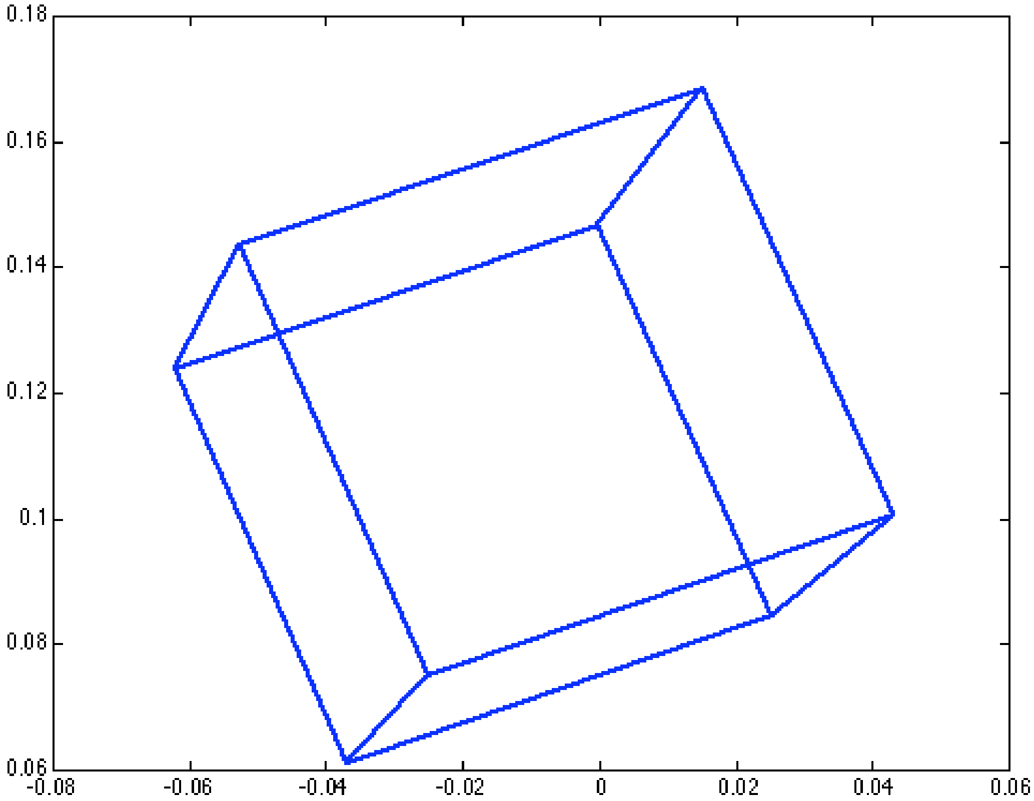
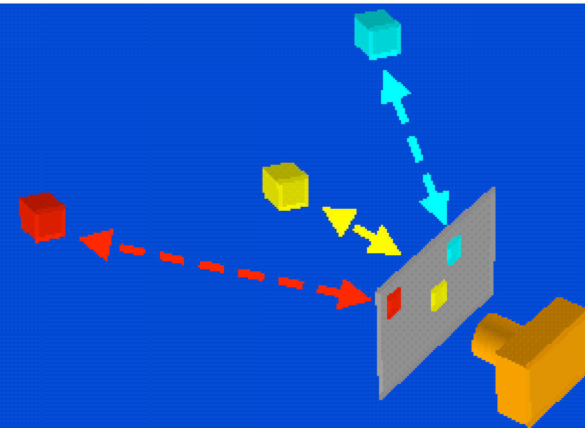
Projected image of a cube



Classical Calibration

Know 3D coords, 2D coords

- **Find projection matrix Π**



$$d\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

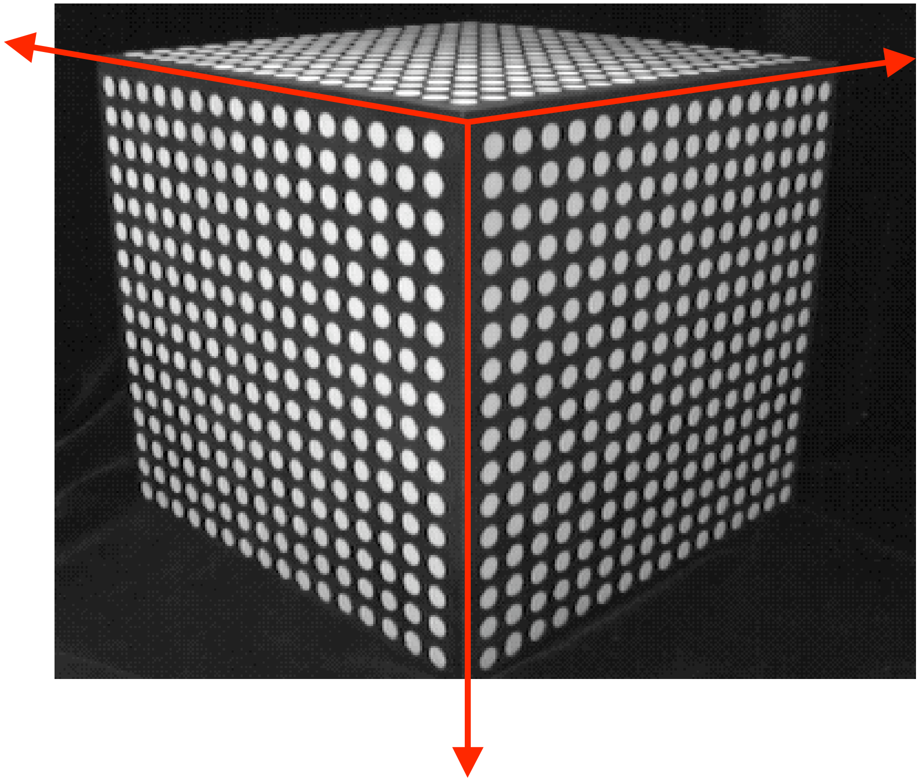
$$\mathbf{u} = \Pi \mathbf{X}$$

11 unknowns (up to scale)

2 equations per point
(eliminate d)

6 points is sufficient

Camera Calibration



Take a known set of points.

Typically 3 orthogonal planes.

