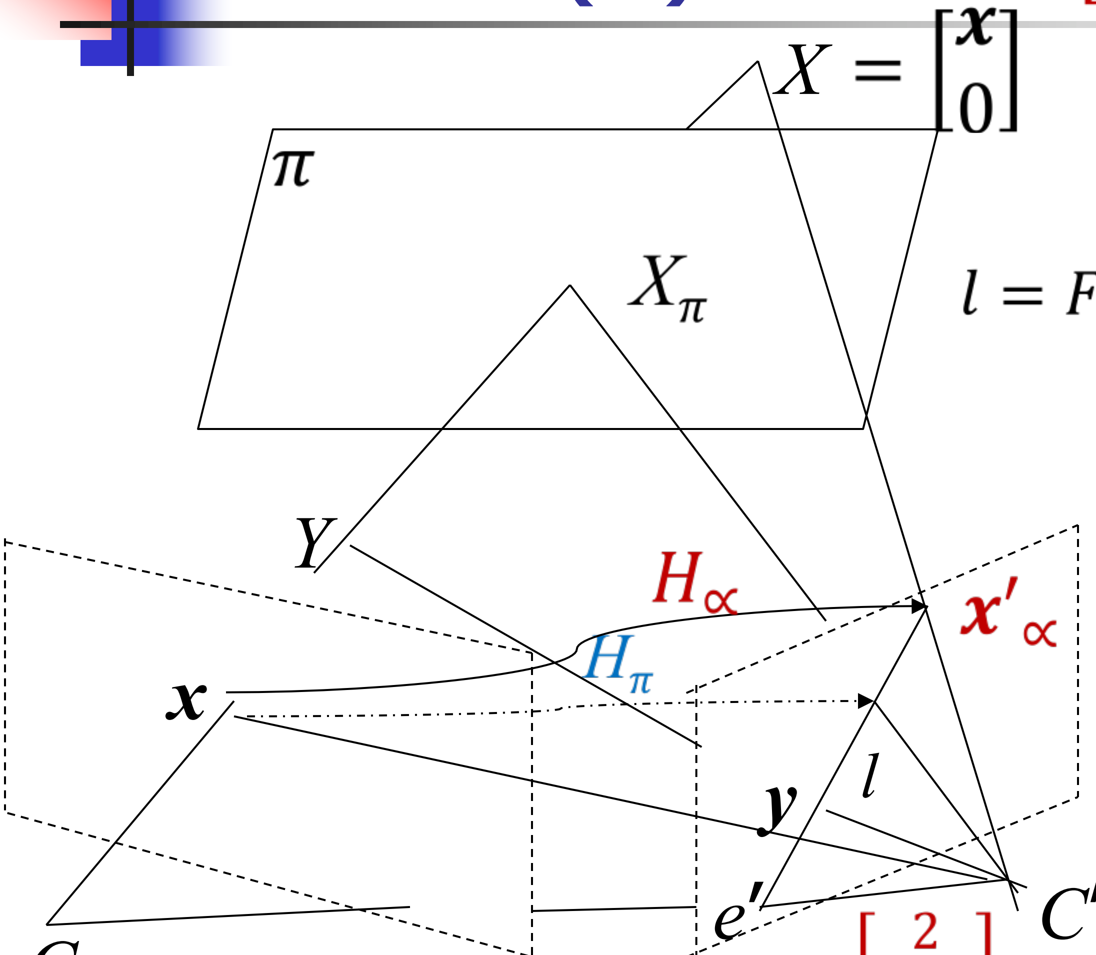
$$F = [e']_\times H_\propto$$

$$[e']_\times = \begin{bmatrix} 0 & -1 & 1.67 \\ 1 & 0 & -2 \\ -1.67 & 2 & 0 \end{bmatrix} \qquad F = \begin{bmatrix} -5.41 & -1.91 & 4.40 \\ 4.67 & 2.21 & -4.52 \\ 2.99 & 0.09 & -1.19 \end{bmatrix}$$

(b) Given an image point (15,20) of the reference camera (P), compute the epipolar line and its two end image points of P′.

Ans. 2(b)

$$F=\begin{bmatrix}-5.41 & -1.91 & 4.40\\ 4.67 & 2.21 & -4.52\\ 2.99 & 0.09 & -1.19\end{bmatrix}$$



$$e'=\begin{bmatrix}2\\1.67\\1\end{bmatrix}$$

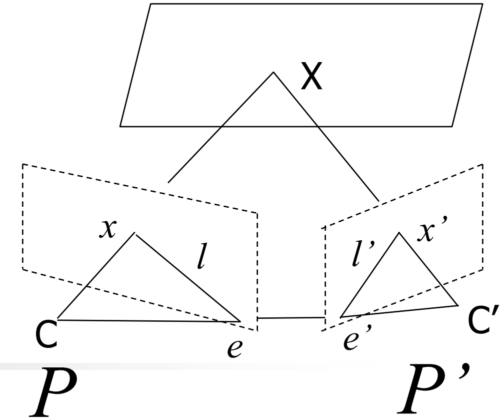
$$l=F\begin{bmatrix}15\\20\\1\end{bmatrix}=\begin{bmatrix}-114.92\\109.76\\45.44\end{bmatrix}\equiv\begin{bmatrix}-2.53\\2.42\\1\end{bmatrix}$$

$$x'_\propto=H_\propto\begin{bmatrix}15\\20\\1\end{bmatrix}$$

$$H_\propto=\begin{bmatrix}9.33 & 2.07 & -4.62\\ 9.33 & 1.78 & -4.48\\ 2.33 & -0.07 & -0.05\end{bmatrix}$$

$$x'_\propto=\begin{bmatrix}176.81\\171.24\\33.52\end{bmatrix}\equiv\begin{bmatrix}5.27\\5.11\\1\end{bmatrix}$$

Fundamental Matrix: Properties



$$x'^T F x = x^T F^T x' = 0, \forall (x', x)$$

Transpose:
If F is fundamental matrix of (P,P'), F^T for (P',P).

Epipolar lines: For x, epipolar line $l'=Fx$.
For x', epipolar line $l=F^T x'$.

F is a correlation.
→ Rank deficient.
→ Inverse does not exit.

Epipoles: $e'^T(Fx) = e^T(F^T x') = (Fe)^T x' = 0$

$e'^T F=0$ ➔ e' is left NULL vector of F.

$Fe=0$ ➔ e is the right NULL vector of F.

Rank deficient:
$\det(F)=0$ and F is a projective element → 7 d.o.f.

Ex. 3

- Consider the following fundamental matrix.

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

a) Given two points (5,8) and (7,-5) in the left image compute the corresponding epipolar lines in the right image.
b) Compute the right epipole.
c) Compute the left epipole.

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

Ans.3

- l_1=F.[5 6 1]^T= [172 57 1]^T
- l_2=F.[7 -5 1]^T=[80 150 -370]^T
- e_R=l_1 x l_2 =[-21240 63720 21240]^T

  =[-1 3 1]^T

- e_L= Right zero of F

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

Ans.3 (contd.)

- e_L= Right zero of F

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

$$\begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \end{bmatrix} \begin{bmatrix} e_{L1} \\ e_{L2} \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} e_{L1} \\ e_{L2} \end{bmatrix} = - \begin{bmatrix} 20 & 12 \\ 8 & -7 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 59 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

Ans.3 (contd.)

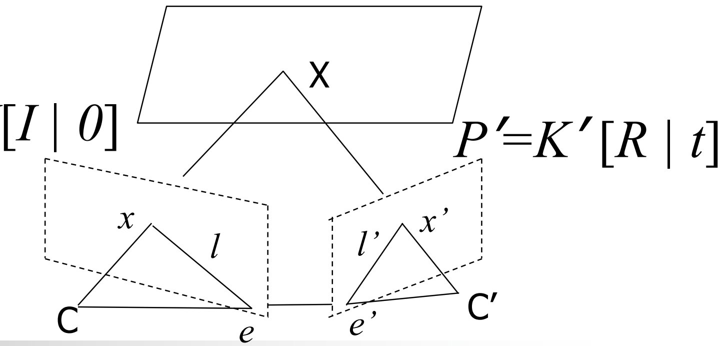
- e_R= Right zero of F^T

$$F^T = \begin{bmatrix} 20 & 8 & -4 \\ 12 & -7 & 33 \\ 0 & 59 & -177 \end{bmatrix}$$

$$\begin{bmatrix} 20 & 8 & -4 \\ 12 & -7 & 33 \end{bmatrix} \begin{bmatrix} e_{R1} \\ e_{R2} \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} e_{R1} \\ e_{R2} \end{bmatrix} = -\begin{bmatrix} 20 & 8 \\ 12 & -7 \end{bmatrix}^{-1} \begin{bmatrix} -4 \\ 33 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Essential Matrix



Stereo geometry for calibrated cameras.

$x_c = K^{-1}x$

$x'^T_c E x_c = 0$

$x'_c = K'^{-1}x'$

Coordinates in calibrated image planes.

