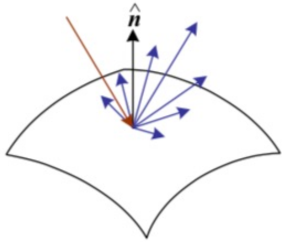


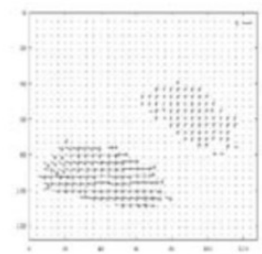
2. Image Formation

3. Image Processing

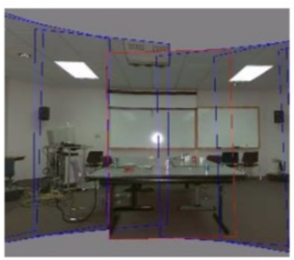
4. Features

5. Segmentation

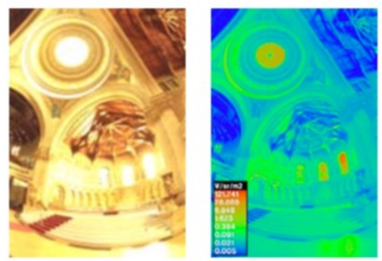
6-7. Structure from Motion



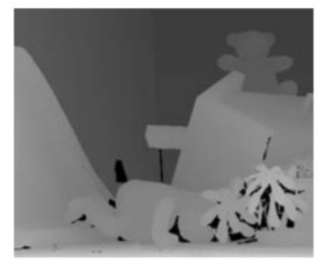
8. Motion



9. Stitching

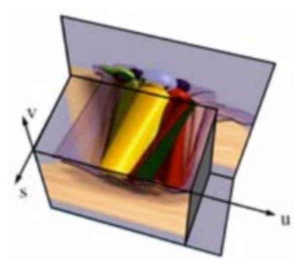


10. Computational Photography



11. Stereo

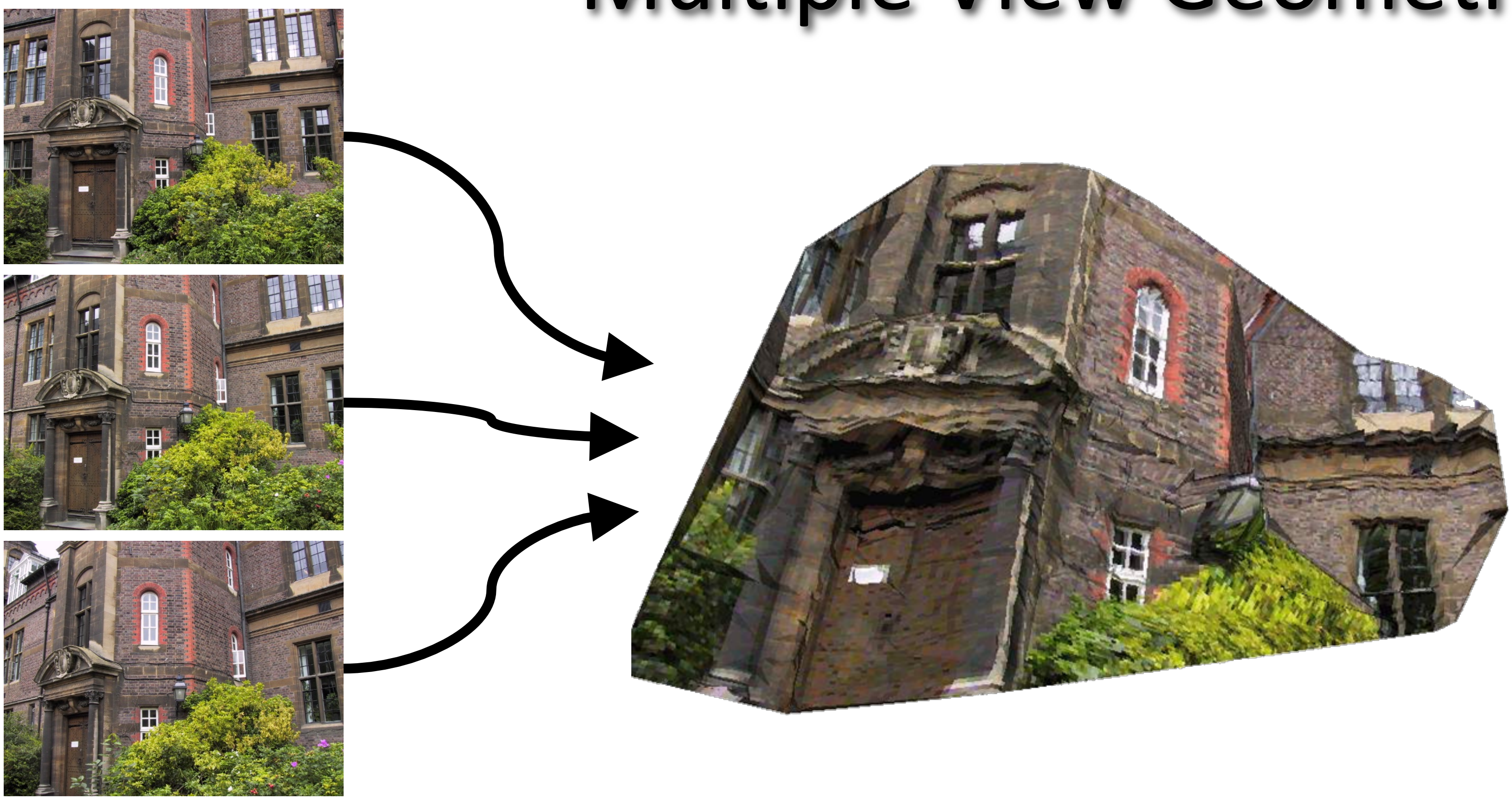
12. 3D Shape



13. Image-based Rendering

14. Recognition

Multiple View Geometry

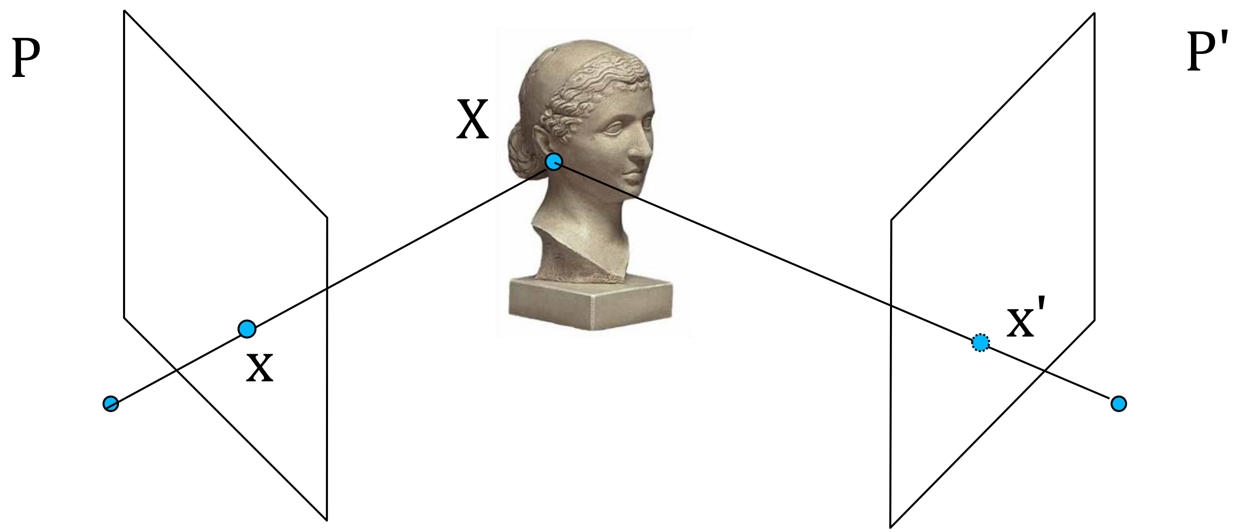


Frank Dellaert

Outline

- Intro
- Camera Review
- Stereo triangulation
- Geometry of 2 views
  - Essential Matrix
  - Fundamental Matrix
- Estimating E/F from point-matches

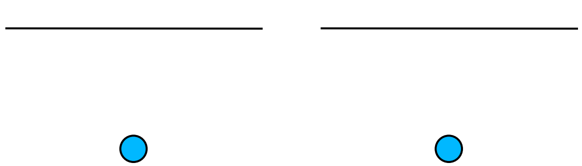
Why Consider Multiple Views?



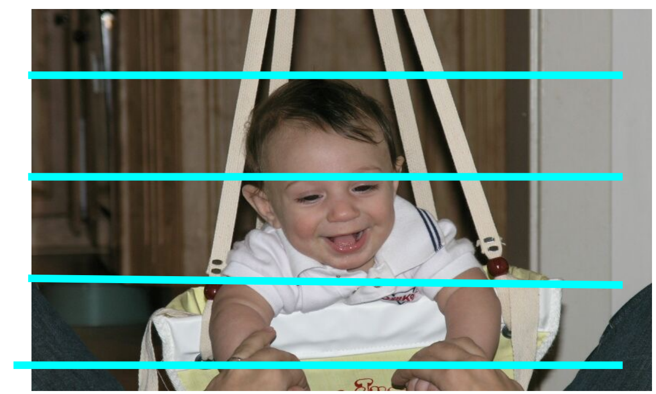
Answer: To extract 3D structure via triangulation.

