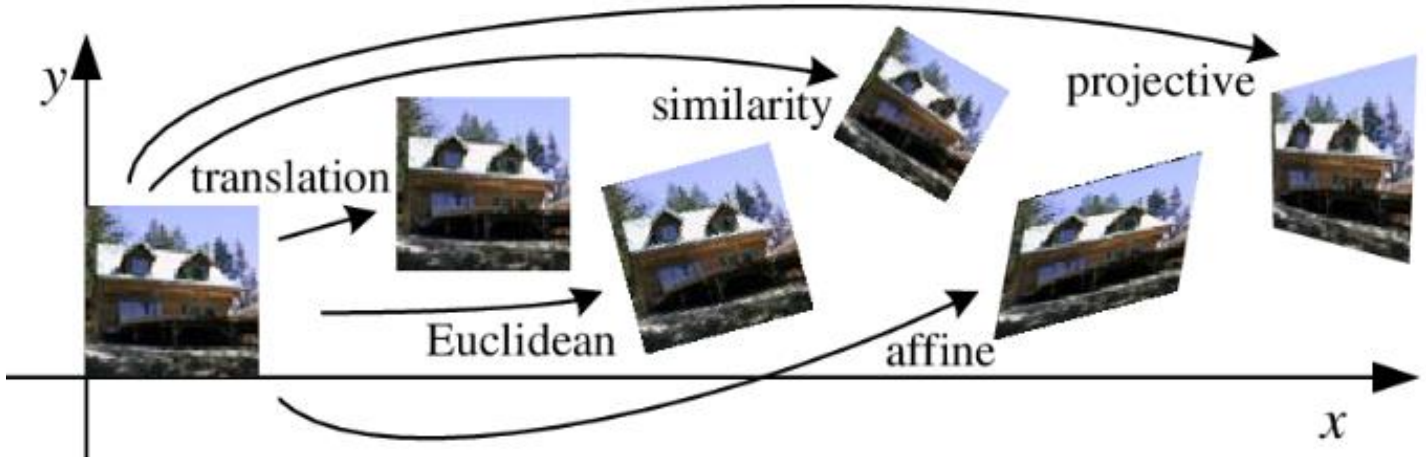
3D Transformations and Camera Calibration

2D Transformations

- The geometric transformations we studies so far apply only for 2D images
- 3D volume images occur in different domain and are specifically important in the medical field (e.g., MRI and CT scans)
- 3D transformations are also used in designing and calibrating cameras where the goal is learning a function that describe how a camera translates a world-view into an image



Review: 2D Transformations

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

translation

$$\begin{bmatrix} \boldsymbol{x}' \\ \boldsymbol{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{s}_{\boldsymbol{x}} & 0 & 0 \\ 0 & \boldsymbol{s}_{\boldsymbol{y}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ 1 \end{bmatrix}$$

scaling

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

rotation

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & \beta_x & 0 \\ \beta_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

shearing

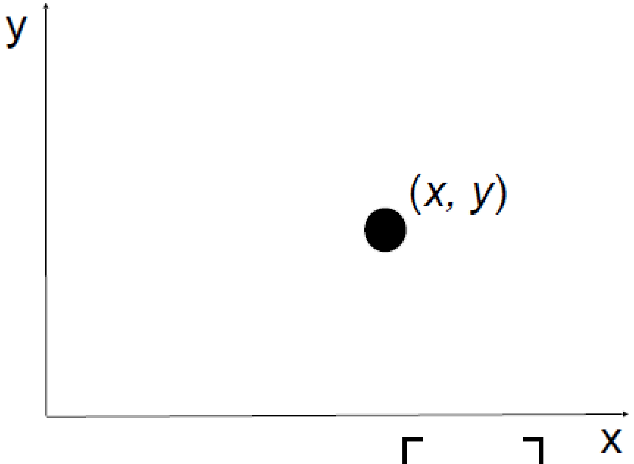
Review: 2D Transformations

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x', y', w') \longrightarrow \left(\frac{x'}{w'}, \frac{y'}{w'}\right)$$

