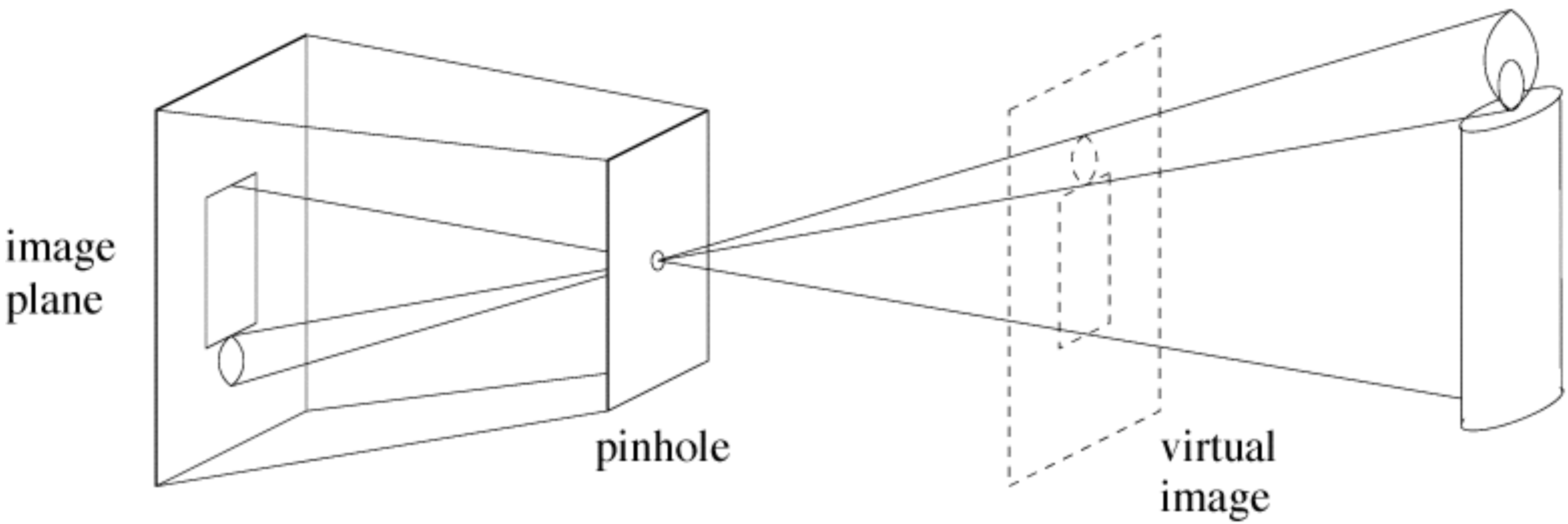
CAMERAS & VISUAL ODOMETRY

CS 3630 Introduction to Robotics and Perception

Frank Dellaert

Pinhole cameras

- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice



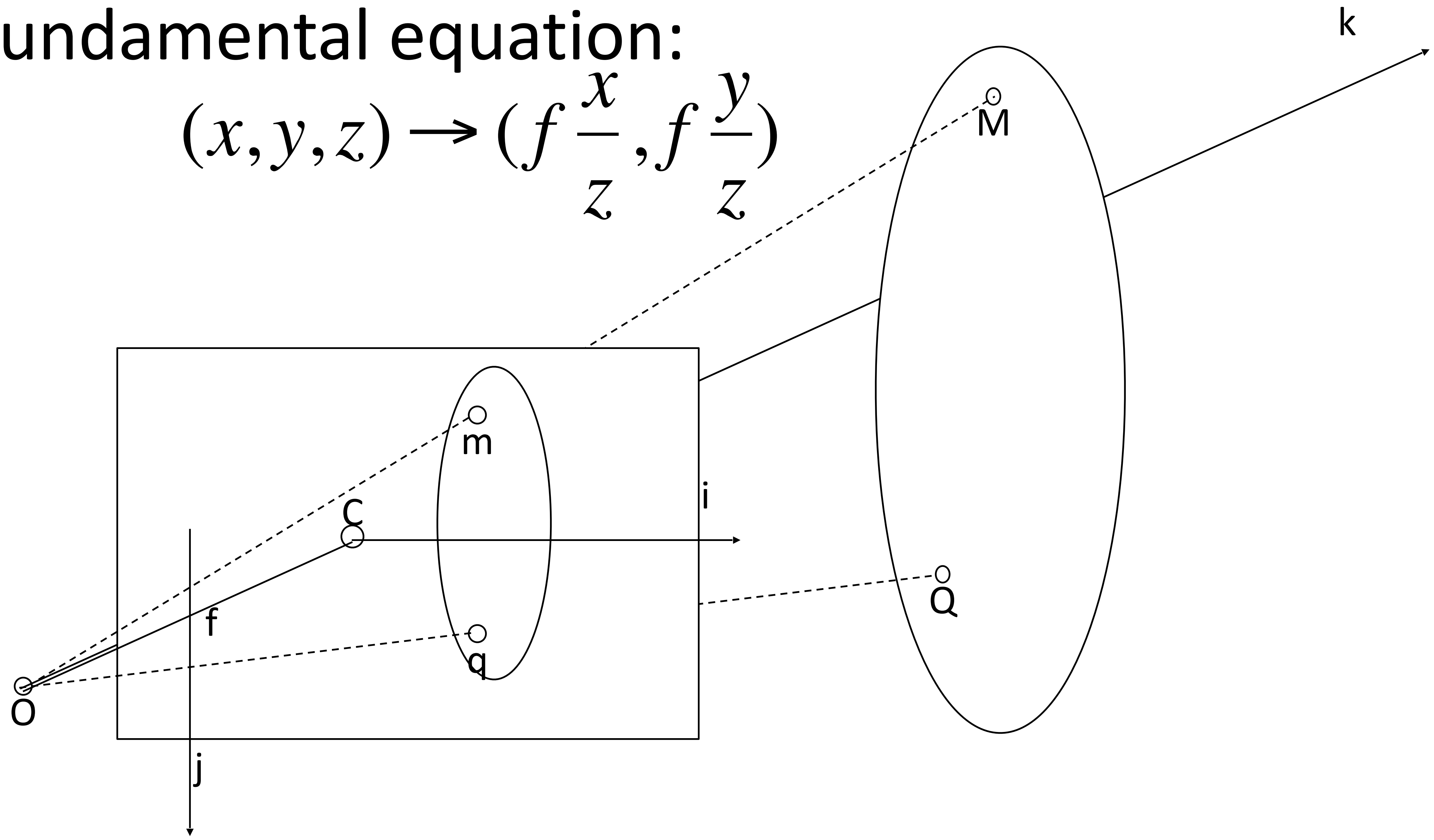
Barcelona CS 4475

Barcelona CS 4475

Pinhole Camera

- Fundamental equation:

$$(x, y, z) \rightarrow (f\frac{x}{z}, f\frac{y}{z})$$



Homogeneous Coordinates

Linear transformation of homogeneous (projective) coordinates

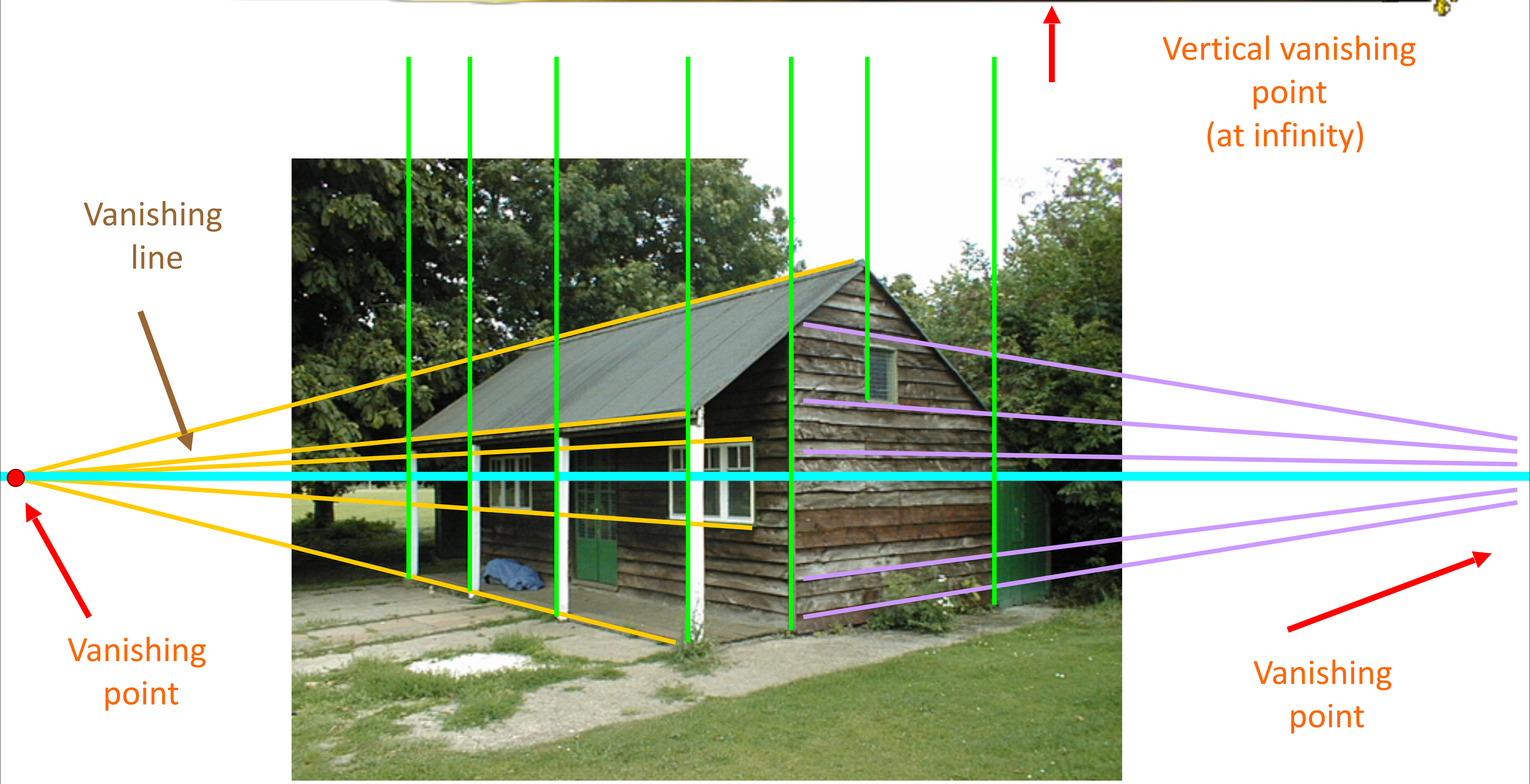
$$m = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = [I \quad 0] M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

Recover image (Euclidean) coordinates by normalizing :

$$\hat{u} = \frac{u}{w} = \frac{X}{Z}$$

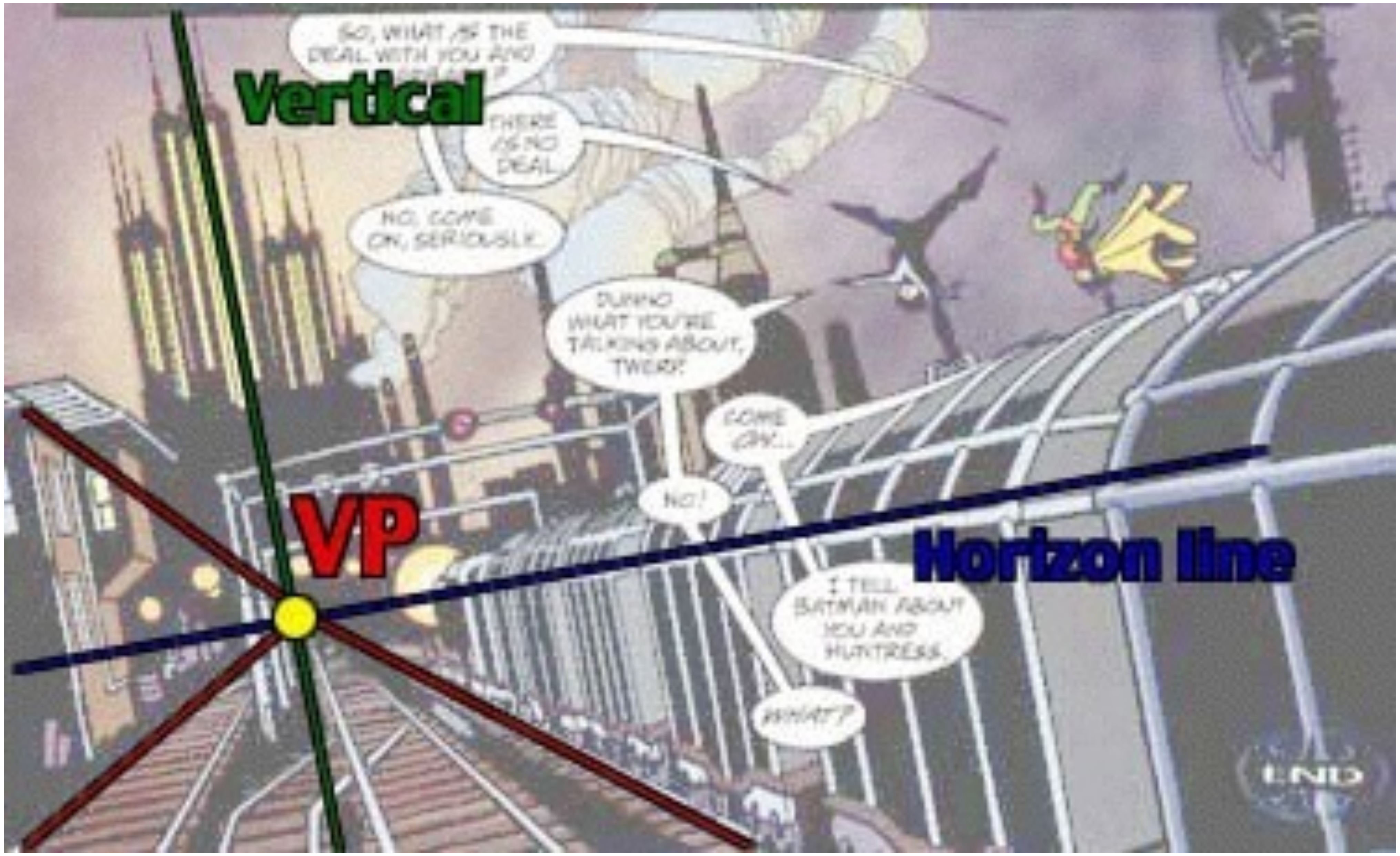
$$\hat{v} = \frac{v}{w} = \frac{Y}{Z}$$

Photographs are Projections



Slide Courtesy Antonio Criminisi

Parallel lines meet



Intrinsic Calibration

3×3 Calibration Matrix K

$$m = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = K[I \quad 0]M = \begin{bmatrix} \alpha & s & u_0 \\ & \beta & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

