THE UNIVERSITY OF TEXAS AT DALLAS

Geometric Primitives and Transformations

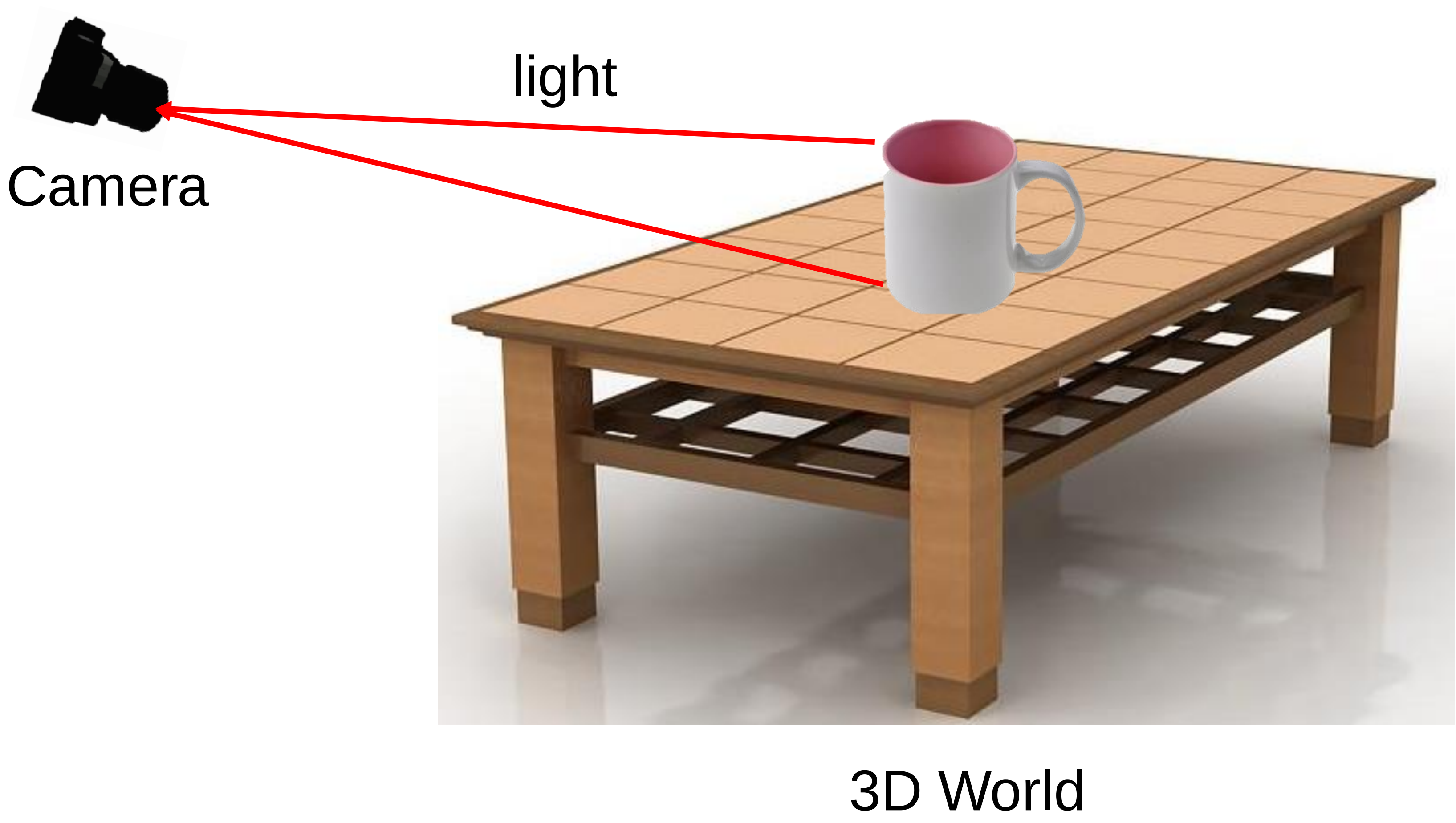
CS 6384 Computer Vision

Professor Yapeng Tian

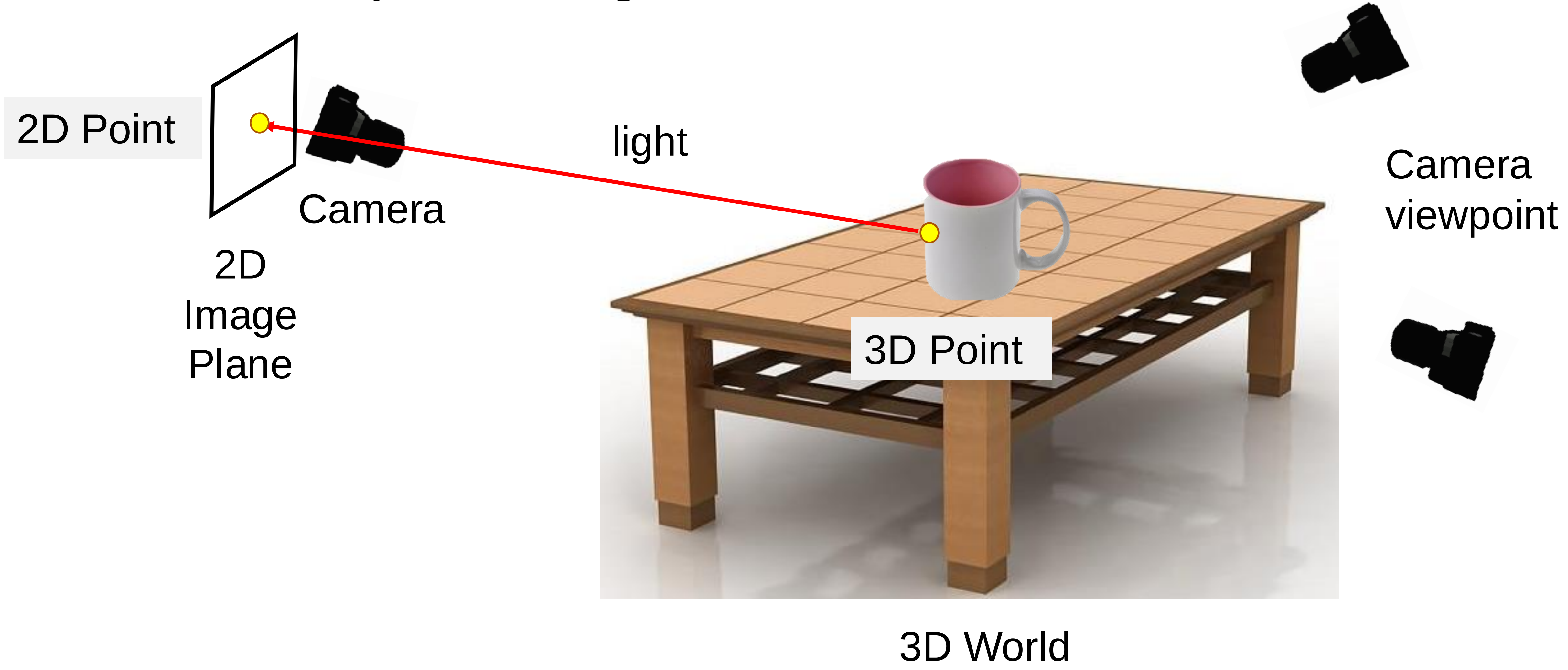
Department of Computer Science

Slides borrowed from Professor Yu Xiang

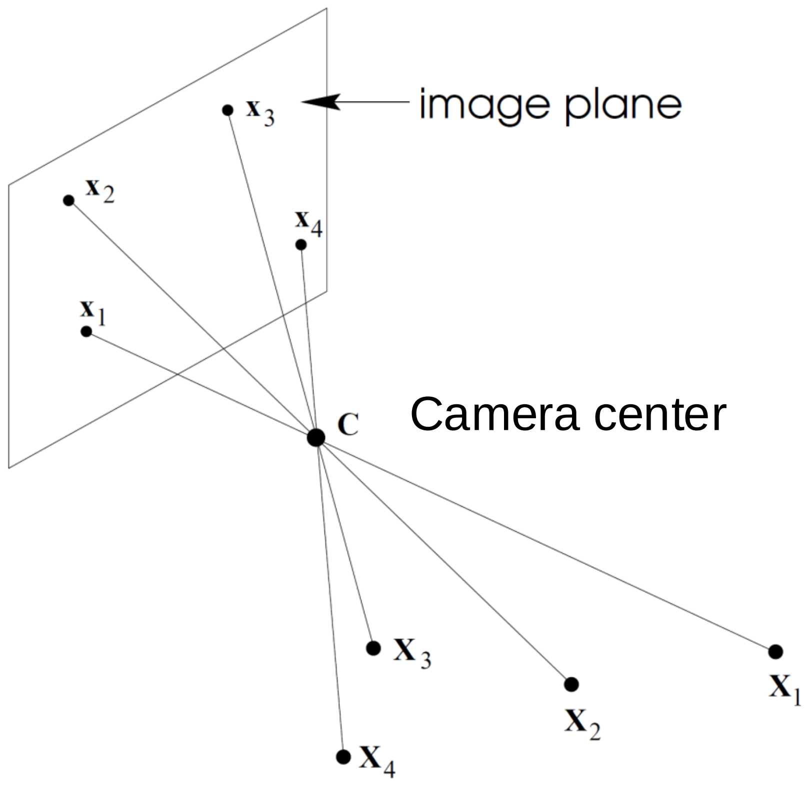