Projective transformation

3D Homogeneous Coordinates

2D

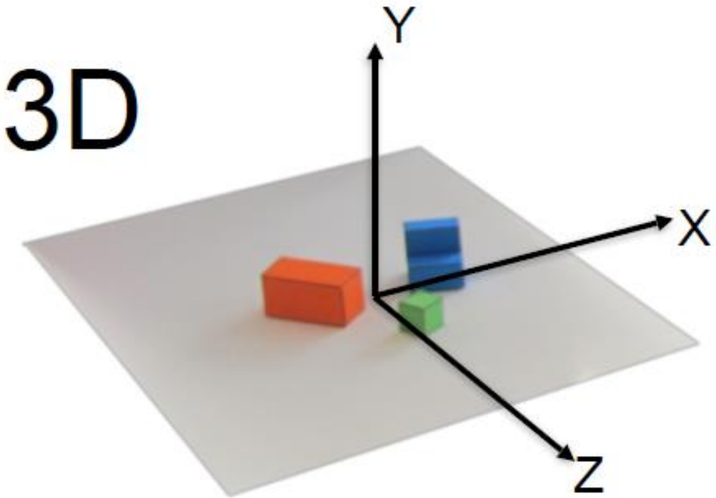


$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

heterogeneous image coordinates

homogeneous image coordinates

3D



$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Homogeneous Coordinates

2D

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

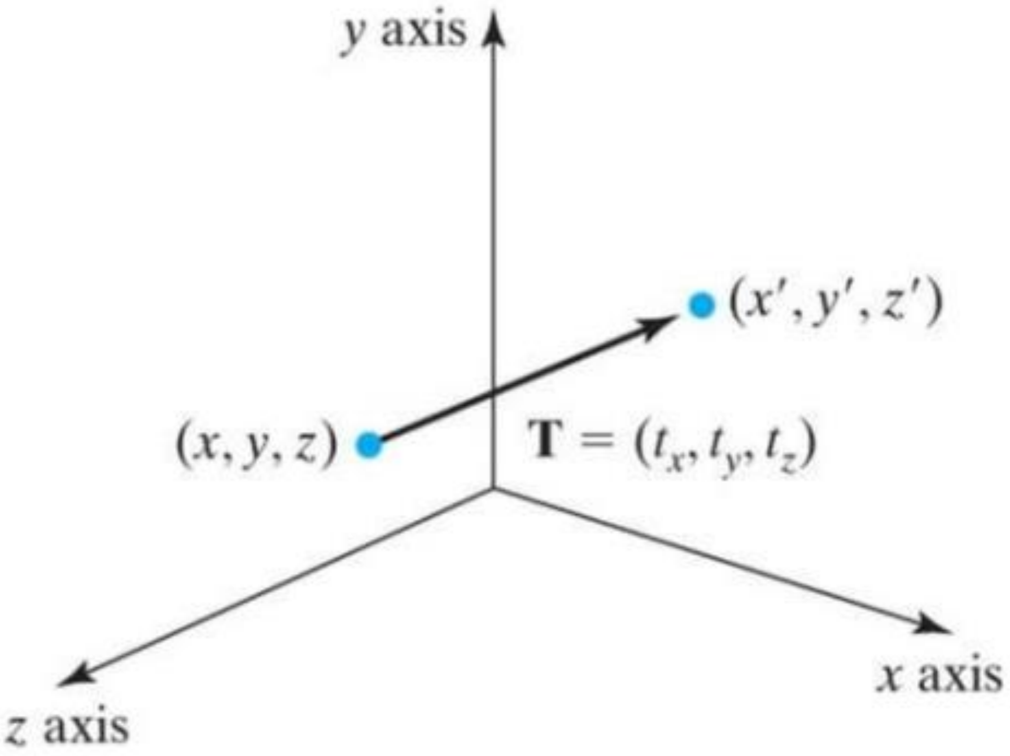
$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