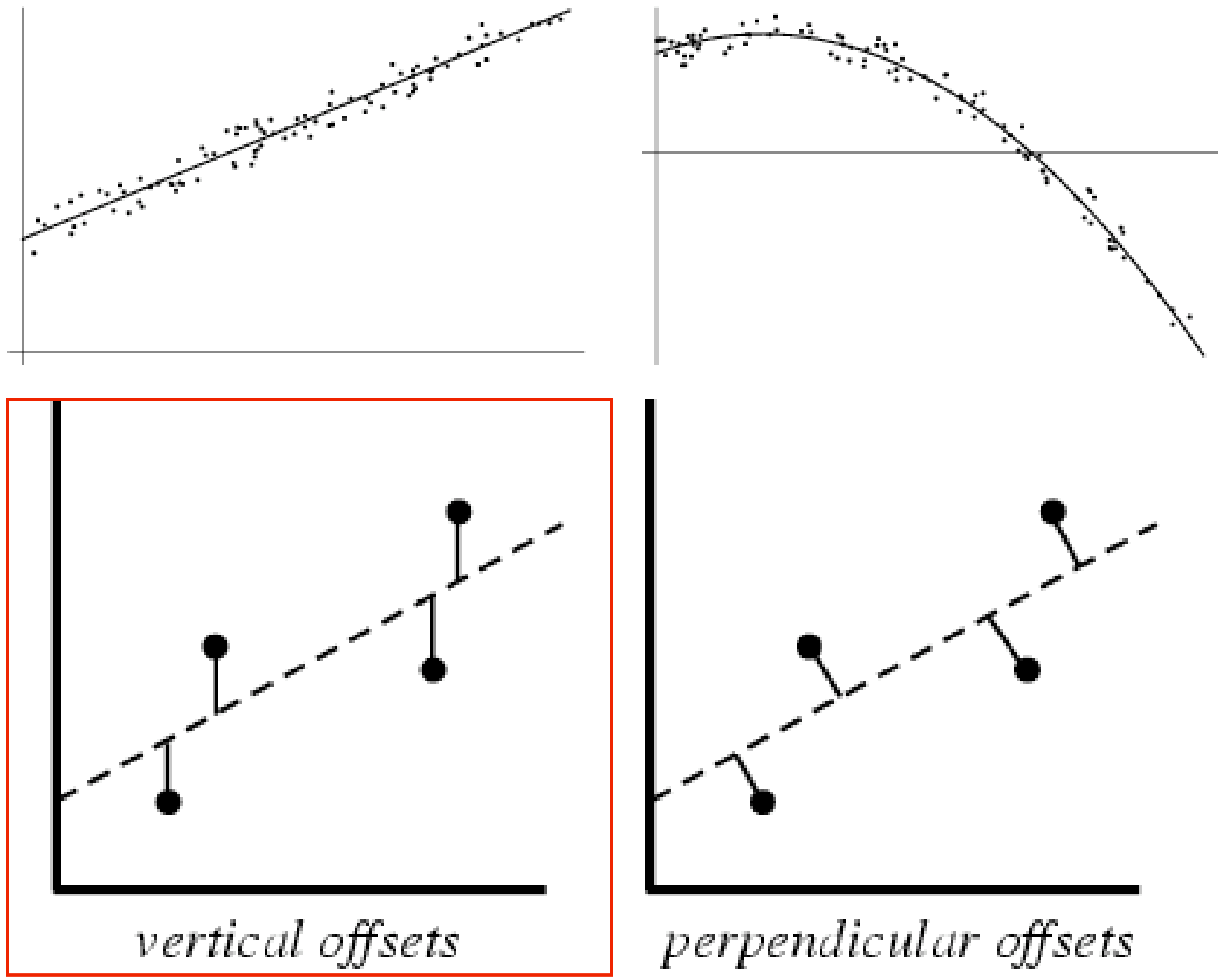
Treat a point in the object as the World origin

Points x1, x2, x3,

Project to y1,y2,y3

Least Square Estimation: Idea

We use:



Least Square Estimation

$$\vec{y}^{\,pred} = A\vec{x} \qquad \textit{prediction}$$

$$E = \sum_{j=1}^{m} (y_j^{measured} - y_j^{pred})^2 \qquad \textit{Error}$$

$$E = \sum_{j=1}^{m} (y_j^{measured} - \sum_{i=1}^{n} a_{ij} x_i)^2 \qquad \textit{Error}$$

$$E = (\vec{y}^{\,meas} - A\vec{x})^t (\vec{y}^{\,meas} - A\vec{x}) \qquad \textit{Error Matrix form}$$

$$\min_x E \quad \Rightarrow \quad \frac{\partial E}{\partial x} = 0 \quad \Rightarrow \quad -2A^t(\vec{y}^{\,meas} - A\vec{x}) = 0$$

$$\Rightarrow \quad A^t \vec{y}^{\,meas} = A^t A \vec{x}$$

$$\Rightarrow \quad (A^t A)^{-1} A^t \vec{y}^{\,meas} = \vec{x}$$

In general, y is a matrix of measurements, and x is a matrix of matched predictors

Converting a Matrix into a vector for Least squares…

$$\begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} = A \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$$

$$\begin{bmatrix} \vec{y}_1 & \cdots & \vec{y}_m \end{bmatrix} = \begin{bmatrix} \vec{a}_1^t \\ \vdots \\ \vec{a}_n^t \end{bmatrix} \begin{bmatrix} \vec{x}_1 & \cdots & \vec{x}_m \end{bmatrix}$$

$$\begin{bmatrix} \vec{y}_1 & \cdots & \vec{y}_m \end{bmatrix} = \begin{bmatrix} \vec{a}_1^t & \cdots & \vec{a}_n^t \end{bmatrix} \begin{bmatrix} \vec{x}_1 & \vec{0} & \cdots & \vec{0} & \vec{0} \\ \vec{0} & \vec{x}_1 & \cdots & \vec{0} & \vec{0} \\ \vec{0} & \vec{0} & \cdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \vec{x}_n & \vec{0} \\ \vec{0} & \vec{0} & \cdots & \vec{0} & \vec{x}_n \end{bmatrix}$$

$$Y = X^t \begin{bmatrix} \vec{a}_1^t & \cdots & \vec{a}_n^t \end{bmatrix}^t$$

Real Calibration Procedure

$$\begin{bmatrix}u_i\\ v_i\\ 1\end{bmatrix} = \frac{1}{z}M\vec{p}_i = \frac{1}{z}\begin{bmatrix}m_1^t\\ m_2^t\\ m_3^t\end{bmatrix}\vec{p} =$$

$$u_i = \frac{m_1^t\vec{p}_i}{m_3^t\vec{p}_i} \Rightarrow \quad m_1^t\vec{p}_i - \left(m_3^t\vec{p}_i\right)u_i = 0 \quad \Rightarrow \quad \left(m_1^t - m_3^t u_i\right)\cdot \vec{p}_i = 0$$

$$v_i = \frac{m_2^t\vec{p}_i}{m_3^t\vec{p}_i} \Rightarrow \quad m_2^t\vec{p}_i - \left(m_3^t\vec{p}_i\right)v_i = 0 \quad \Rightarrow \quad \left(m_2^t - m_3^t v_i\right)\cdot \vec{p}_i = 0$$