How are Images Generated?



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

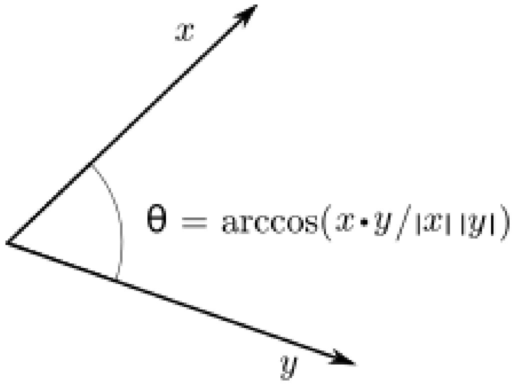
$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

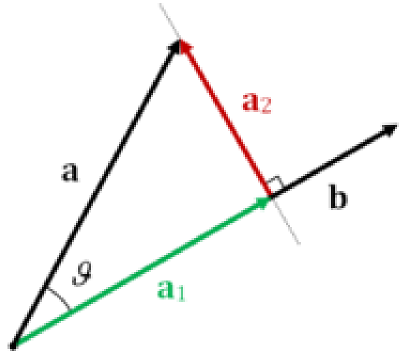
$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Vector Inner Product

Dot product

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \ \|\mathbf{b}\| \cos\theta$$





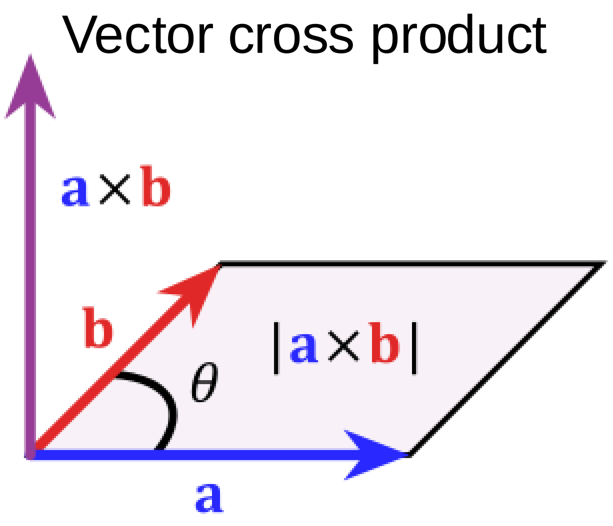
Vector Projection

$$a_1 = \|\mathbf{a}\| \cos\theta = \|\mathbf{a}\| \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \ \|\mathbf{b}\|} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|}$$