3D

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

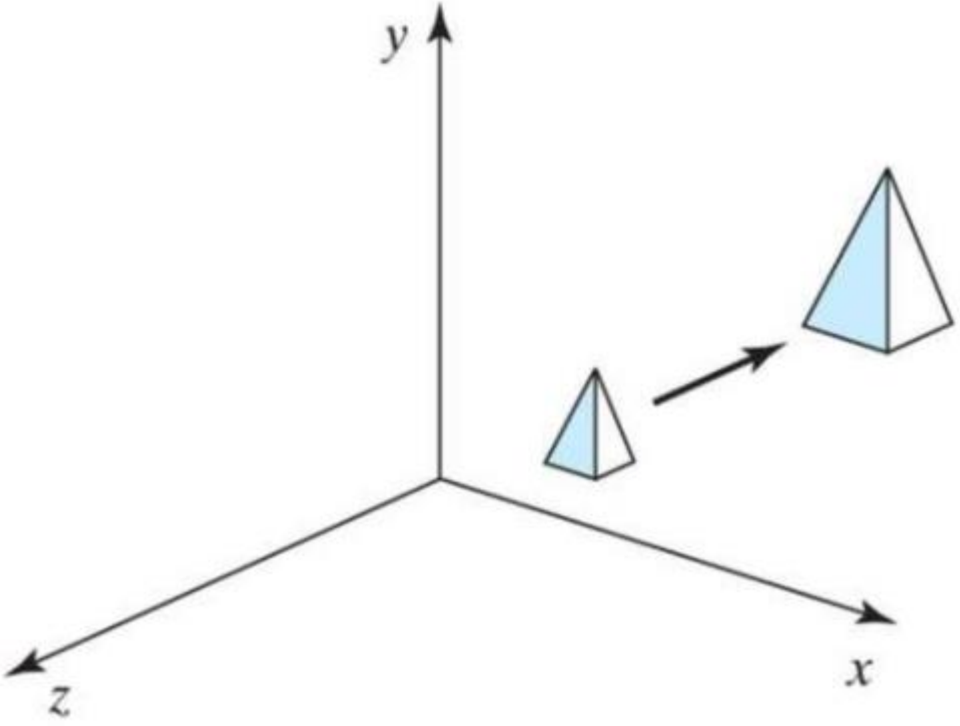
3D Transformations



$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\mathbf{P}' = \mathbf{T} \cdot \mathbf{P}$$

3D Scaling



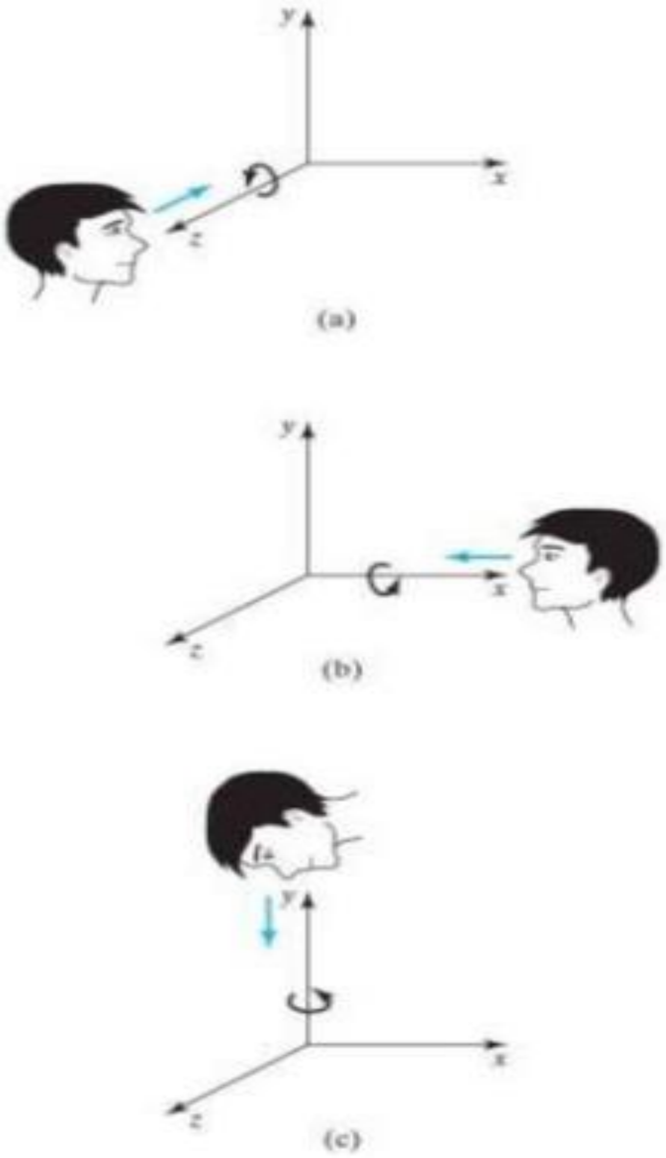
$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\mathbf{P}' = \mathbf{S} \cdot \mathbf{P}$$

