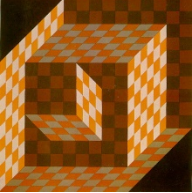


---

## Lecture 2

# Rigid–Body Motion and Imaging Geometry



## OUTLINE

---

### 3-D EUCLIDEAN SPACE & RIGID-BODY MOTION

- Coordinates and coordinate frames
- Rigid-body motion and homogeneous coordinates

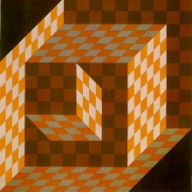
### GEOMETRIC MODELS OF IMAGE FORMATION

- Pinhole camera model

### CAMERA INTRINSIC PARAMETERS

- From metric to pixel coordinates

### SUMMARY OF NOTATION



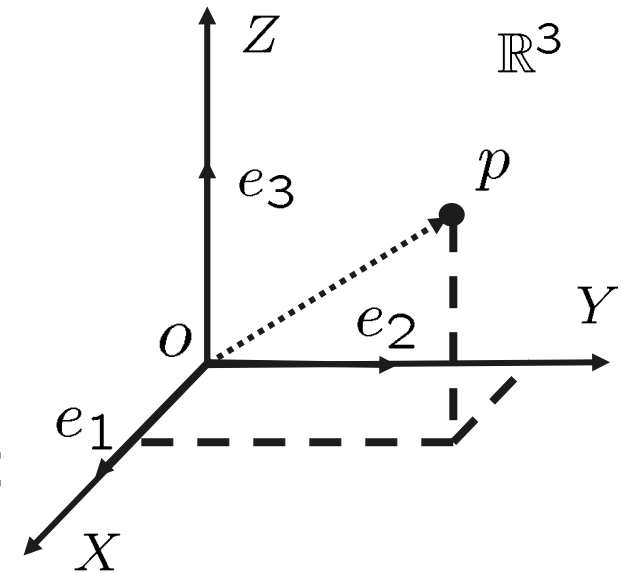
## 3-D EUCLIDEAN SPACE – Cartesian Coordinate Frame

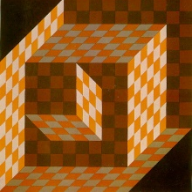
Standard base vectors:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Coordinates of a point  $p$  in space:

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3$$





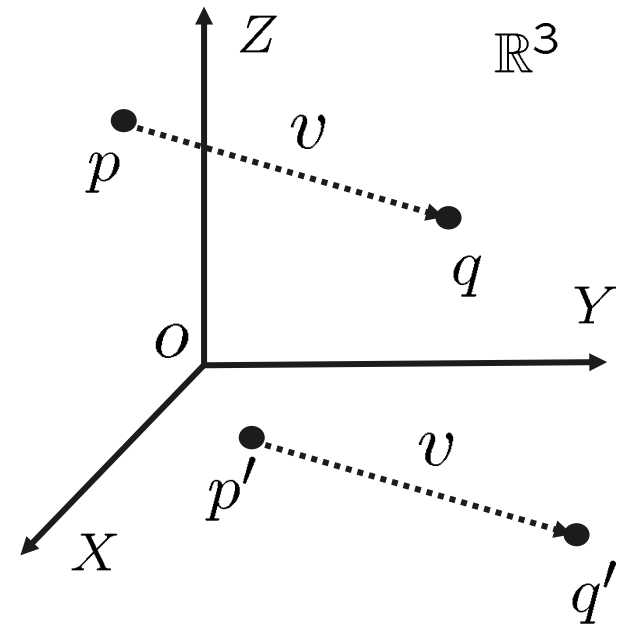
## 3-D EUCLIDEAN SPACE – Vectors

A “free” **vector** is defined by a pair of points  $(p, q)$ :

$$\mathbf{X}_p = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \in \mathbb{R}^3, \quad \mathbf{X}_q = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \in \mathbb{R}^3,$$

Coordinates of the vector  $v$ :

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} X_2 - X_1 \\ Y_2 - Y_1 \\ Z_2 - Z_1 \end{bmatrix} \in \mathbb{R}^3$$