$x'^T F x = 0$

$\rightarrow (K'x'_c)^T F (K x_c) = 0$

$\rightarrow x'^T_c K'^T F K x_c = 0$

$\rightarrow x'^T_c (K'^T F K) x_c = 0$

E

$E = K'^T F K$

$F = K'^{-T} E K^{-1}$

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

$E = [t]_{\times} R$

6 parameters

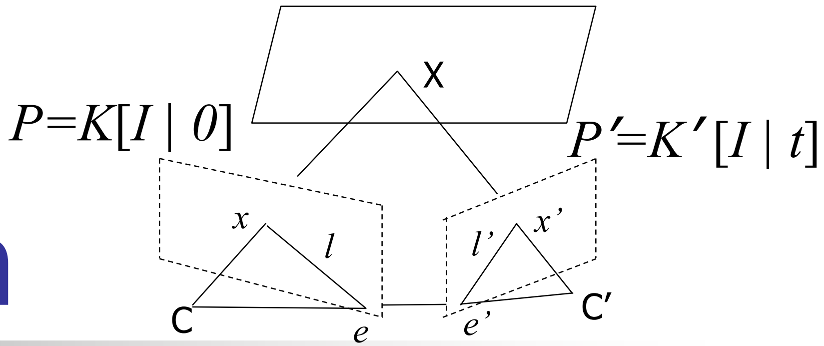
d.o.f.=5

$Ee_c = e'_c E = 0$

Rank: 2

det(E)=0

Pure translation



$F=[e']_x K' I K^{-1}$
$=[e']_x K' K^{-1}$

camera translation $||^{l}$ to x-axis, $e'=[1\ 0\ 0]^T$

$$\rightarrow F=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{bmatrix}$$

$\rightarrow x'^T F x=0$

$\rightarrow y'=y$

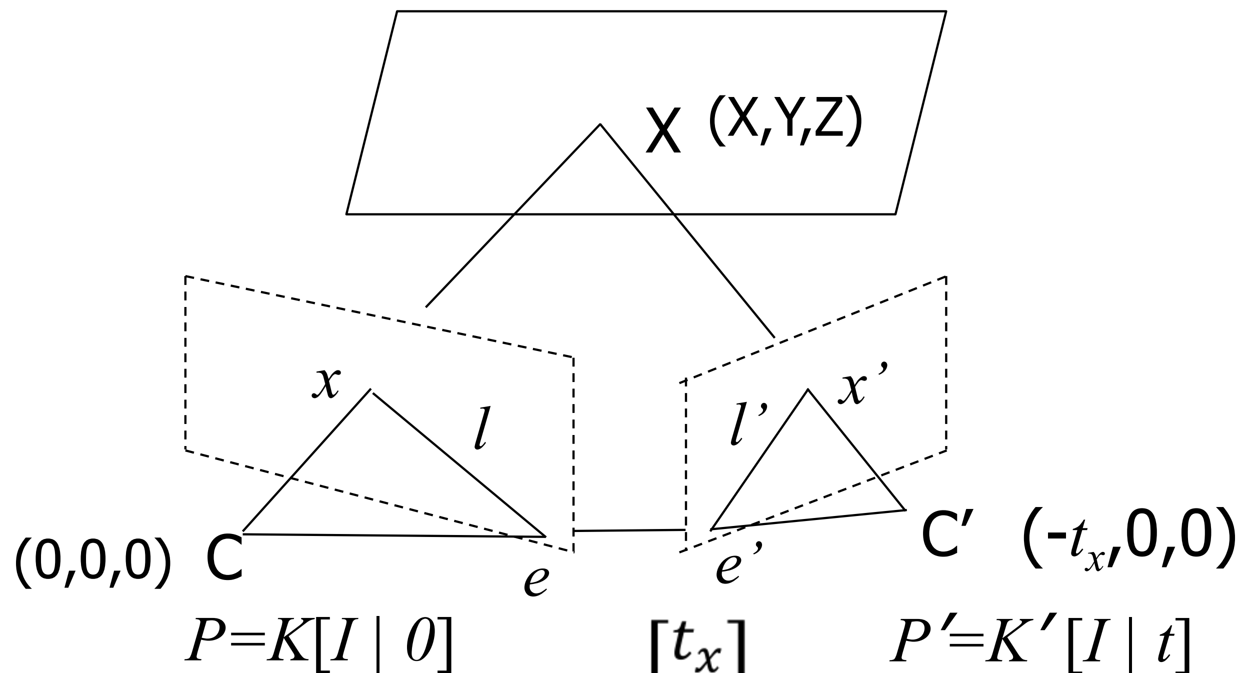
For $K=K'$, $F=[e']_x$

$$[e']_\times=\begin{bmatrix}0 & -e_z & e_y\\ e_z & 0 & -e_x\\ -e_y & e_x & 0\end{bmatrix}$$

Pure Translation: Computing Depth

$$\tilde{X} = ZK^{-1}x$$

$$\tilde{X} + t = ZK'^{-1}x'$$



For $K=K'$

$$ZK^{-1}(x' - x) = (\tilde{X} + t) - \tilde{X} = t \qquad t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix}$$

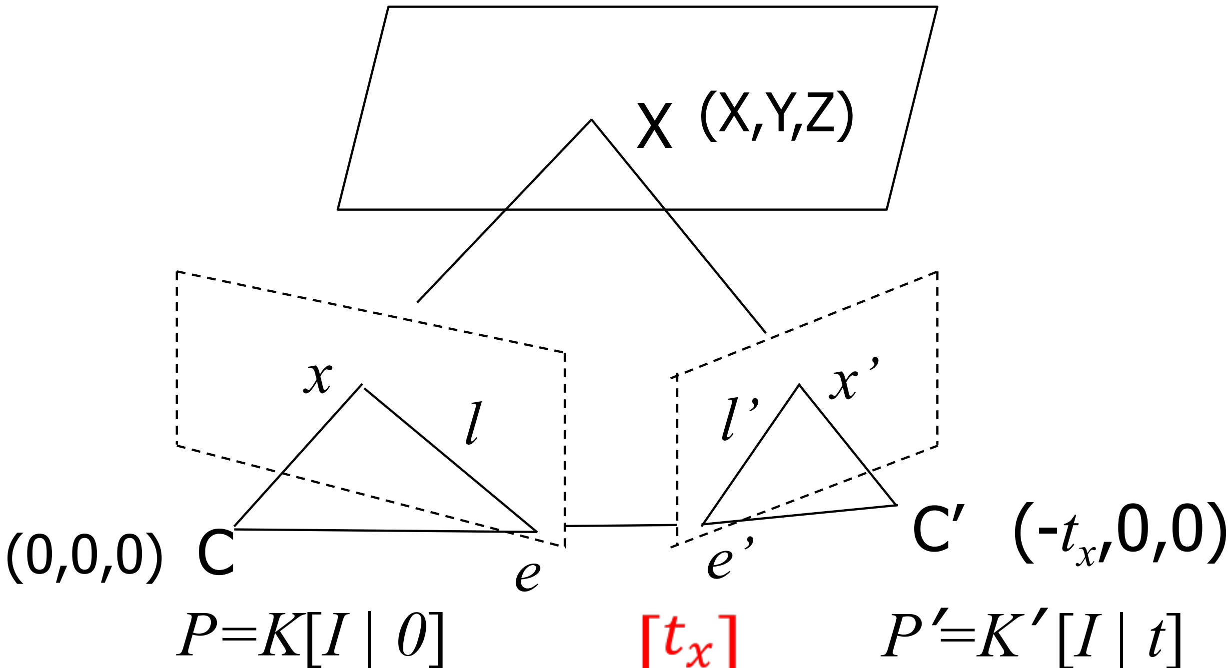
$$\Rightarrow x' = x + \frac{Kt}{Z} \quad \Rightarrow \quad Z = \frac{Kt}{x' - x}$$

Pure Translation: Computing Depth

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Z = \frac{Kt}{x' - x}$$

$$Z = \frac{f t_x}{x' - x}$$



$$t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix}$$

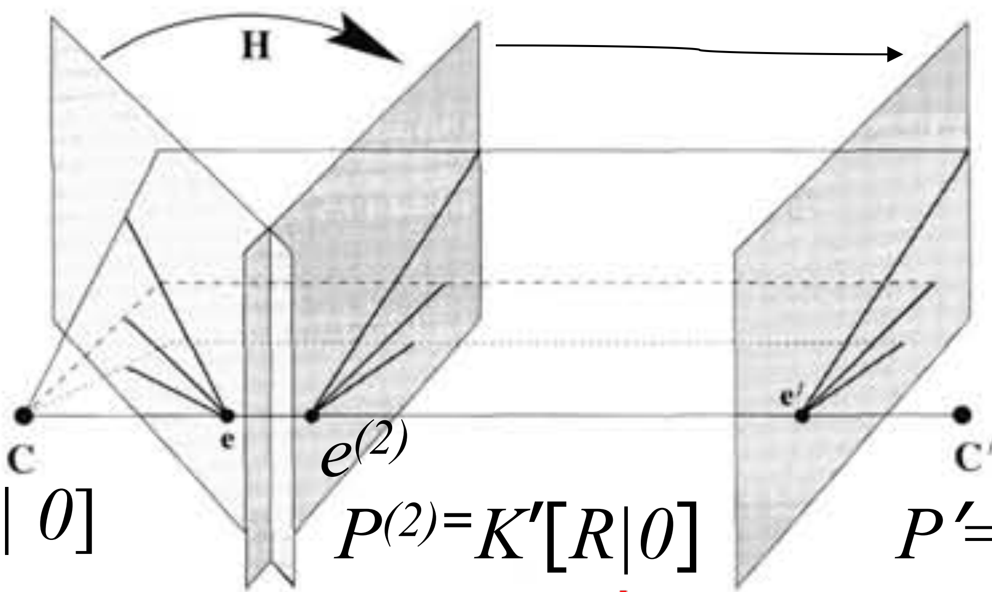
General Motion of Camera

$$x^{(2)} = K'[R|0]X$$
$$= K'R[I|0]X$$
$$= K'RK^{-1}K[I|0]X$$
$$= K'RK^{-1}x$$

Rotational Homography

$H = K'RK^{-1}$

$||^l$ stereo



$$x' = x^{(2)} + \frac{K't}{Z}$$
$$= K'RK^{-1}x + \frac{K't}{Z}$$

$P = K[I \mid 0]$

$P^{(2)} = K'[R|0]$

$P' = K'[R \mid t]$

$$Z = \frac{K't}{x' - K'RK^{-1}x}$$

From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

General Motion of Camera

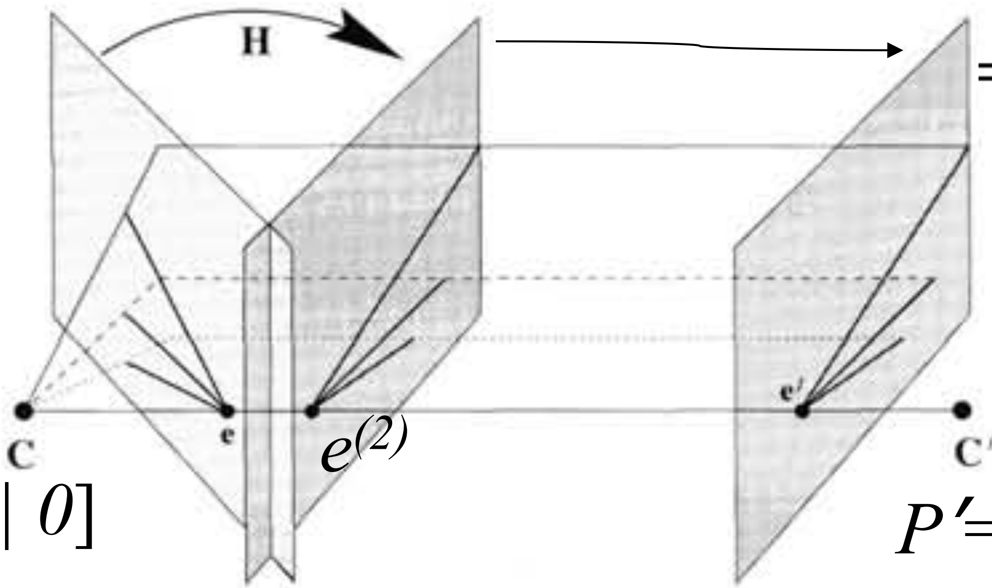
$$x^{(2)} = K[R|0]X$$
$$= KR[I|0]X$$
$$= KRK^{-1}K[I|0]X$$
$$= KRK^{-1}x$$

$H = KRK^{-1}$

Rotational Homography

$||^{l}$ stereo

$$x' = x^{(2)} + \frac{K't}{Z}$$
$$= KRK^{-1}x + \frac{K't}{Z}$$



$P = K[I \mid 0]$

$P' = K'[R \mid t]$

$$Z = \frac{K't}{x' - KRK^{-1}x}$$

From Hartley and Zisserman, “Multiple view geometry in computer vision”, Cambridge Univ. Press (2000)

Ex. 4

- Consider a stereo set-up with P and P′ (camera matrices for left and right camera) as given below.

$$P = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad P' = \begin{bmatrix} 6 & 0 & 0 & 10 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

If the image coordinates of a 3-D point are (6,8) and (9.33,8) in left and right cameras, compute its depth (z-coordinate) in the 3D.