3D Rotations

- Three rotation angles

We define rotations are counterclockwise
When looking from the positive axis direction

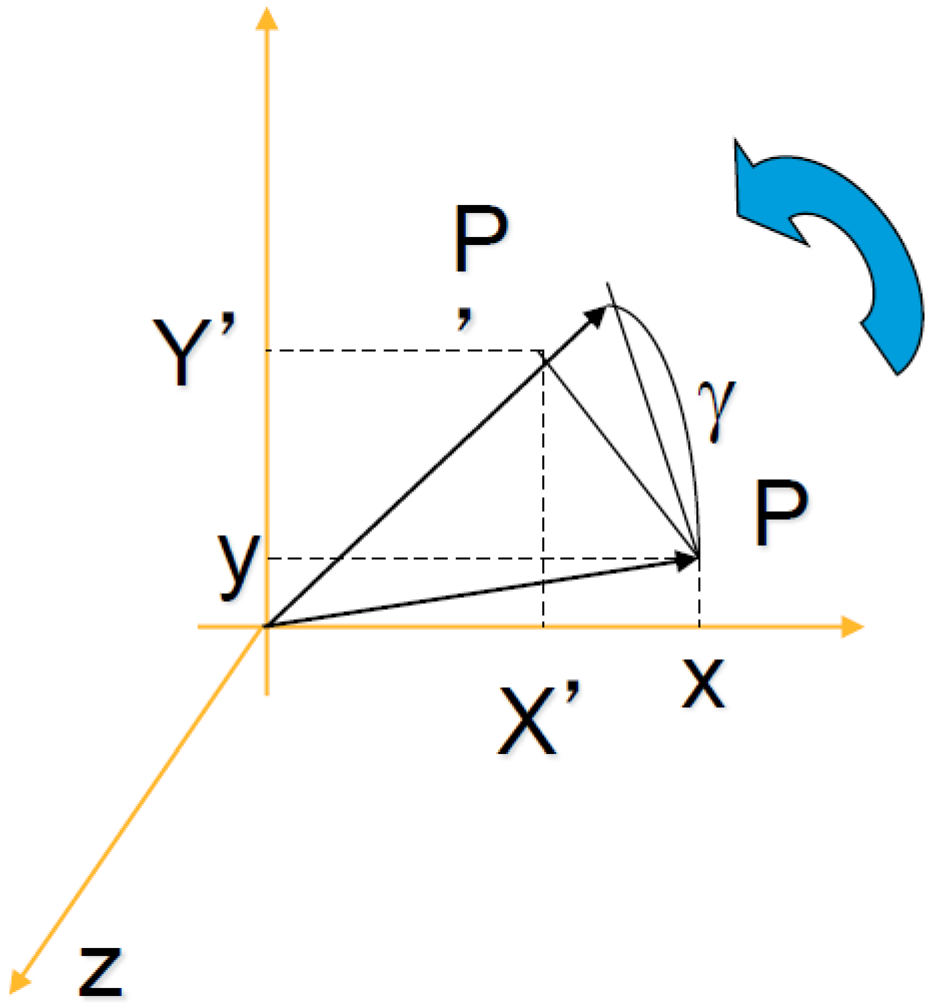


3D Rotation Matrices