## 3-D EUCLIDEAN SPACE – Inner Product and Cross Product

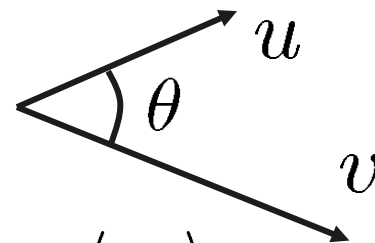
Inner product between two vectors:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$\langle u, v \rangle \doteq u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\cos(\theta) = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

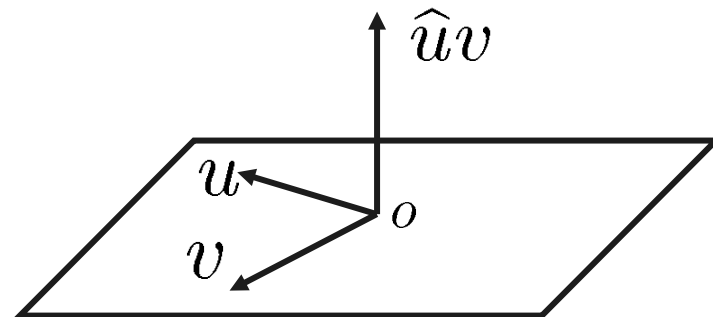
$$\|u\| \doteq \sqrt{u^T u} = \sqrt{u_1^2 + u_2^2 + u_3^2}$$

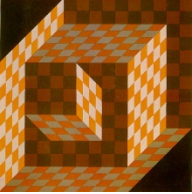


Cross product between two vectors:

$$u \times v \doteq \hat{u}v, \quad u, v \in \mathbb{R}^3$$

$$\hat{u} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$



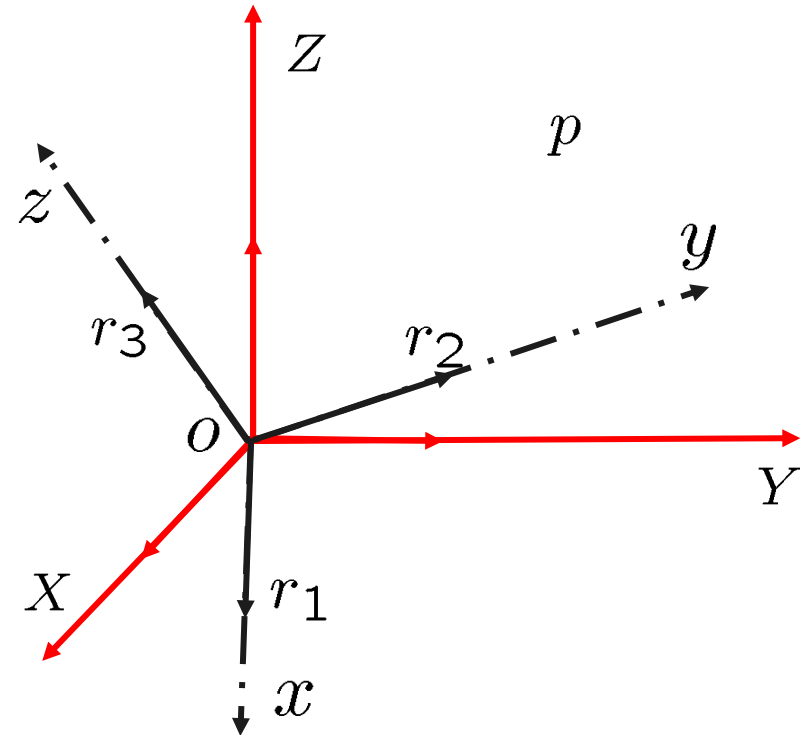


## RIGID-BODY MOTION – Rotation

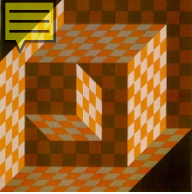
Rotation matrix:

$$R \doteq [r_1, r_2, r_3] \in \mathbb{R}^{3 \times 3}$$

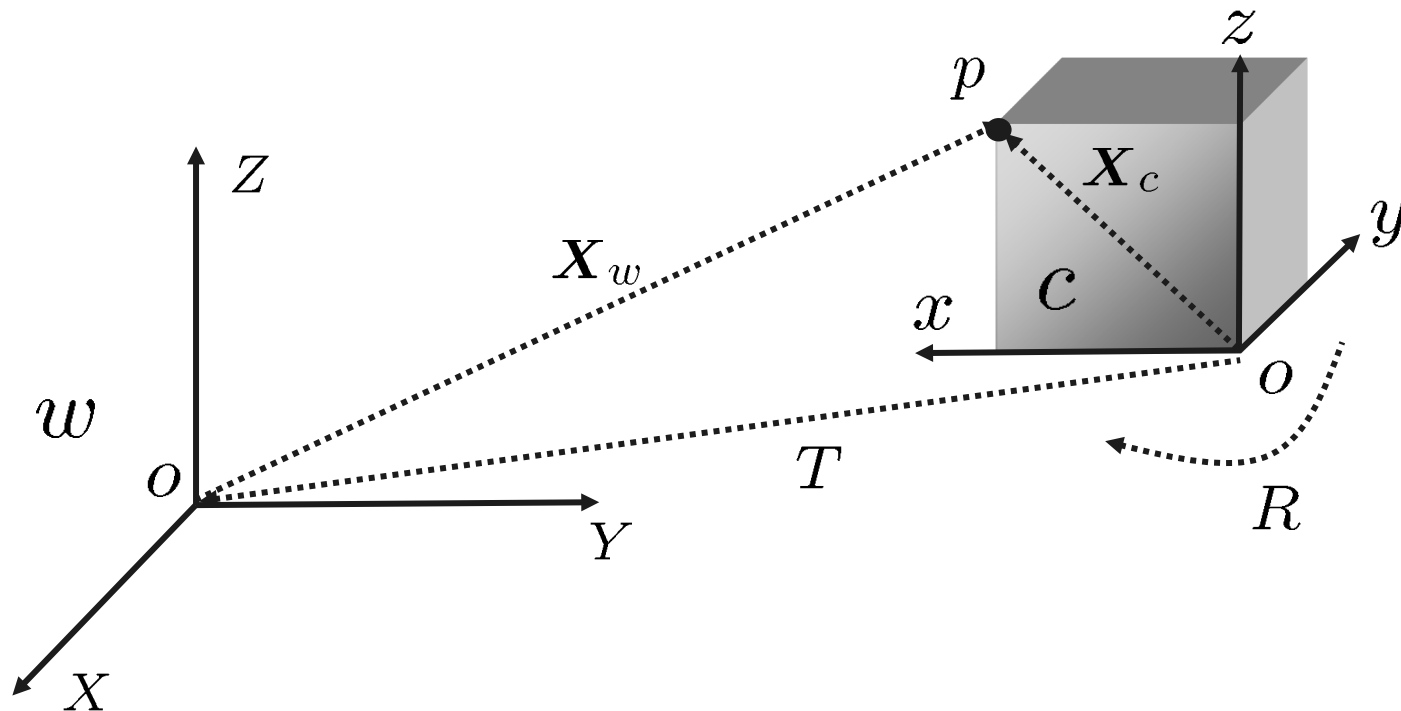
$$R^T R = I, \det(R) = +1$$



Coordinates are related by:  $X_c = R X_w$

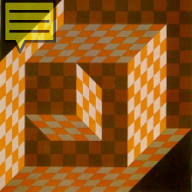


## RIGID-BODY MOTION – Rotation and Translation



Coordinates are related by:  $X_c = R X_w + T$ ,

Velocities are related by:  $\dot{X}_c = \hat{\omega} X_c + v$ .



## RIGID-BODY MOTION - Homogeneous Coordinates

---

3-D coordinates are related by:  $\mathbf{X}_c = R\mathbf{X}_w + T$ ,

Homogeneous coordinates:

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

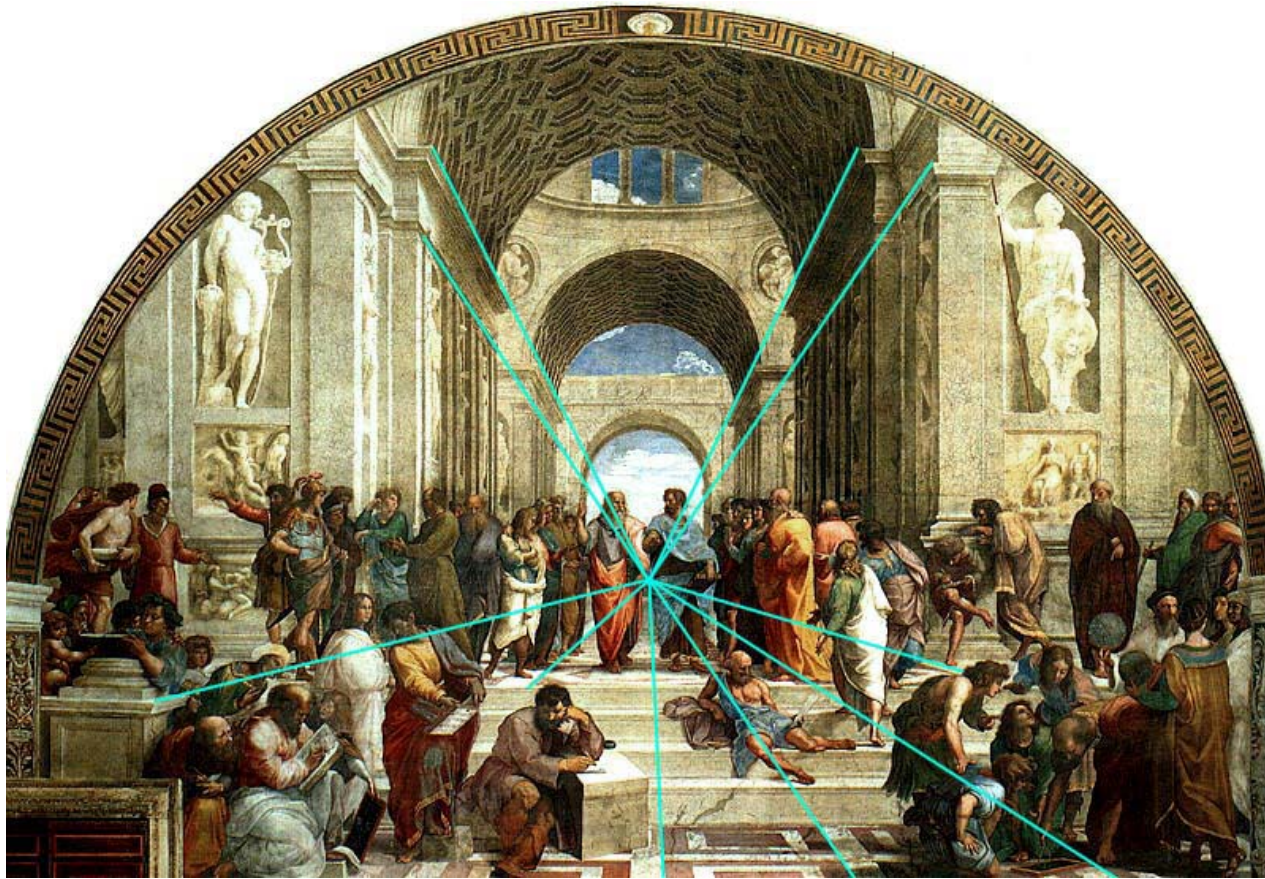
Homogeneous coordinates/velocities are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \\ 1 \end{bmatrix} = \begin{bmatrix} \hat{\omega} & v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$



