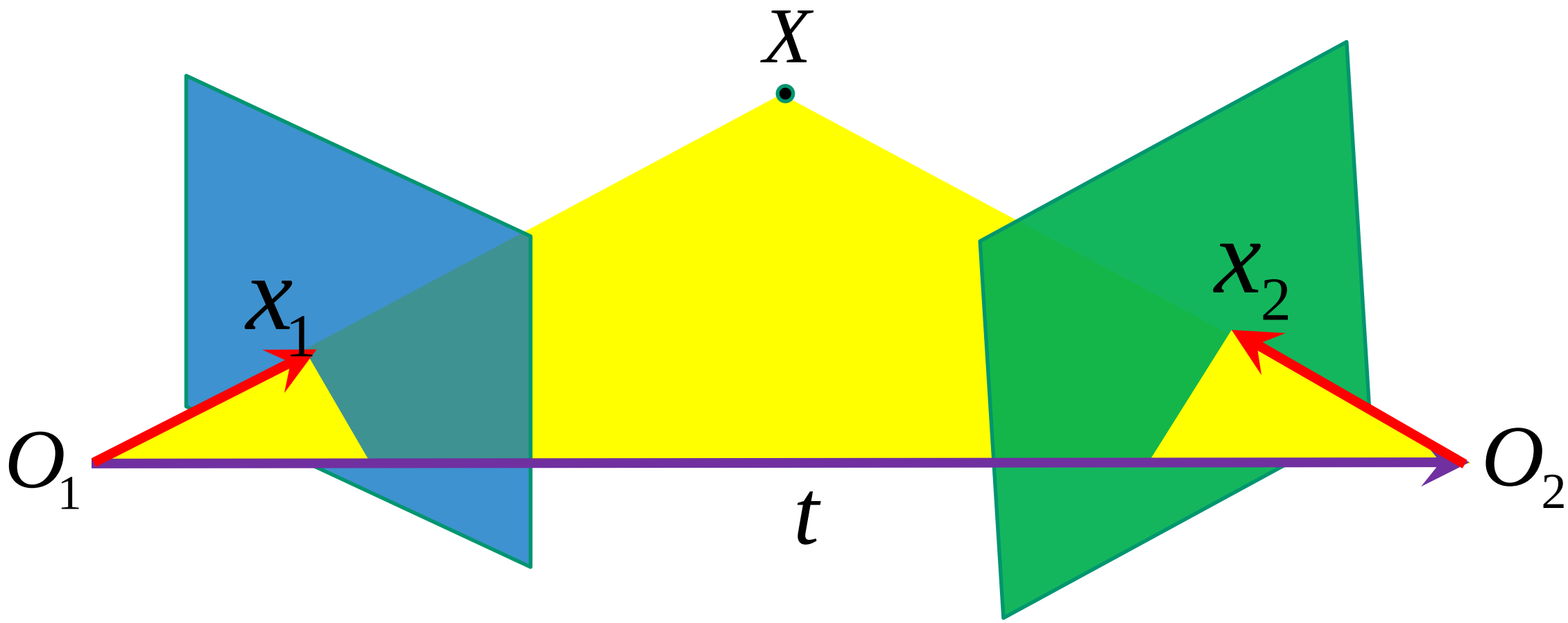


If these three vectors are coplanar x_1 t x_2 then

$$x_2 \bullet \left(t \times x_1 \right) = 0$$

For this to work, everything needs to be in the same coordinate system (camera1, camera2, or world)

Cross Products

$$\begin{bmatrix} a_x & a_y & a_z \end{bmatrix} \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} = \text{zero}$$

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

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

$$= \mathbf{A}_{\text{skew}} \mathbf{v}$$

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

MATLAB Extrinsics Matrix

Converts world coordinates to camera coordinates

$$\begin{bmatrix} x_c & y_c & z_c \end{bmatrix} = \begin{bmatrix} x_w & y_w & z_w \end{bmatrix} \mathbf{R} + \mathbf{T}$$