$$\mathbf{a}_1 = a_1 \hat{\mathbf{b}} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|} \frac{\mathbf{b}}{\|\mathbf{b}\|}$$

https://en.wikipedia.org/wiki/Dot_product

Vector Cross Product



$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \sin(\theta) \, \mathbf{n}$$

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}$$
$$= (a_2 b_3 - a_3 b_2)\mathbf{i} - (a_1 b_3 - a_3 b_1)\mathbf{j} + (a_1 b_2 - a_2 b_1)\mathbf{k}$$

https://en.wikipedia.org/wiki/Cross_product

2D Lines

A line in a 2D plane $ax + by + c = 0$ $\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

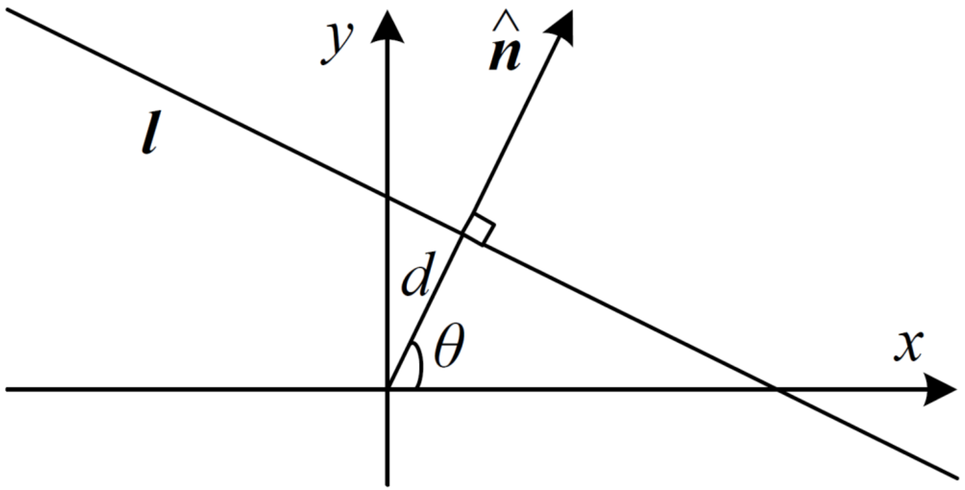
It is parameterized by $\mathbf{l} = (a, b, c)^T$ Homogeneous Coordinates

$k(a, b, c)^T$ represents the same line for nonzero k

Line equation

$$\mathbf{x}^T \mathbf{l} = 0 \qquad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

2D Lines



$$\mathbf{l} = (a, b, c)$$

Normalize by $\sqrt{a^2 + b^2}$

$$\mathbf{l} = (\hat{n}_x, \hat{n}_y, d) = (\hat{\mathbf{n}}, d)$$

Normal vector $\|\hat{\mathbf{n}}\| = 1$

Distance to the origin d

$$\hat{\mathbf{n}} = (\hat{n}_x, \hat{n}_y) = (\cos\theta, \sin\theta)$$

polar coordinates (θ, d)

Intersection of 2D Lines

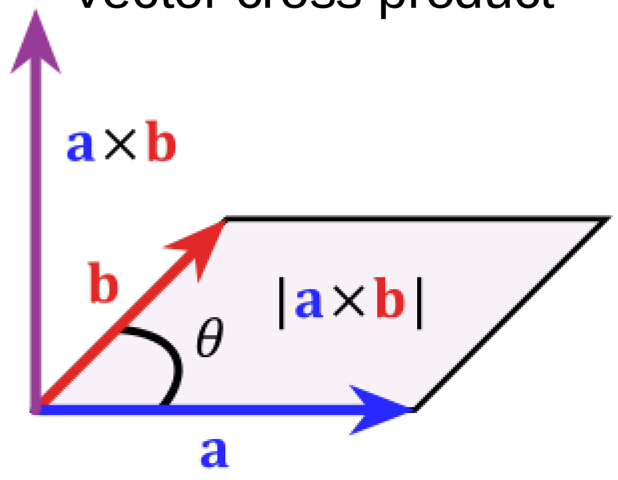
$$\mathbf{l} = (a, b, c)^T \qquad \mathbf{l}' = (a', b', c')^T$$