Ans. 4

- P=K[I|0] and P'=K[I|t]

$$K = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad t = \begin{bmatrix} 10/6 \\ 0 \\ 0 \end{bmatrix}$$

- **x**'=**x**+(**Kt**)/Z
- Z=(6x10/6)/(x'-x)

  =10/(9.33-6)=3

Estimation of Fundamental Matrix

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

$$\mathrm{x'}^{\mathrm{T}}\mathrm{Fx} = 0$$

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

$$[x'x, x'y, x', y'x, y'y, y', x, y, 1][f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}]^{\mathrm{T}} = 0$$

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

$$\mathrm{Af} = 0$$

- Solution up to scale.
- Minimum 8 point correspondences.
- Use of DLT (for 7 point correspondences from linear combination of smallest and second smallest eigen vectors.

Estimation of Fundamental Matrix

$$\begin{bmatrix} x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \\ x_2x_2' & y_2x_2' & x_2' & x_2y_2' & y_2y_2' & y_2' & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx_n' & y_nx_n' & x_n' & x_ny_n' & y_ny_n' & y_n' & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

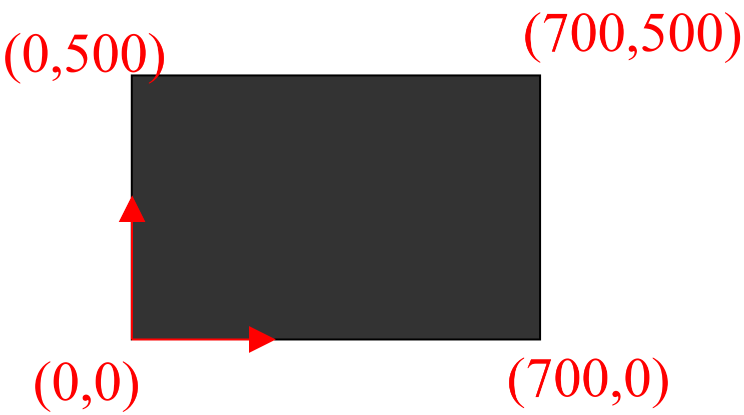
~10000 ~10000 ~100 ~10000 ~10000 ~100 ~100 ~100 1

!

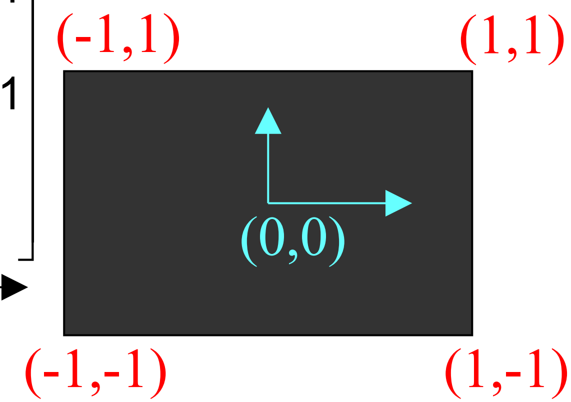
Orders of magnitude difference
Between column of data matrix
$\rightarrow$ least-squares yield poor results

The normalized 8-point algorithm

Transform image to
~[-1,1] x [-1,1]



$$\begin{bmatrix} \frac{2}{700} & 0 & -1 \\ & \frac{2}{500} & -1 \\ & & 1 \end{bmatrix}$$



Least squares yields good results
(Hartley, PAMI´97)

The singularity constraint

$$\det F = 0 \quad \text{rank } F = 2$$

SVD from linearly computed F matrix (rank 3)

$$F = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T + U_3 \sigma_3 V_3^T$$

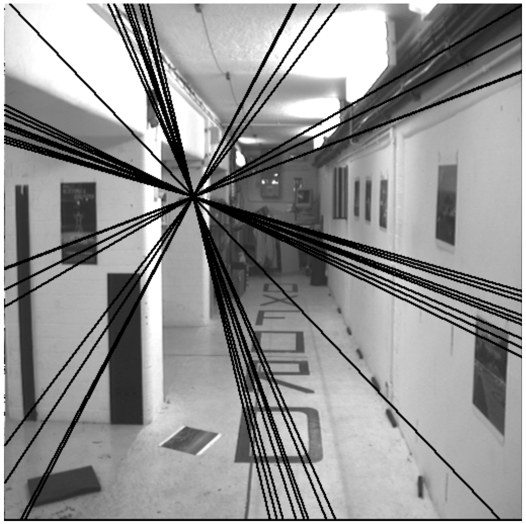
$$\min \| F - F' \|_F$$

Compute closest rank-2 approximation

$$F' = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T$$

The singularity constraint

Nonsingular F



Singular F

Non-singular F causes epipolar lines not converging.

From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

The singularity constraint for Essential Matrix

$$\det(E)=0$$

Estimate $\hat{E}$ by any of the techniques used for F.

Perform SVD of $\hat{E}$.

$\hat{E}=UDV^T$ Where $D=\text{diag}(a,b,c)$ $a \geq b \geq c$

For essential matrix, two singular values are the same.

$$\Rightarrow \hat{E} = U\widehat{D}V^T \qquad where\ \widehat{D} = \left(\frac{a+b}{2}, \frac{a+b}{2}, 0\right)$$

The minimum case – 7 point correspondences

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_7 x_7 & x'_7 y_7 & x'_7 & y'_7 x_7 & y'_7 y_7 & y'_7 & x_7 & y_7 & 1 \end{bmatrix} \mathrm{f} = 0$$

$$\mathrm{A} = \mathrm{U}_{7x7} \mathrm{diag}\left(\sigma_1, \ldots, \sigma_7, 0, 0\right) \mathrm{V}_{9x9}^{\mathrm{T}}$$

Last two column vectors of V also provide F1 and F2.

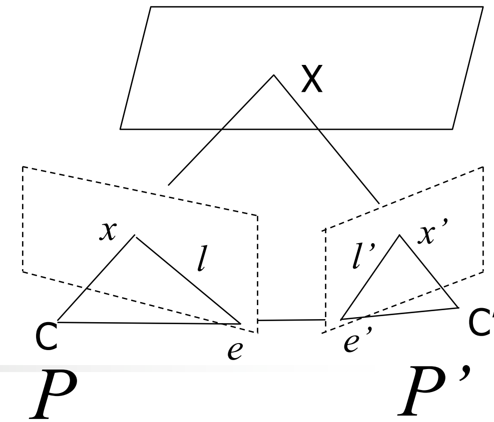
$F_1, F_2 \rightarrow$ Eigen vectors corresponding to two zero's. The solution is $F_1 + \lambda F_2$.

But $F_1 + \lambda F_2$ not automatically rank 2.

Solve for λ from $\det(F_1 + \lambda F_2) = 0$.

As it is a cubic polynomial, there are 1 or 3 solutions.

Parametric representation of F



Over parameterization: $F=[t]_{\times}M$ ➔ $\{t,M\}$ → 12 params.

Epipolar parameterization:

$$F=\begin{bmatrix} a & b & \alpha a+\beta b \\ c & d & \alpha c+\beta d \\ e & f & \alpha e+\beta f \end{bmatrix}$$

$$\{a,b,c,d,e,f,\alpha,\beta\}$$

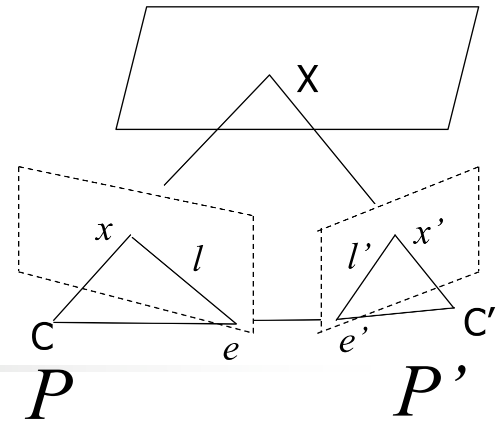
Left epipole $e=[\alpha \quad \beta \quad -1]^T$

Both epipoles as parameters $\{a,b,c,d,\alpha,\beta,\alpha',\beta'\}$

$$F=\begin{bmatrix} a & b & \alpha a+\beta b \\ c & d & \alpha c+\beta d \\ \alpha' a+\beta' c & \alpha' b+\beta' d & \alpha\alpha' a+\alpha\beta' c+\alpha'\beta b+\beta\beta' d \end{bmatrix}$$

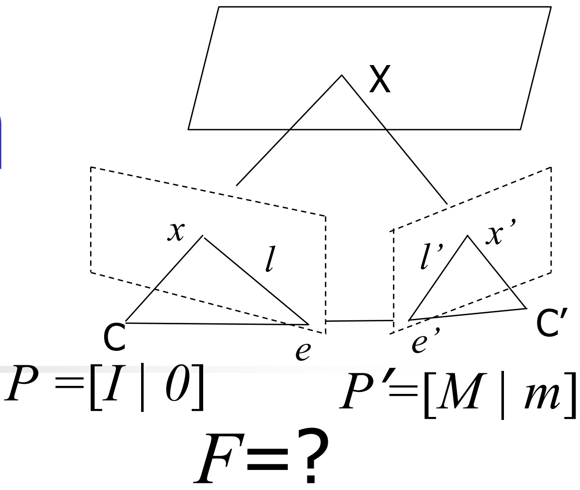
Epipoles $e=[\alpha \quad \beta \quad -1]^T$ $e'=[\alpha' \quad \beta' \quad -1]^T$