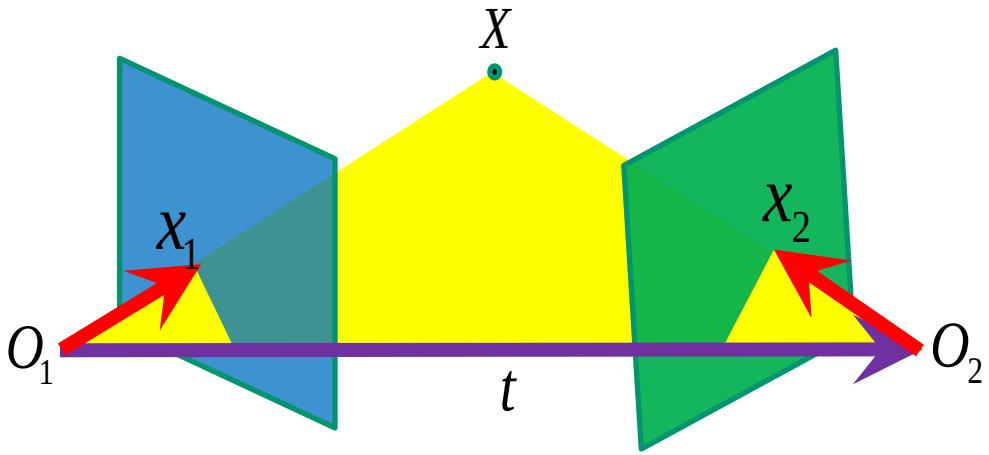
T is world origin in camera coordinates

camera origin in world coordinates

$$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} x_w & y_w & z_w \end{bmatrix} \mathbf{R} + \mathbf{T}$$

$$\begin{bmatrix} x_w & y_w & z_w \end{bmatrix} = -\mathbf{T}\mathbf{R}^{\mathrm{T}}$$

Essential Matrix



Assume row vectors (1x3)

Use world as reference

$$\mathbf{T}_W = -\mathbf{T}\mathbf{R}^{\mathrm{T}} \qquad x_{1W} = x_1 R_1^{\mathrm{T}}$$

$$t = \mathbf{T}_{W2} - \mathbf{T}_{W1} \qquad x_{2W} = x_2 R_2^{\mathrm{T}}$$

$$x_{2W} \bullet \left(\mathbf{T}_w \times x_{1w}\right) = 0$$

$$x_{2W}\mathbf{T}_{skew}x_{1W}^{\mathrm{T}} = 0$$

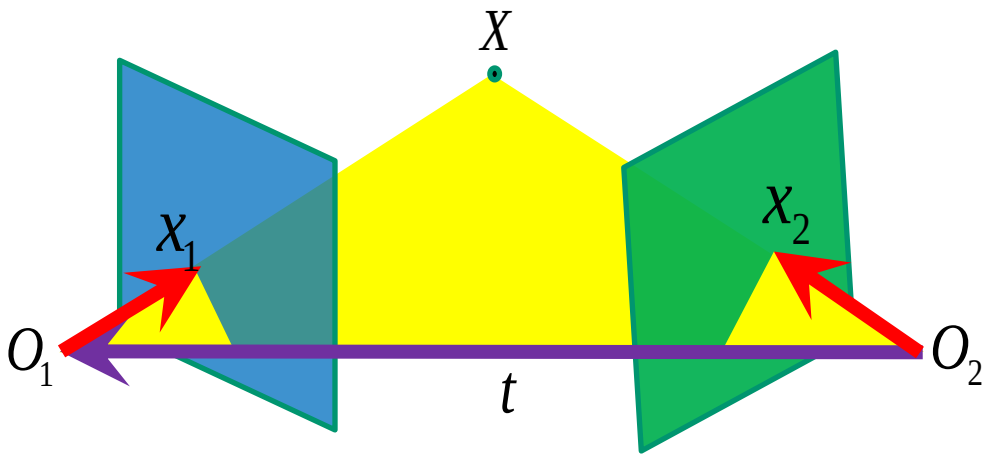
$$\mathcal{E} = R_2^{\mathrm{T}}\mathbf{T}_{skew}R_1 \qquad x_2 R_2^{\mathrm{T}}\mathbf{T}_{skew}R_1 x_1^{\mathrm{T}} = 0$$