Stereo Rig

Top View

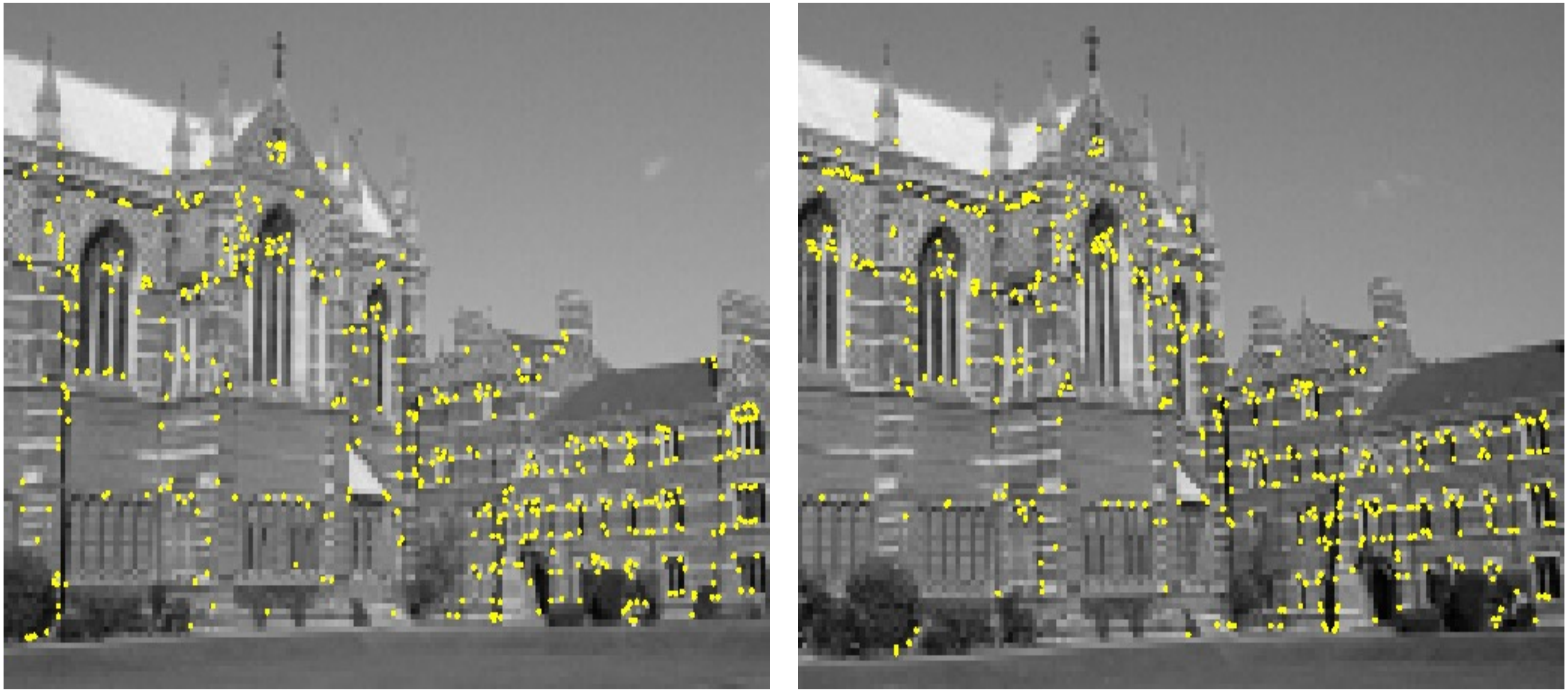


Matches on Scanlines



Convenient when searching for correspondences.

Feature Matching !



Real World Challenges

Bad News: Good correspondences are hard to find

- Good news: Geometry constrains possible correspondences.
  - 4 DOF between x and x'; only 3 DOF in X.
  - Constraint is manifest in the **Fundamental matrix**

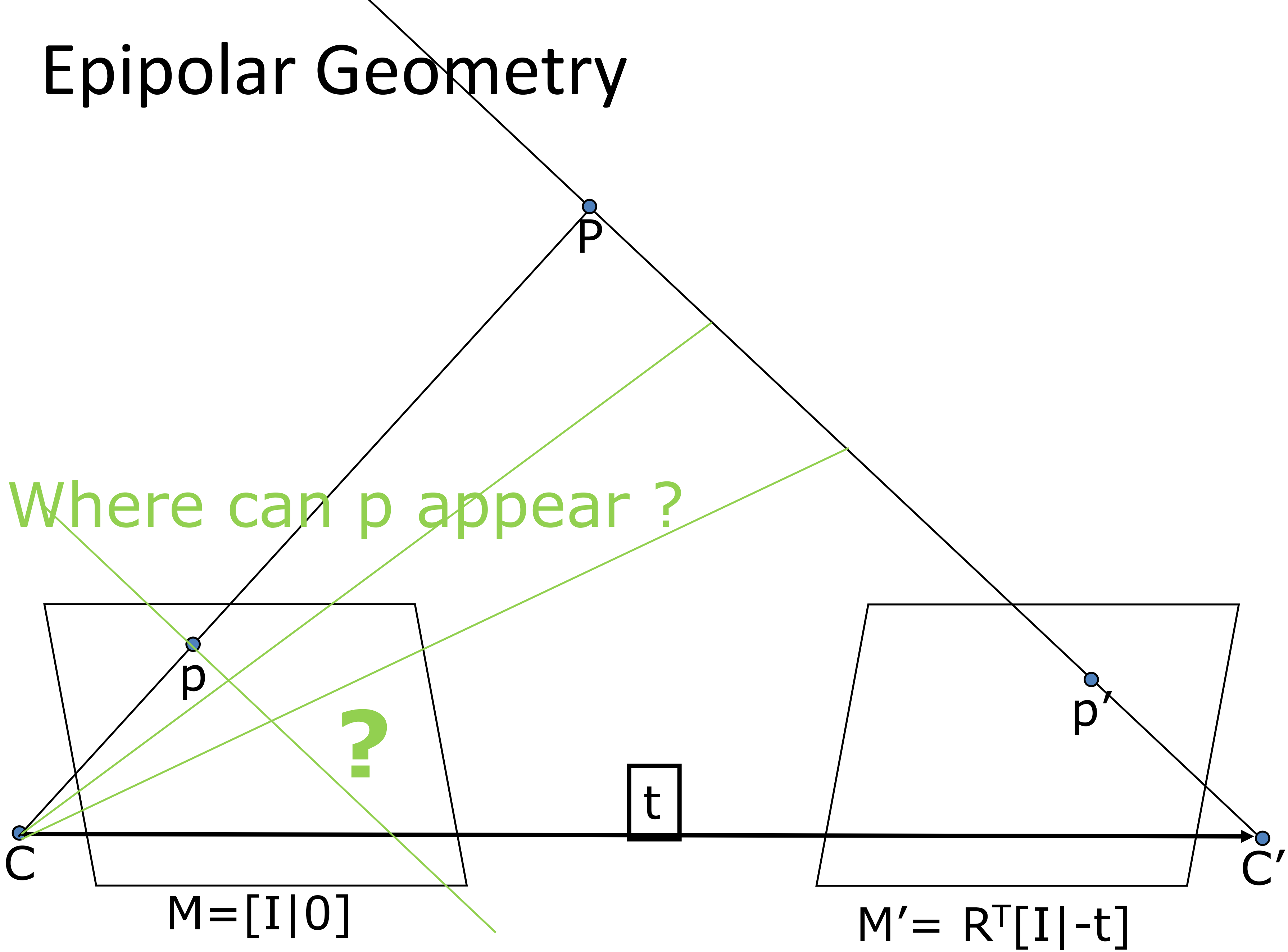
$$x'^T F x = 0.$$

  - F can be calculated either from camera matrices or a set of good correspondences.

Geometry of 2 views ?