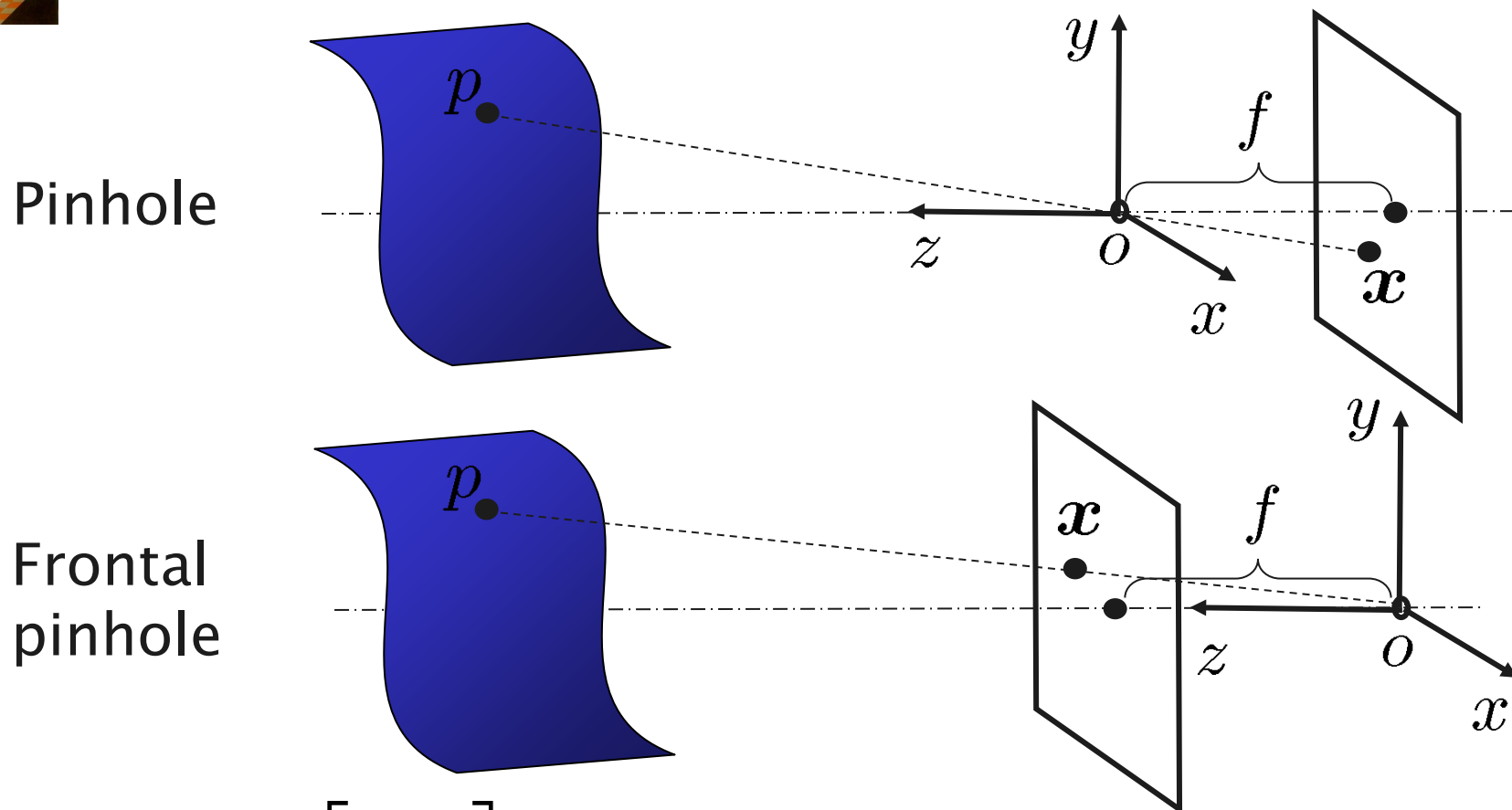
## IMAGE FORMATION – Perspective Imaging

“The Scholar of Athens,” Raphael, 1518





## IMAGE FORMATION – Pinhole Camera Model



$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$



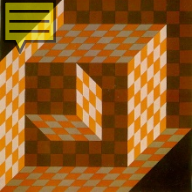
## IMAGE FORMATION – Pinhole Camera Model

$$\text{2-D coordinates } \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Homogeneous coordinates

$$\mathbf{x} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \quad \mathbf{X} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