Retrieving the camera matrices from F



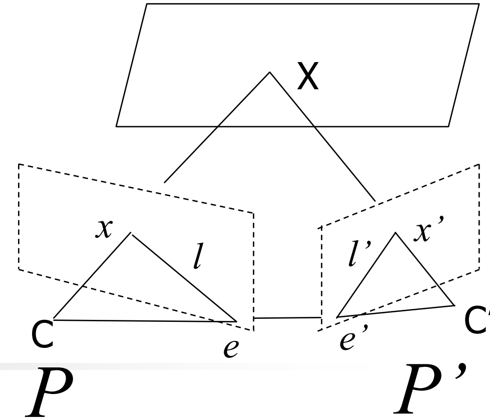
- F only depends on projective properties of P and P'.
- Independent of choice of world frame.
- (P,P') → F (unique)
- F→ (P,P') (?)
- Given a homography H (4x4 non-singular matrix) in P^3, if (P,P') → F, then $(PH,P'H)$→ F.
  Proof: PX ←→ $P'X$
  ➔ $(PH)(H^{-1}X)$ ←→ $(P'H)\ (H^{-1}X)$
- *F does not uniquely map to (P,P').*

Retrieving the camera matrices from F



- P=[I | 0] & P'=[M | m] → F=$[m]_x M$.
- If F derived from both (P_1, P_1') and (P_2, P_2'), there exists 4x4 H s.t. $P_2 = P_1 H$ & $P_2' = P_1' H$.
- d.o.f. of P + d.o.f. of P'=22
- d.o.f. of H = 15
- d.o.f. of F=22-15=7

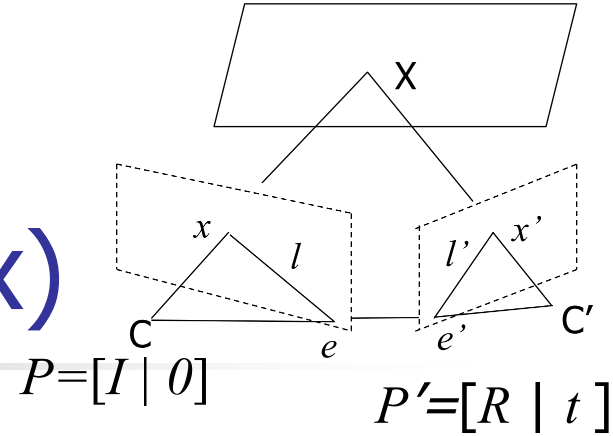
Retrieving the camera matrices from F



- F corresponds to (P,P'), iff $P'^{T}FP$ is *skew symmetric*.

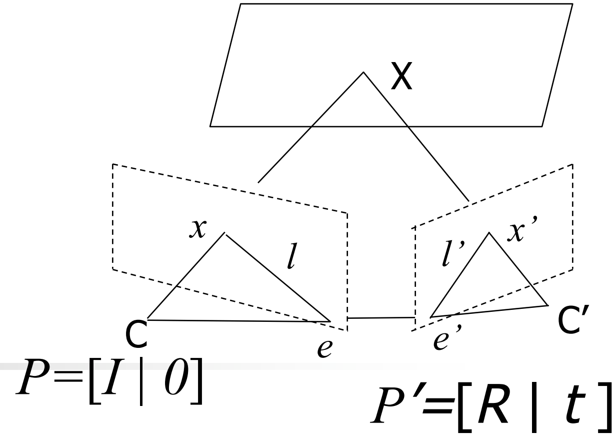
  Proof: For a skew symmetric matrix S, $X^{T}SX=0$, for all X.
  Now, $X^{T}P'^{T}FPX=(P'X)^{T}F(PX)=x'^{T}Fx=0$ (for any X in P^{3}, as F is the fundamental matrix). …
- F corresponds to $P=[I \mid 0]$ & $P'=[SF \mid e']$, where e' is the right epipole of F s.t. $e'^{T}F=0$.
- A good choice of $S=[e']_{x}$.
- F → $([I \mid 0], [[e']_{x}F \mid e'])$
  ←→ $([I \mid 0], [[e']_{x}F+e'\mathbf{v}^{T} \mid ke'])$

The camera matrices from E (Essential matrix)



- E is an essential matrix iff two of its singular values are equal and the third one is zero.
- $E=[t]_x R$
- $[t]_x$ and R can be computed through decomposition of E s.t. $E=SR$, where S is a skew symmetric matrix and R is orthogonal.

Decomposition of E (Essential matrix)

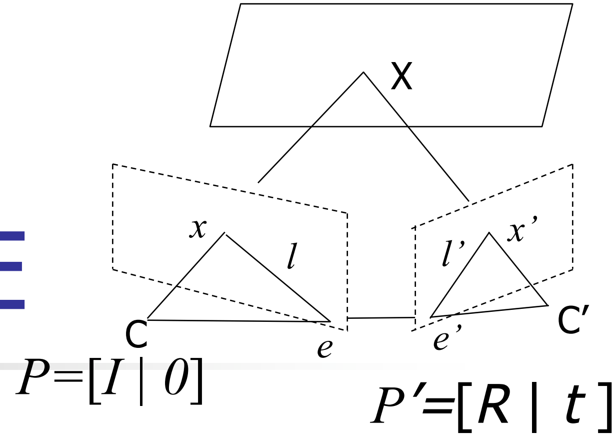


- SVD of $E=U\,\mathrm{diag}(1,1,0)\,V^T$
- Two possible decomposition of $E=SR$
- $S=UZU^T$ and $R=UWV^T$ or UW^TV^T

$$z=\begin{bmatrix}0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \qquad W=\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}$$

- Any skew symmetric matrix S can be decomposed as
  $S=kUZU^T$
- W is orthogonal and $Z=\mathrm{diag}(1,1,0)W$.

Camera matrices from E



- SVD of $E=U\,\mathrm{diag}(1,1,0)\,V^T$
- Two possible decomposition of $E=SR$
- $S=UZU^T$ and $R=UWV^T$ or UW^TV^T

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

$P'=[UWV^T \mid +u_3]$ or $[UWV^T \mid -u_3]$

$[UW^TV^T \mid +u_3]$ or $[UW^TV^T \mid -u_3]$

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

Last column of U.

Out of the four only one is valid for viewing a point from both the cameras. It is sufficient to test a single point for the above.

Ex. 5

- Check whether the following fundamental matrix and the camera matrices P and P′ (for left and right cameras) are compatible.

$$F = \begin{bmatrix} 20 & 12 & 0 \\ 8 & -7 & 59 \\ -4 & 33 & -177 \end{bmatrix}$$

$$P = \begin{bmatrix} 7 & 4 & -6 & 3 \\ 8 & -1 & 2 & -5 \\ 9 & -10 & 4 & 1 \end{bmatrix} \qquad P' = \begin{bmatrix} 6 & 4 & -6 & 10 \\ 8 & -5 & 2 & -7 \\ 9 & -10 & 6 & 2 \end{bmatrix}$$

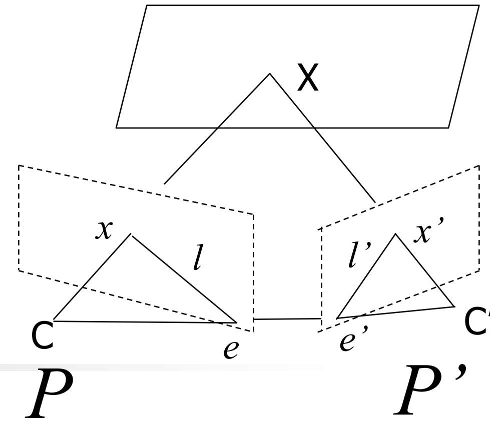
Ans. 5

- $P'^T FP=$

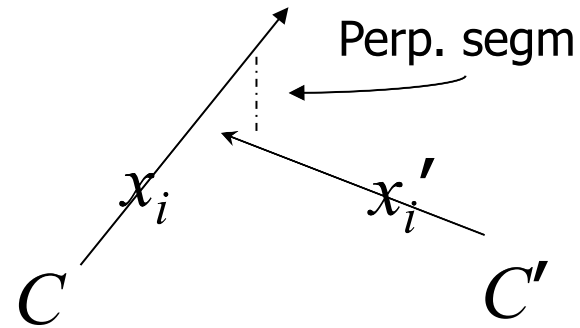
$$\begin{bmatrix} -6549 & 11489 & -4746 & -2242 \\ 11859 & -14183 & 4926 & 2950 \\ -8496 & 8816 & -2784 & -1888 \\ -4071 & 7979 & -3414 & -1534 \end{bmatrix}$$

- Not a skew symmetric matrix.
- Not compatible.

Computing scene points (structure)



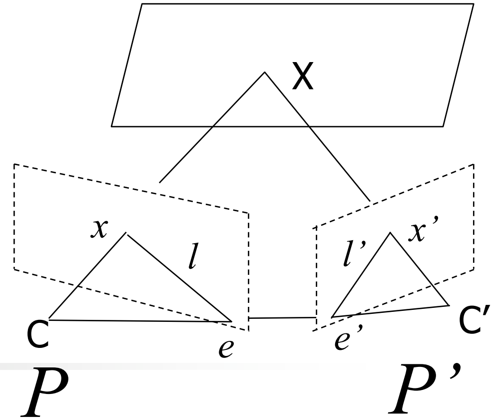
Given $x_i \leftarrow\rightarrow x_i'$, compute X.



1. Compute F.
2. Compute P and P'.
3. For each (x_i, x_i') compute X by triangulation.
   1. Compute intersection of Cx_i and $C'x_i'$.
   2. Compute segment perpendicular to both.
   3. Get the mid-point.

Not projective invariant, i.e. $(PH, P'H)$ does not give $H^{-1}X$.

Minimizing Reprojection Error



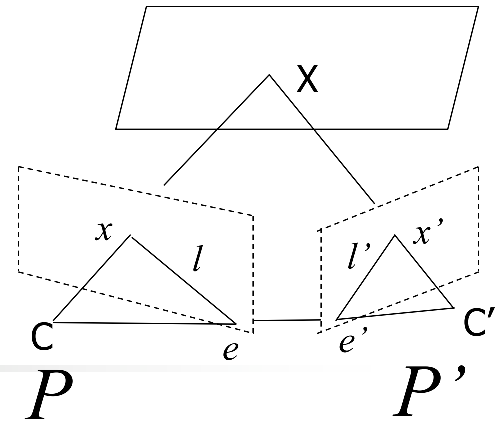
Given $x_i \leftarrow\rightarrow x_i'$, compute X.