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

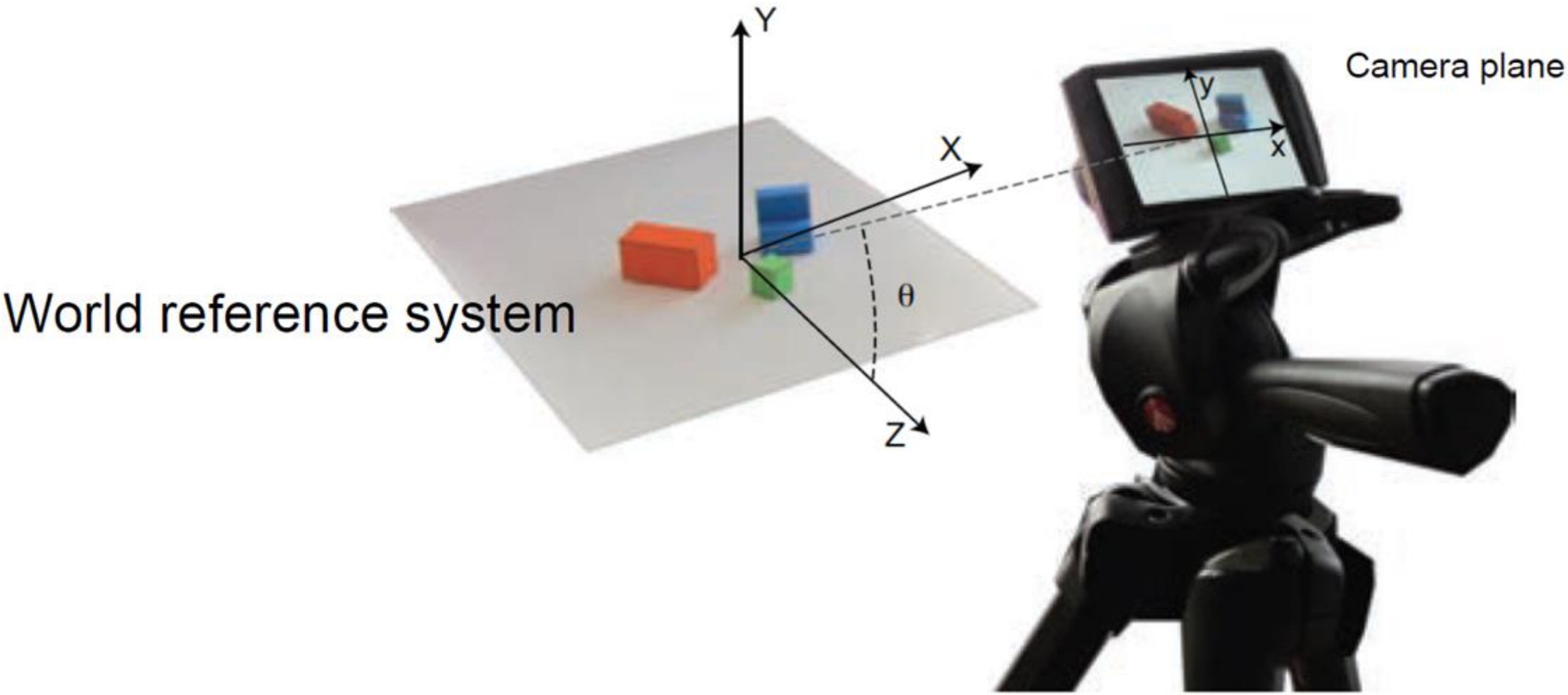
$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Perspective Projection