- What if we do not know R,t ?
- Caveat:
  - My exposition uses different R, t
  - but more intuitive (IMHO)
  - I use $[R^T|-R^Tt] = R^T[I|-t]$ camera matrices
  - Szeliski uses $[R|t]$

Epipolar Geometry



Epipolar Lines

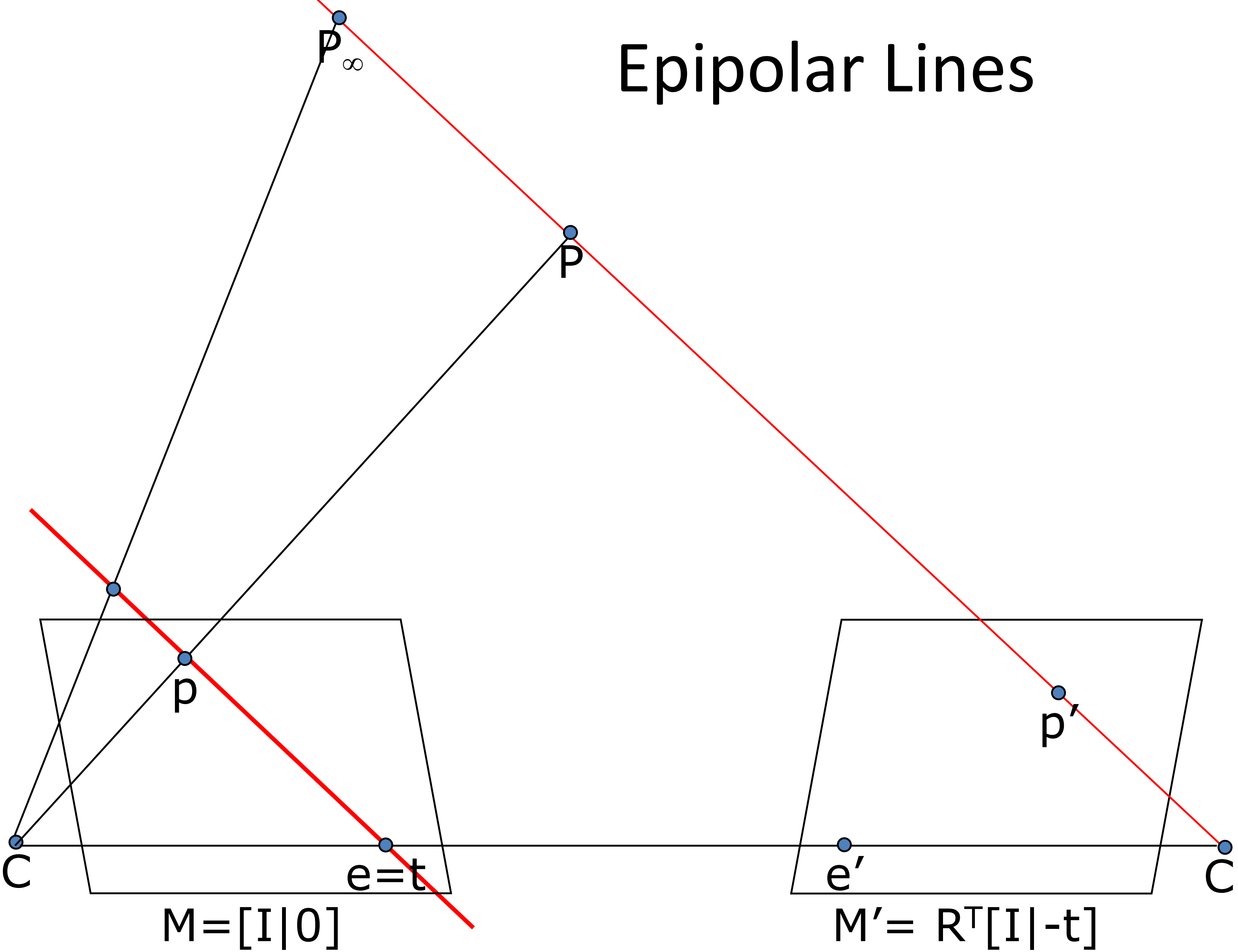


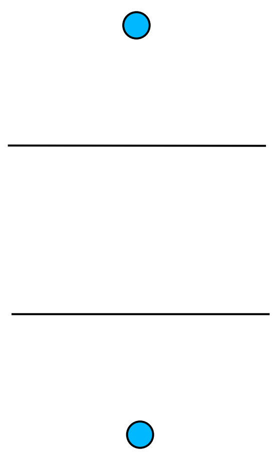
Image of Camera Center

$M=[I|0]$

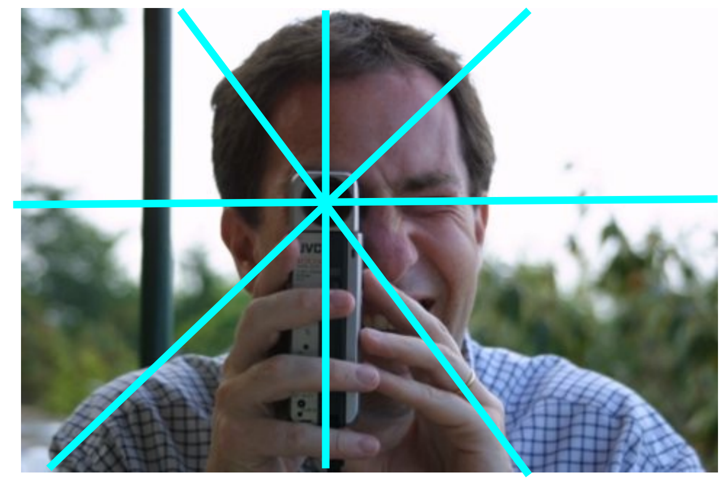
$M' = R^T[I|-t]$

Example:
Cameras Point at Each Other

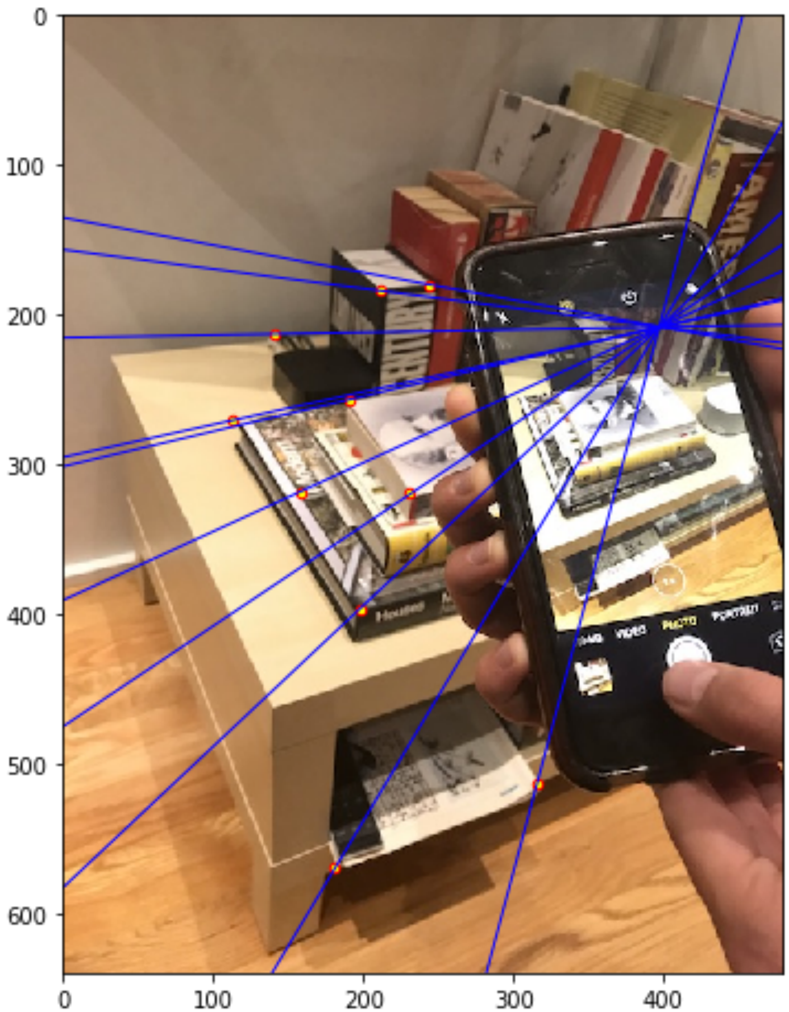
Top View

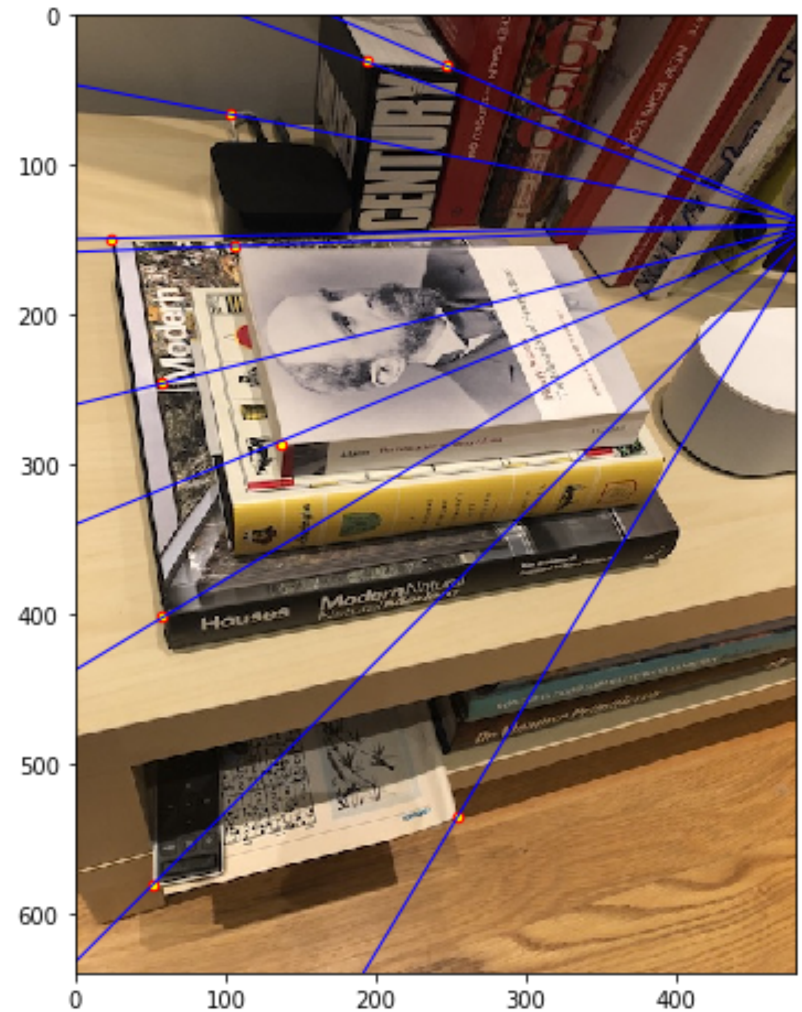