The intersection is $\mathbf{x} = \mathbf{l} \times \mathbf{l}'$

$$\mathbf{l} \cdot (\mathbf{l} \times \mathbf{l}') = \mathbf{l}' \cdot (\mathbf{l} \times \mathbf{l}') = 0$$

$$\mathbf{l}^T \mathbf{x} = \mathbf{l}'^T \mathbf{x} = 0$$

Vector cross product



$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \sin(\theta) \, \mathbf{n}$$

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Vector dot product

A scalar $\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta$

A Line Joining two Points

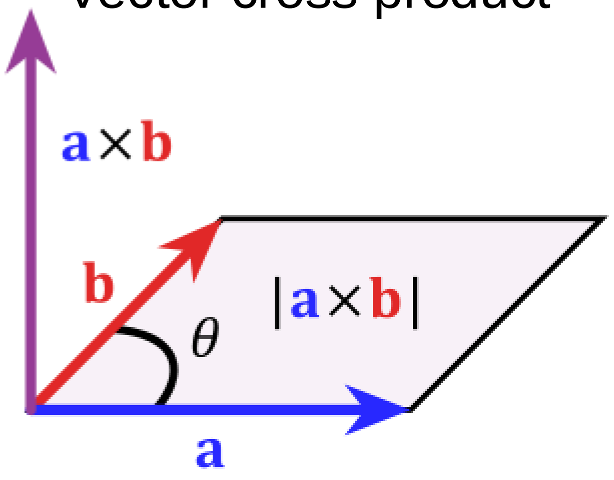
$$\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \mathbf{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

$$\mathbf{l} = \mathbf{x} \times \mathbf{x}'$$

$$\mathbf{x} \cdot (\mathbf{x} \times \mathbf{x}') = \mathbf{x}' \cdot (\mathbf{x} \times \mathbf{x}') = 0$$

$$\mathbf{x}^T \mathbf{l} = \mathbf{x}'^T \mathbf{l} = 0$$

Vector cross product



$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \sin(\theta) \, \mathbf{n}$$

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Vector dot product

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta$$

3D Plane

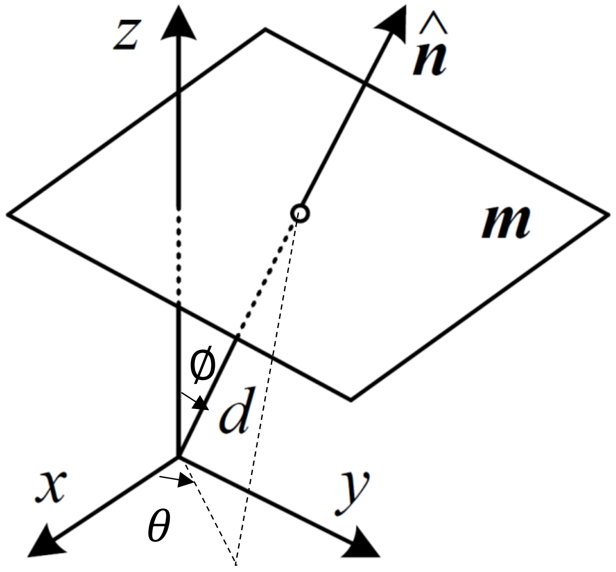
A 3D plane equation $ax + by + cz + d = 0$

It is parameterized by (a, b, c, d)