1. Estimate $\hat{X}$ s.t. $P\hat{X} = \hat{x}$ and $P'\hat{X} = \widehat{x'}$.
2. Minimize the reprojection error (E_{rp}).

$$E_{rp} = d(x, \hat{x})^2 + d\left(x', \widehat{x'}\right)^2$$
$$\text{subject to } x'^T F x = 0$$

Projective invariant.

Linear triangulation methods



Given $x_i \leftarrow\rightarrow x_i'$, compute X.

$$x \times PX = 0$$
$$x' \times P'X = 0$$

4 equations,
3 unknowns.

$[A]_{4x4}X=0$

Minimize $||AX||$
subject to $||X||=1$.

Use DLT.

Not projective invariant.

Generalize to multiview correspondences.

$$x_1 \leftarrow\rightarrow x_2 \leftarrow\rightarrow x_3$$
$$P_1 \quad P_2 \quad P_3$$

$\Rightarrow$

$$x_1 \times P_1 X = 0$$
$$x_2 \times P_2 X = 0$$
$$x_3 \times P_3 X = 0$$

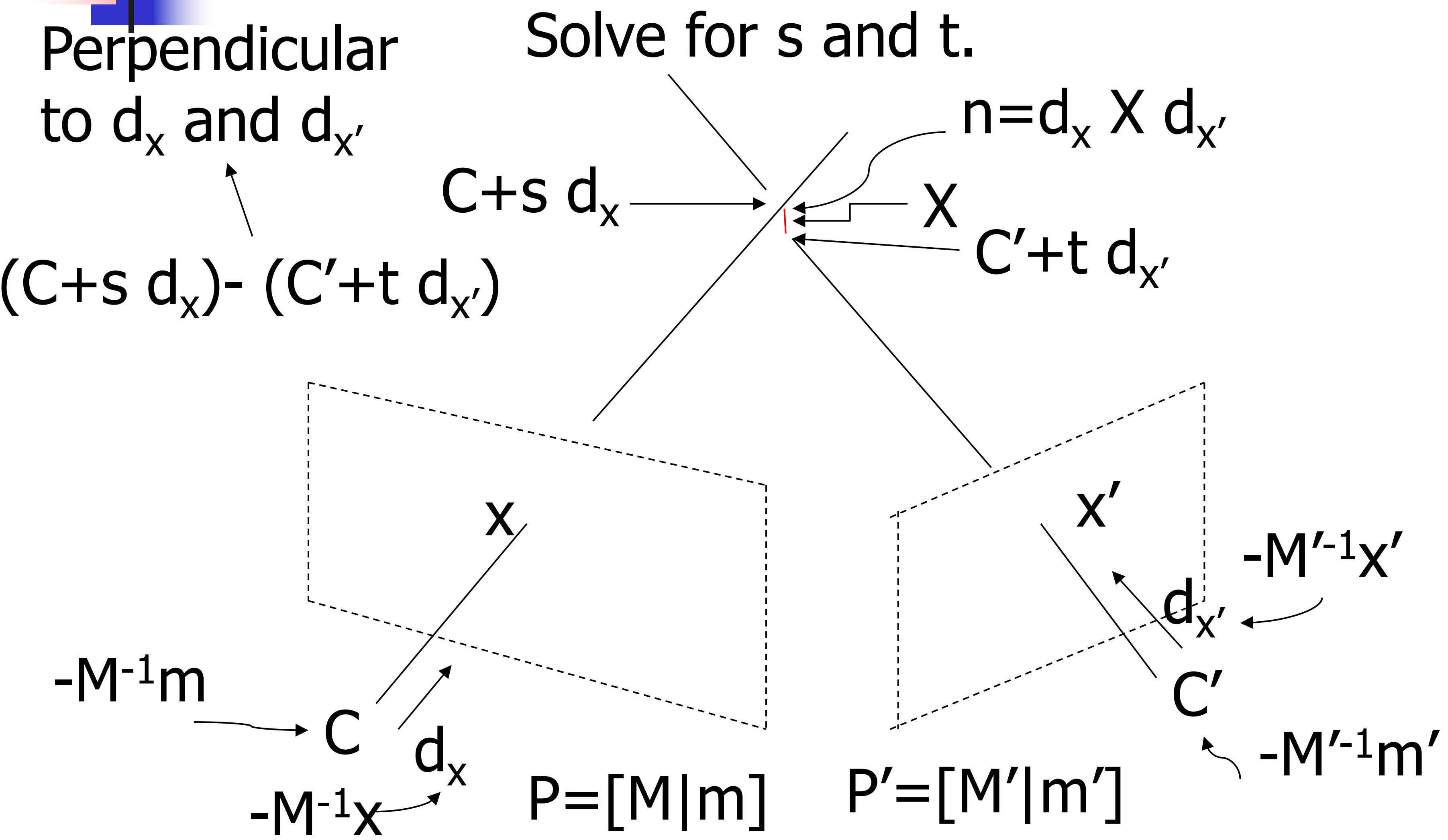
6 equations
3 unknowns.

Ex. 4

- Suppose P and P' (for left and right cameras). Images of a scene point are formed at (0,3.5), and (2/3, -1/3), respectively. Find the 3D coordinate of the scene point.

$$P = \begin{bmatrix} 7 & 4 & -6 & 3 \\ 8 & -1 & 2 & -5 \\ 9 & -10 & 4 & 1 \end{bmatrix} \quad P' = \begin{bmatrix} 6 & 4 & -6 & 10 \\ 8 & -5 & 2 & -7 \\ 9 & -10 & 6 & 2 \end{bmatrix}$$

Computation of structure



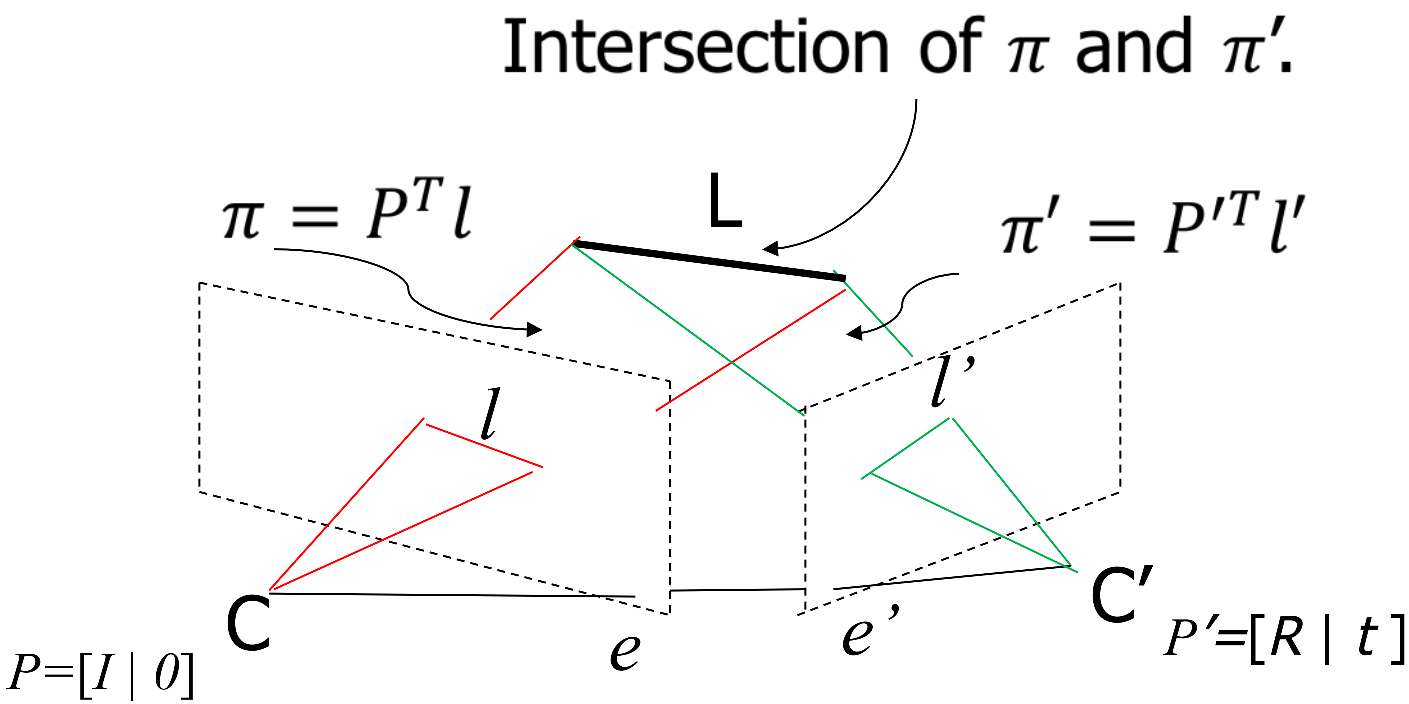
$$P = \begin{bmatrix} 7 & 4 & -6 & 3 \\ 8 & -1 & 2 & -5 \\ 9 & -10 & 4 & 1 \end{bmatrix} \quad P' = \begin{bmatrix} 6 & 4 & -6 & 10 \\ 8 & -5 & 2 & -7 \\ 9 & -10 & 6 & 2 \end{bmatrix}$$

4 Ans.

- $P=[M|m]$ $P'=[M'|m']$
- $C=-M^{-1}m= [0.35\ 1\ 1.62]^T$
- $C'=-M'^{-1}m'=[13\ 35\ \ 38]^T$
- $x=[0\ 3.5\ 1]^T \quad x'=[2/3\ -1/3\ 1]^T$
- $d_x= M^{-1}\ x=[0.32 \quad 0.47 \quad 0.69]$
- $d_{x'}= M'^{-1}\ x'=[1.36 \quad 3.67 \quad 3.91]$
- $d_x \times d_{x'}= [0.7 \quad 0.33 \quad -0.55]$

- $(C+s\ d_x)-(C'+t\ d_{x'})=[-12.85\ \ -33.93\ \ -36.58]+s[0.32\ \ 0.47\ \ 0.69]-t[1.36\ \ 3.67\ \ 3.91]$
- $0.32(-12.85+0.3s-1.36t)+0.47(-33.93+0.47s-3.67t)+0.69(-36.58+0.69s-3.91t)=0$ –(1)
- $1.36(-12.85+0.3s-1.36t)+3.67(-33.93+0.47s-3.67t)+3.91(-36.58+0.69s-3.91t)=0$ – (2)
- Solve (1) and (2) to get s and t, and the point.

Line reconstruction



$$L = \begin{bmatrix} \pi \\ \pi' \end{bmatrix}$$

A convenient way of representing 3D line.

