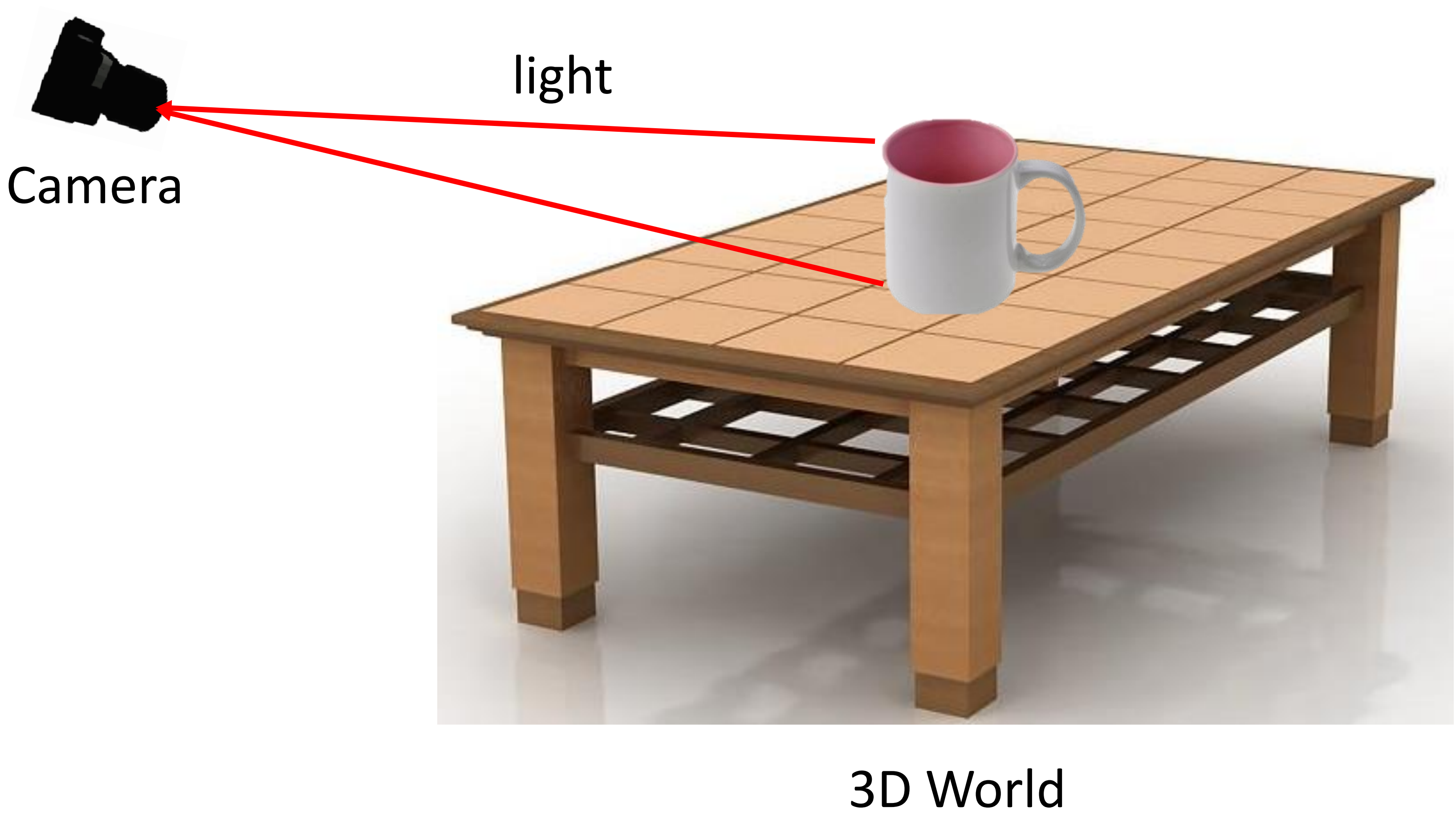
Geometric Primitives and Transformations