Collecting all corresponding points into a matrix

$$\square \overset{def}{=} \begin{bmatrix} \vec{p}_1^t & 0\,0\,0\,0 & -u_1\vec{p}_1^t \\ 0\,0\,0\,0 & \vec{p}_1^t & -v_1\vec{p}_1^t \\ \cdots & \cdots & \cdots \\ \vec{p}_n^t & 0\,0\,0\,0 & -u_n\vec{p}_n^t \\ 0\,0\,0\,0 & \vec{p}_n^t & -v_n\vec{p}_n^t \end{bmatrix} \qquad \vec{m} = \begin{bmatrix} m_1\\ m_2\\ m_3\end{bmatrix}$$

$$\square\vec{m} = 0 \quad \Rightarrow \quad E = (\square\vec{m} - 0)^t(\square\vec{m} - 0) = \vec{m}^t\square^t\square\vec{m}$$

To minimize the error, find

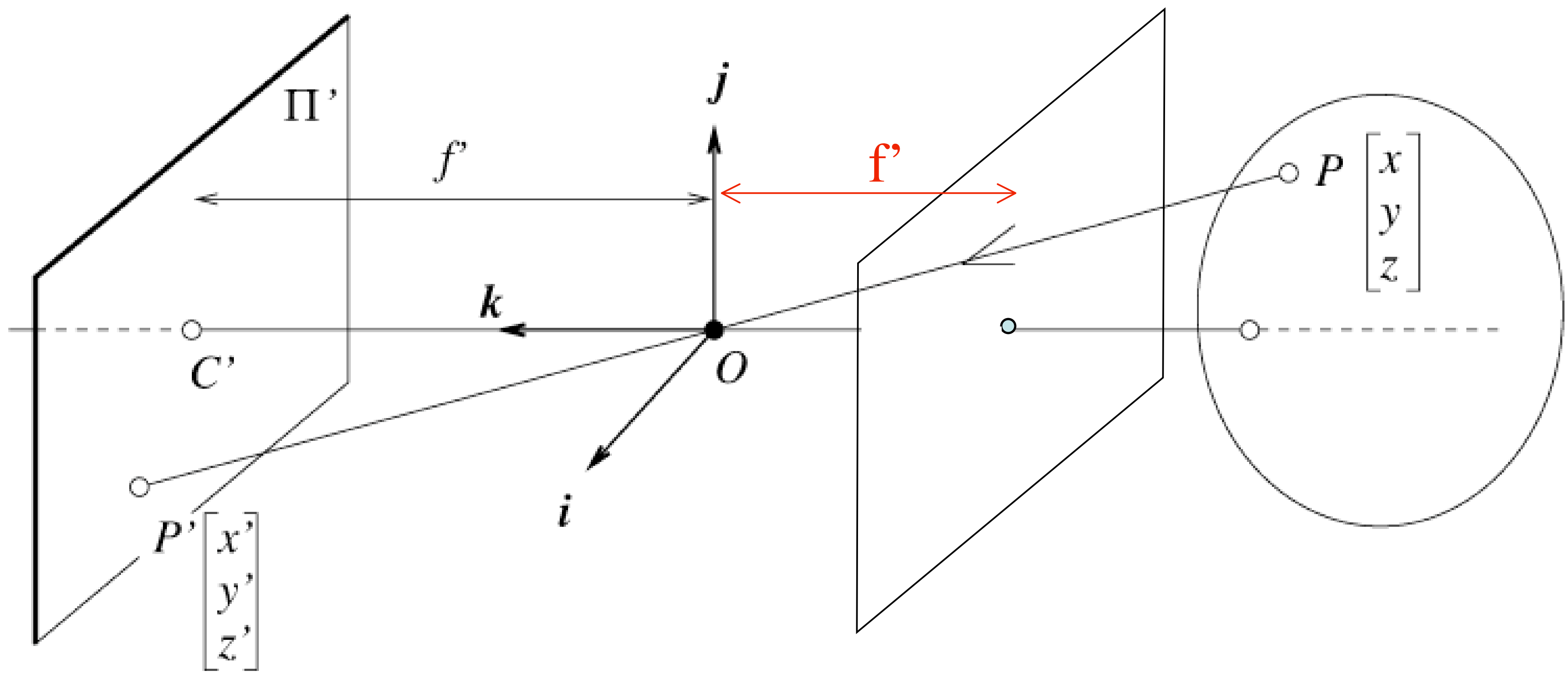
$$\square\vec{m} = 0\vec{m} \quad \Rightarrow \quad \vec{m} = Null(\square)$$