skew

Recover image (Euclidean) coordinates by normalizing :

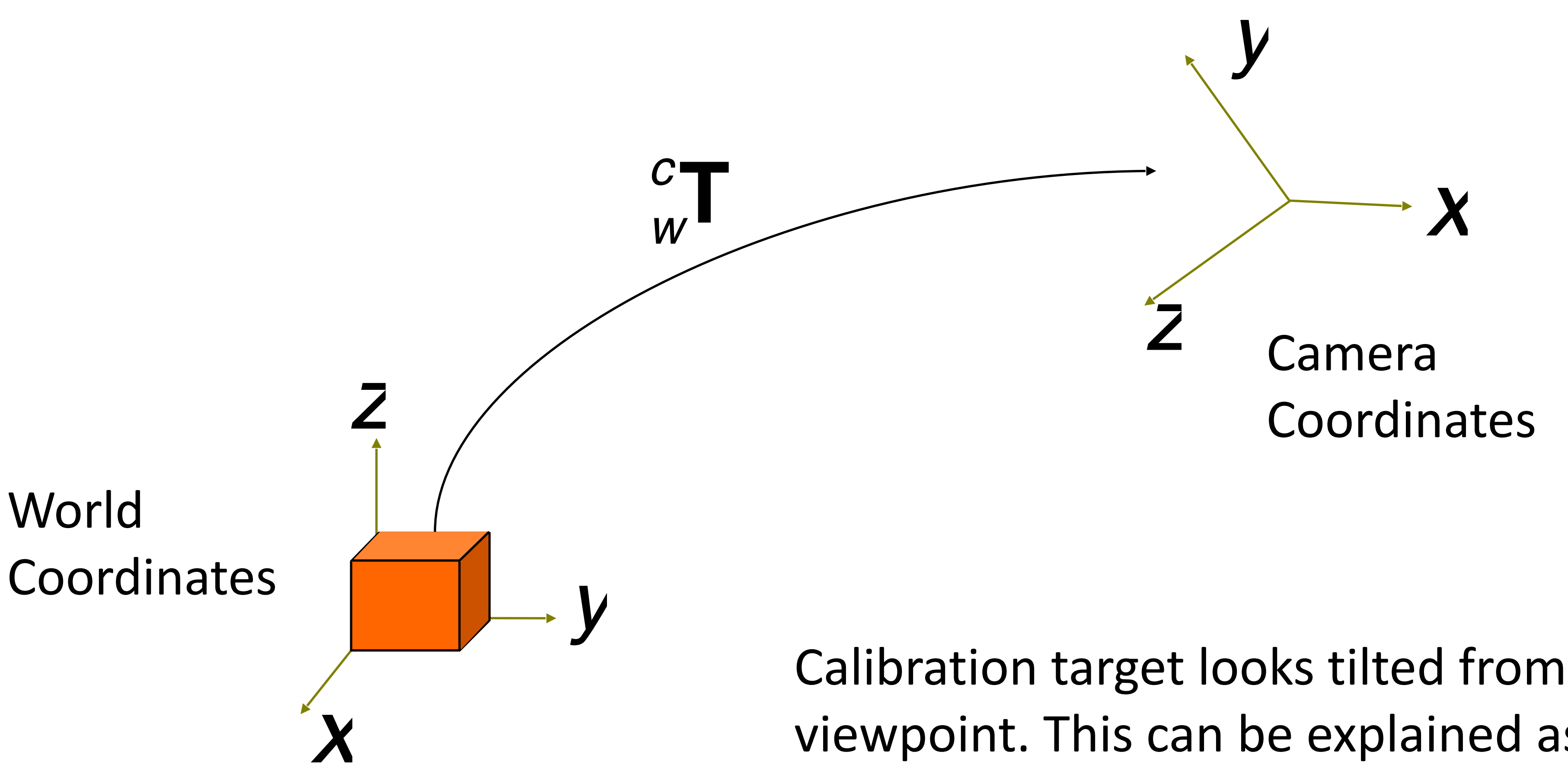
$$\hat{u} = \frac{u}{w} = \frac{\alpha X + sY + u_0}{Z}$$

$$\hat{v} = \frac{v}{w} = \frac{\beta Y + v_0}{Z}$$

5 Degrees of Freedom !

Camera Pose

In order to apply the camera model, objects in the scene must be expressed in *camera coordinates*.



Calibration target looks tilted from camera viewpoint. This can be explained as a difference in coordinate systems.

Projective Camera Matrix

$$Camera = Calibration \times \Pr ojection \times Extrinsics$$

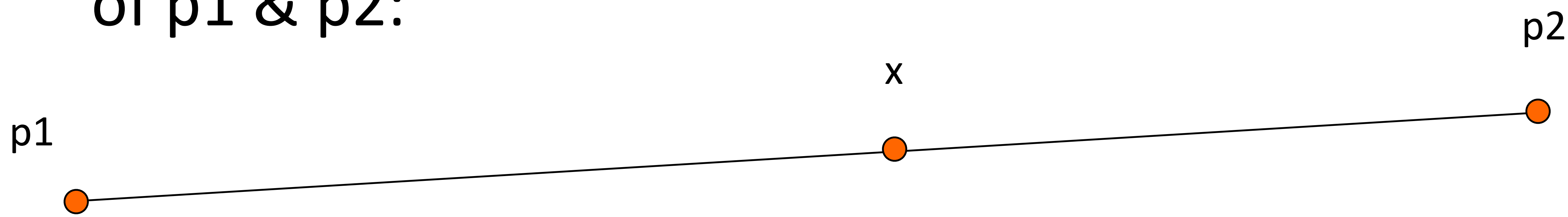
$$m = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ & \beta & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

$$= K\begin{bmatrix} R & t \end{bmatrix} M = PM$$

5+6 DOF = 11 !

Joining two Points?

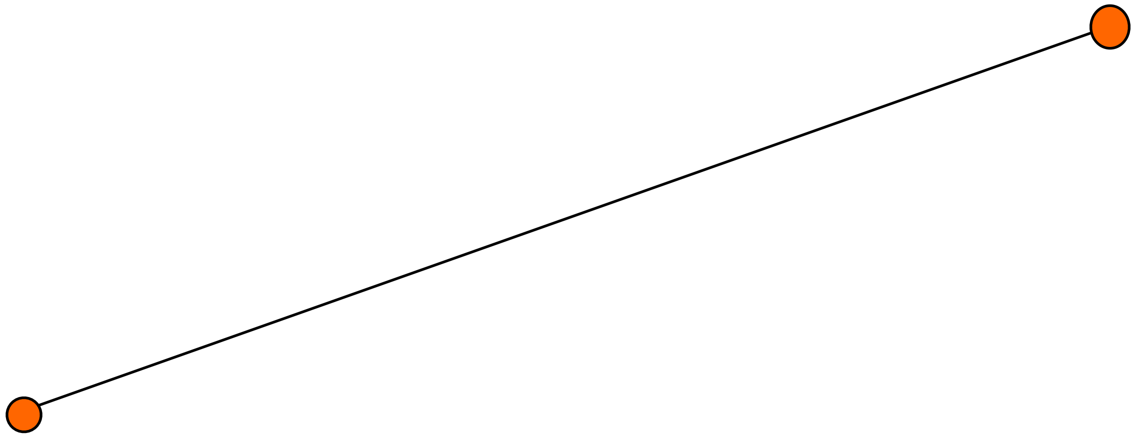
- 2D Line between two 2D points
  - Point x on l must be linear combination of p1 & p2:



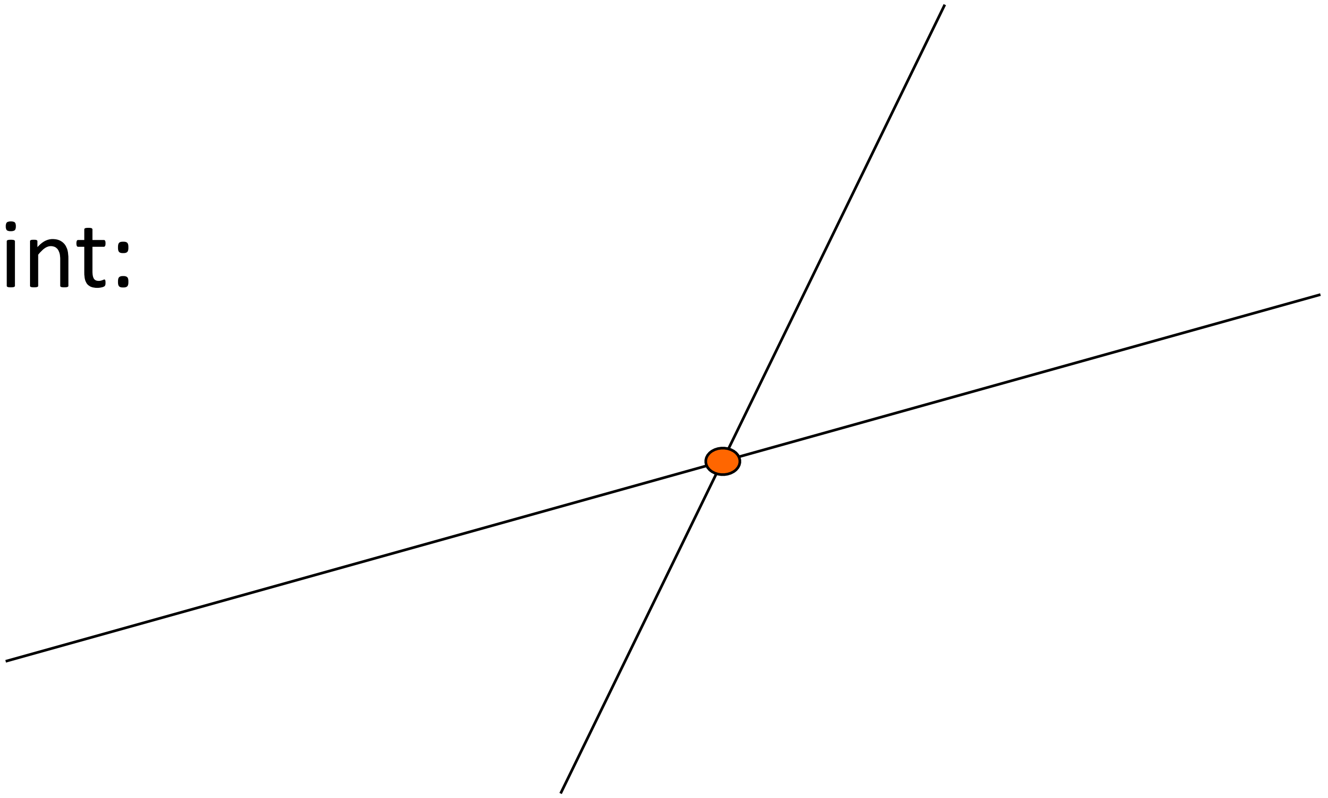
$$\begin{vmatrix} x^T \\ p_1^T \\ p_2^T \end{vmatrix} = \begin{vmatrix} x & y & w \\ x_1 & y_1 & w_1 \\ x_2 & y_2 & w_2 \end{vmatrix} = 0 \Leftrightarrow x = p_1 \times p_2$$

Join = cross product !