Essential Matrix
[Longuet-Higgins 1981]

$$x_2 \mathcal{E} x_1^{\mathrm{T}} = 0$$

Essential Matrix

Use camera2 as reference



$$\mathbf{T}_W = -\mathbf{T}\mathbf{R}^{\mathrm{T}} \qquad x_{1,2} = x_1 R_1^{\mathrm{T}} R_2$$

$$t = \left(\mathbf{T}_{W2} - \mathbf{T}_{W1}\right)\mathbf{R}_2 \qquad x_2 = x_2$$

$$x_2 \bullet \left(t \times x_{1,2}\right) = 0$$

Assume row vectors (1x3)

$$x_2 \mathbf{T}_{skew} x_{1,2}^{\mathrm{T}} = 0$$

$$\mathcal{E} = \mathbf{T}_{skew} R_2^{\mathrm{T}} R_1$$

$$x_2 \mathbf{T}_{skew} \left(R_2^{\mathrm{T}} R_1\right) x_1^{\mathrm{T}} = 0$$

Essential Matrix
[Longuet-Higgins 1981]

$$x_2 \mathcal{E} x_1^{\mathrm{T}} = 0$$

Extract T and R from Essential Matrix

How do we recover *t* and *R*? Answer: SVD of $\mathcal{E}$

$$\mathcal{E} = USV^t$$

S diagonal

U, V orthogonal and $\det() = 1$ (rotation)

$$R = UWV^t \quad \text{or} \quad R = UW^tV^t \qquad \vec{t} = u_3 \quad \text{or} \quad \vec{t} = -u_3$$

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

$u_3 = \mathrm{U}(:,3)$

Ideally S has two identical non-zero values and one zero value

Reconstruction Ambiguity

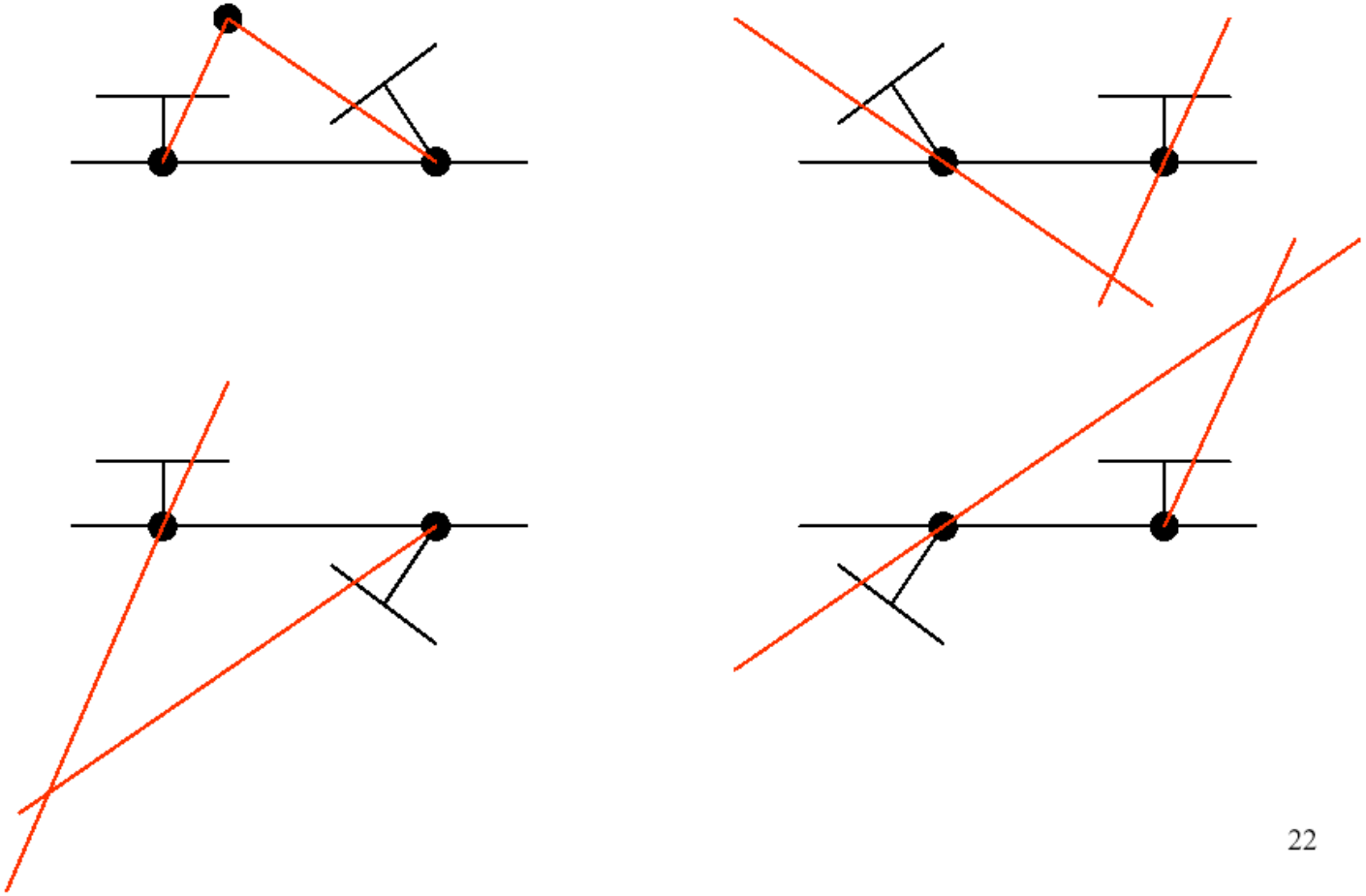
So we have 4 possible combinations of translations and rotations giving 4 possibilities for $\boldsymbol{M}_2^{norm} = [\boldsymbol{R} \mid \boldsymbol{t}]$