Reshape $\vec{m}$ into M, and can extract parameters from M

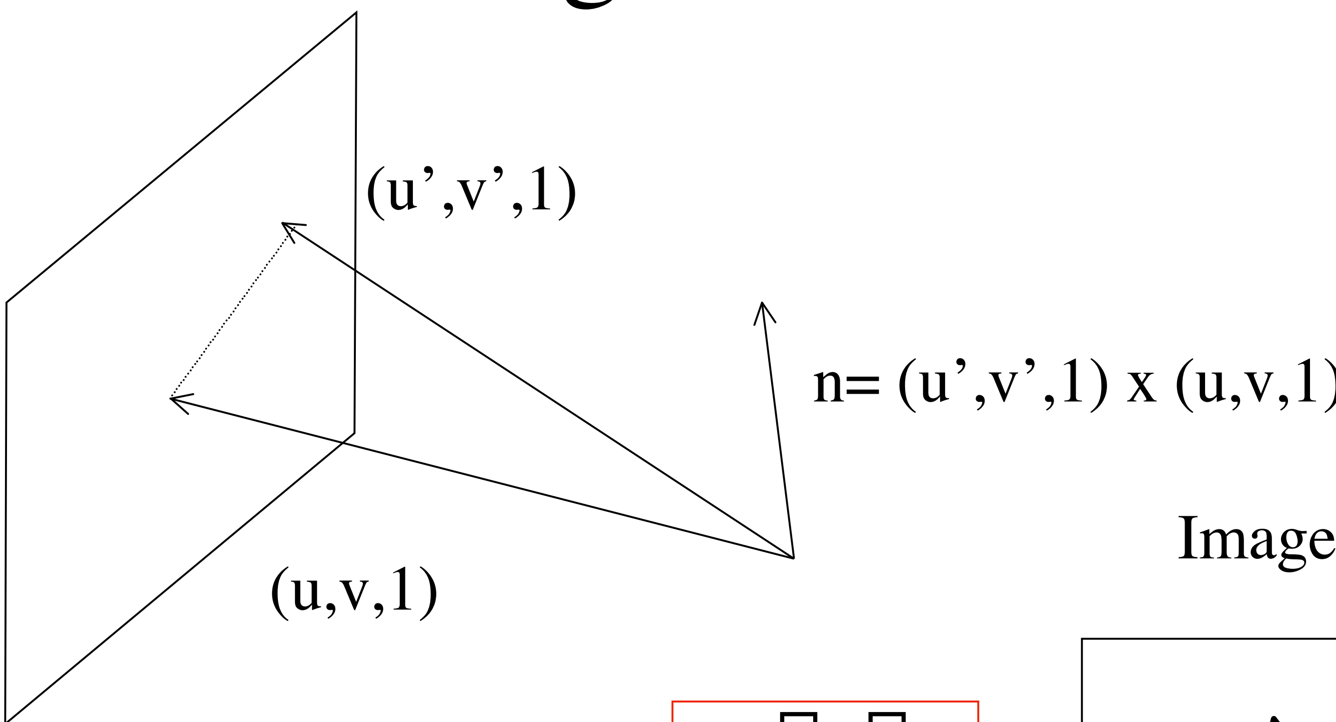
Multi-view geometry

- How can we recover information about the structure of an object?
  - Take multiple views of the object.
  - Each image supplies a constraint on the location of points in space
  - Intersecting the constraints provides a solution