Normal vector and distance

$$\mathbf{m} = (\hat{n}_x, \hat{n}_y, \hat{n}_z, d) = (\hat{\mathbf{n}}, d)$$

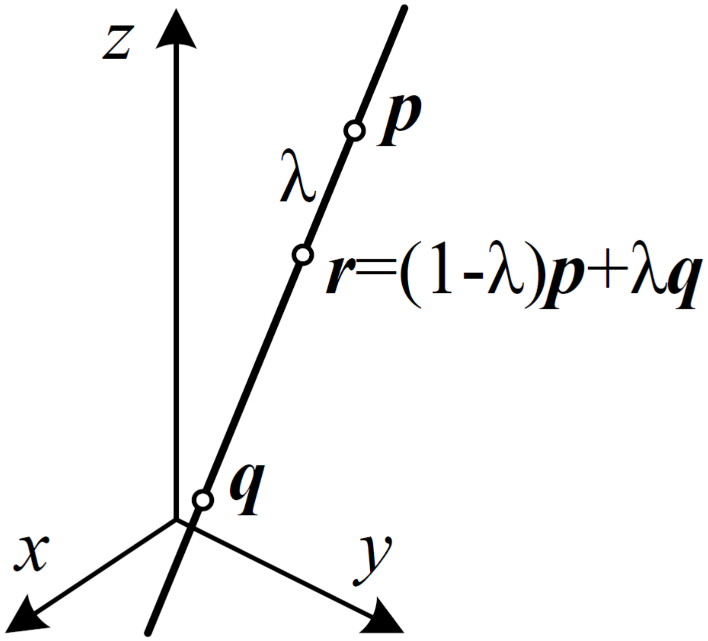
$$\hat{\mathbf{n}} = (\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi)$$



3D Lines

Any point on the line is a linear combination of two points

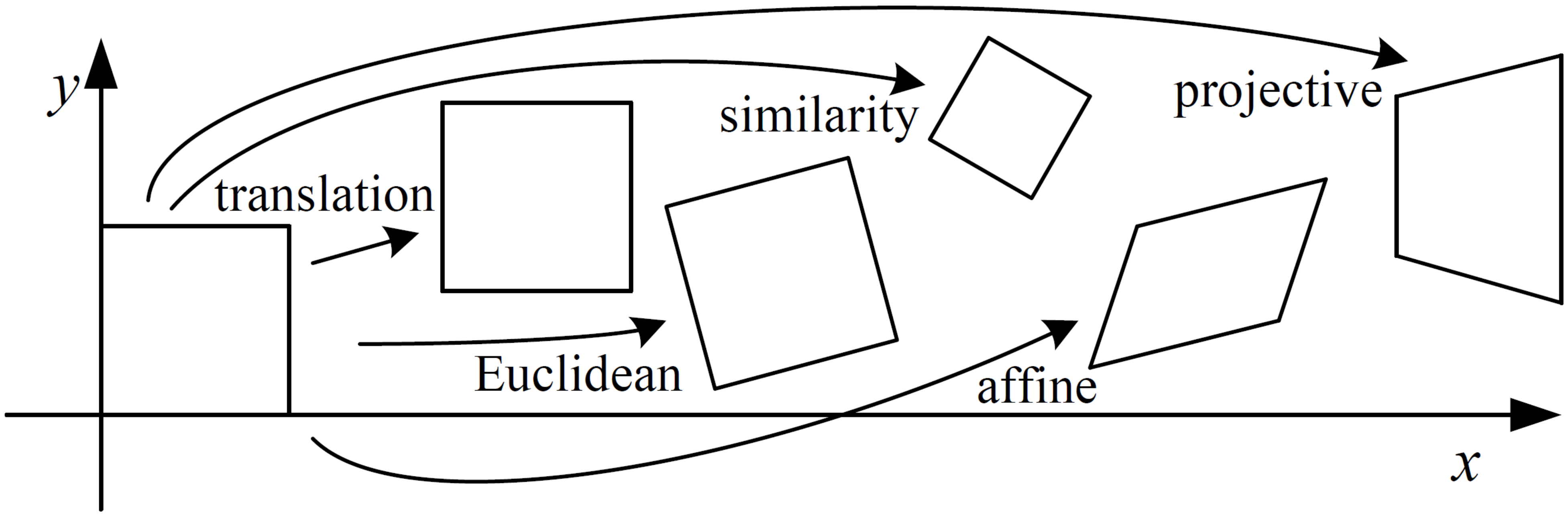
$$\mathbf{r} = (1 - \lambda)\mathbf{p} + \lambda\mathbf{q}$$