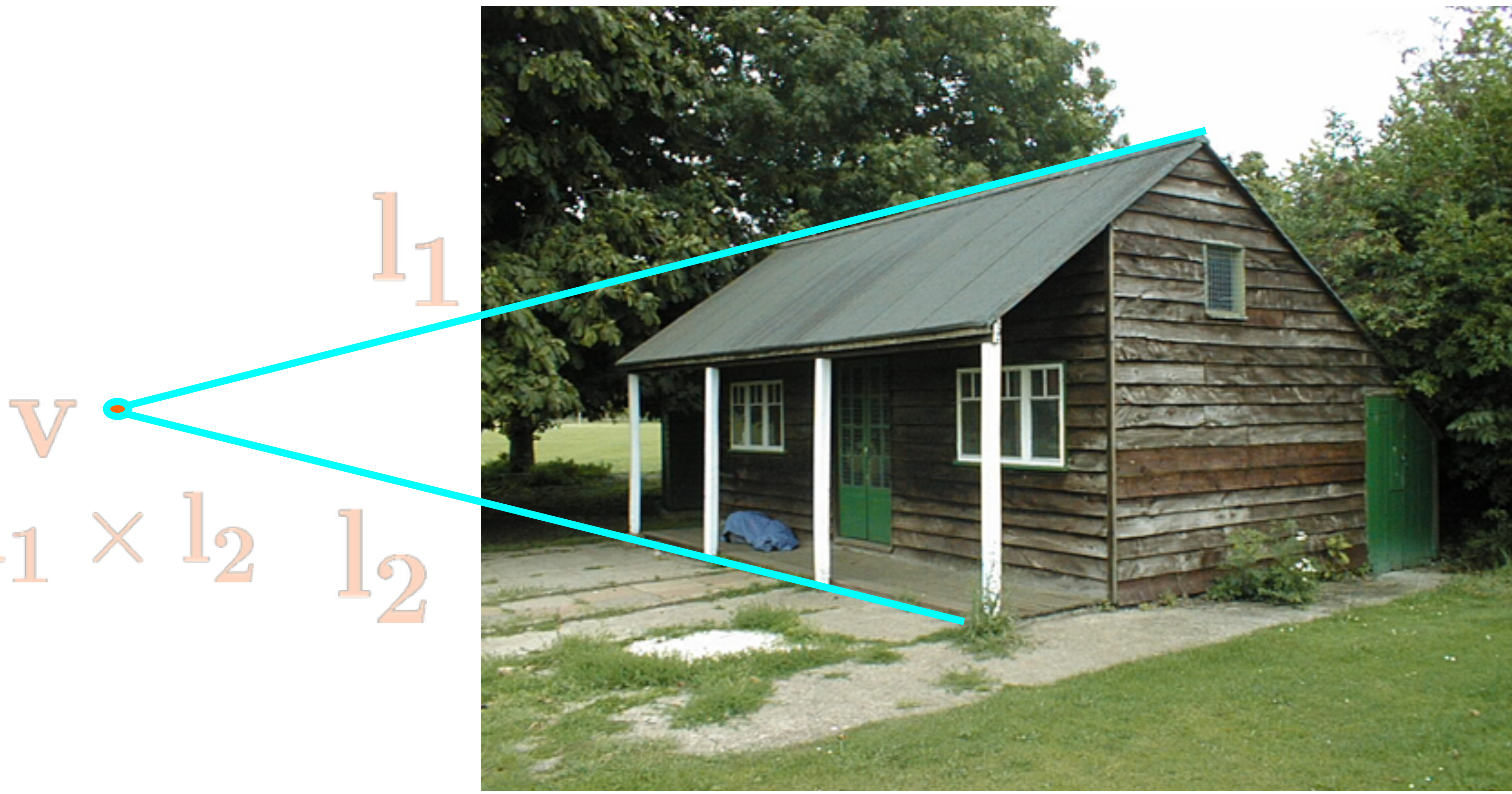
- Join of two points is a line:
  $l = p_1 x p_2$



- ALSO:
  Join of two lines is a point:
  $p = l_1 \text{ x } l_2$

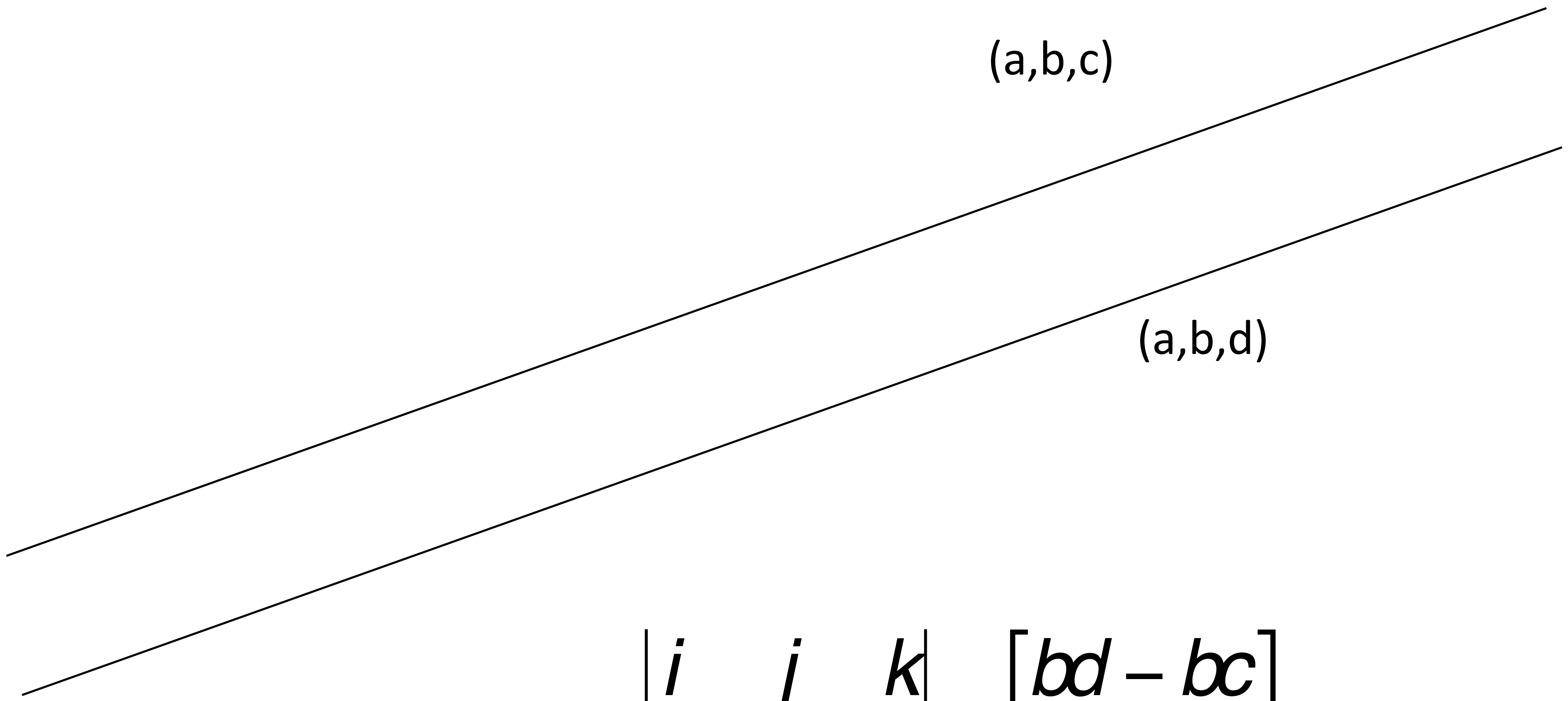


Automatic estimation of vanishing points and lines



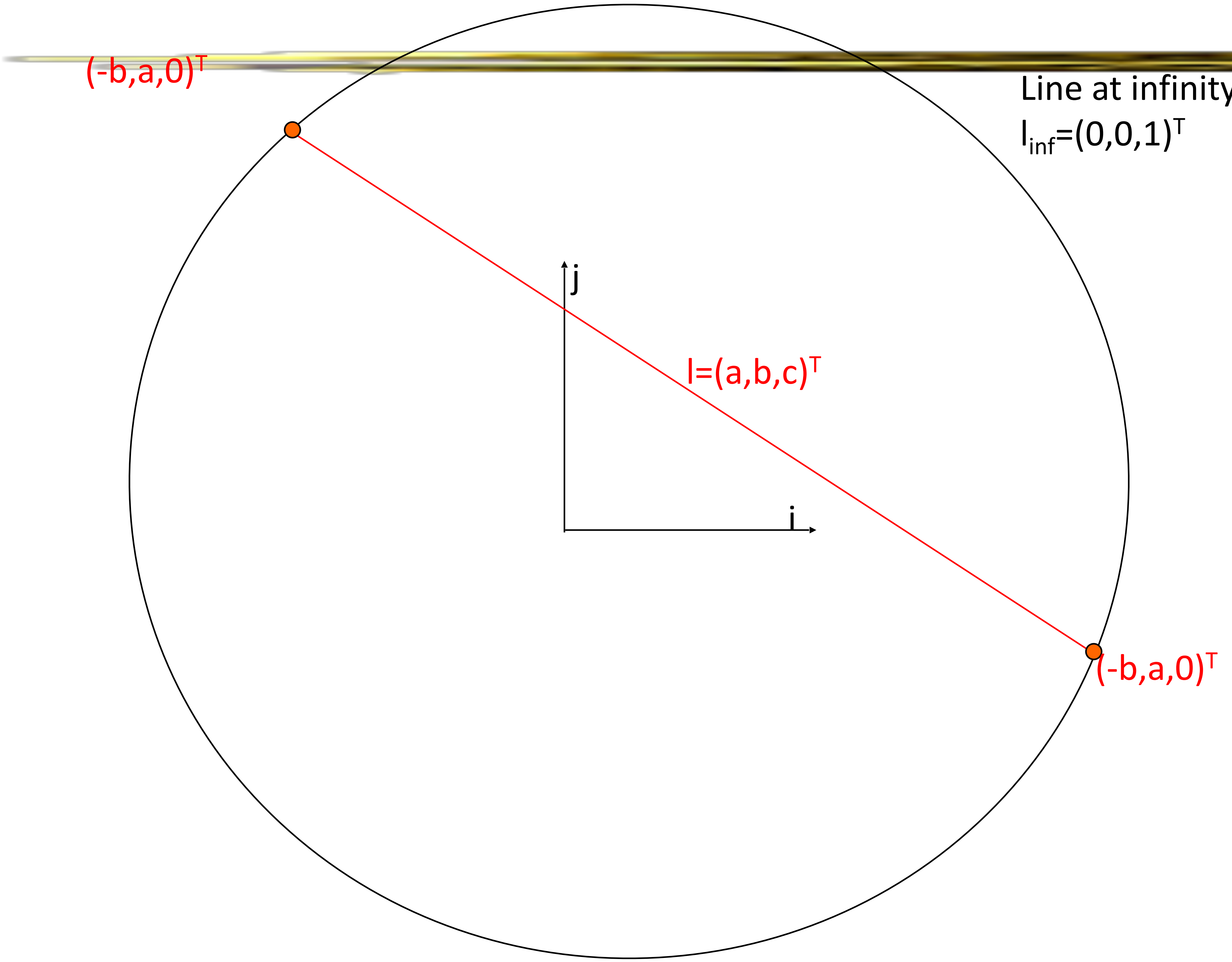
$$\mathbf{v} = \mathbf{l}_1 \times \mathbf{l}_2$$

What if we were to join two parallel lines ???



$$p = \begin{vmatrix} i & j & k \\ a & b & c \\ a & b & d \end{vmatrix} = \begin{bmatrix} bd - bc \\ ca - ad \\ 0 \end{bmatrix} = (d - c) \cdot \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}$$