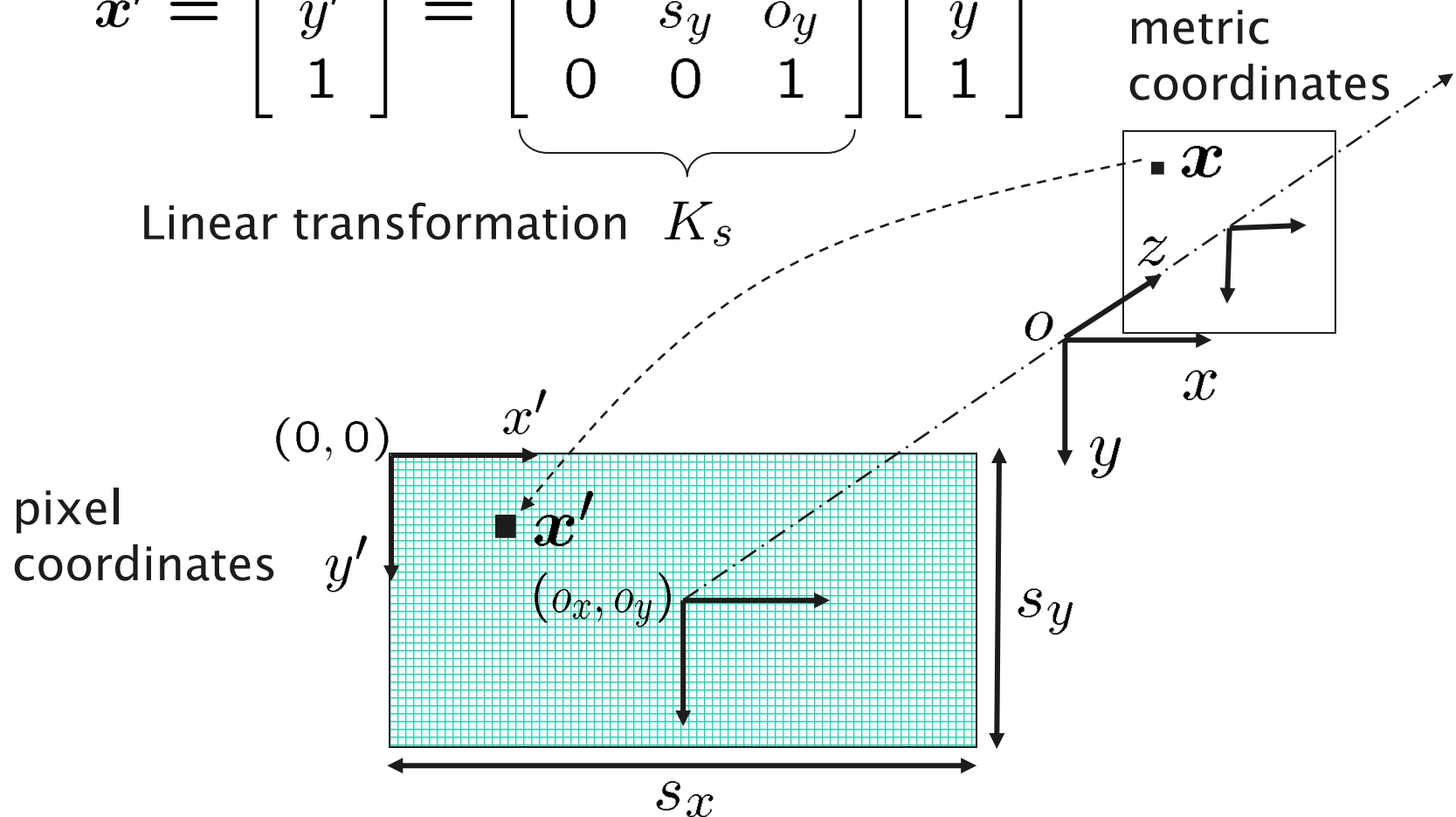
$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

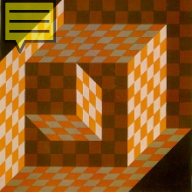


## CAMERA PARAMETERS – Pixel Coordinates

$$\mathbf{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Linear transformation } K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation  $K_s$