Epipolar Lines



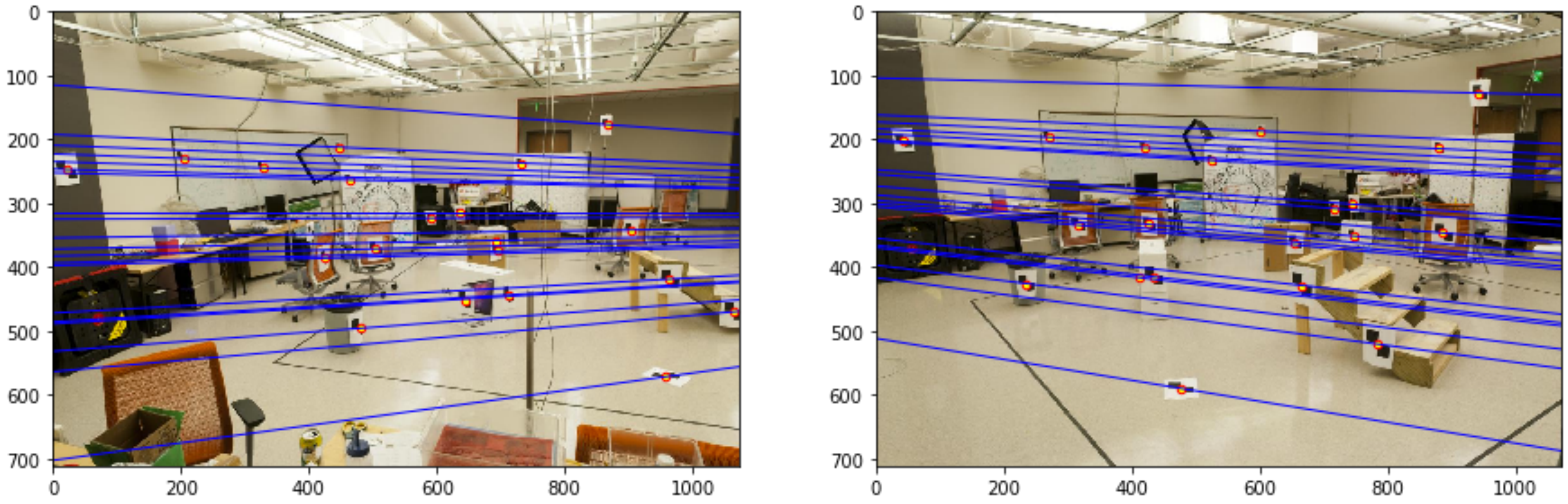
Epipoles





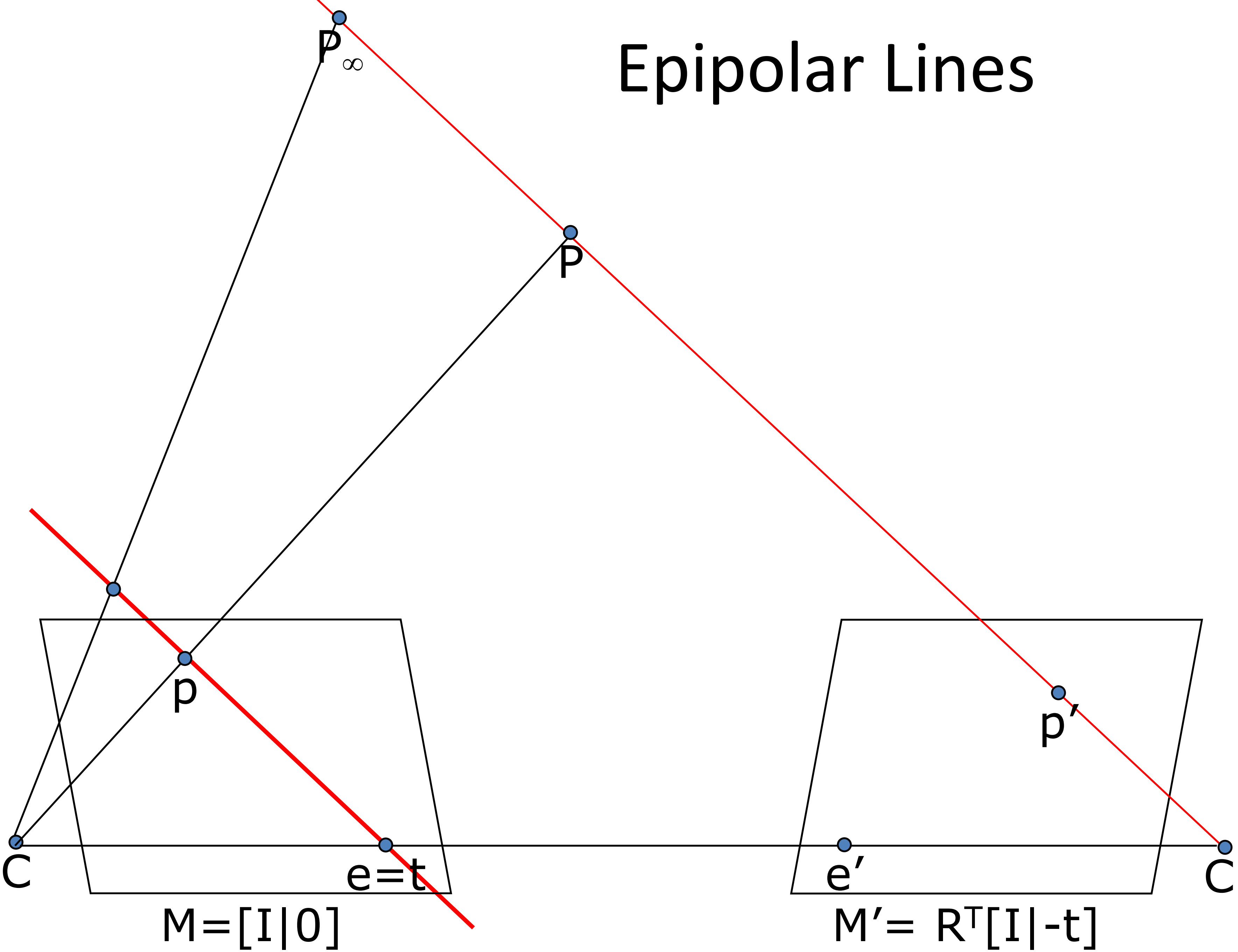
- Epipoles inside the image: zoom-like setup.

Epipoles



- Epipoles in near-stereo config.

Epipolar Lines



Epipoles

- Camera Center C' in first view:

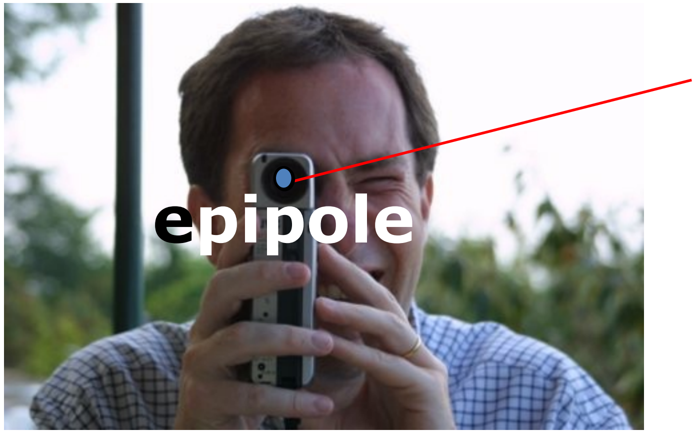
$$e = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix} = t$$

- Origin C in second view:

$$e' = \begin{bmatrix} R^T & -R^T t \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -R^T t$$