CS 4391 Introduction Computer Vision

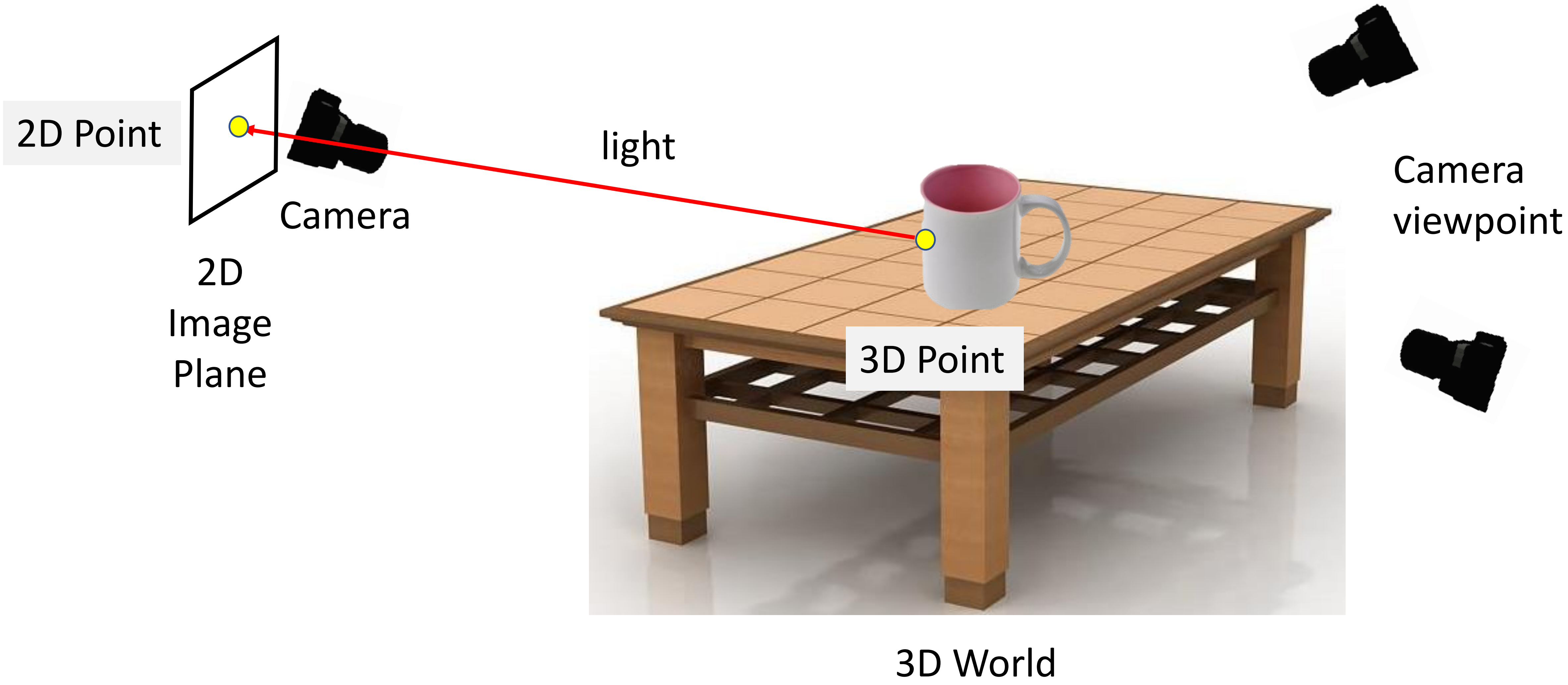
Professor Yu Xiang

The University of Texas at Dallas

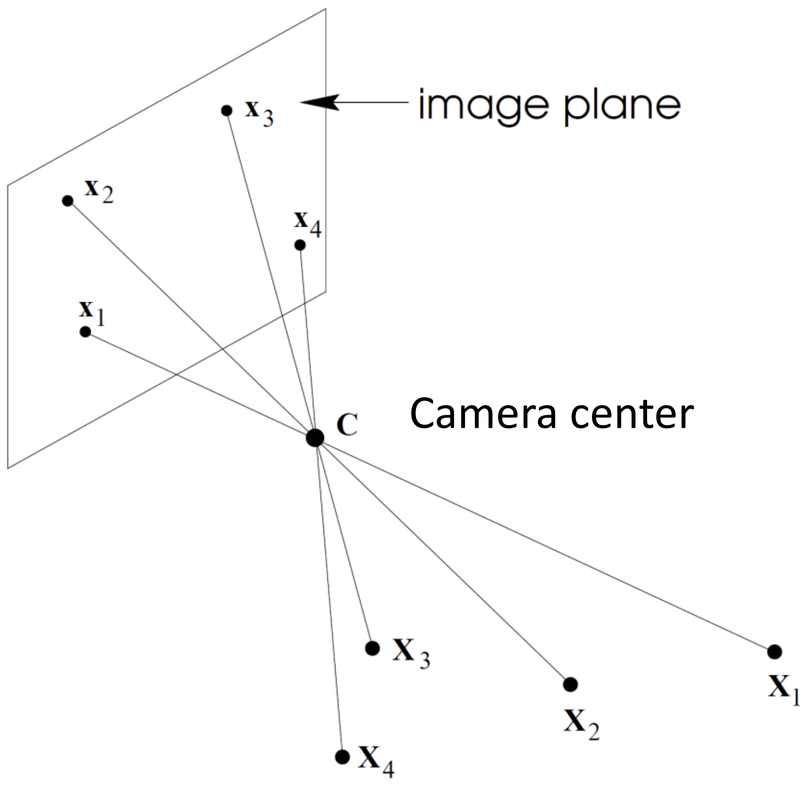
How are Images Generated?



Geometry in Image Generation



2D Points and 3D Points



- A 2D point is usually used to indicate pixel coordinates of a pixel

$$\mathbf{x} = (x, y) \in \mathcal{R}^2 \qquad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

- A 3D point in the real world

$$\mathbf{X} = (x, y, z) \in \mathcal{R}^3 \qquad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Homogeneous Coordinates