## CAMERA PARAMETERS – Calibration Matrix and Camera Model

Pinhole camera

$$\lambda \mathbf{x} = K_f \Pi_0 \mathbf{X}$$

Pixel coordinates

$$\mathbf{x}' = K_s \mathbf{x}$$

$$\lambda \mathbf{x}' = K_s K_f \Pi_0 \mathbf{X} = \underbrace{\begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Calibration matrix  
(intrinsic parameters)

$$K = K_s K_f \quad \Pi_0$$

Projection matrix

$$\Pi = [K, 0] \in \mathbb{R}^{3 \times 4}$$

Camera model

$$\lambda \mathbf{x}' = K \Pi_0 \mathbf{X} = \Pi \mathbf{X}$$



## IMAGE FORMATION – Image of a Point

Homogeneous coordinates of a 3-D point  $p$

$$\mathbf{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$

Homogeneous coordinates of its 2-D image

$$\mathbf{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$$

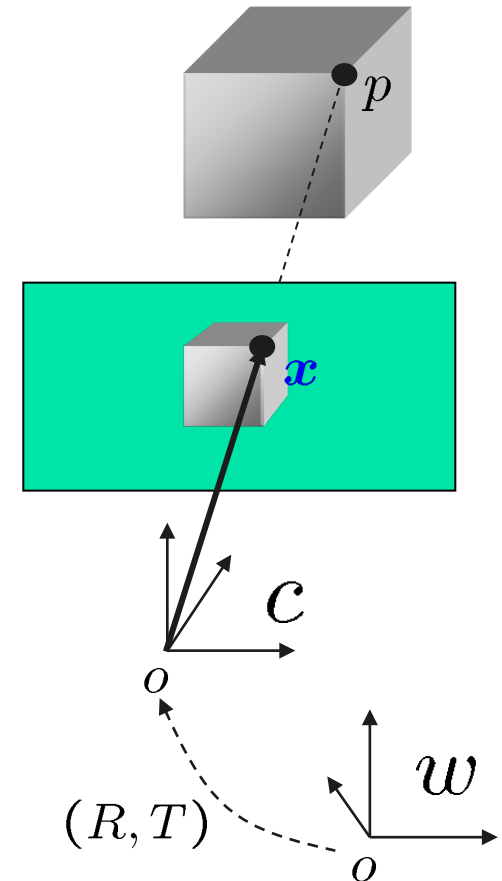
Projection of a 3-D point to an image plane

$$\lambda \mathbf{x} = \Pi \mathbf{X}$$

$$\lambda \in \mathbb{R}, \quad \Pi = [R, T] \in \mathbb{R}^{3 \times 4}$$

$$\lambda \mathbf{x}' = \Pi \mathbf{X}$$

$$\lambda \in \mathbb{R}, \quad \Pi = [KR, KT] \in \mathbb{R}^{3 \times 4}$$





## IMAGE FORMATION – Image of a Line

Homogeneous representation of a 3-D line  $L$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} + \mu \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}, \quad \mu \in \mathbb{R}$$

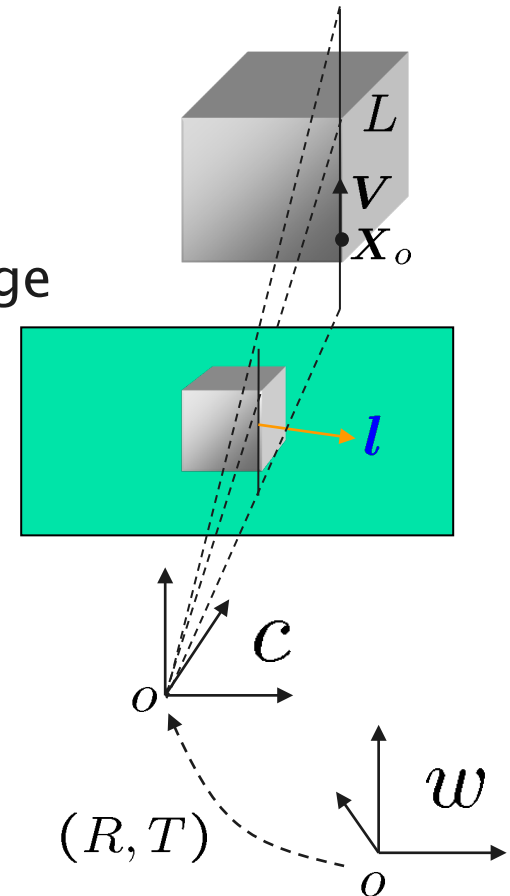
Homogeneous representation of its 2-D image