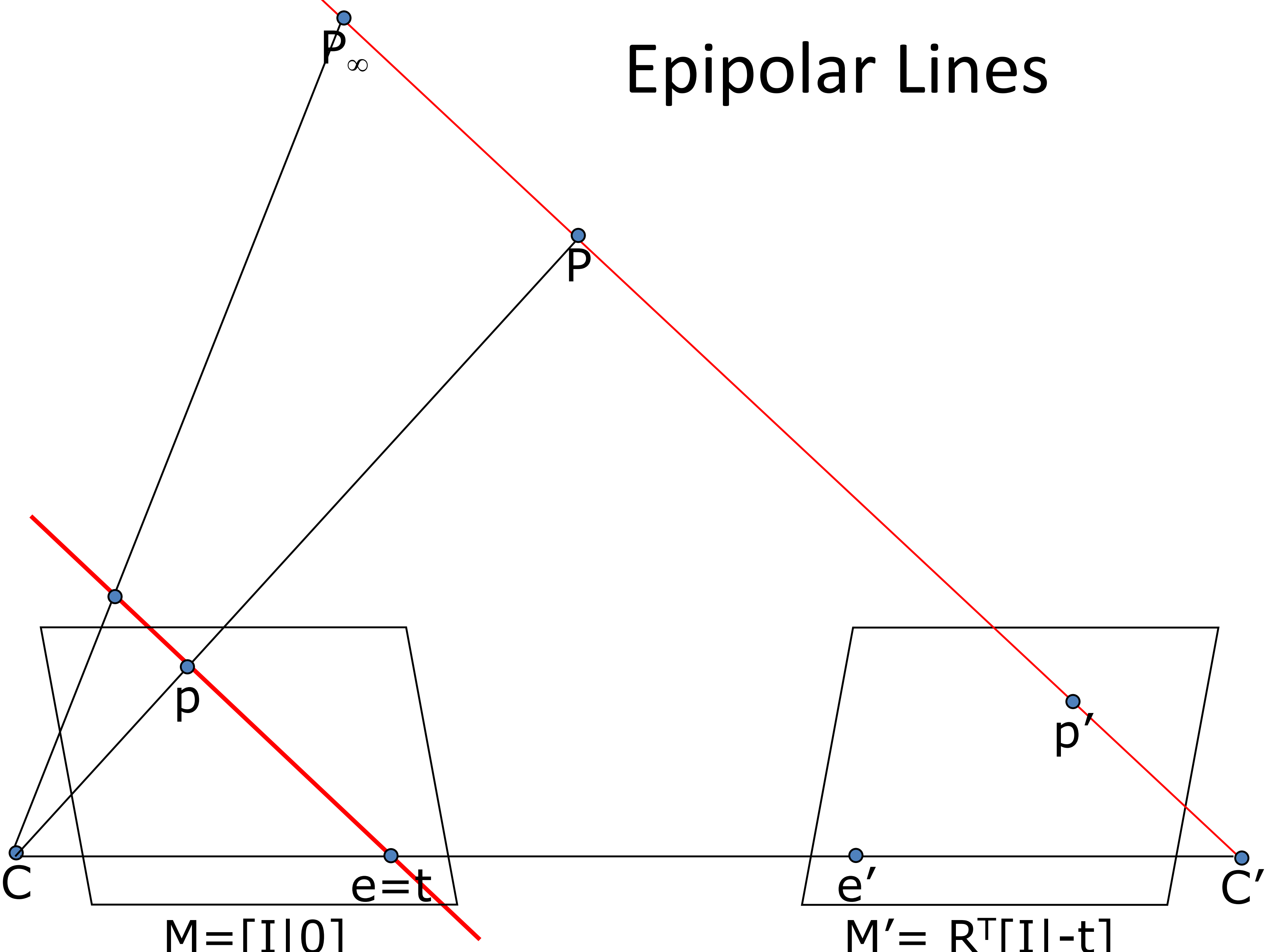
Image of Camera Ray ?

$M=[I|0]$



$M' = R^T[I|-t]$

Epipolar Lines



Point at infinity

- Given p', what is corresponding point at infinity [x 0] ?
- Answer for any camera M'=[A|a]:

$$p' = \begin{bmatrix} A & a \end{bmatrix} * \begin{bmatrix} x \\ 0 \end{bmatrix} = Ax \Rightarrow x = A^{-1} p'$$

- A^{-1} = Infinite homography
- In our case $M' = [R^T | -R^T t]$: $x = Rp'$

Sidebar: Infinite Homographies

- Homography between
  - image plane
  - plane at infinity
- Navigation by the stars:
  - Image of stars =
    function of rotation R only !
  - Traveling on a sphere rotates viewer

https://www.usni.org/magazines/naval-history-magazine/2020/june/navigating-sextant

Epipolar Line Calculation

1) Point 1 = epipole e=t

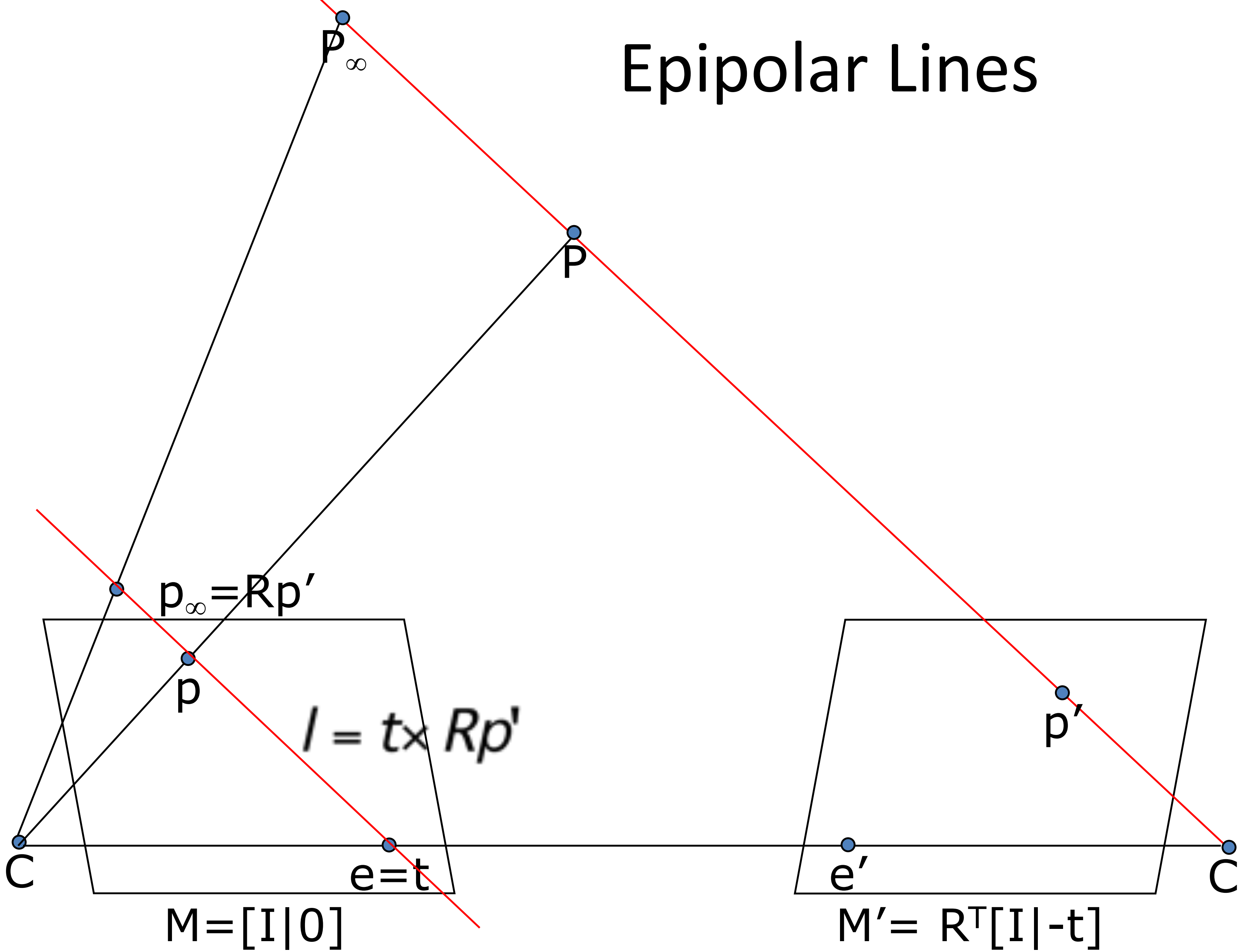
2) Point 2 = point at infinity

$$p_\infty = \begin{bmatrix} I & 0 \end{bmatrix} * \begin{bmatrix} Rp' \\ 0 \end{bmatrix} = Rp'$$