From Pinhole to Picture Plane

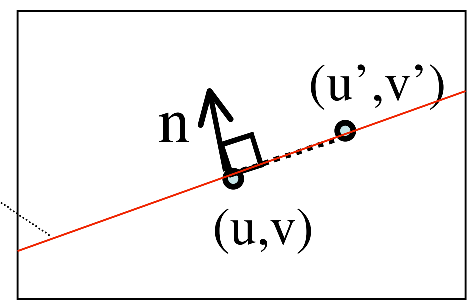


Understanding Homogeneous Image Coordinates

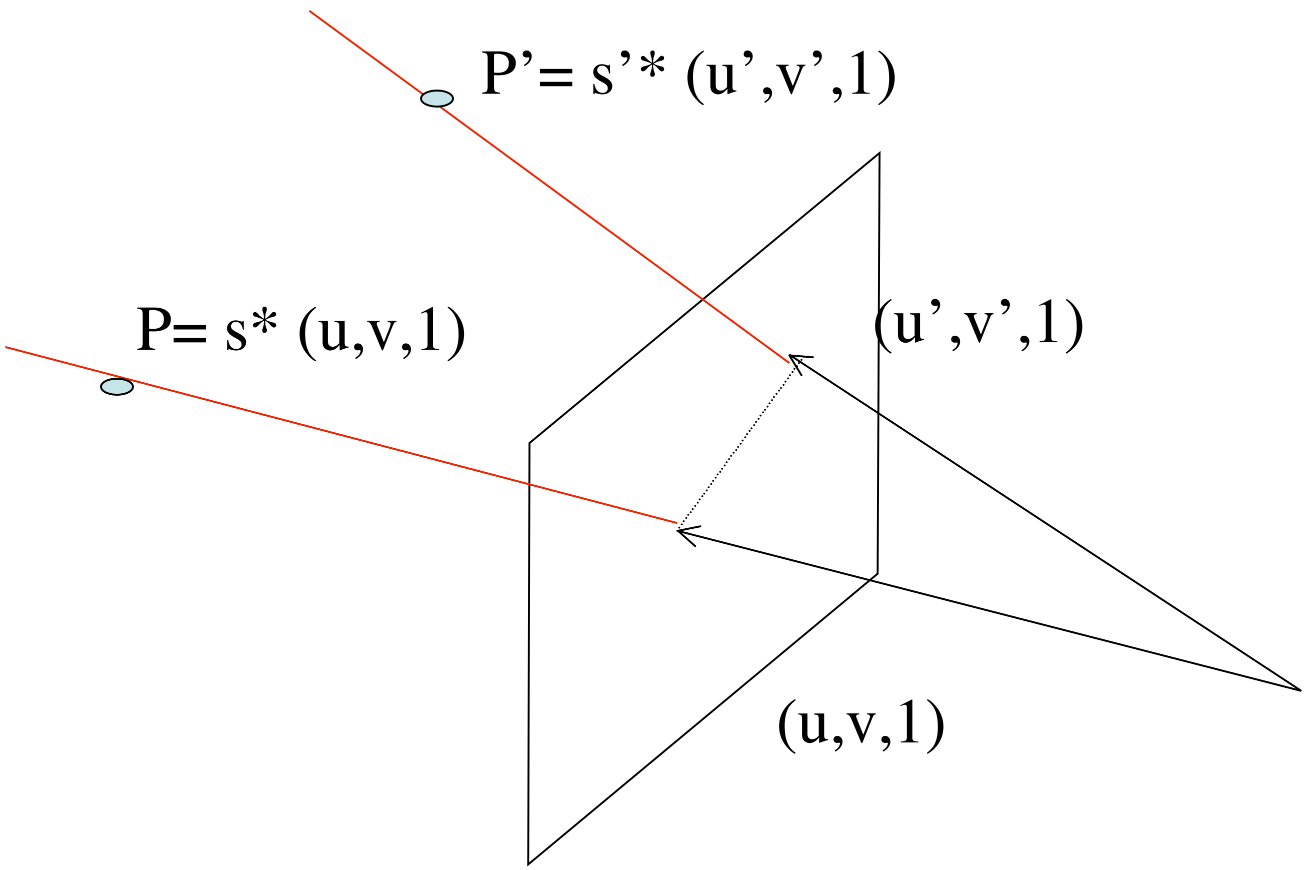


$$l = \begin{bmatrix} u_3 \\ v_3 \end{bmatrix} \cdot \vec{n}$$

Image

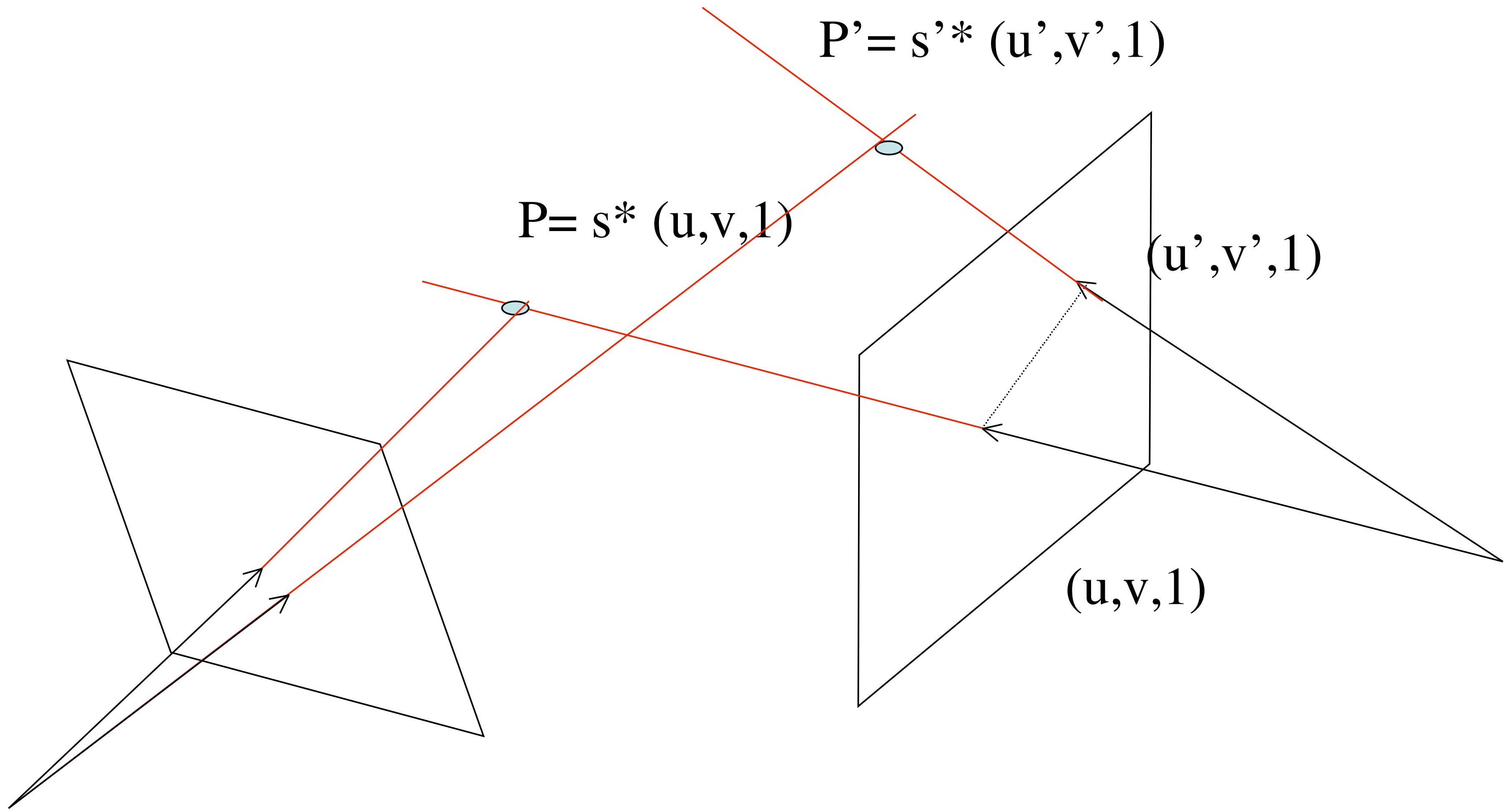


Constraint supplied by one view