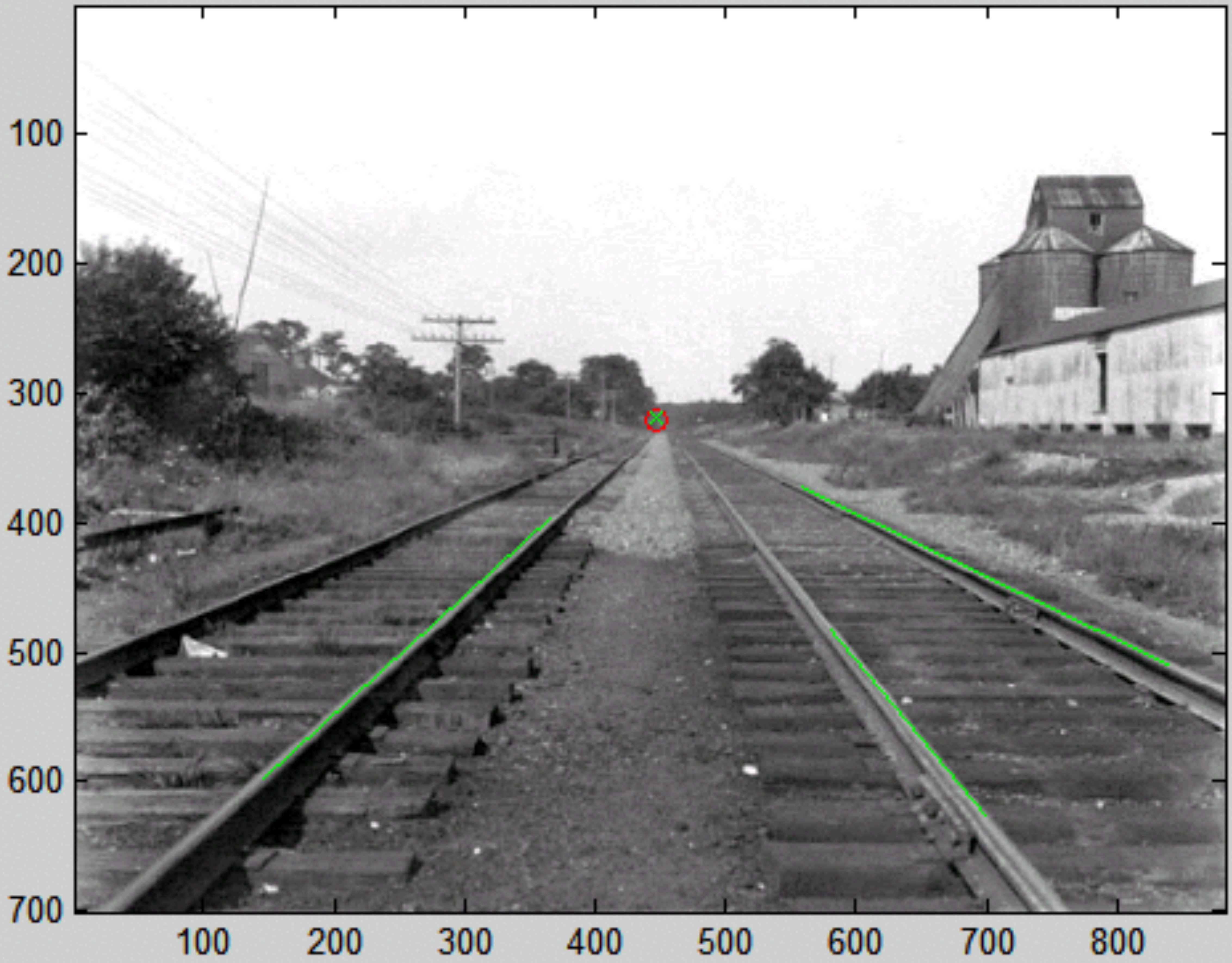
Points at Infinity !



Joining Multiple Lines: Homogenous Equations

- we want $l^T p=0$ for all lines:

$$Ap = \begin{bmatrix} l_1^T \\ l_2^T \\ l_3^T \\ l_4^T \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$



cameras03.m

Solving Homogeneous Equations

- Nice explanation:

$$Ap = \begin{bmatrix} l_1^T \\ l_2^T \\ l_3^T \\ l_4^T \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left|Ap\right|^2 = \left|USV^T p\right|^2$$

$$= p^T V S^T S V^T p$$

write

$$x = V^T p$$

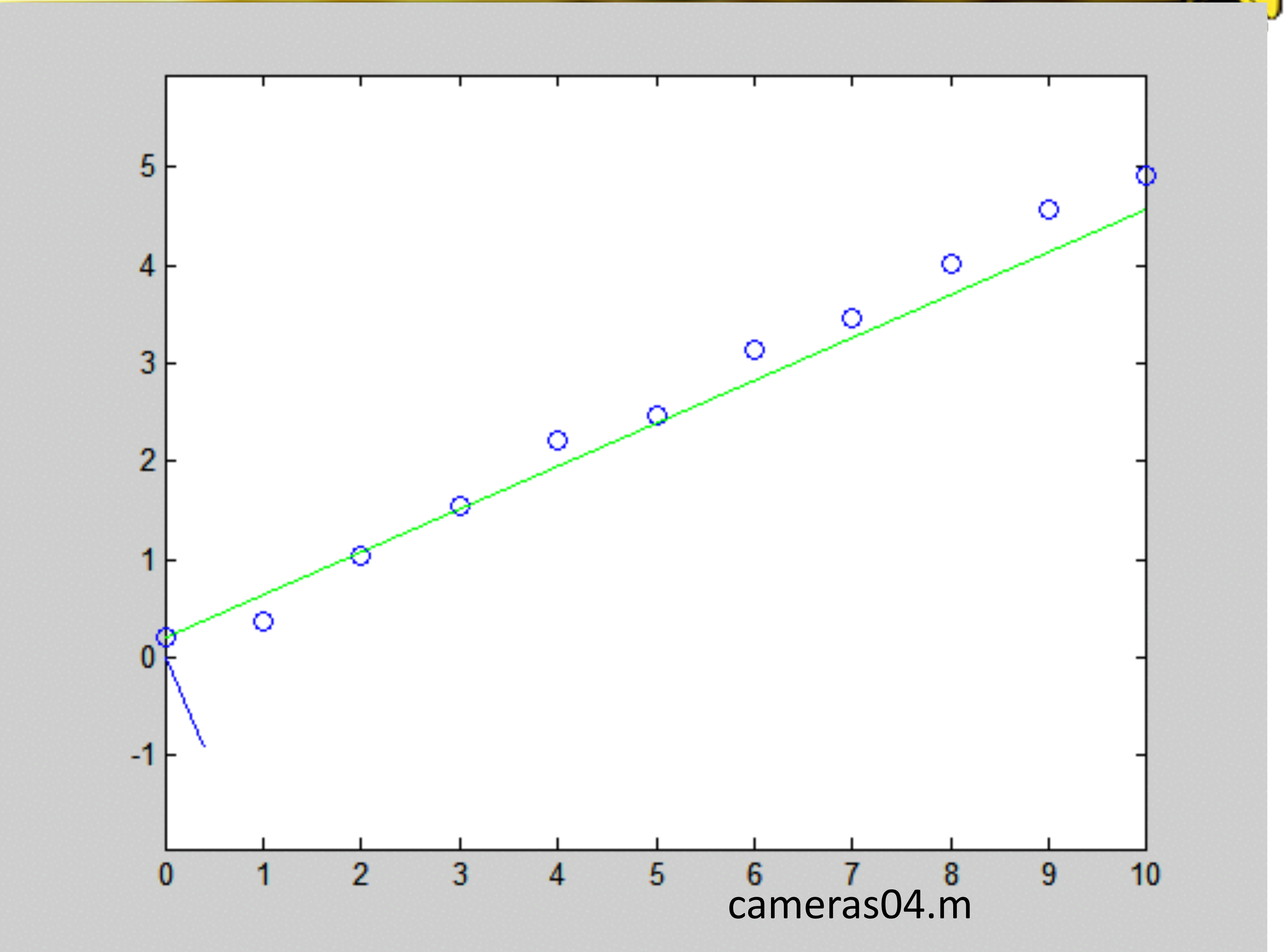
$$\arg\min \left|Ap\right|^2 \ s.t. \left|p\right| = 1$$

$$\arg\min x^T S^2 x$$

$$\Leftrightarrow x = (0,0,\ldots,1)^T$$

Applied to line fitting

$$\begin{bmatrix} x_1^T \\ x_2^T \\ x_3^T \\ x_4^T \\ x_5^T \\ x_6^T \\ x_7^T \\ x_8^T \\ x_9^T \\ x_{10}^T \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$



Caveat: Normalization matters !

Projective Camera Matrix

$$Camera = Calibration \times \Pr ojection \times Extrinsics$$

$$m = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ & \beta & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

$$= K \begin{bmatrix} R & t \end{bmatrix} M = PM$$

5+6 DOF = 11 !

Projective Camera Matrix

$$m = K\begin{bmatrix} R & t \end{bmatrix} M = PM$$

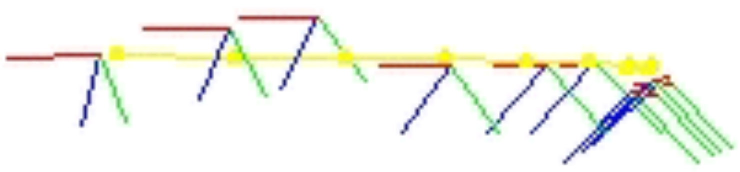
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

5+6 DOF = 11 !