3) Epipolar line = join of points 1 and 2

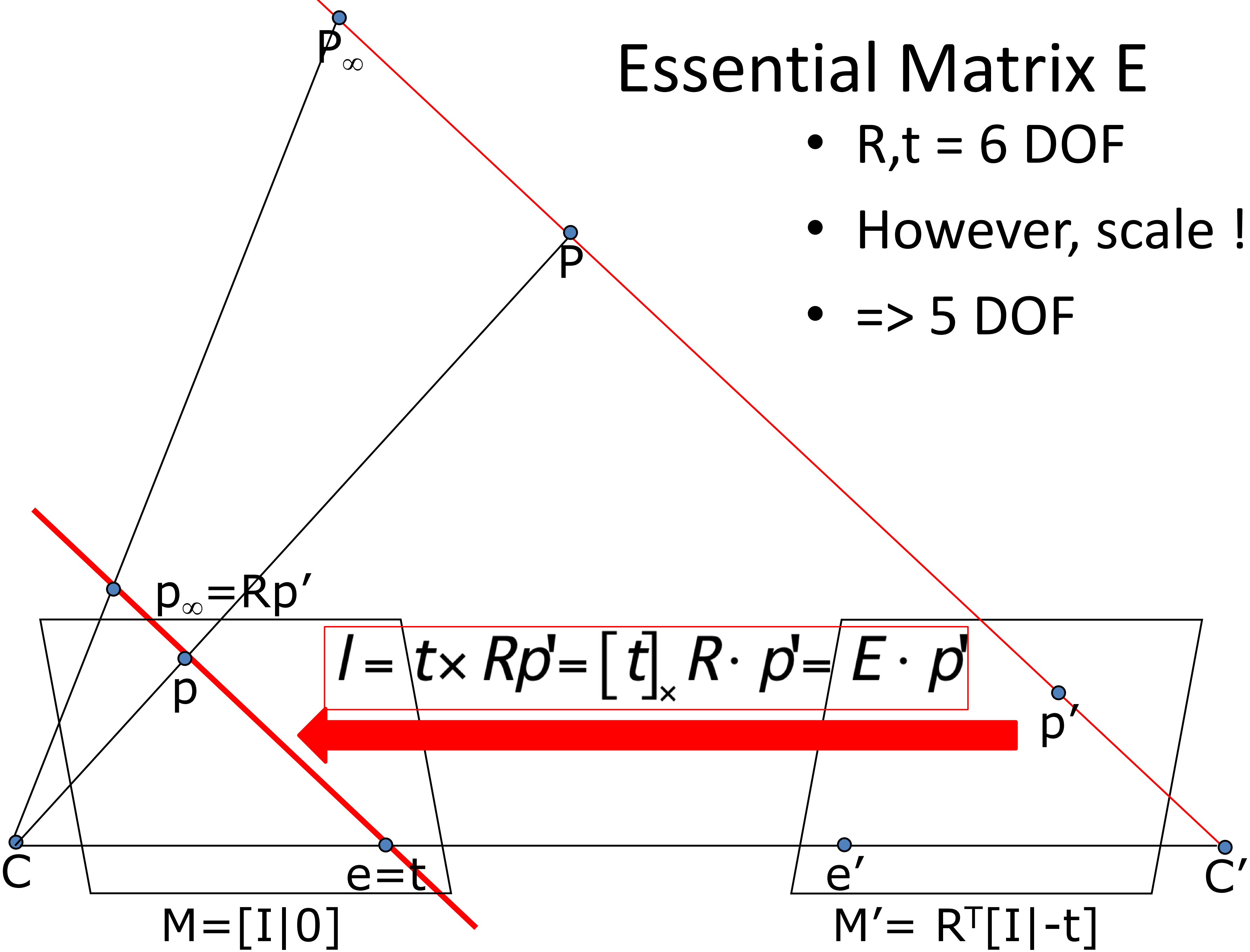
$$l = t \times Rp'$$

Epipolar Lines



Essential Matrix E

- R,t = 6 DOF
- However, scale !
- => 5 DOF



Essential Matrix

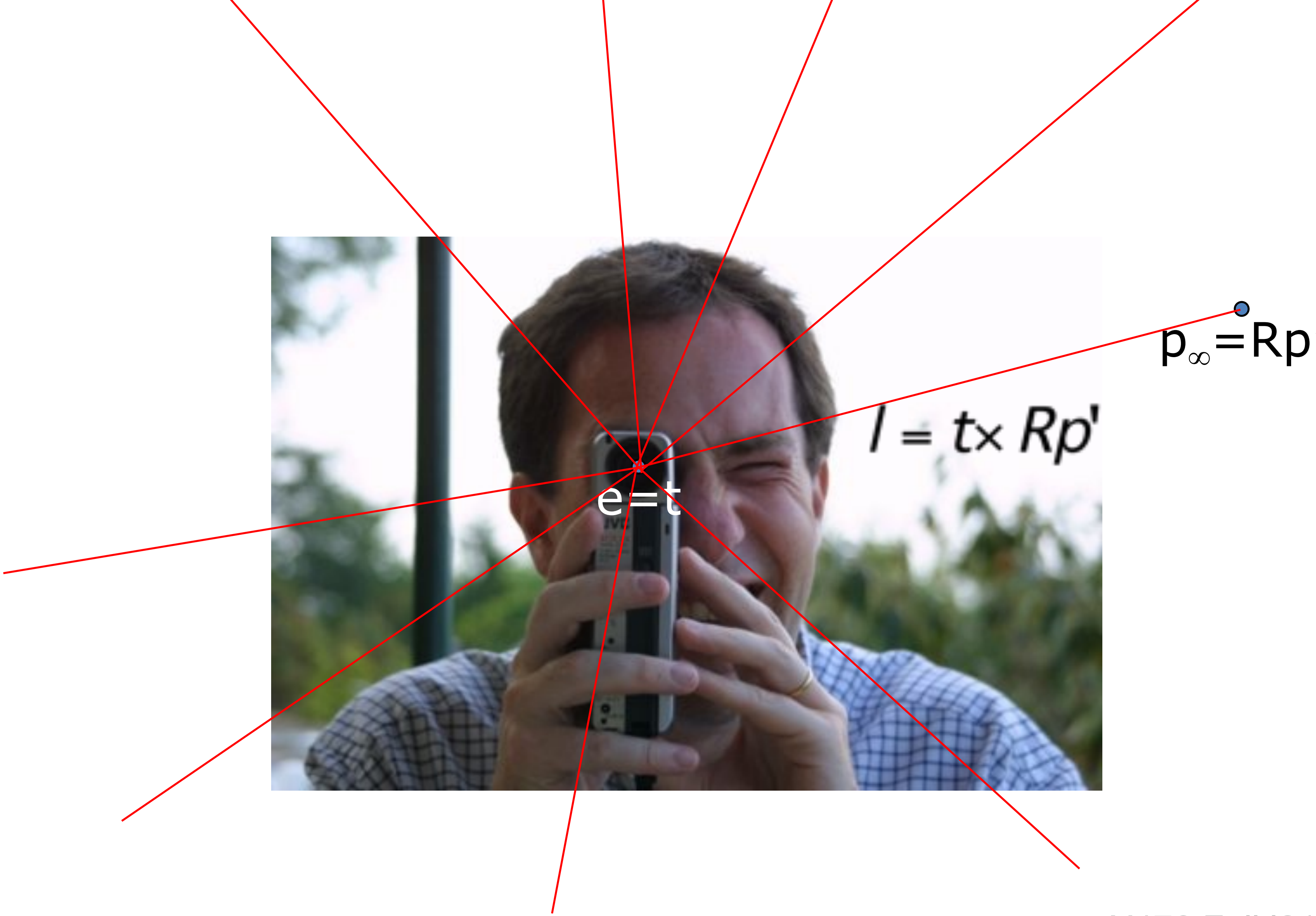
- mapping from p' to l

$$l = t \times Rp' = [t]_{\times} R \cdot p' = E \cdot p'$$

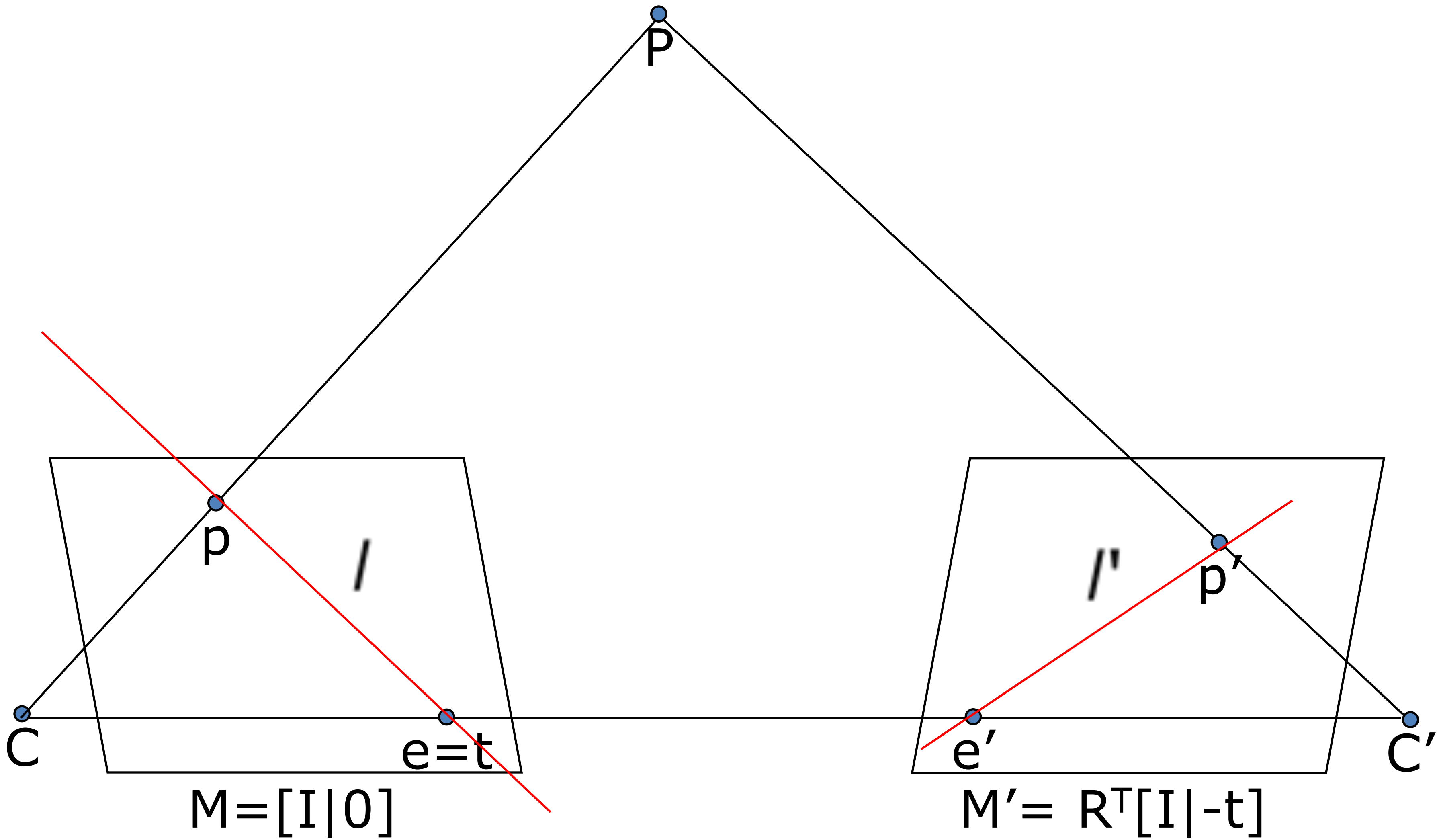
- E = 3*3 matrix
- Because p is on l, we have