Plane Induced Homography

Proof:

$$x' = P'X = [A|a]X$$

Now, $x = PX = [I\ |0]X$

So any point in $\overrightarrow{CX}$ is $X = \begin{bmatrix} x \\ \rho \end{bmatrix}$

When it intersects π, $\pi^T \begin{bmatrix} x \\ \rho \end{bmatrix} = 0.$

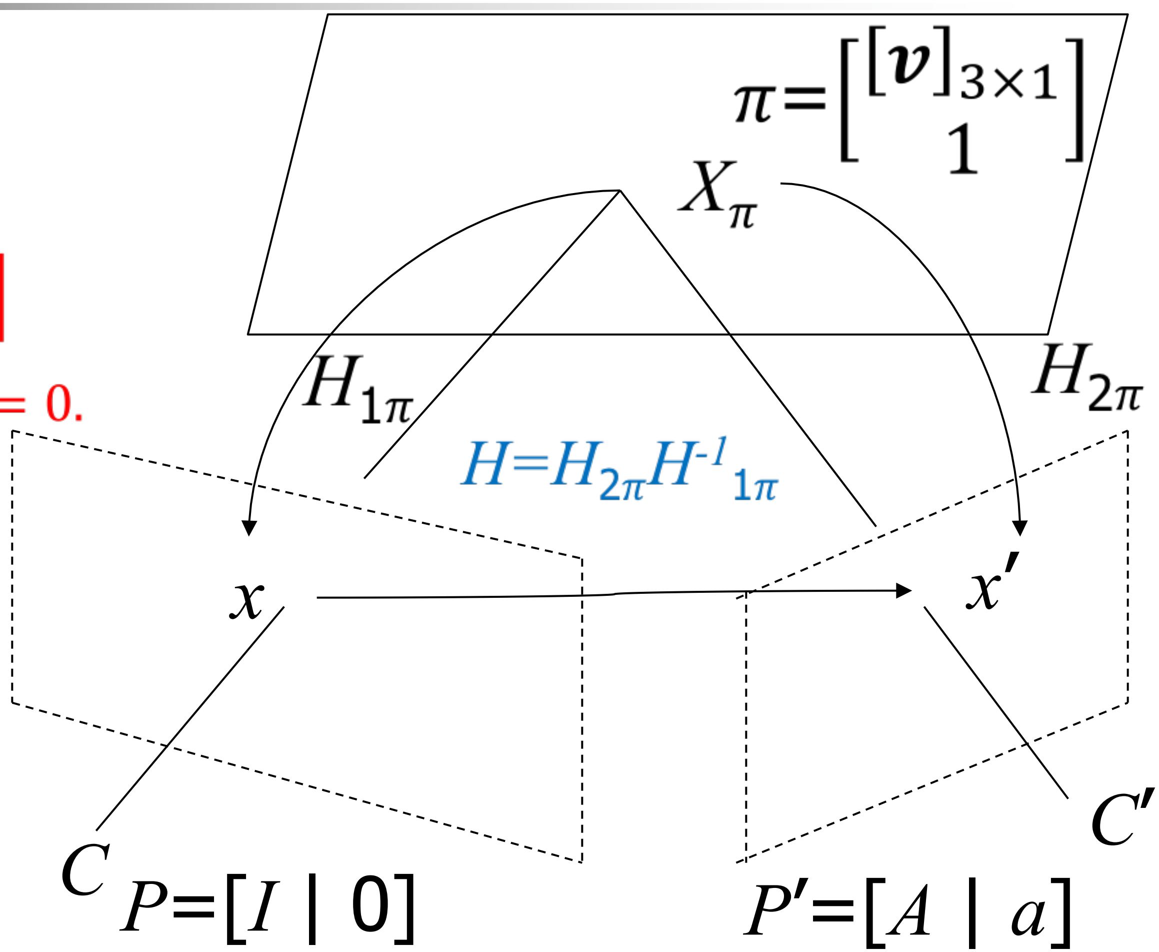
$$\Rightarrow\ \boldsymbol{v}^T x + \rho = 0$$

$$\Rightarrow\ \rho = -\boldsymbol{v}^T x$$

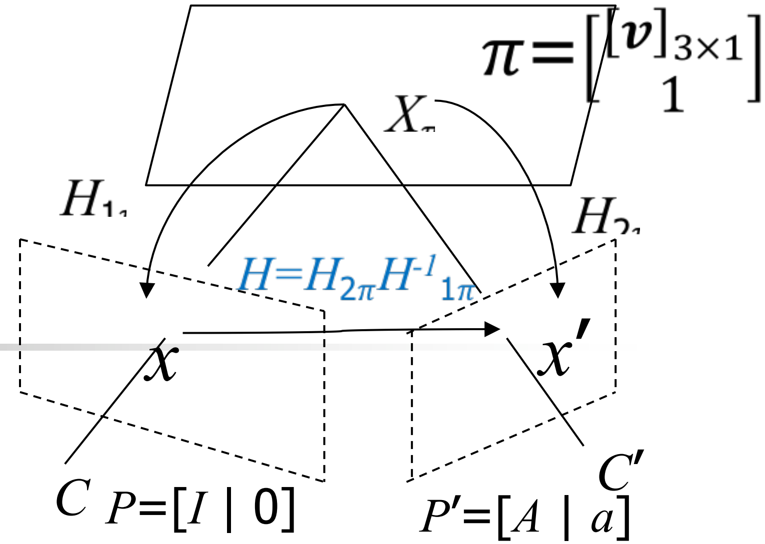
So, $x' = P'X = [A|a] \begin{bmatrix} x \\ -\boldsymbol{v}^T x \end{bmatrix}$

$= Ax - a\ \boldsymbol{v}^T x$

$= \underbrace{(A - a\ \boldsymbol{v}^T)}_{H}x$



Plane induced homography

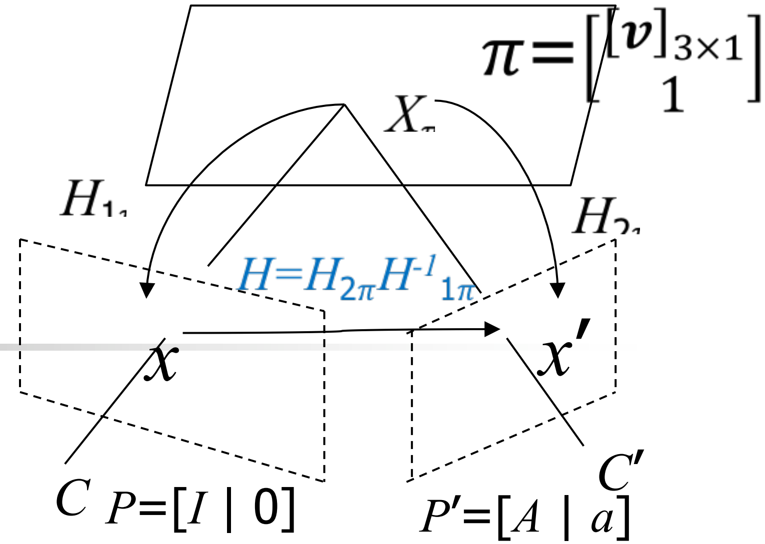


A transformation H between two stereo images is plane induced homography if F is decomposed into $[e']_x H$.

Hence, $P=[I|0]$ & $P'=[H|e']$.

Given $P=[I|0]$, $P'=[A|a]$, & a plane induced homography H, the plane can be recovered by solving $kH=A-a\mathbf{v}^T$, (linear equations for unknowns k and v).

Plane induced homography



A transformation H between two stereo images is plane induced homography if F is decomposed into $[e']_x H$. Hence, $P=[I|0]$ & $P'=[H|e']$.

Plane at infinity

H is the transformation w.r.t. plane $[0\ 0\ 0\ 1]^T$ in the camera coordinate.

Given $P=[I|0]$, $P'=[A|a]$, & a plane induced homography H, the plane can be recovered by solving $kH=A-a\mathbf{v}^T$, (linear equations for unknowns k and v).

Homography compatible stereo geometry

H is compatible iff $H^T F$ is skew symmetric, i.e.

$$H^T F + F^T H = 0$$

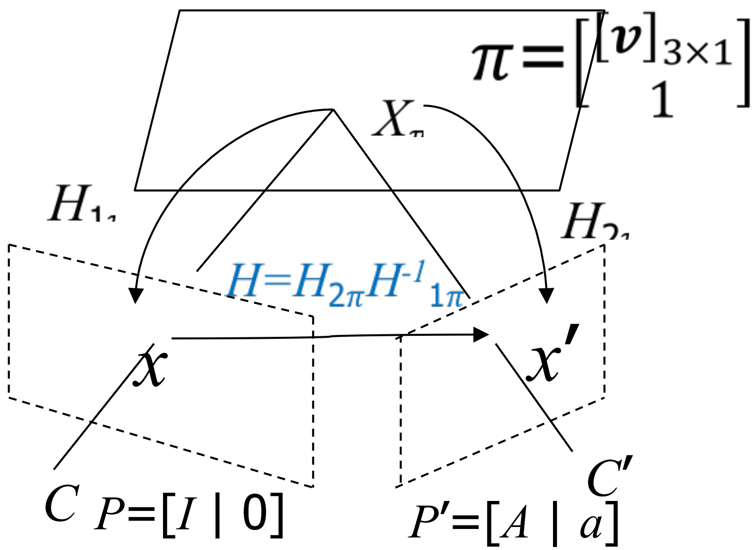
$$x'^T F x = 0$$

And, $x' = Hx$

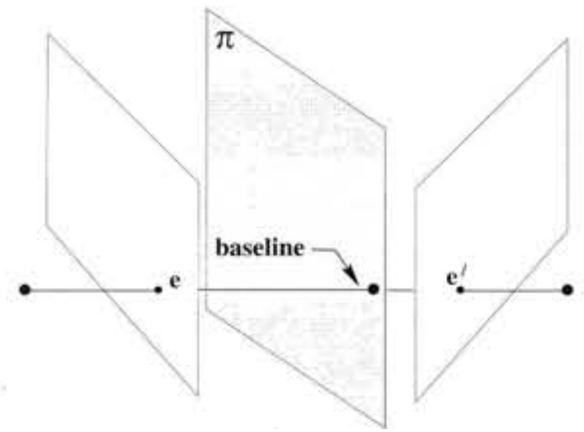
$$\Rightarrow (Hx)^T F x = 0$$

$$\Rightarrow x^T H^T F x = 0$$

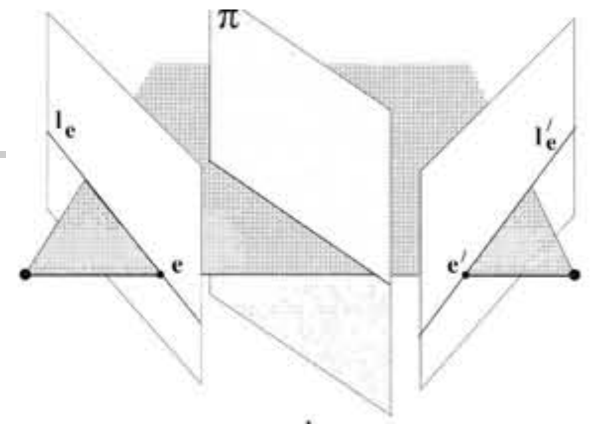
As this is true for all x, $H^T F$ is skew symmetric.