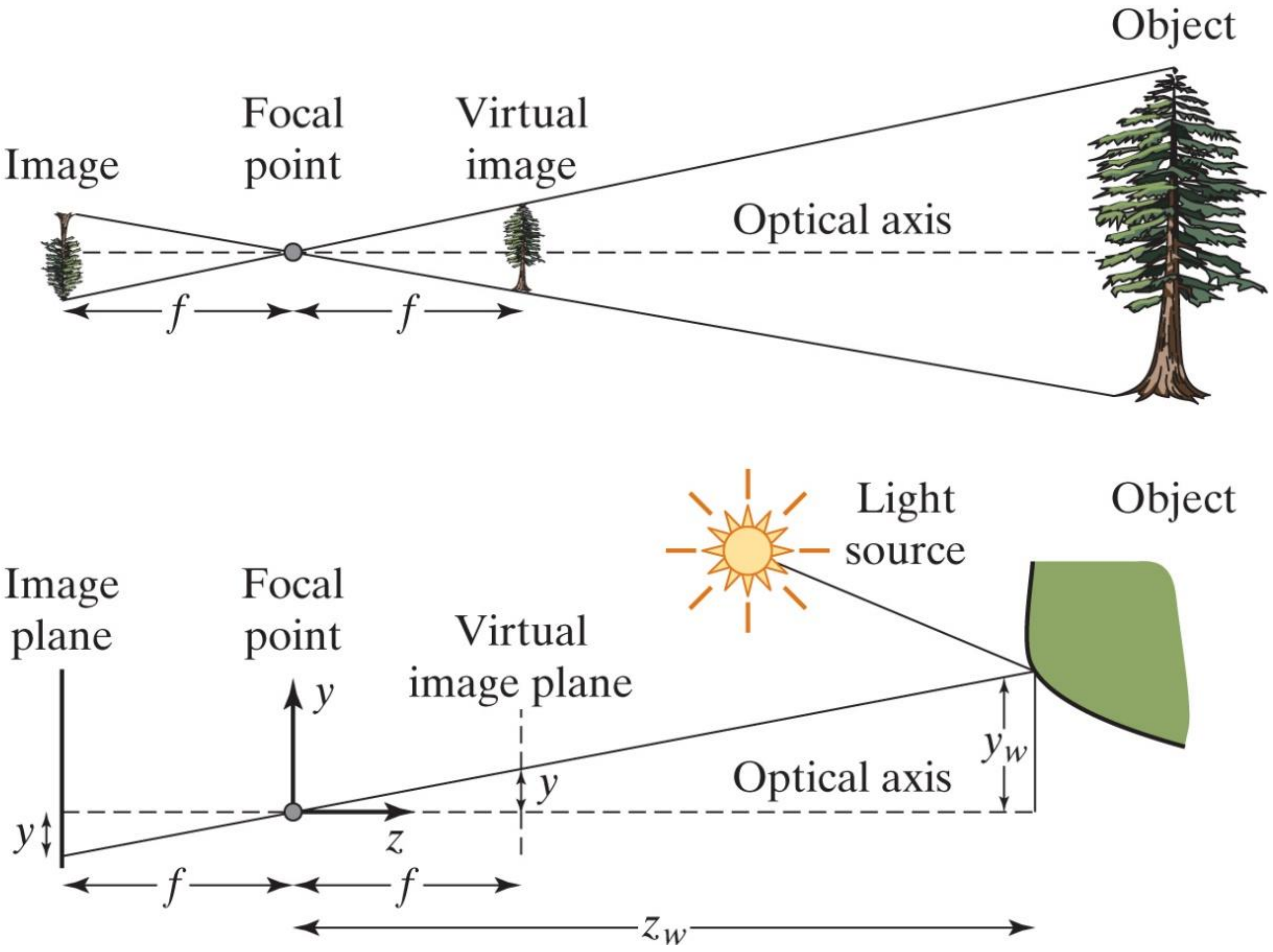
- Related to 3D transformations is camera calibration where we try to find the parameters that move data from world coordinates (measured in units of length like meters) to image coordinates (measured in pixels)
- These transformations take a 3D input points (X, Y, Z) and output a 2D projection (x, y)