Structure and Motion

Book Sequence

Book Structure & Motion

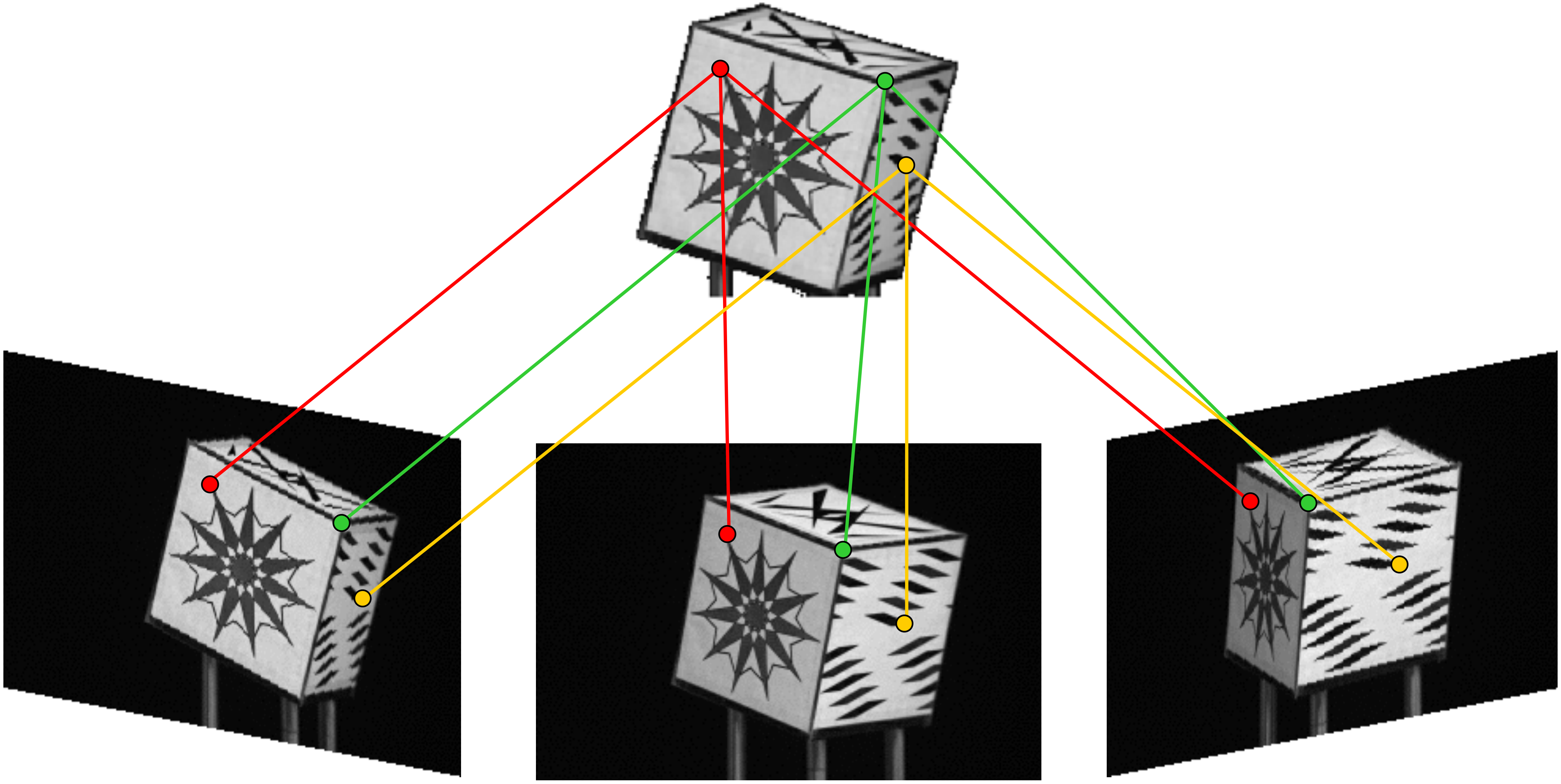


2 Problems !

Correspondence

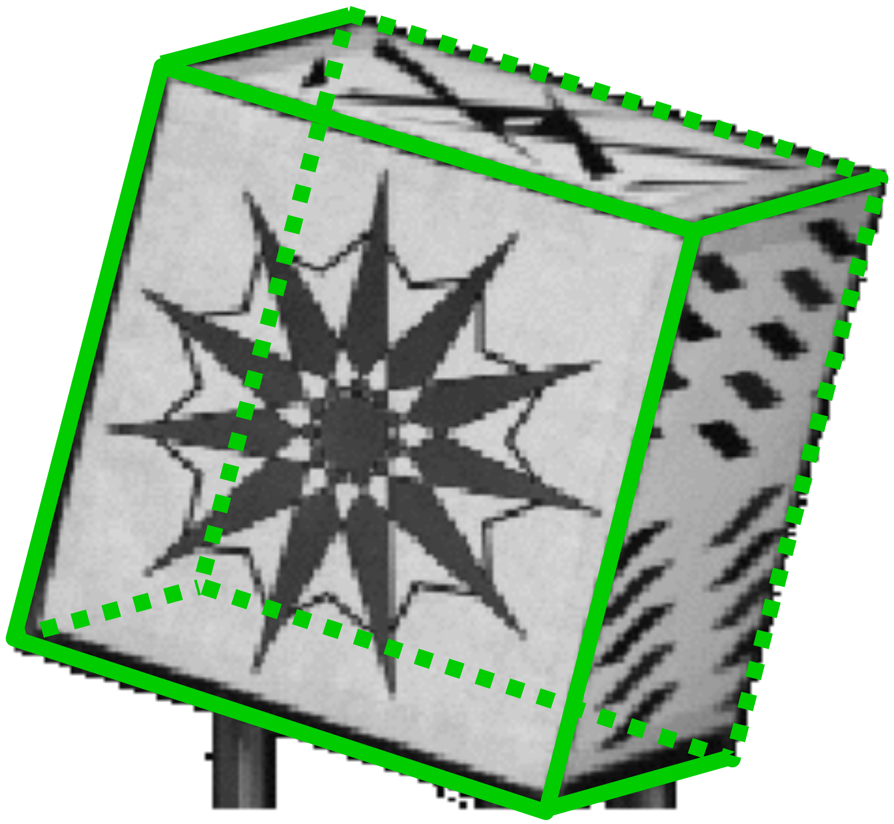
Optimization

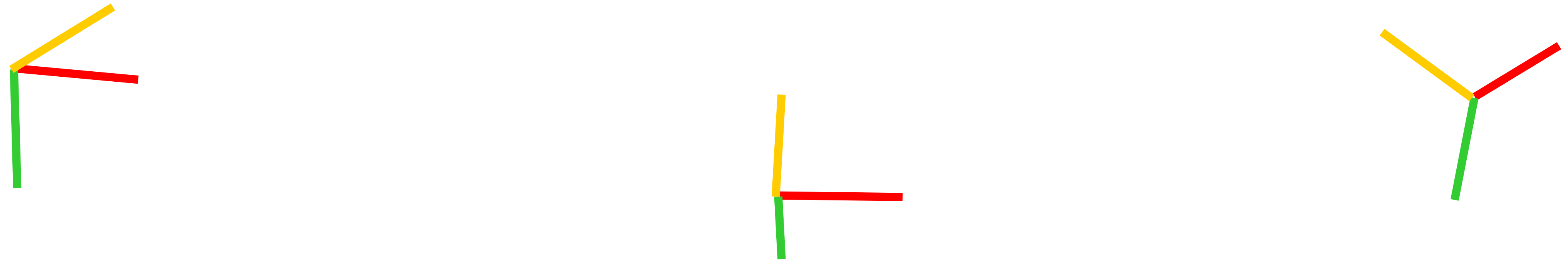
A Correspondence Problem



An Optimization Problem

- Find the most likely structure and motion Θ

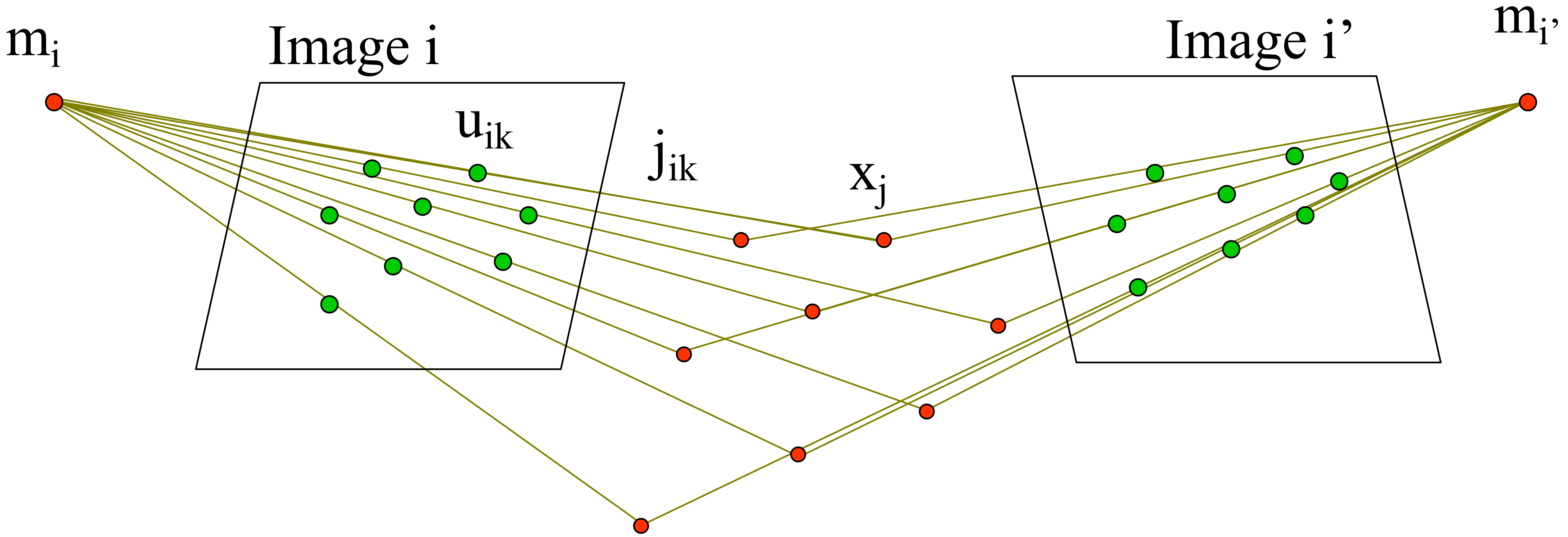




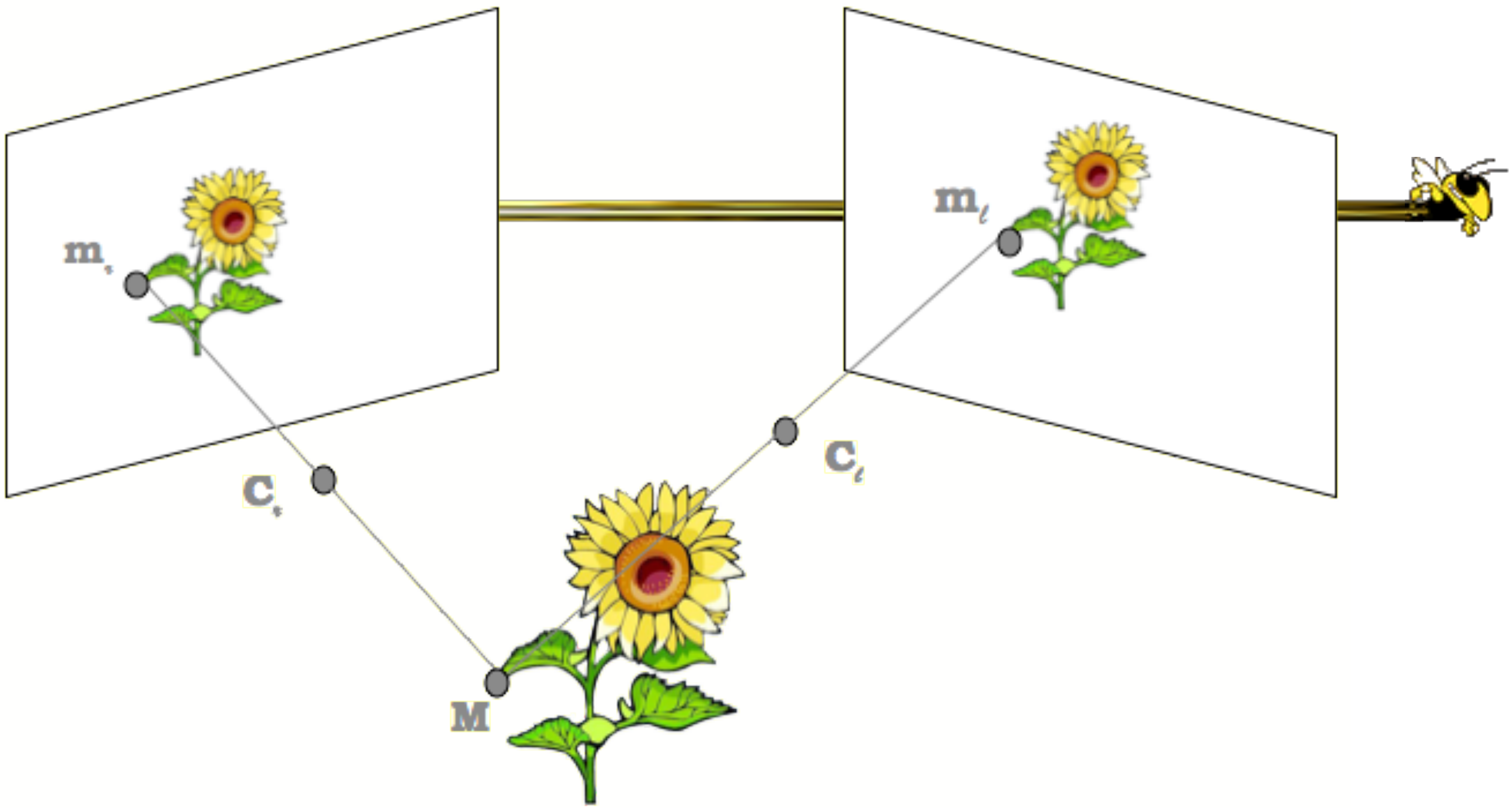
Optimization

=Non-linear Least-Squares !

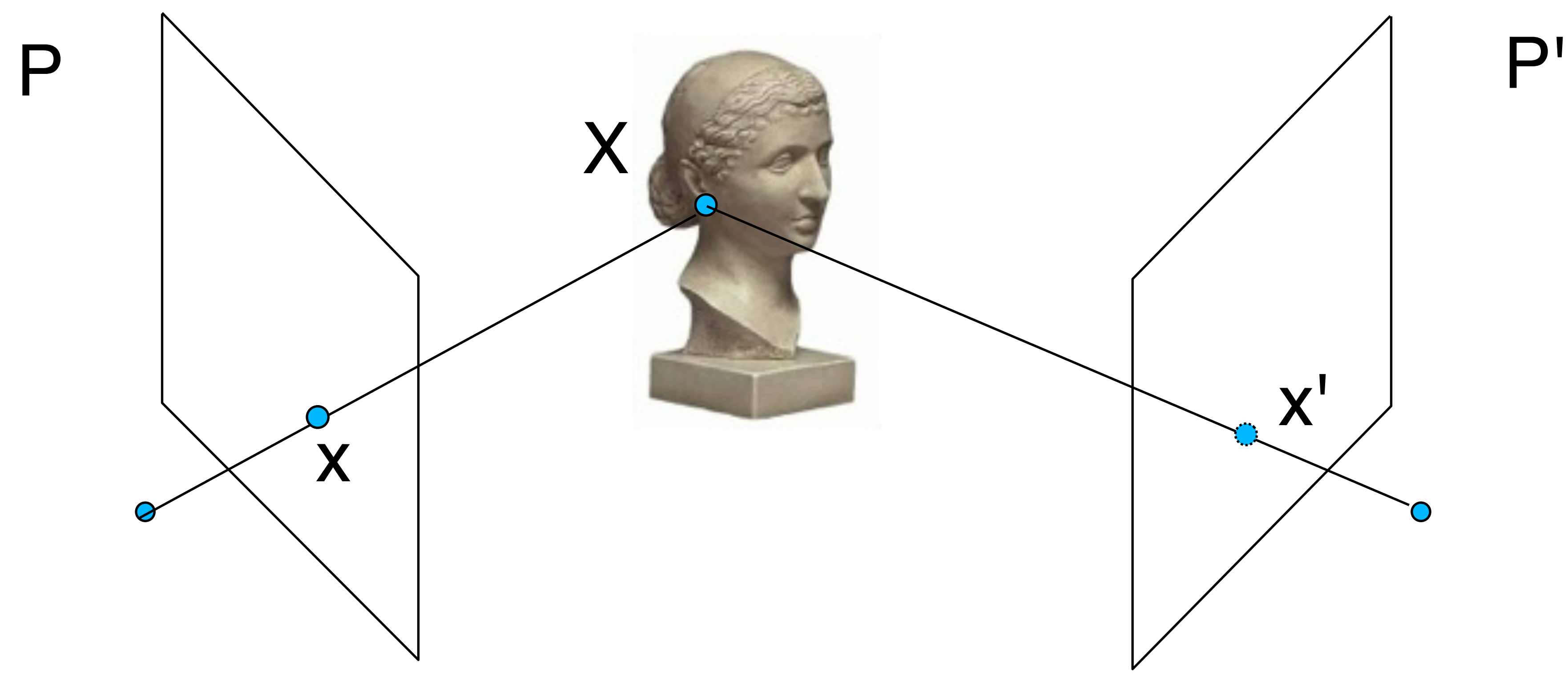
$$\sum_{i=1}^{m} \sum_{k=1}^{K_i} \|\mathbf{u}_{ik} - \mathbf{h}(\mathbf{m}_i, \mathbf{x}_{\mathbf{j}_{ik}})\|^2$$



Two View Geometry



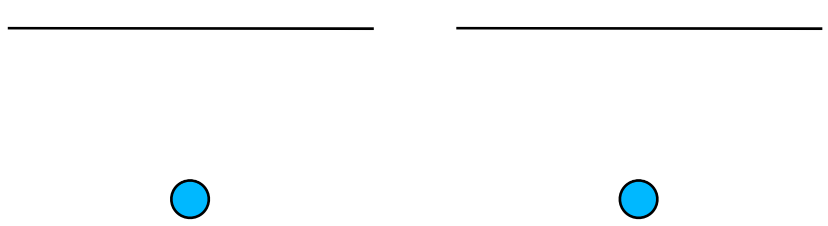
Why Consider Multiple Views?