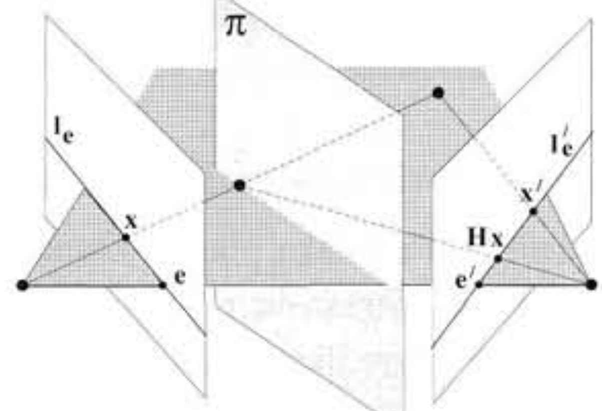
Plane induced *H* and epipolar constraints



Epipoles mapped by H, as e'=He, *since they are images of the point on the plane where the baseline intersects it.*



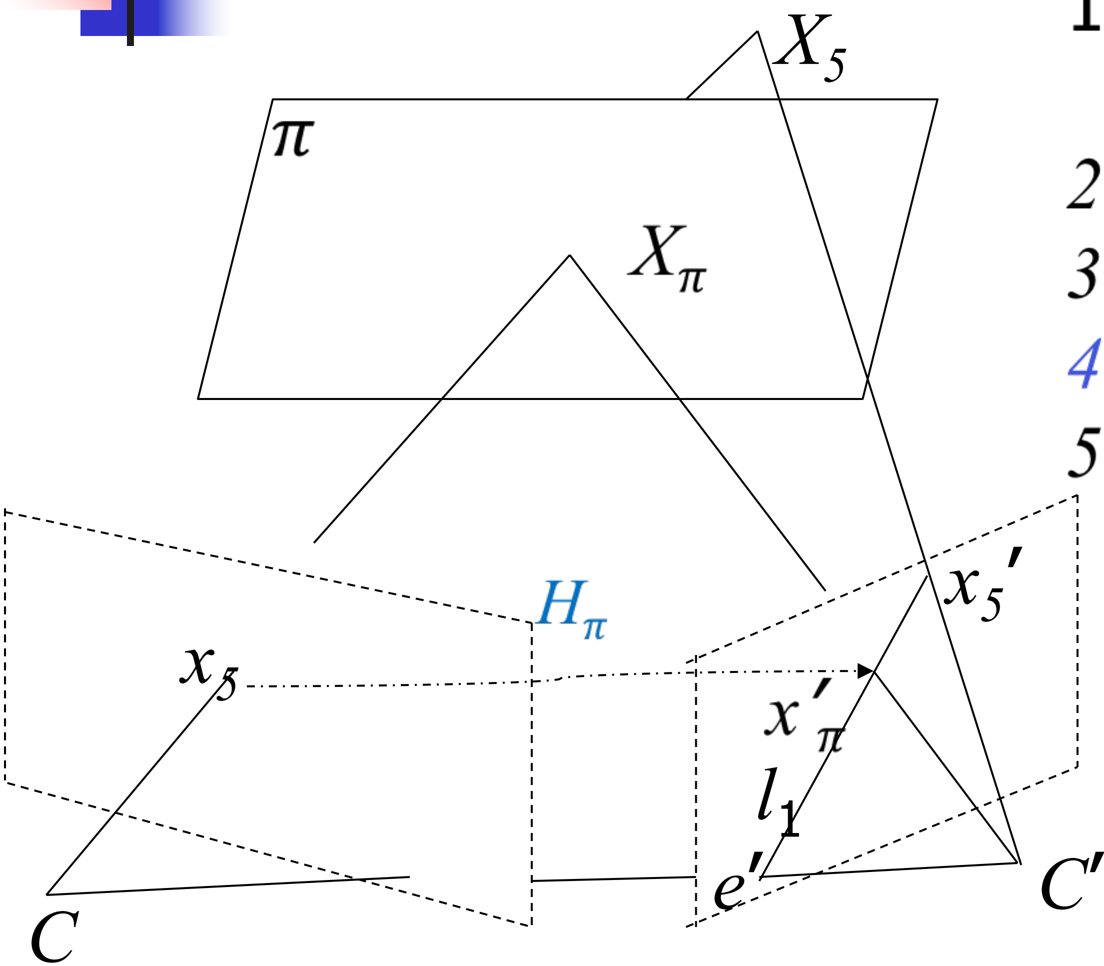
Epipolar lines are mapped by the homography as $H^{-T}l'_e = l_e$.



Hx *lies on epipolar line* l_e' *so* $l'_e = Fx = x' \times (Hx)$.

From Hartley and Zisserman, "Multiple view geometry in computer vision", Cambridge Univ. Press (2000)

Computing F from 6 points out of which 4 are coplanar



1. Use 4 coplanar points to compute H_π.
2. $l_1 = H_\pi(x_5) \times x_5'$
3. $l_2 = H_\pi(x_6) \times x_6'$
4. $e' = l_1 \times l_2$
5. $F = [e']_\times H_\pi$.

Given F and 3 point correspondences, compute H.

First Method

1. Obtain $(P=[I|0], P'=[A|a])$ from F and construct 3 scene points, X_1, X_2, & X_3.
2. Obtain plane $(v^T, 1)^T$.
3. Compute $H=A-av^T$

Second Method

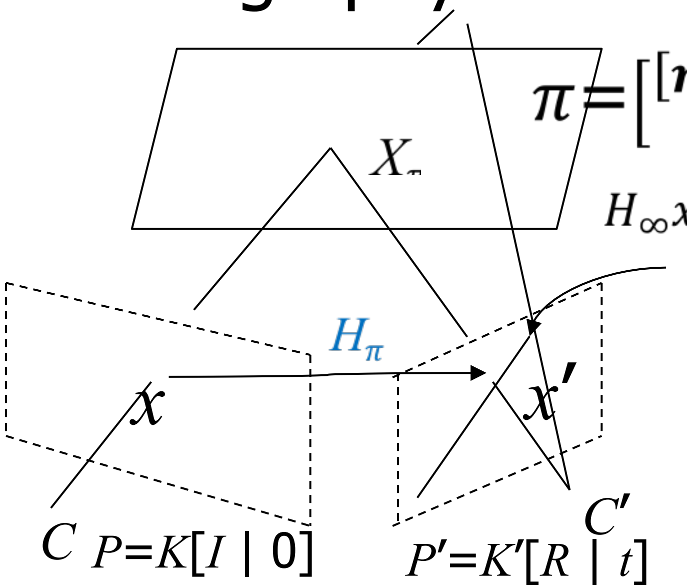
1. Obtain (e,e') from F.
2. Use 3 correspondences + (e,e'), to obtain H.

Any 3 points can bipartition the image space, w.r.t. the plane formed by them.

Infinite Homography

$H = (A - a\,\boldsymbol{v}^T)$
For P=[I |0], P'=[A | a]
and plane=[$\boldsymbol{v}^T$ 1]^T

Homography induced by the plane at α.



$$\pi = \begin{bmatrix} [\boldsymbol{n}]_{3\times1} \\ d \end{bmatrix}$$

$$H_\pi = (K'R - e'\boldsymbol{v}^T)K^{-1}$$

$$\text{where } v = \frac{n}{d} \text{ and } e' = K't$$

$$H_\pi = K'\left(R - \frac{\boldsymbol{tn}^T}{d}\right)K^{-1}$$

$$= K'RK^{-1} - K'\frac{\boldsymbol{tn}^T}{d}K^{-1}$$

As $d \to \infty$, $H_\infty = K'RK^{-1}$

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

Vanishing point

$$= H_\infty x + \frac{K't}{Z}$$

As $Z \to \infty$,
x' is the image of point on π_∞.