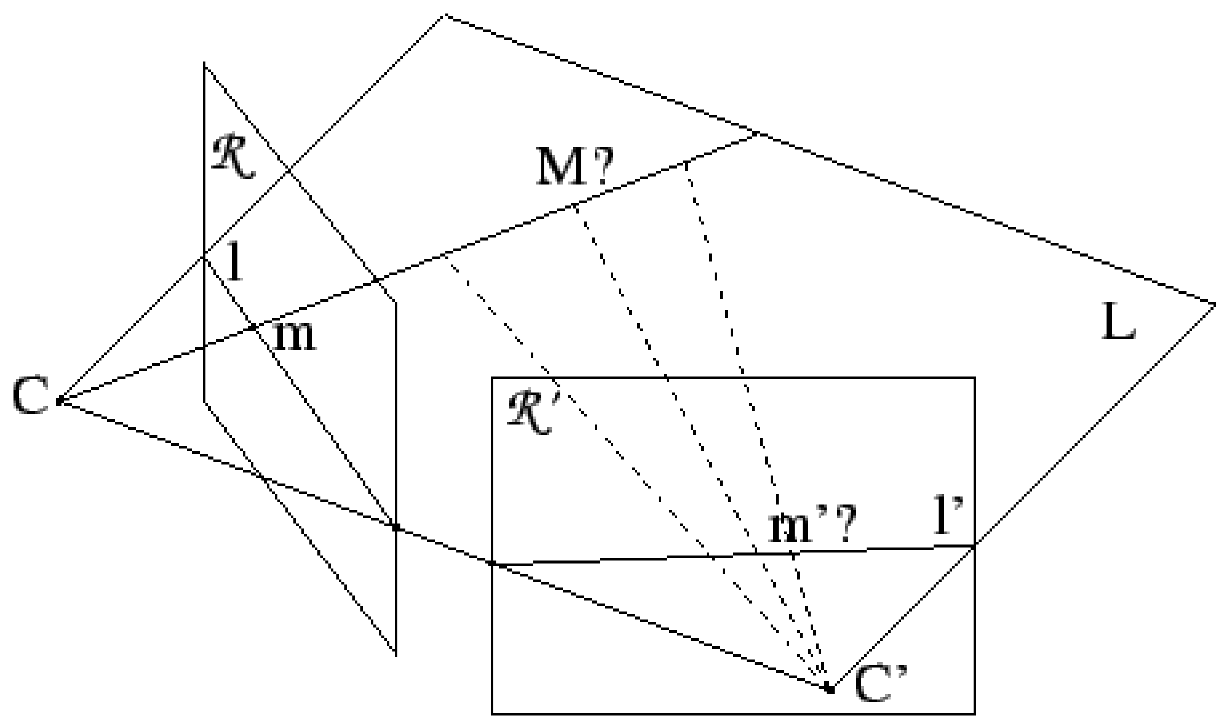
Projective ambiguity

Constraint supplied by two views



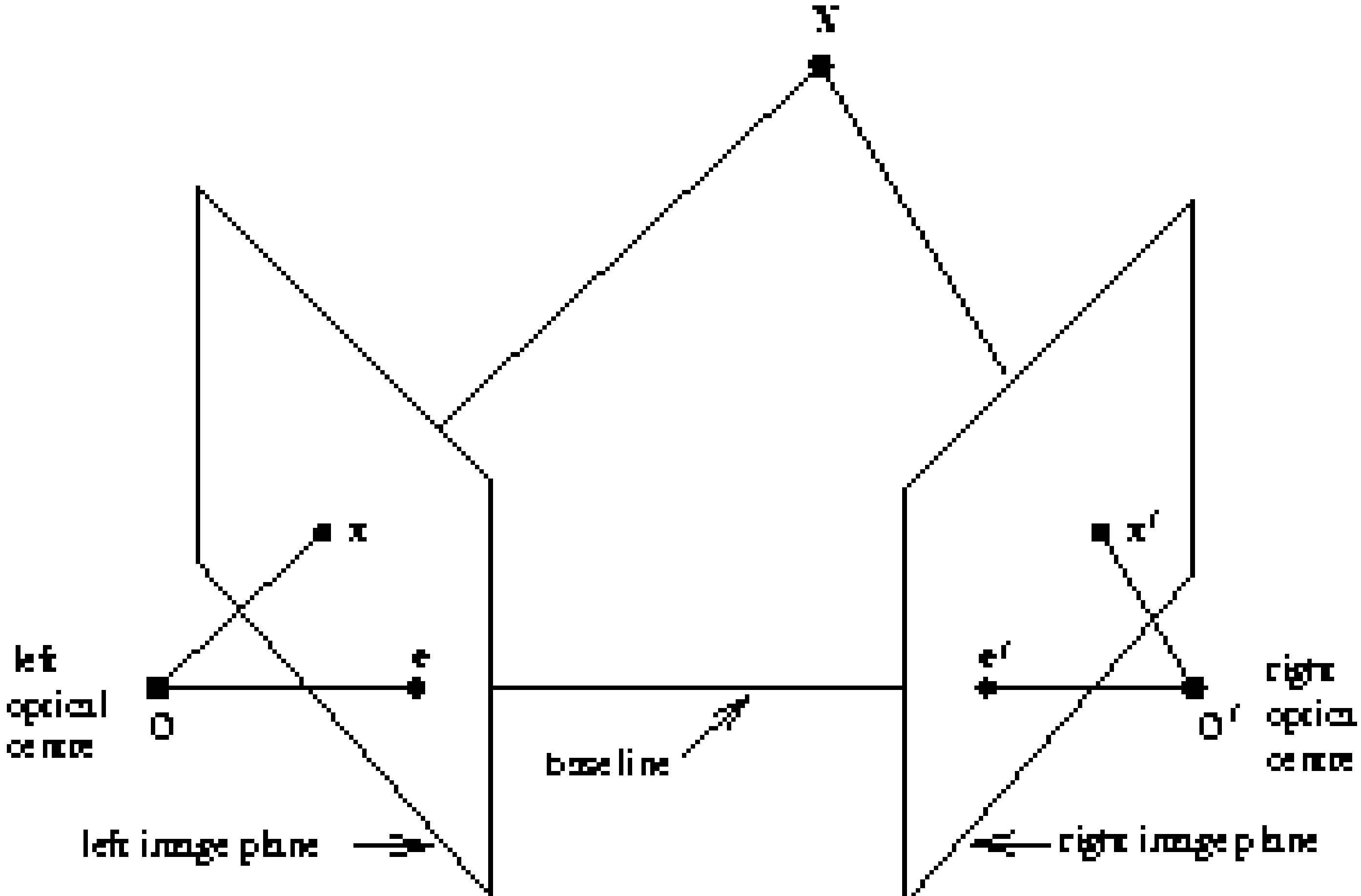
Relations between image coordinates

Given coordinates in one image, and the tranformation
Between cameras, T = [R t], what are the image coordinates
In the other camera's image.