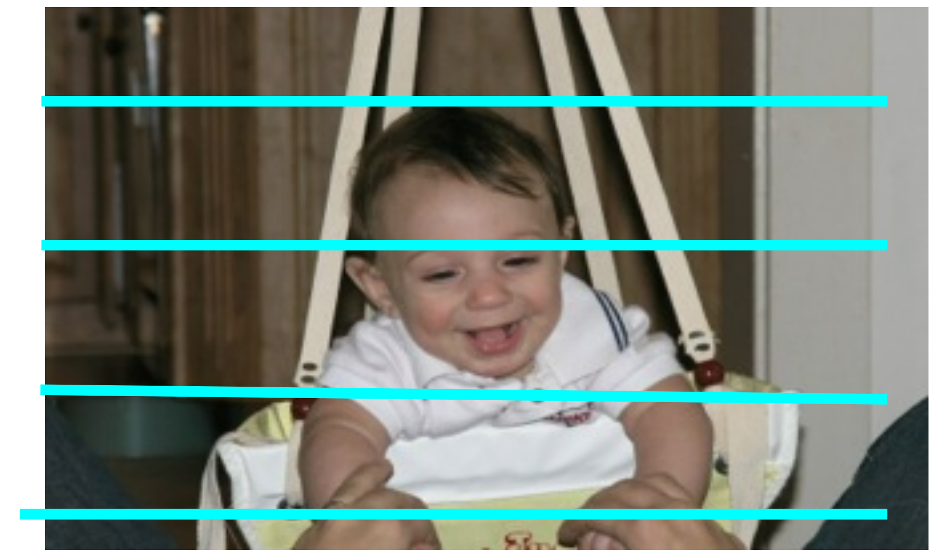
Answer: To extract 3D structure via triangulation.

Stereo Rig

Top View

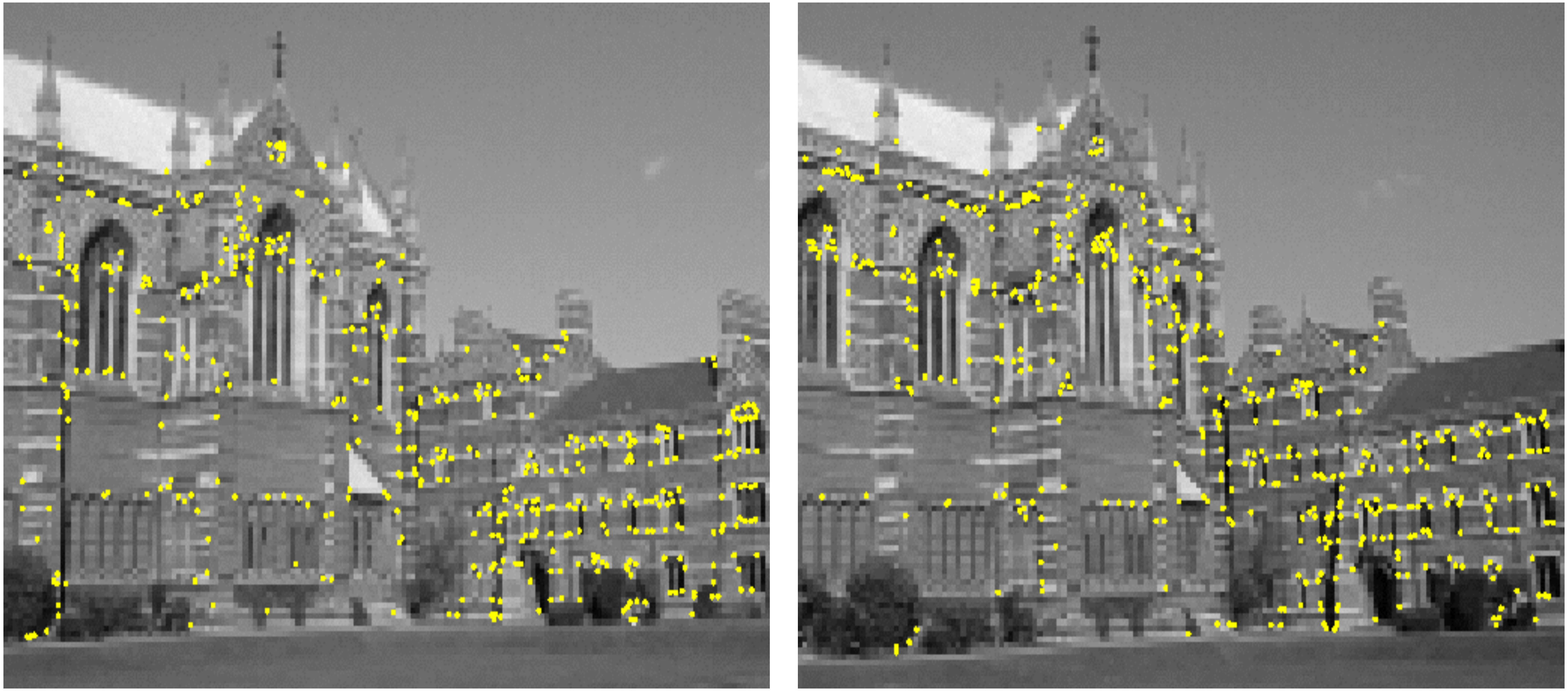


Matches on Scanlines



•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**

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

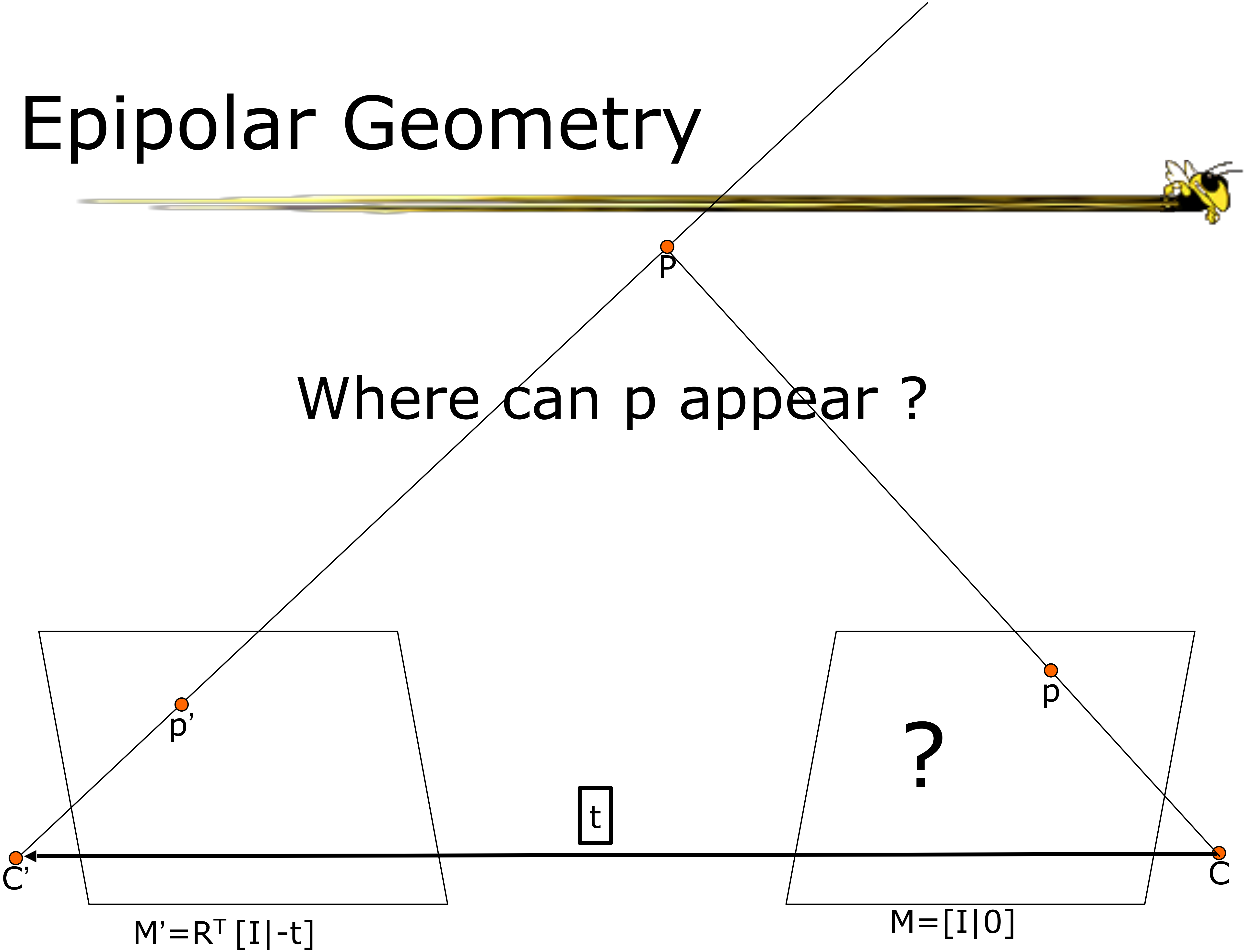
  - F can be calculated either from camera matrices or a set of good correspondences.

Geometry of 2 views ?

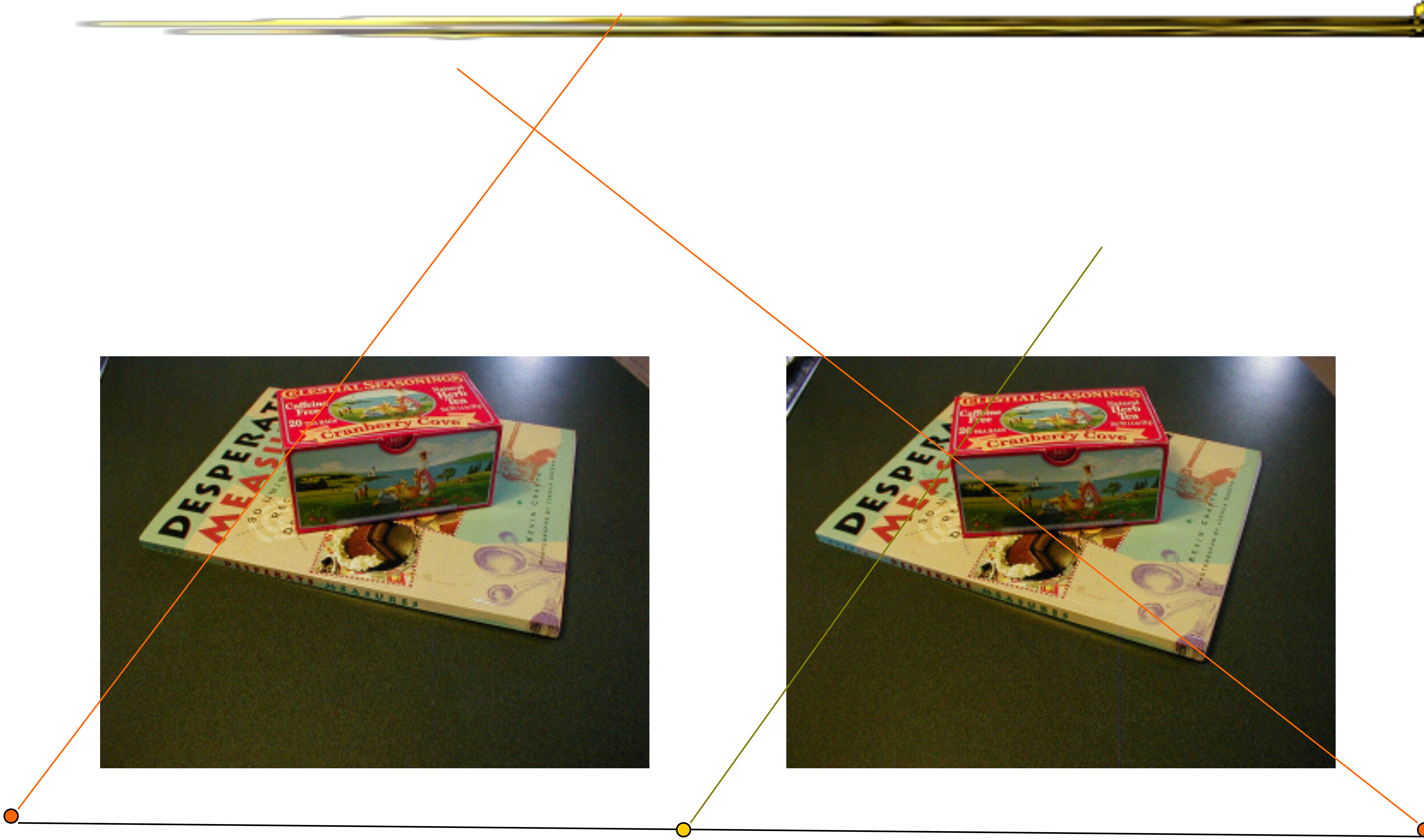
- What if we do not know R,t ?
- Caveat:
  - My exposition follows book conventions
  - but more intuitive (IMHO)