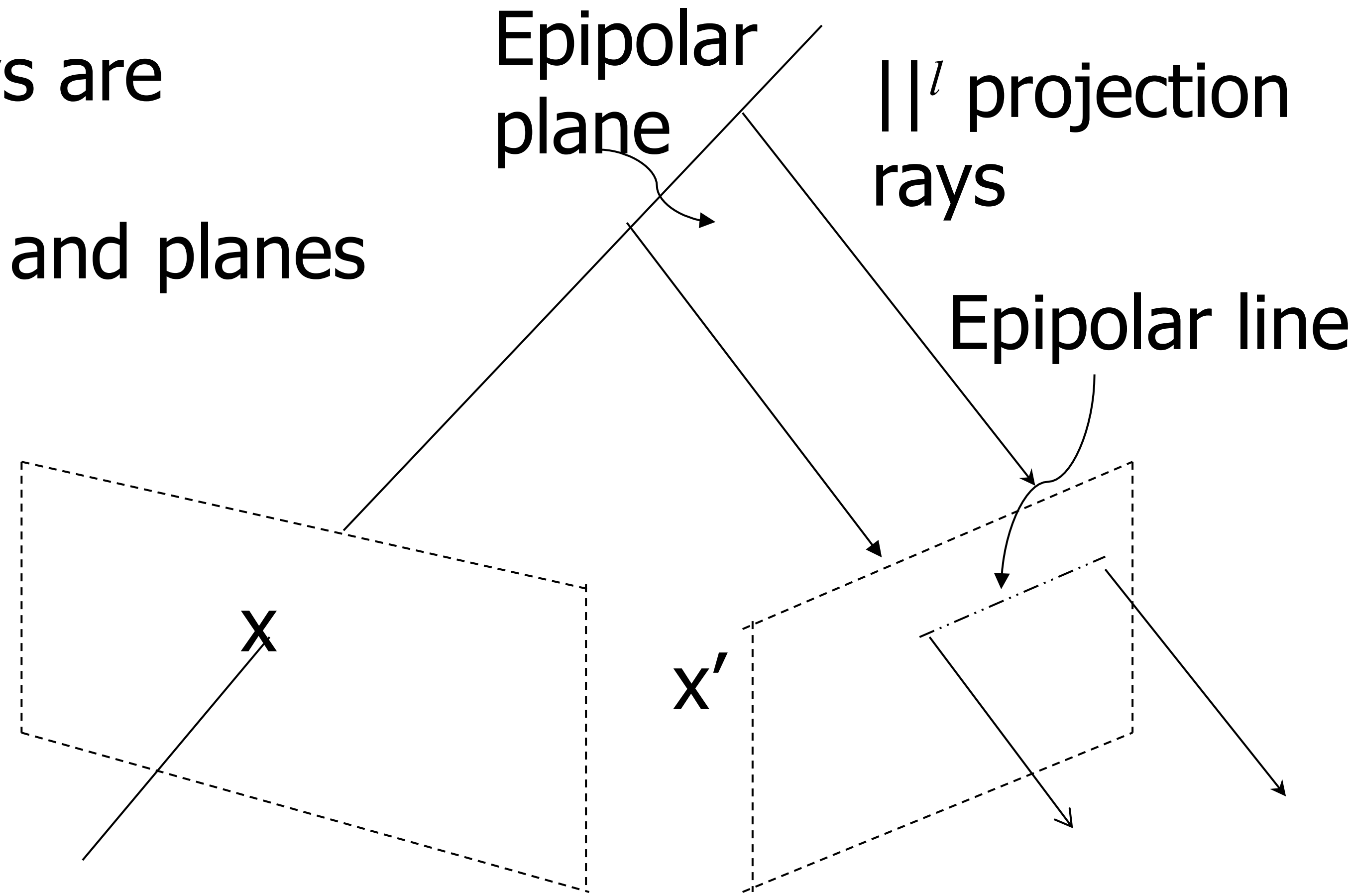
H_α and Vanishing points

- H_α maps vanishing points between two images.
- H_α can be computed by identifying three non-collinear vanishing points given F or from 4 vanishing points.
- Let $P=[M\,|m]$, $P'=[M'\,|\,m']$, $X=[x_\alpha{}^T\,0]^T$ (a point at infinity).

$$x=PX=M\,x_\alpha$$
$$x'=P'X=M'\,x_\alpha$$
$$x'=M'M^{-1}x \rightarrow H_\alpha=M'M^{-1}$$

Affine epipolar geometry

- Projection rays are parallel.
- Epipolar lines and planes are parallel.



Form of F:

$$\begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ e & d & c \end{bmatrix}$$

a,b,c,d,e all non-zero.

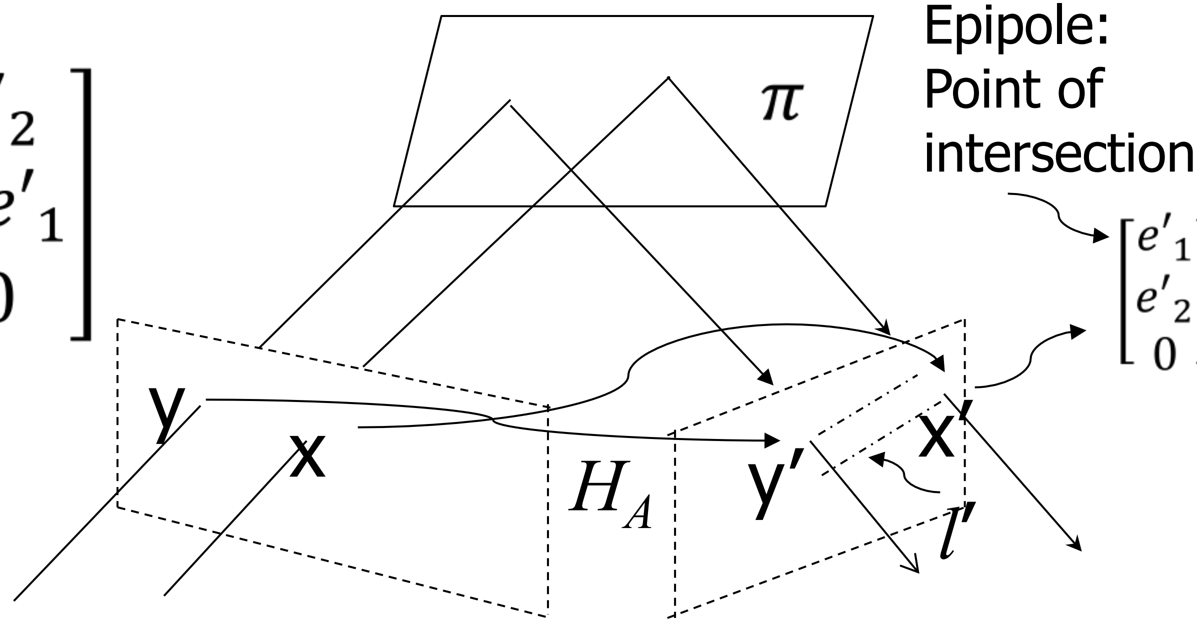
As epipolar lines are parallel, epipoles are in the form $[e_1\ e_2\ 0]^T$

Affine stereo

$$l' = e' \times H_A x$$
$$= [e']_\times H_A x$$
$$\begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ e & d & c \end{bmatrix} \Longrightarrow F_A = [e']_\times H_A$$

$$[e']_\times = \begin{bmatrix} 0 & 0 & e'_2 \\ 0 & 0 & -e'_1 \\ -e'_2 & e'_1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{0} & \boldsymbol{b} \\ -\boldsymbol{b}^T & 0 \end{bmatrix}$$

$$H_A = \begin{bmatrix} A & \boldsymbol{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$



F_A: d.o.f.: 4

$$F_A = \begin{bmatrix} 0 & \boldsymbol{b} \\ -\boldsymbol{b}^T A & -\boldsymbol{b}^T \boldsymbol{t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & e'_2 \\ 0 & 0 & -e'_1 \\ e & d & c \end{bmatrix}$$

Left epipole: $[-d \; e \; 0]^T$

right epipole: $[-b \; a \; 0]^T$

Estimating F_A

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ e & d & c \end{bmatrix}$$

Epipolar lines:

$$l' = F_A\boldsymbol{x} = [a \quad b \quad ex + dy + c]^T$$
$$l = {F_A}^T\boldsymbol{x}' = [e \quad d \quad ax' + by' + c]^T$$

Point correspondence: *Reduced to a single linear equation.*

$$x'^T F_A x = 0 \Longrightarrow ax' + by' + ex + dy + c = 0$$

$$[A]_{N\times5}\boldsymbol{f}_{5\times1} = 0$$ Solve using DLT.

Minimum 4 point correspondences required to get F_A.

Singularity constraint is satisfied by the structure of F_A.

Estimating F_A (another approach)

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ e & d & c \end{bmatrix}$$

1. Compute H_A using 3 point-correspondences.
2. $l' = H_A x_4' \times x_4'$ (say, $[l_1\ l_2\ l_3]^T$)
3. Get e' from l' as $[l_2\ -l_1\ 0]$
4. $F_A = [e']_{\times} H_A$

Summary

- Epipolar geometry in a stereo imaging system.
  - epipoles, scene point, corresponding image points, and camera centers lie on the same plane.
- Fundamental matrix (F) characteristics and unique to the stereo set-up.
  - Transformation invariant.
  - 3x3 singular matrix with d.o.f. 7 and rank 2.
  - Given an image point x, its epipolar line $l=Fx$.
  - For any pair of correspondence point (x,x'), $x'^T F x=0$.
  - Given epipoles (e,e'), $Fe=e'^T F=0$
  - Given camera matrices (P,P'), F is unique.
    - $P=[I|0]$, $P'=[M|m]$ ➔ $F=[m]_x M$.
    - $P=[M|m]$, $P'=[M'|m']$ ➔ $F=[m'-M'M^{-1}m]_x M'M^{-1}$

Summary

- Given a homography H (4x4 non-singular matrix) in P^3, if (P,P′) → F, then (PH,P′H)→ F.
- Given a fundamental matrix F, there exist a family of stereo setups (pairs of camera matrices).
  - $([I \mid 0], [[e']_{\times} F + e' \mathbf{v}^T \mid ke'])$, $\mathbf{v}$: any arbitrary 3 vector, k : a scalar constant.
- Given camera matrices (P,P′) and a corresponding pair of image points (x,x′), it is possible to reconstruct the respective 3-D scene point X.

Summary

- Fundamental matrix of calibrated cameras is called Essential Matrix E.
- $E=[t]_x R$, given $([I|0], [R|t])$.
- E is an essential matrix iff two of its singular values are equal and the third one is zero.
- $[t]_x$ and R can be computed through decomposition of E s.t. $E=SR$, where S is a skew symmetric matrix and R is orthogonal.

Summary

- SVD of $E=U$ diag(1,1,0) V^T
- Two possible decomposition of $E=SR$
- $S=UZU^T$ and $R=UWV^T$ or UW^TV^T

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

P'=[UWV^T | $+u_3$] or [UWV^T | $-u_3$]
[UW^TV^T | $+u_3$] or [UW^TV^T | $-u_3$]

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

One of the above is valid in viewing a point from both the cameras.

Summary

- Given a set of pairs of corresponding points, it is possible to estimate F.
  - Minimum 7 pairs required.
- Parametric representation of F.

$$F = \begin{bmatrix} a & b & \alpha a + \beta b \\ c & d & \alpha c + \beta d \\ \alpha' a + \beta' c & \alpha' b + \beta' d & \alpha\alpha' a + \alpha\beta' c + \alpha'\beta b + \beta\beta' d \end{bmatrix}$$

$$Epipoles: \cancel{e' = [\alpha \quad \beta \quad -1]^T \qquad e = [\alpha' \quad \beta' \quad -1]^T}$$

$$e = [\alpha \quad \beta \quad -1]^T \quad e' = [\alpha' \quad \beta' \quad -1]^T$$

Summary

- Given a set of pairs of corresponding points, it is possible to estimate camera matrices and scene points up to projective (4x4 homography matrix) ambiguity.
  - $(PH)(H^{-1}X) \leftarrow\rightarrow (P'H)(H^{-1}X)$
- Given a pair of corresponding lines l and l', and camera matrices (P,P') possible to reconstruct respective 3D line L.
  - Intersection of planes $P^T l$ and $P'^T l'$

Summary

- A plane induces homography between corresponding image points in a stereo set-up.
  - Given a plane $(\mathbf{v}^T, 1)$, and camera matrices $([I|0],[A|a])$, $H=(A-a\mathbf{v}^T)$
    - Homography at infinity: Plane at infinity $(\mathbf{0}^T,1)$ induces $H=A$.
- Affine epipolar geometry simplifies the structure of fundamental matrix.
  - Right epipole: $[-d\ e\ 0]^T$
  - Left epipole: $[-b\ a\ 0]^T$

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ e & d & c \end{bmatrix}$$