$$\boxed{p^T E p' = 0}$$

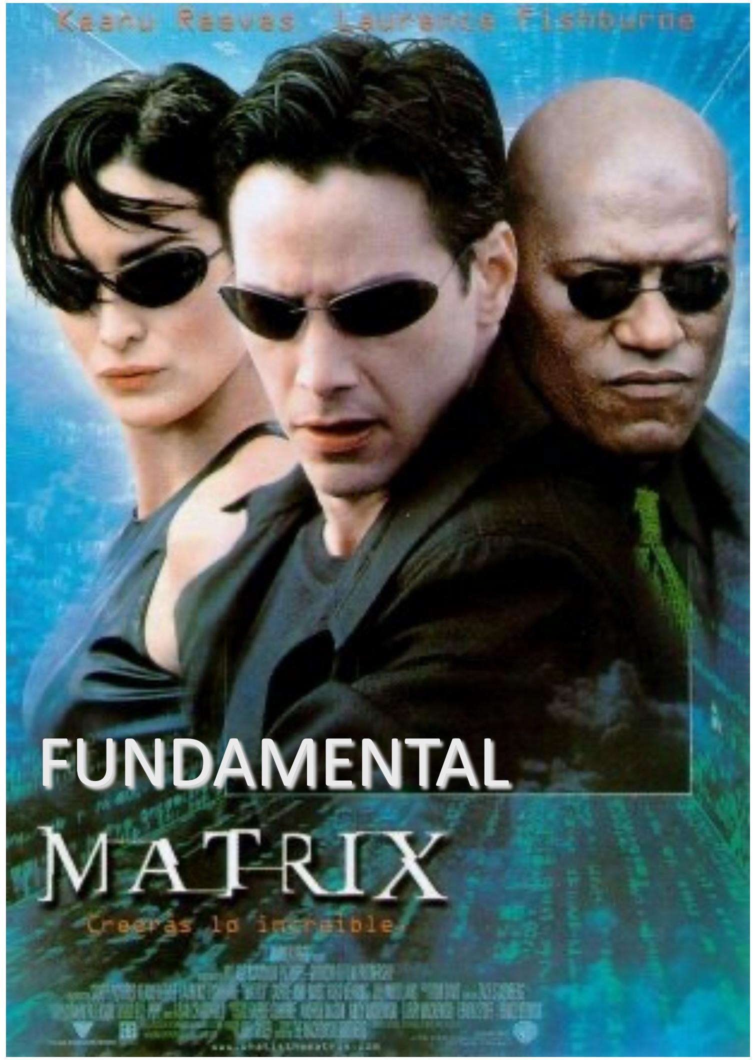
Epipolar lines



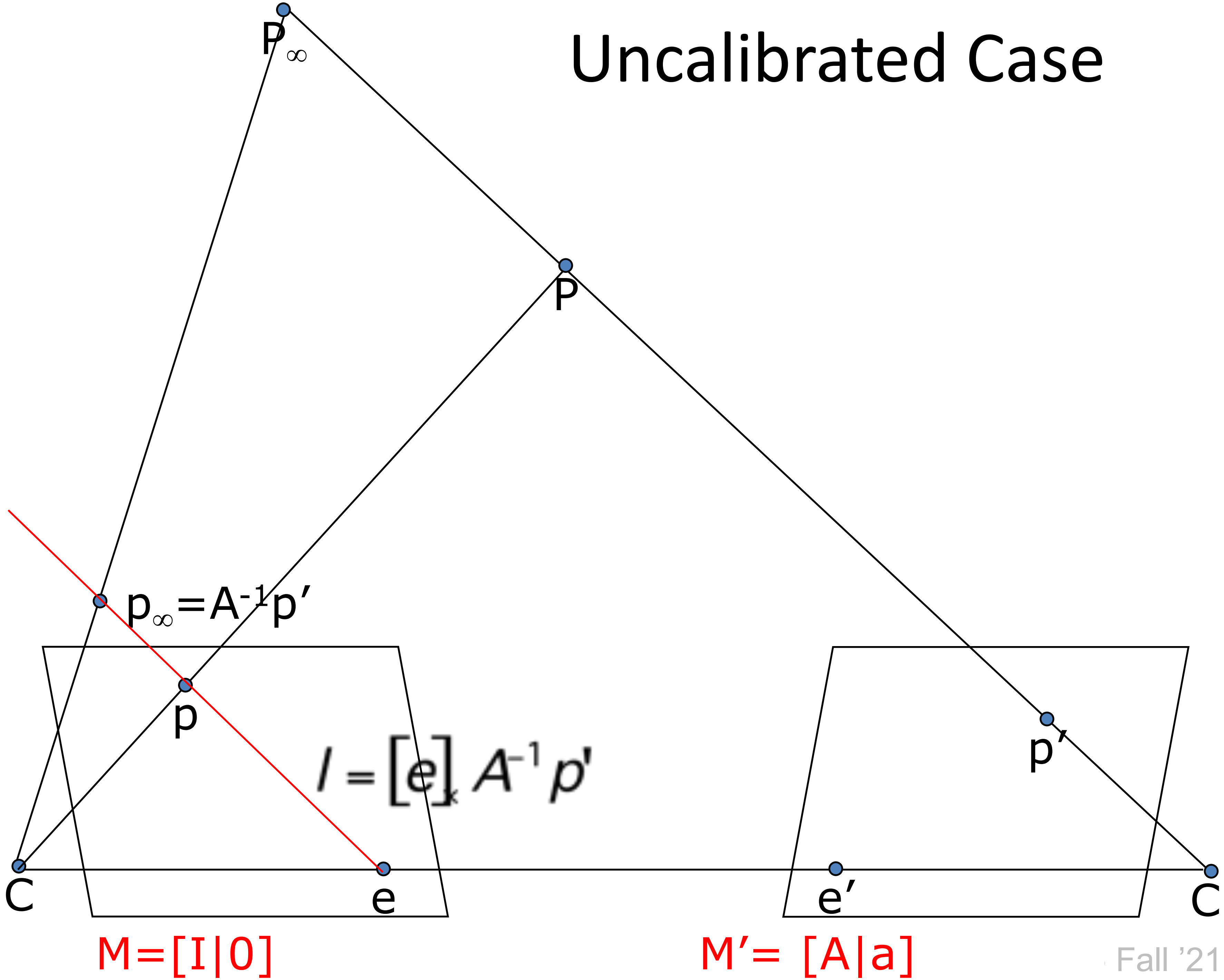
Epipolar Plane



Fundamental Matrix

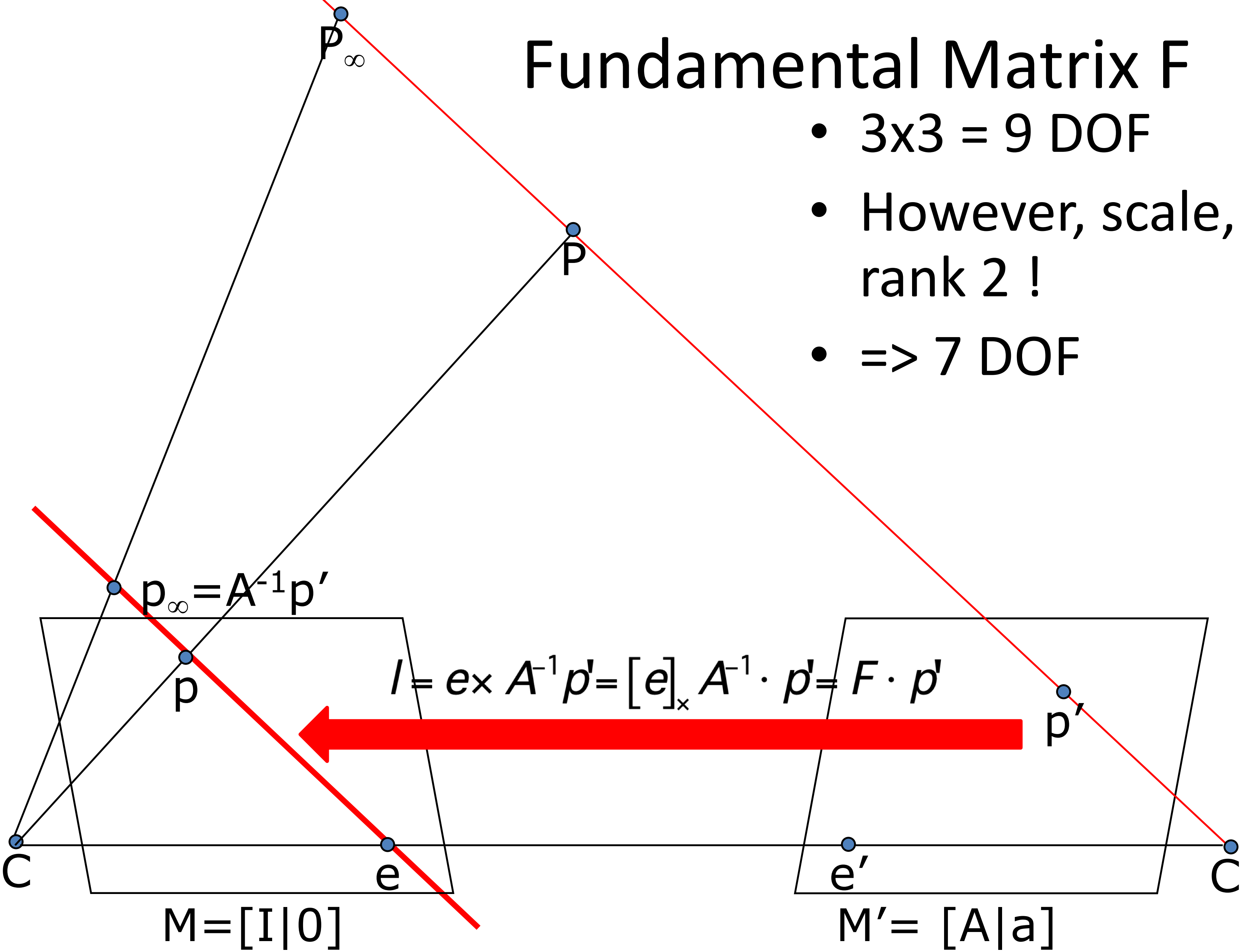


Uncalibrated Case



Fundamental Matrix F

- 3x3 = 9 DOF
- However, scale, rank 2 !
- => 7 DOF



Fundamental Matrix

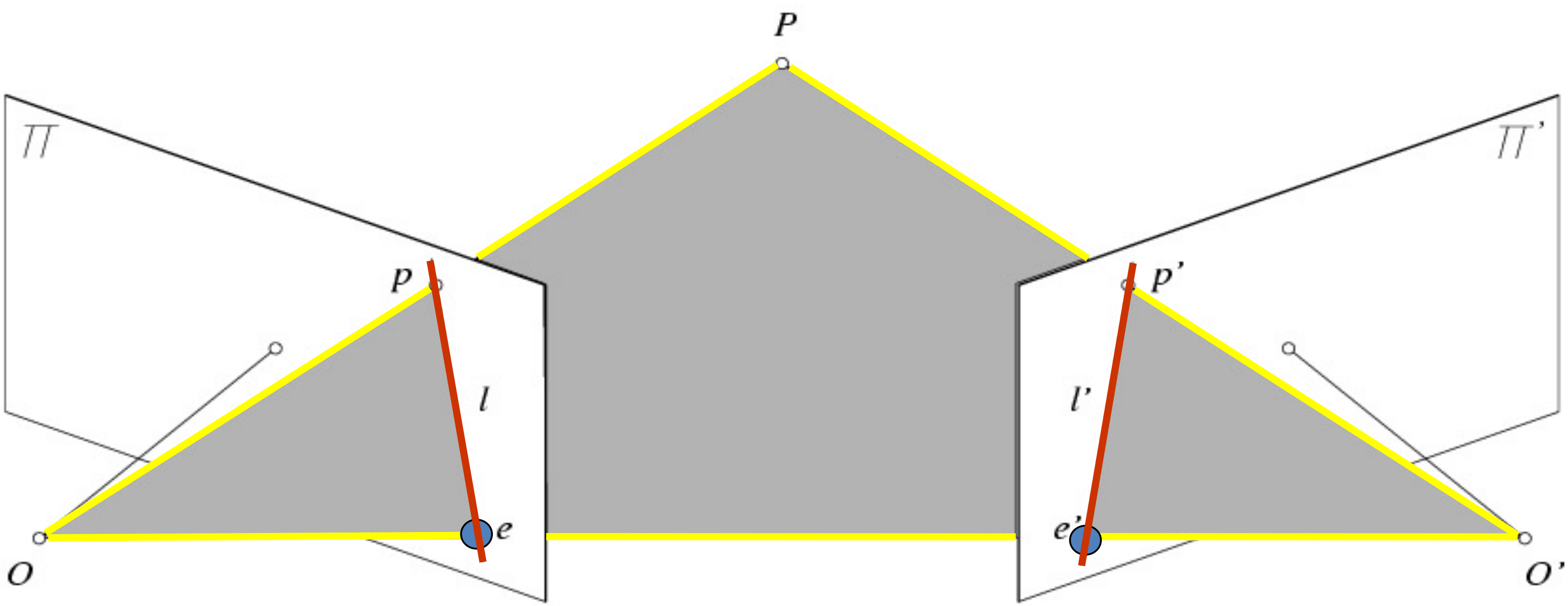
- mapping from p’ to l

$$l = e \times A^{-1} p' = [e]_{\times} A^{-1} \cdot p' = F \cdot p'$$

- F = 3*3 matrix
- Because p is on l, we have

$$\boxed{p^T F p' = 0}$$

Uncalibrated Case, Forsyth & Ponce Version



$$\hat{\mathbf{p}}^T \mathcal{E} \hat{\mathbf{p}}' = 0$$

$$\mathbf{p} = \mathcal{K} \hat{\mathbf{p}}$$

$$\mathbf{p}' = \mathcal{K}' \hat{\mathbf{p}}'$$

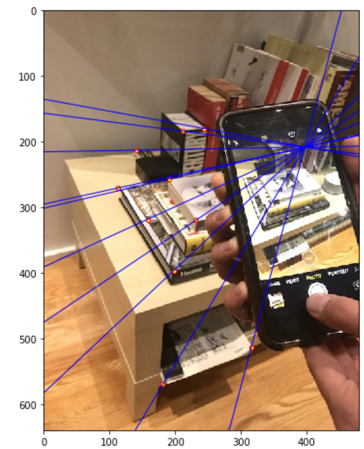
$$\mathbf{p}^T \mathcal{F} \mathbf{p}' = 0 \quad \text{with} \quad \mathcal{F} = \mathcal{K}^{-T} \mathcal{E} \mathcal{K}'^{-1}$$

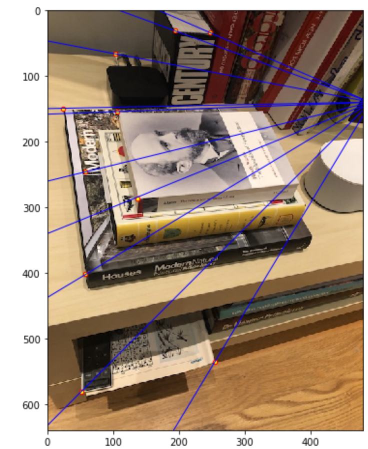
Fundamental Matrix
(Faugeras and Luong, 1992)

Properties of the Fundamental Matrix

- $\mathcal{F}p'$ is the epipolar line associated with p'.
- $\mathcal{F}^{\mathcal{T}}p$ is the epipolar line associated with p.
- $\mathcal{F}^{\mathcal{T}}e=0$ and $\mathcal{F}e'=0$.
- $\mathcal{F}$ is singular.

Non-Linear Least-Squares Approach (Luong et al., 1993)





Minimize

$$\sum_{i=1}^{n}[\mathrm{d}^2(\boldsymbol{p}_i, \mathcal{F}\boldsymbol{p}'_i) + \mathrm{d}^2(\boldsymbol{p}'_i, \mathcal{F}^T\boldsymbol{p}_i)]$$