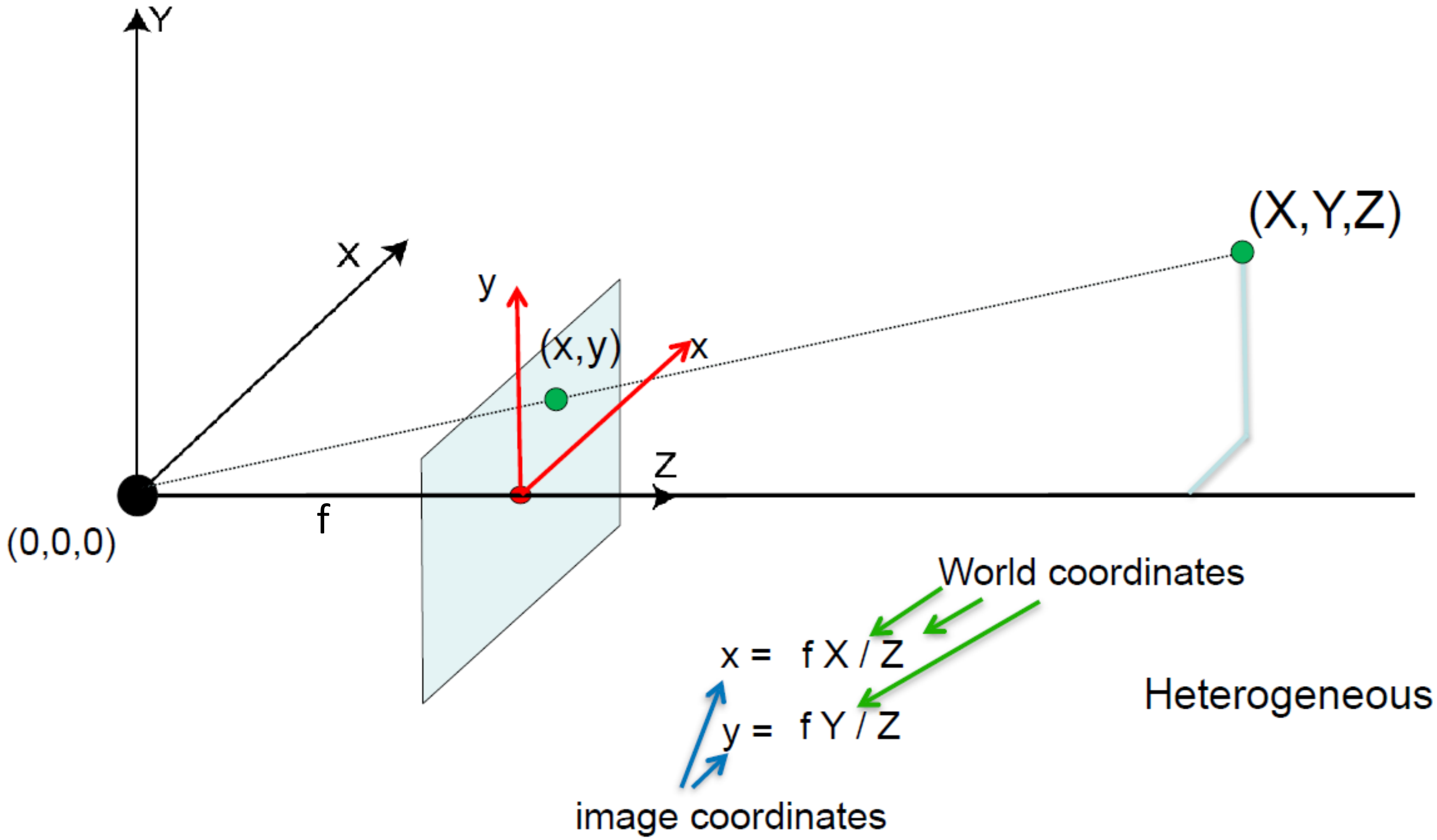
Y
X
Z
θ
World reference system
y
x
Camera plane

Focal length and Aperture

Figure 2.23 Perspective projection caused by a pinhole camera, showing the focal point (pinhole), image plane, focal length, and optical axis. The light rays emitted by the light source reflect off the surface in the world and pass through the aperture to form an upside-down image on the image plane. This is mathematically equivalent to producing a rightside-up image on the virtual image plane in front of the focal point.