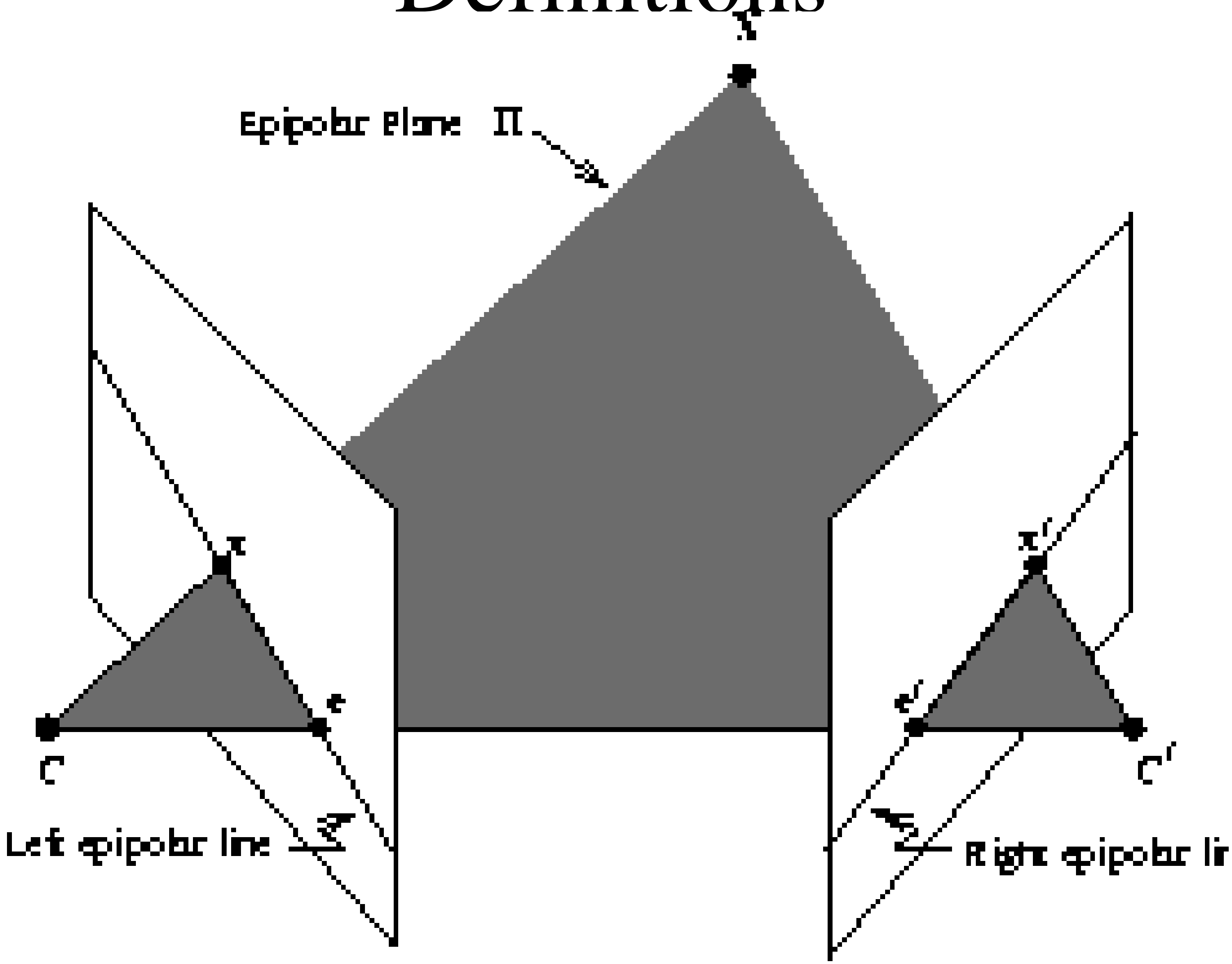
Epipolar Geometry

The fundamental **geometric** relationship between two perspective cameras.



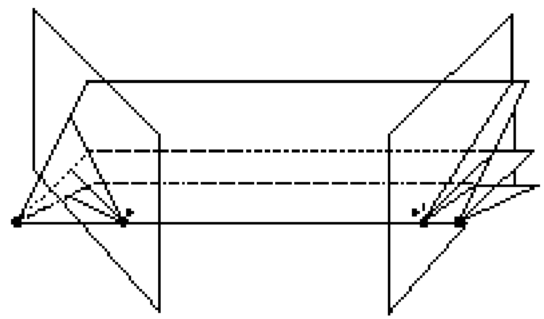
The **epipole**: is the *point* of intersection of the line joining the optical centres---the *baseline*---with the image plane. The epipole is the image in one camera of the optical centre of the other camera.