  - Different from Hartley & Zisserman !
  - F&P use R^T [I|-t] camera matrices
  - H&Z uses [R|t]

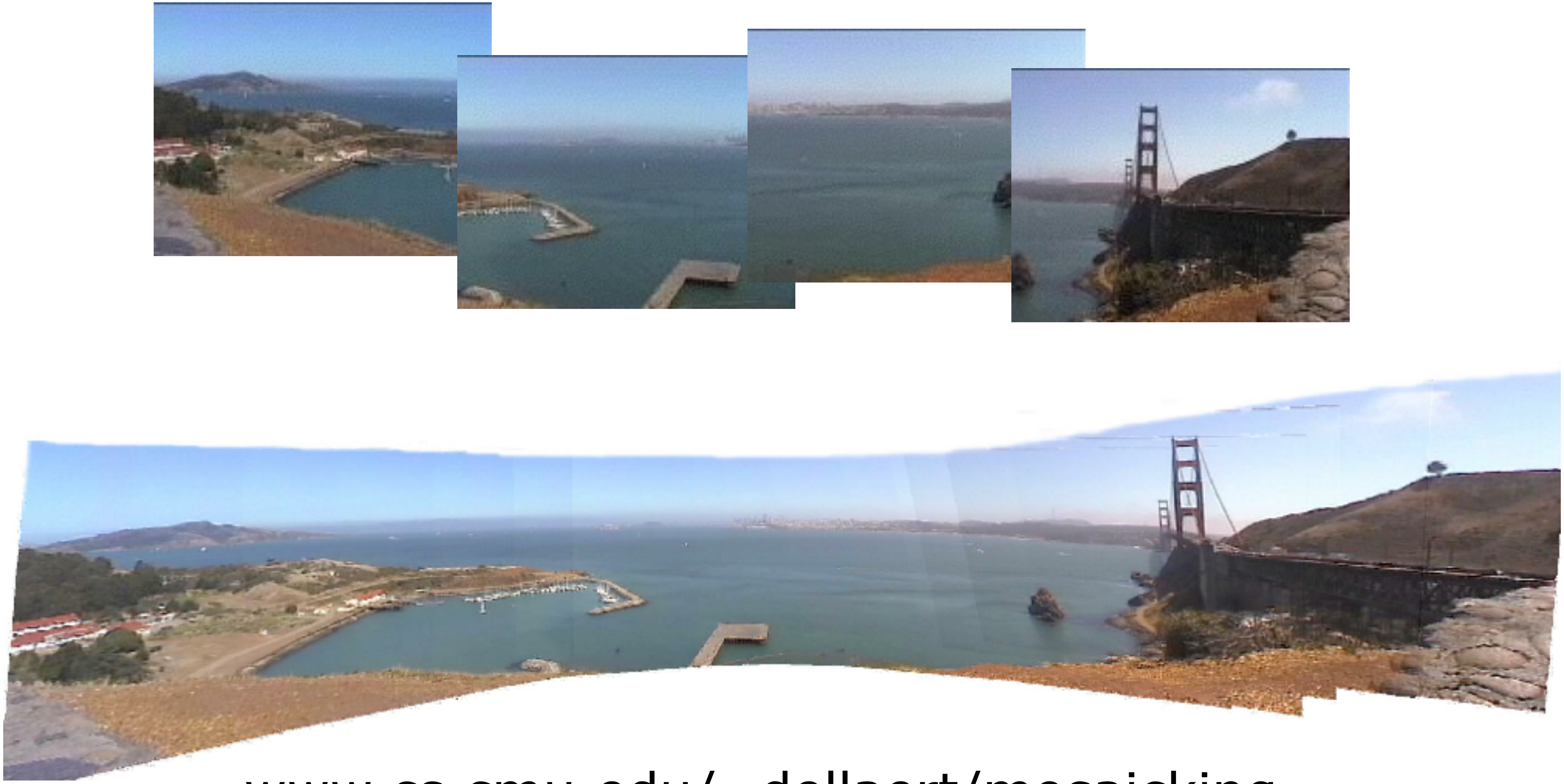
Epipolar Geometry



Fundamental Matrix



Mosaicking: Homography



www.cs.cmu.edu/~dellaert/mosaicking

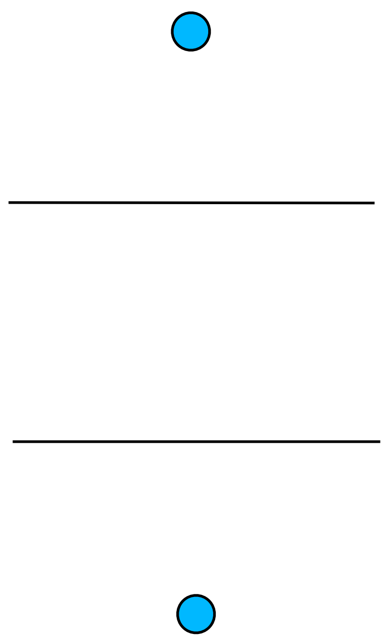
Image of Camera Center

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

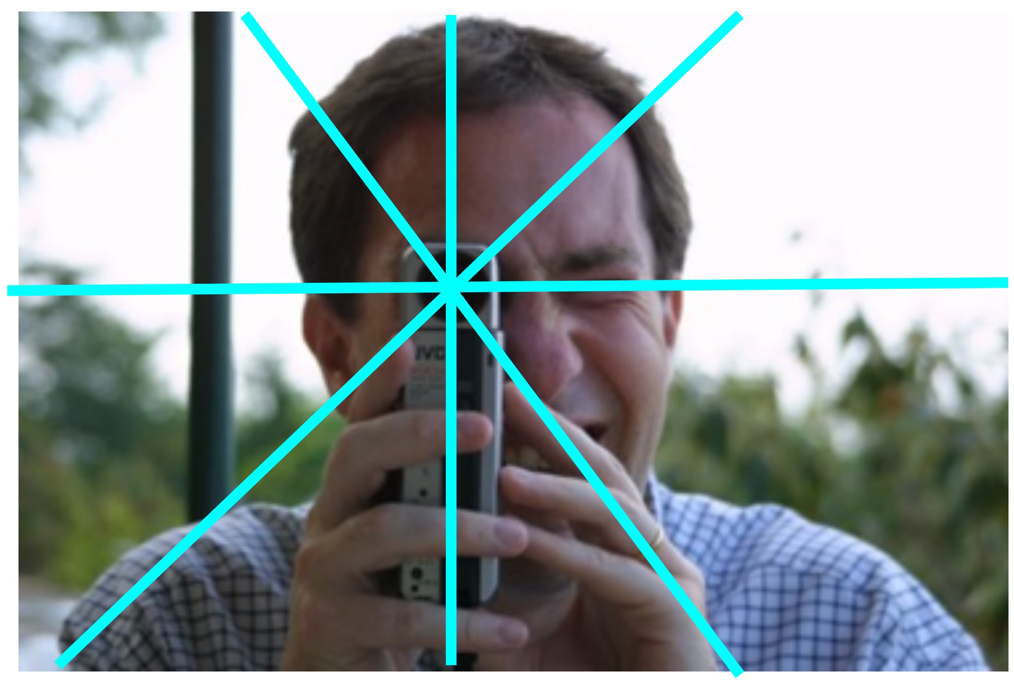
$M = [I|0]$

Example:
Cameras Point at Each Other

Top View



Epipolar Lines



Epipoles

- Camera Center C’ in first view:

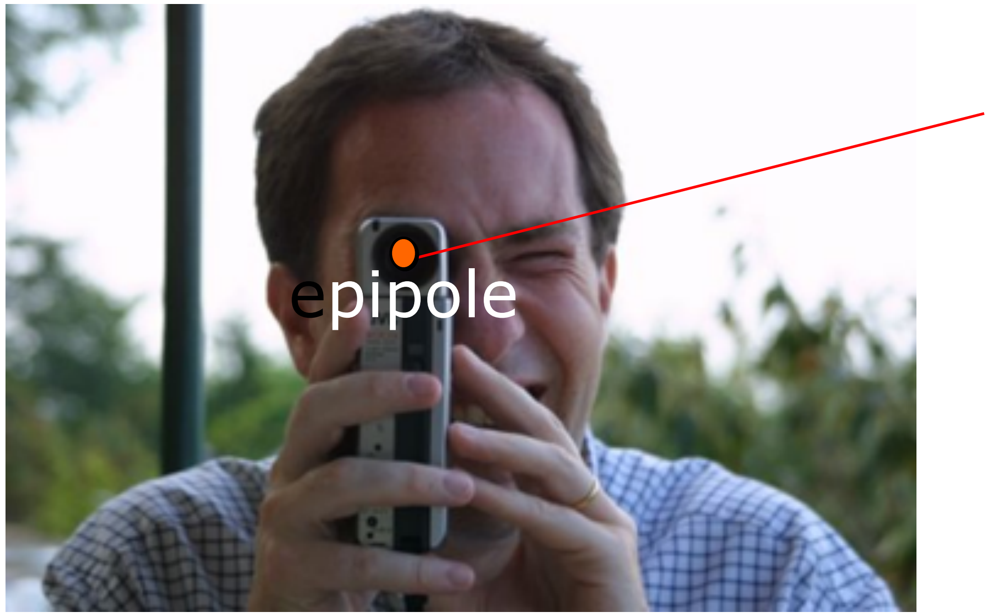
$$e = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix} = t$$

- Origin C in second view:

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

Image of Camera Ray ?

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



$M = [I|0]$

Point at infinity

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

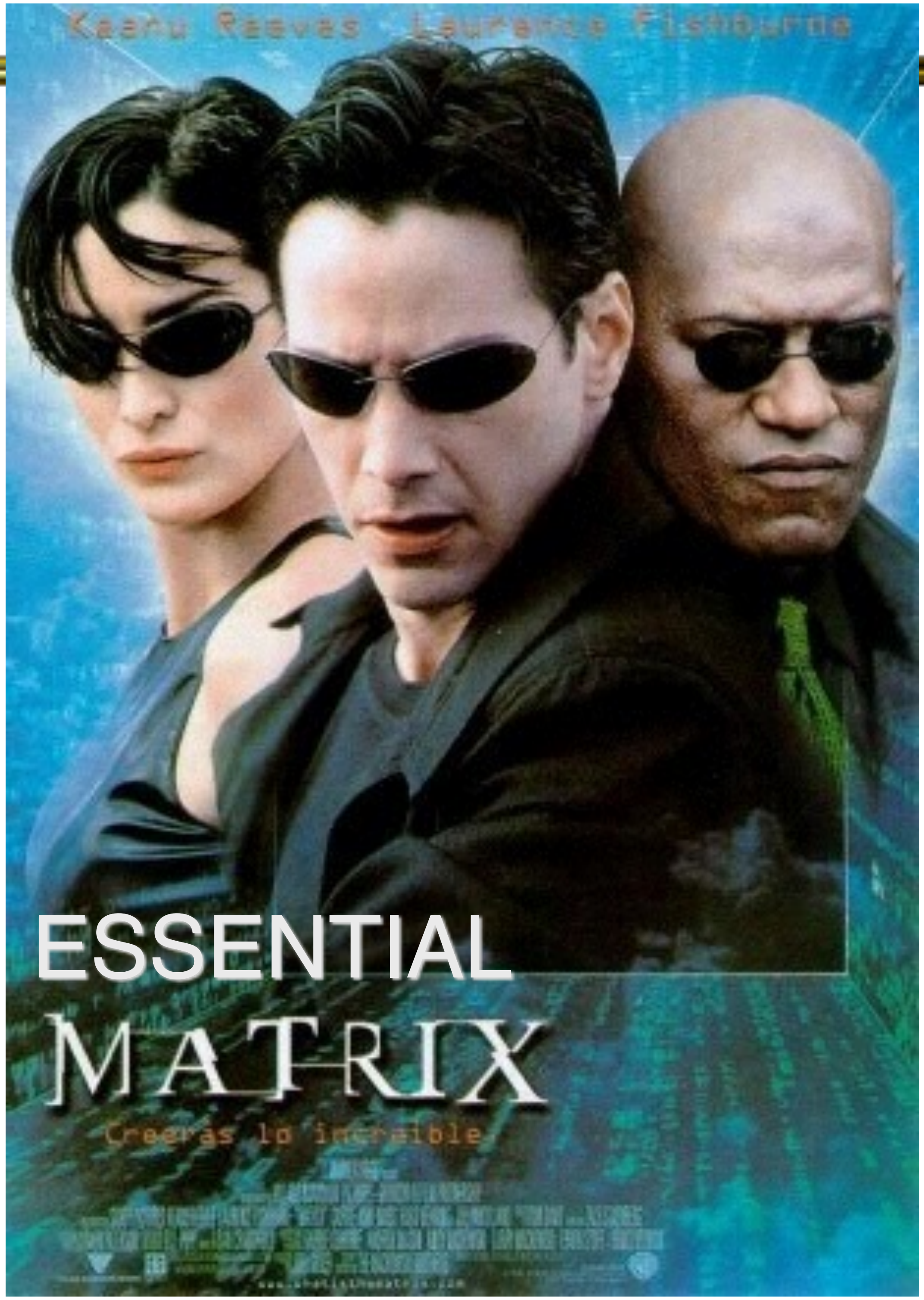
$$p' = \begin{bmatrix} A & a \end{bmatrix} * \begin{bmatrix} x \\ 0 \end{bmatrix} = Ax \Rightarrow x = A^{-1} p'$$