Perspective Projection



Heterogeneous coordinates

Perspective Projection

Heterogeneous coordinates

World coordinates

$$x = f X / Z$$

$$y = f Y / Z$$

image coordinates

Homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \overset{?}{\longleftrightarrow} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Perspective Projection

Heterogeneous coordinates

World coordinates

$x = fX/Z$

$y = fY/Z$

image coordinates

Homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Perspective Projection

Heterogeneous coordinates

World coordinates

$$x = f X / Z$$

$$y = f Y / Z$$

image coordinates

Homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z/f \end{bmatrix}$$

Perspective Projection

Heterogeneous coordinates

World coordinates

$$x = f X / Z$$

$$y = f Y / Z$$

image coordinates

Homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z/f \end{bmatrix}$$

Going back to heterogeneous coordinates:

$$\begin{bmatrix} X \\ Y \\ Z/f \end{bmatrix} \longrightarrow (f\,X/Z,\ f\,Y/Z)$$

Two Equivalent Representations

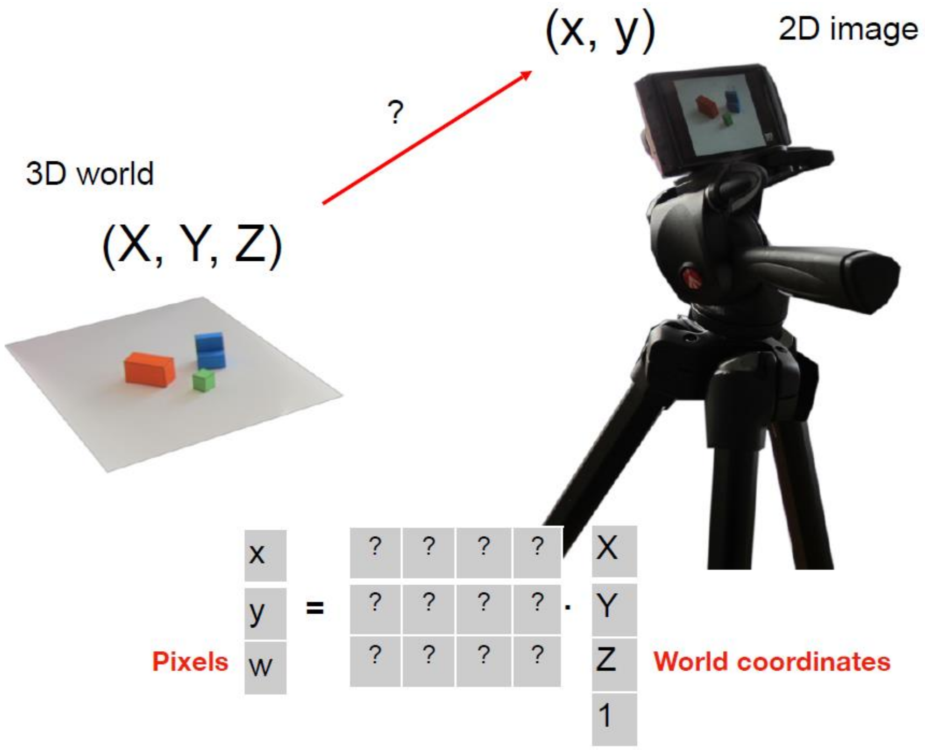
$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z/f \end{bmatrix}$$

$$(f\, X/Z,\ f\, Y/Z)$$

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$

Camera Calibration

- Camera calibration is the process of identifying the parameters of the camera that are responsible for converting the 3D scene into a 2D image



Camera Calibration

- The goal in camera calibration is often to find the intrinsic properties. A widely used algorithm is Zhang's calibration algorithm.

$$\underbrace{\begin{bmatrix} x \\ y \\ w \end{bmatrix}}_{\substack{\text{2D Image} \\ \text{Coordinates}}} = \underbrace{\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\substack{\text{Intrinsic properties} \\ \text{(Optical Centre, scaling)}}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}}_{\substack{\text{Extrinsic properties} \\ \text{(Camera Rotation} \\ \text{and translation)}}} \underbrace{\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}}_{\substack{\text{3D World} \\ \text{Coordinates}}}$$

Camera Calibration