1. $\boldsymbol{M}_2^{norm} = [\boldsymbol{U}\boldsymbol{W}^{\boldsymbol{t}}\boldsymbol{V}^{\boldsymbol{t}} \mid \boldsymbol{t}]$
2. $\boldsymbol{M}_2^{norm} = [\boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{\boldsymbol{t}} \mid \boldsymbol{t}]$
3. $\boldsymbol{M}_2^{norm} = [\boldsymbol{U}\boldsymbol{W}^{\boldsymbol{t}}\boldsymbol{V}^{\boldsymbol{t}} \mid -\boldsymbol{t}]$
4. $\boldsymbol{M}_2^{norm} = [\boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{\boldsymbol{t}} \mid -\boldsymbol{t}]$

Maybe if t is negative then R should be negative?

Which one is right?

- We can determine which of these is correct by looking at their geometric interpretation.



Which one is right?

- The correct pair will have our data points in front of both cameras.
- How do we choose the correct pair?
- Procedure:
  - Take a test point from data
  - Backproject to find 3D location
  - Determine the depth of 3D point in both cameras
  - Choose the camera pair that has a positive depth for both cameras.

Fundamental Matrix

The Fundamental matrix is a generalization of the Essential matrix, where the assumption of calibrated cameras is removed

$$p = x\mathbf{K}$$

$$x_2 \mathcal{E} x_1^{\mathrm{T}} = 0$$

The *p* vector gives image coordinates (in pixels) from the *x* vector, which contains normalized camera coordinates.

$$p\,\mathrm{inv}(\mathbf{K}) = x$$

$$p_2\,\mathrm{inv}(\mathbf{K}_2)\,\mathcal{E}\,\mathrm{inv}(\mathbf{K}_1)^{\mathrm{T}}\,p_1 = 0$$

$$\mathcal{F} = \mathrm{inv}(\mathbf{K}_2)\,\mathcal{E}\,\mathrm{inv}(\mathbf{K}_1)^{\mathrm{T}}$$

$$p_2 \mathcal{F} p_1^{\mathrm{T}} = 0$$