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

homogeneous image coordinates

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

homogeneous scene coordinates

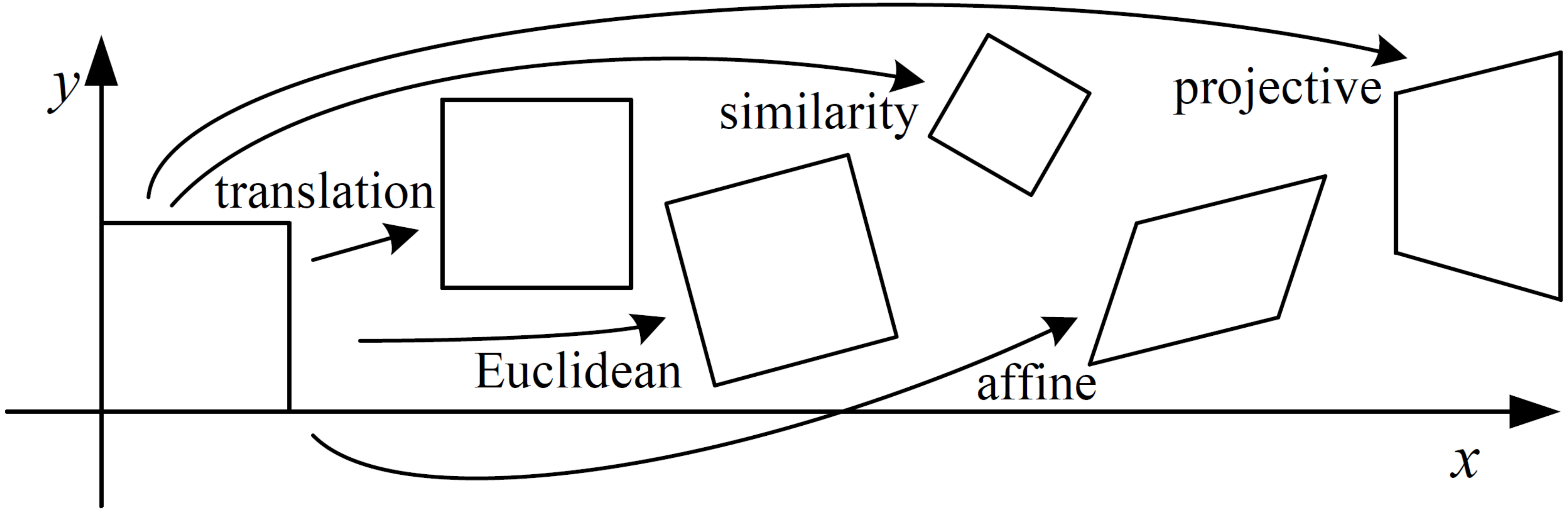
Up to scale

Conversion

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

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

2D Transformations



2D Translation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

$$\mathbf{x}' = \mathbf{x} + \mathbf{t}$$

$$\mathbf{x}' = \begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}}$$

$$2 \times 3$$

augmented vector $\bar{\mathbf{x}} = (x, y, 1)$

Homogeneous coordinate

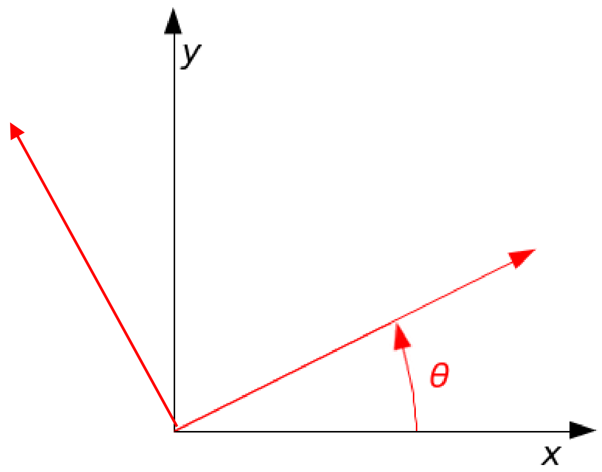
$$\bar{\mathbf{x}}' = \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \bar{\mathbf{x}}$$

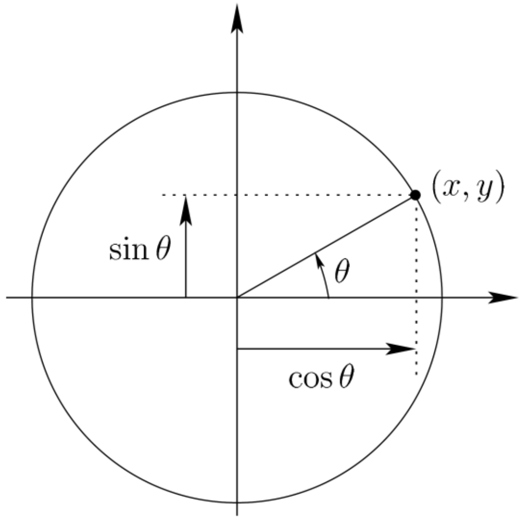
$$3 \times 3$$

2D Rotation

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

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$





$$[\hat{\mathrm{x}}_\mathrm{b} \ \ \hat{\mathrm{y}}_\mathrm{b}] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

orthonormal rotation matrix

$$\mathbf{R}\mathbf{R}^T = \mathbf{I} \text{ and } |\mathbf{R}| = 1$$

2D Euclidean Transformation

- 2D Rotation + 2D translation

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t} \qquad \mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$x' = x\cos\theta - y\sin\theta$$
$$y' = x\sin\theta + y\cos\theta$$

2D Euclidean Transformation

- 2D Rotation + 2D translation

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t} \qquad \mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\mathbf{x}' = \underbrace{\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}}_{2 \times 3} \bar{\mathbf{x}}$$

$$\bar{\mathbf{x}} = (x, y, 1)$$

- Degree of freedom (DOF)
  - The maximum number of logically independent values
  - 2D Rotation?
  - 2D Euclidean transformation?

2D Similarity Transformation

- Scaled 2D rotation + 2D translation

$$\mathbf{x}' = s\mathbf{R}\mathbf{x} + \mathbf{t} \qquad \mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\mathbf{x}' = \begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}} = \begin{bmatrix} a & -b & t_x \\ b & a & t_y \end{bmatrix} \bar{\mathbf{x}} \qquad \bar{\mathbf{x}} = (x, y, 1)$$