with respect to the coefficients of $\mathcal{F}$, using an appropriate rank-2 parameterization.

d(p, l) = point to line distance
= $(ax + by + cw)/\mathrm{sqrt}(a^2 + b^2)$

$d^2(p, l) = |ax + by + cw|^2/(a^2 + b^2)$

The Eight-Point Algorithm (Longuet-Higgins, 1981)

$$(u, v, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad (uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

$$\begin{pmatrix}
u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' \\
u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' \\
u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' \\
u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4' & v_4' \\
u_5u_5' & u_5v_5' & u_5 & v_5u_5' & v_5v_5' & v_5 & u_5' & v_5' \\
u_6u_6' & u_6v_6' & u_6 & v_6u_6' & v_6v_6' & v_6 & u_6' & v_6' \\
u_7u_7' & u_7v_7' & u_7 & v_7u_7' & v_7v_7' & v_7 & u_7' & v_7' \\
u_8u_8' & u_8v_8' & u_8 & v_8u_8' & v_8v_8' & v_8 & u_8' & v_8'
\end{pmatrix}
\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix}
= - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Minimize:

$$\sum_{i=1}^{n} (\boldsymbol{p}_i^T \mathcal{F} \boldsymbol{p}_i')^2$$

under the constraint

$|\mathcal{F}|^2 = 1.$

The Normalized Eight-Point Algorithm (Hartley, 1995)

- Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels:

$$q_i = T\, p_i \qquad q_i' = T'\, p_i'.$$

- Use the eight-point algorithm to compute $\mathcal{F}$ from the points q_i and q'_i.

- Enforce the rank-2 constraint.

- Output $T^{-1} \mathcal{F}\, T'$.