$$\mathbf{l} = [a, b, c]^T \in \mathbb{R}^3$$

Projection of a 3-D line to an image plane

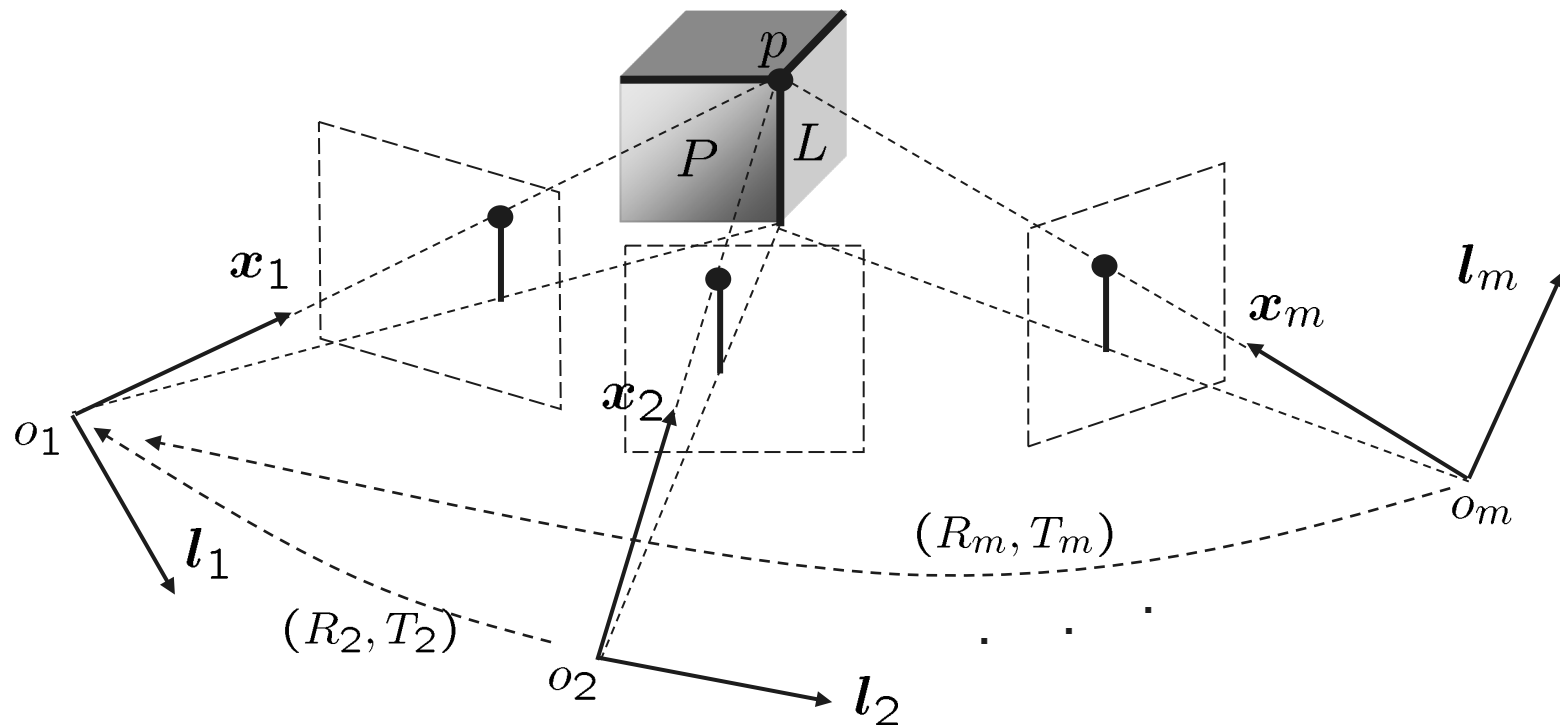
$$\mathbf{l}^T \mathbf{x} = \mathbf{l}^T \Pi \mathbf{X} = 0$$

$$\Pi = [KR, KT] \in \mathbb{R}^{3 \times 4}$$





## SUMMARY OF NOTATION – Multiple Images



1. Images are all “**incident**” at the corresponding features in space;
2. Features in space have many types of **incidence relationships**;
3. Features in space have many types of **metric relationships**.