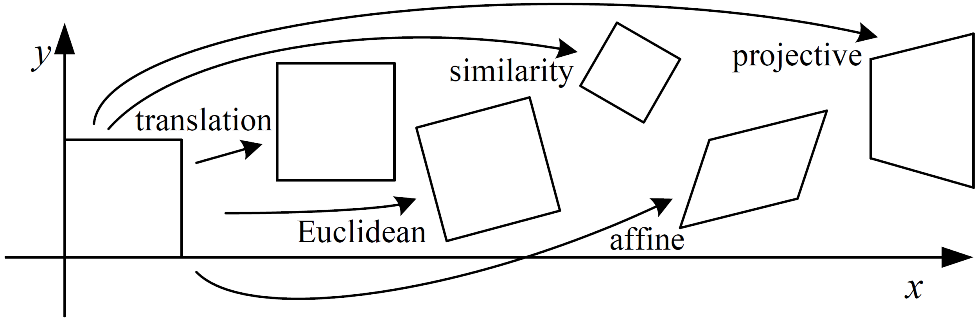
The similarity transform preserves angles between lines.

2D Affine Transformation

- Arbitrary 2x3 matrix

$$\mathbf{x}' = \mathbf{A}\bar{\mathbf{x}} \qquad \bar{\mathbf{x}} = (x, y, 1)$$

$$\mathbf{x}' = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \end{bmatrix} \bar{\mathbf{x}}$$



Parallel lines remain parallel under affine transformations.

2D Projective Transformation

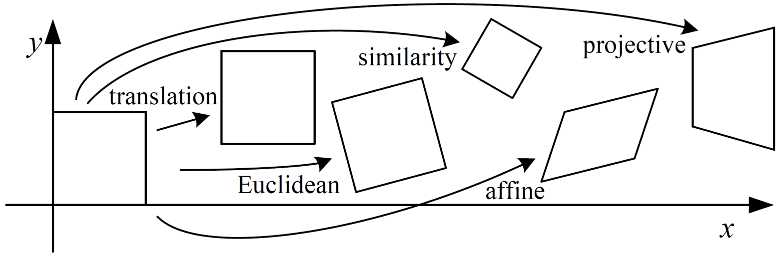
- Also called perspective transform or homography

$$\tilde{\mathbf{x}}' = \tilde{\mathbf{H}}\tilde{\mathbf{x}}$$ homogeneous coordinates

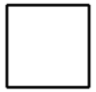
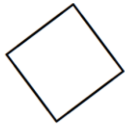
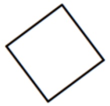
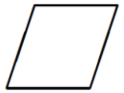
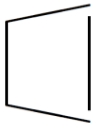
3×3 $\tilde{\mathbf{H}}$ is only defined up to a scale

$$x' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}} \quad \text{and} \quad y' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}}$$

Perspective transformations preserve straight lines



Hierarchy of 2D Transformations

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \mathbf{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\mathbf{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	

3D Translation

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

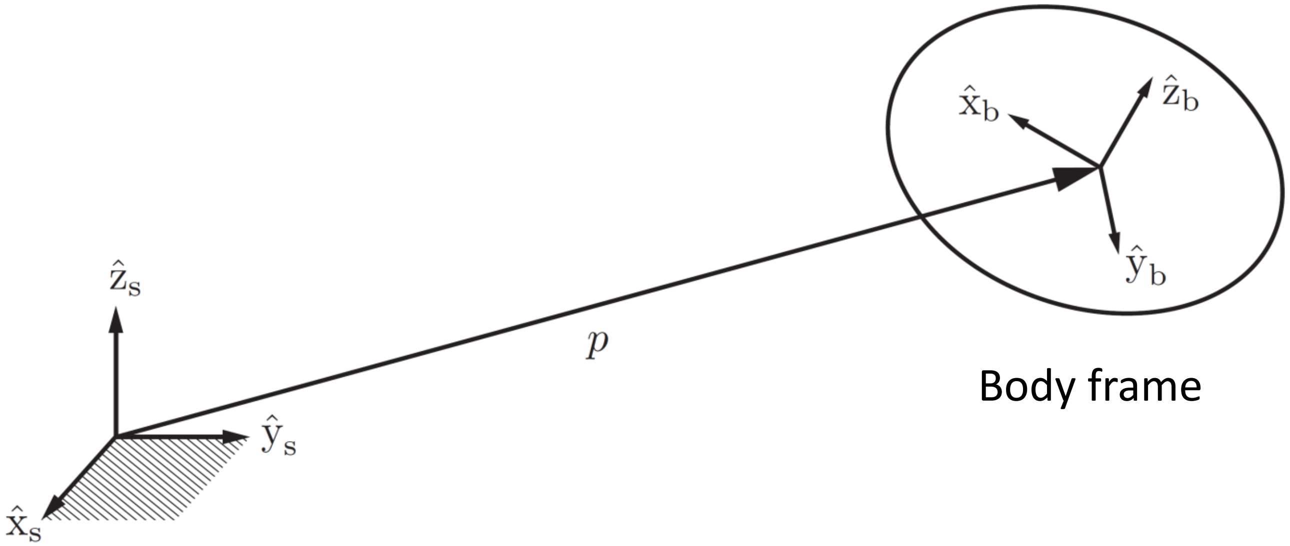
$$\mathbf{x}' = \mathbf{x} + \mathbf{t}$$