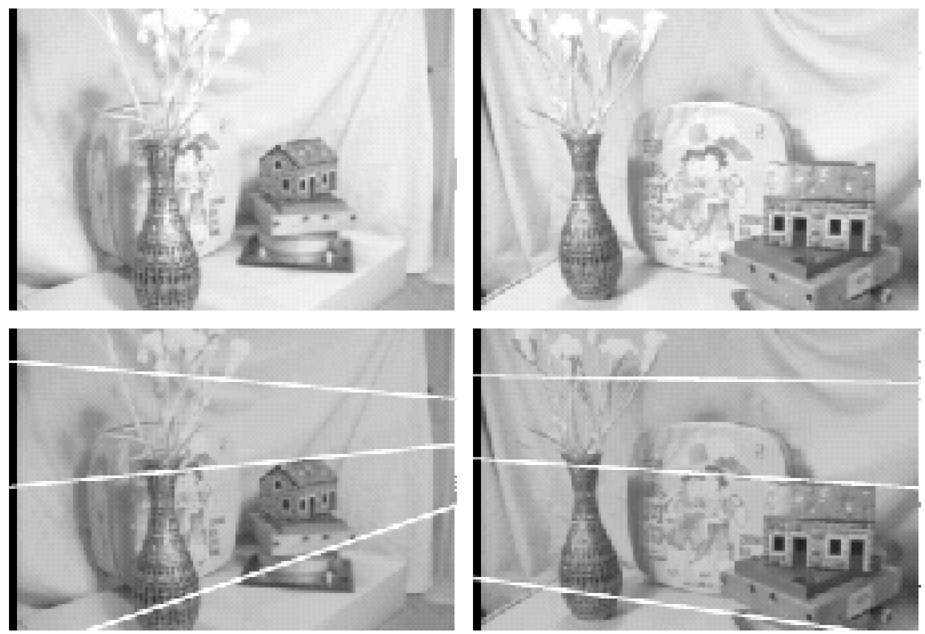
Definitions



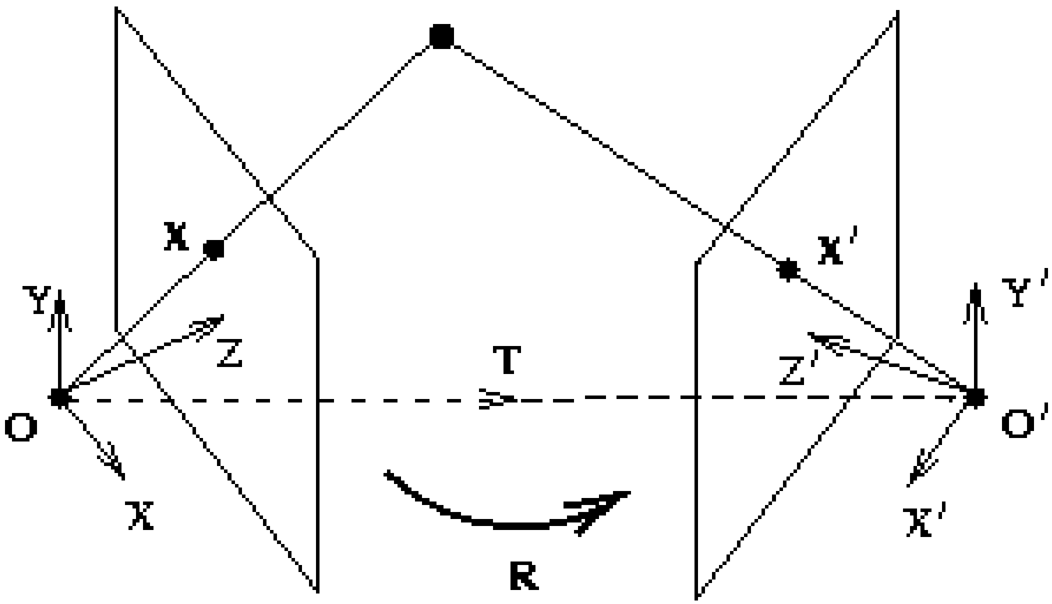
Epipolar Example

Converging cameras





Essential Matrix: Relating between image coordinates



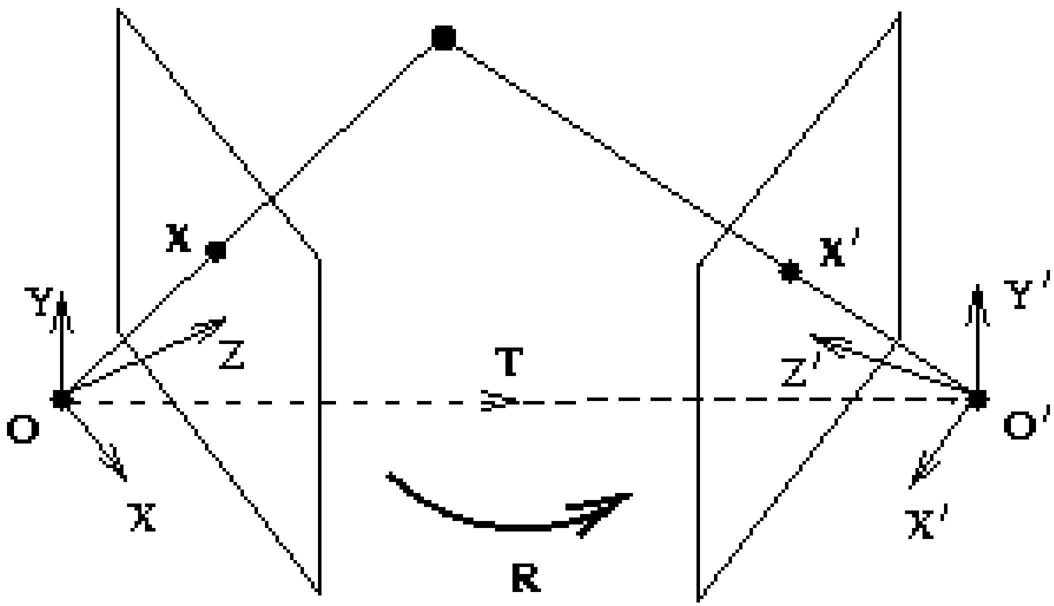
$$\vec{Ox} \quad \vec{O'x'} \quad \vec{OO'}$$

Are Coplanar, so:

$$\vec{O'x'} \bullet \left[\vec{Ox} \times \vec{OO'} \right] = 0$$

camera coordinate systems, related by a rotation **R** and a translation **T**:

$$x' = \begin{bmatrix} R & t \\ 000 & 1 \end{bmatrix} x$$



$$\mathbf{a} \times \mathbf{v} = \begin{bmatrix} a_y v_z - a_z v_y \\ a_z v_x - a_x v_z \\ a_x v_y - a_y v_x \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

$$= \mathbf{A}\mathbf{v}$$

$$x' = \begin{bmatrix} R & t \\ 000 & 1 \end{bmatrix} x$$

$$\vec{O'x'} \cdot \left[\vec{Ox} \times \vec{OO'} \right] = 0$$

$$x' \cdot \left(\vec{t} \times Rx \right) = 0$$

$$x' \cdot \left(\mathcal{E} x \right) = 0$$

$$x' \mathcal{E} x = 0$$

$$\mathcal{E} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} R$$

What does the Essential matrix do?

It represents the normal to the epipolar line *in the other image*

$$n = E x$$

The normal defines a line in image 2:

$$x'_{on\ epipolar\ line} \Rightarrow n \cdot x' = 0$$

$$n_1 x_1 + n_2 x_2 + n_3 1 = 0$$

$$(y = mx + b) \Rightarrow b = -n_3, \quad m = -\frac{n_1}{n_2}$$

What if cameras are uncalibrated? Fundamental Matrix

Choose world coordinates as Camera 1.
Then the extrinsic parameters for camera 2 are just **R** and **t**
However, intrinsic parameters for both cameras are unknown.
Let C_1 and C_2 denote the matrices of intrinsic parameters. Then the pixel coordinates measured are not appropriate for the Essential matrix.
Correcting for this distortion creates a new matrix: the Fundamental Matrix.

$$x'_{measured} = C_2 x' \qquad x_{measured} = C_1 x$$

$$(x')^t E\, x = 0 \Rightarrow \left(C_2^{-1} x'_{measured}\right)^t E \left(C_1^{-1} x_{measured}\right) = 0$$

$$(x'_{measured})^t F\, x_{measured} = 0$$

$$F = C_2^{-t} E\, C_1^{-1}$$

$$C = \begin{bmatrix} -f \cdot s_u & 0 & u_0 \\ 0 & -f \cdot s_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Computing the fundamental Matrix

Computing : I Number of Correspondences Given perfect image points (no noise) in general position. Each point correspondence generates one constraint on

$$\begin{bmatrix} x'_i & y'_i & 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = 0$$