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

$$x = Rp'$$

$$M' = \left[R_w^c \mid R_w^c t \right] \qquad x = R_c^w p^c$$

Essential Matrix



Epipolar Line Calculation

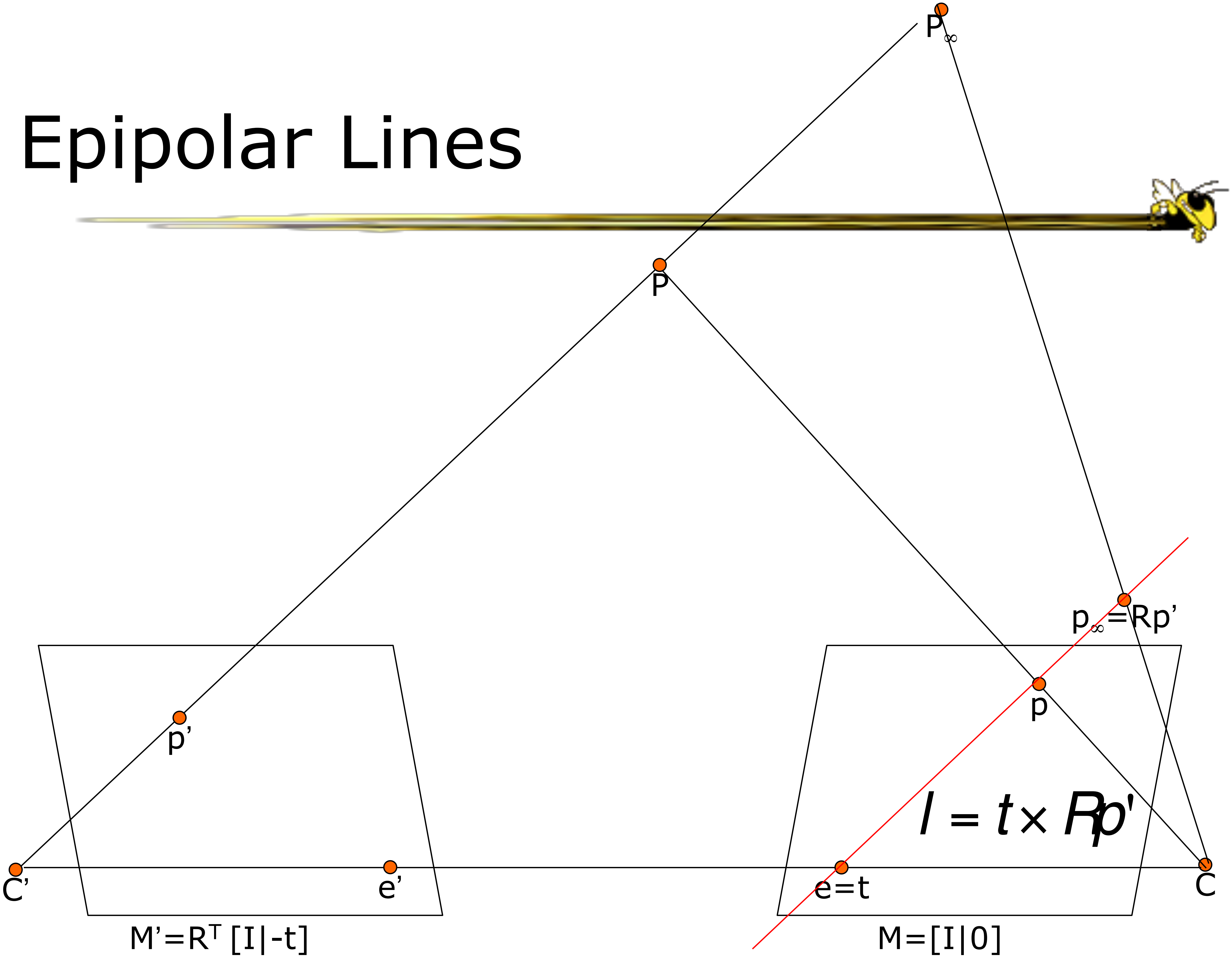
1) Point 1 = epipole e=t
2) Point 2 = point at infinity

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

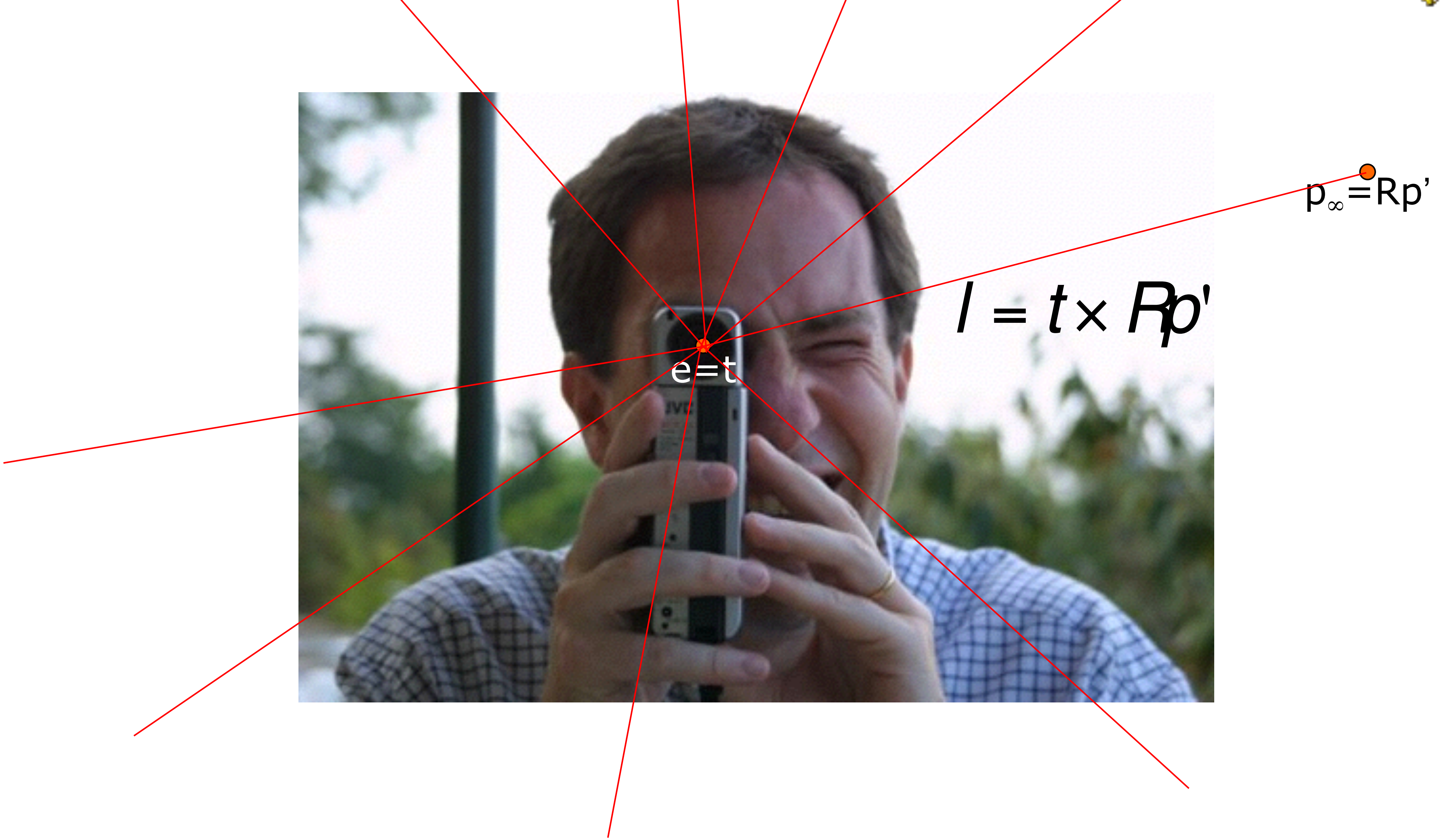
3) Epipolar line = join of points 1 and 2

$$l = t \times Rp'$$

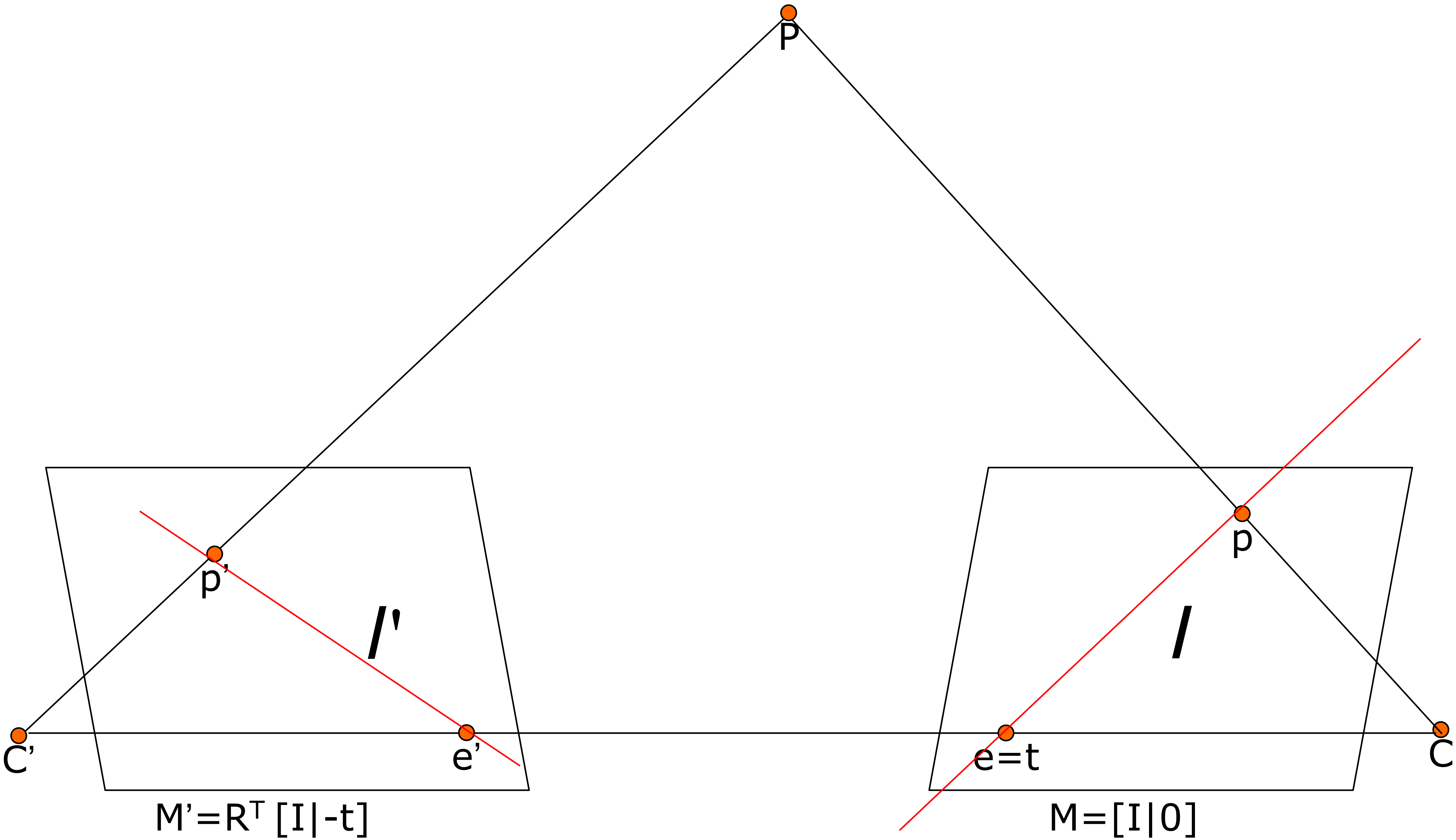
Epipolar Lines



Epipolar lines



Epipolar Plane



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

$$p^T E p' = 0$$

E's Degrees of Freedom

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

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

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

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

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

Minimize:

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

under the constraint

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