$$\mathbf{x}' = \begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}}$$

$$3 \times 4$$

augmented vector $\bar{\mathbf{x}} = (x, y, z, 1)$

3D Rotation



Body frame

Fixed frame

- Axes of the body frame

$$
\begin{aligned}
\hat{x}_b &= r_{11}\hat{x}_s + r_{21}\hat{y}_s + r_{31}\hat{z}_s, \\
\hat{y}_b &= r_{12}\hat{x}_s + r_{22}\hat{y}_s + r_{32}\hat{z}_s, \\
\hat{z}_b &= r_{13}\hat{x}_s + r_{23}\hat{y}_s + r_{33}\hat{z}_s.
\end{aligned}
$$

$$
R = [\hat{x}_b \ \ \hat{y}_b \ \ \hat{z}_b] = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}
$$

Write as column vectors

Rotation matrix

We will focus on 3D rotations in next lectures.

3D Euclidean Transformation SE(3)

- 3D Rotation + 3D translation

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}$$

$$\mathbf{x}' = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}}$$

$$3 \times 4$$

$$\bar{\mathbf{x}} = (x, y, z, 1)$$

orthonormal rotation matrix

$$\mathbf{R}\mathbf{R}^T = \mathbf{I} \text{ and } |\mathbf{R}| = 1$$

$$3 \times 3$$

3D Similarity Transformation

- Scaled 3D rotation + 3D translation

$$\mathbf{x}' = s\mathbf{R}\mathbf{x} + \mathbf{t}$$

$$\mathbf{x}' = \underset{3 \times 4}{\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}} \bar{\mathbf{x}} \qquad \bar{\mathbf{x}} = (x, y, z, 1)$$

This transformation preserves angles between lines and planes.

3D Affine Transformation

$$\mathbf{x}' = \mathbf{A}\bar{\mathbf{x}} \qquad \bar{\mathbf{x}} = (x, y, z, 1)$$

$$\mathbf{x}' = \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \end{bmatrix} \bar{\mathbf{x}}$$

$$3 \times 4$$

Parallel lines and planes remain parallel under affine transformations.

3D Projective Transformation

- Also called 3D perspective transform or homography

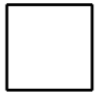
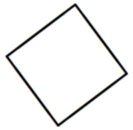
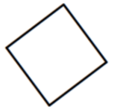
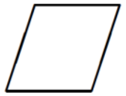
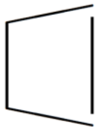
$$\tilde{\mathbf{x}}' = \tilde{\mathbf{H}}\tilde{\mathbf{x}}$$

homogeneous coordinates

4×4 $\tilde{\mathbf{H}}$ is only defined up to a scale

- Perspective transformations preserve straight lines

3D Transformations

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix}_{3\times 4}$	3	orientation	
rigid (Euclidean)	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}_{3\times 4}$	6	lengths	
similarity	$\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}_{3\times 4}$	7	angles	
affine	$\begin{bmatrix} \mathbf{A} \end{bmatrix}_{3\times 4}$	12	parallelism	
projective	$\begin{bmatrix} \tilde{\mathbf{H}} \end{bmatrix}_{4\times 4}$	15	straight lines	

Further Reading

- Section 2.1, Computer Vision, Richard Szeliski
- Chapter 2 and 3, Multiple View Geometry in Computer Vision, Richard Hartley and Andrew Zisserman