Using a line direction

$$\mathbf{r} = \mathbf{p} + \lambda\hat{\mathbf{d}}$$

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

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

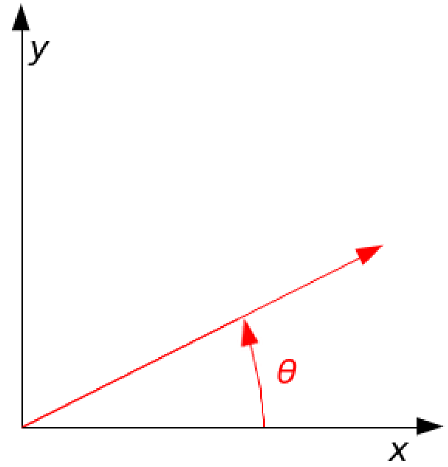
$$3 \times 3$$

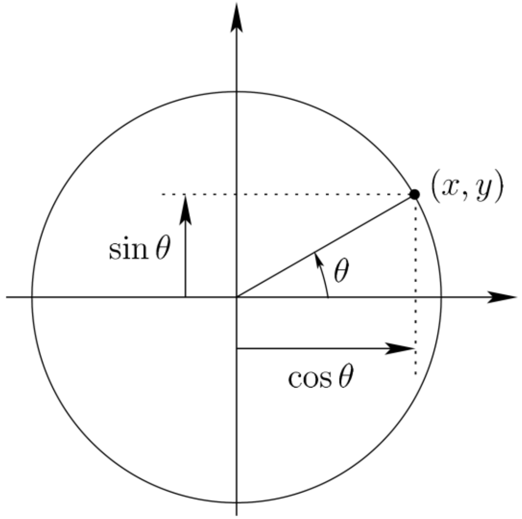
2D Euclidean Transformation

2D Rotation + 2D translation

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

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





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

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

orthonormal rotation matrix

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

2D Euclidean Transformation

2D Rotation + 2D translation

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

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

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

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

2D Similarity Transformation

Scaled 2D rotation + 2D translation

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

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

$$\mathbf{x}' = \begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}} = \begin{bmatrix} a & -b & t_x \\ b & a & t_y \end{bmatrix} \bar{\mathbf{x}}$$

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

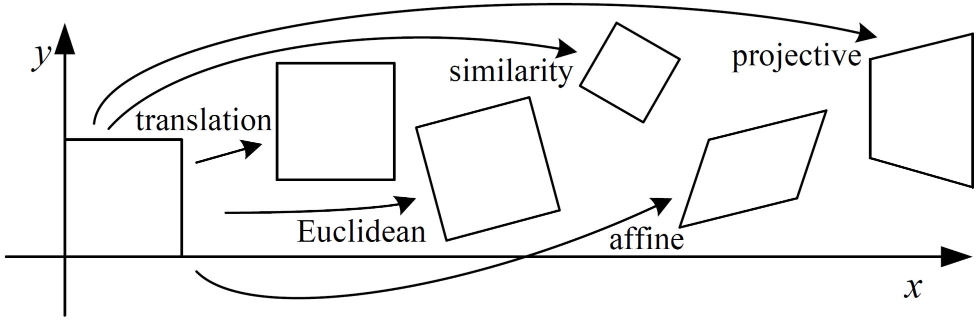
The similarity transform preserves angles between lines.

2D Affine Transformation

Arbitrary 2x3 matrix

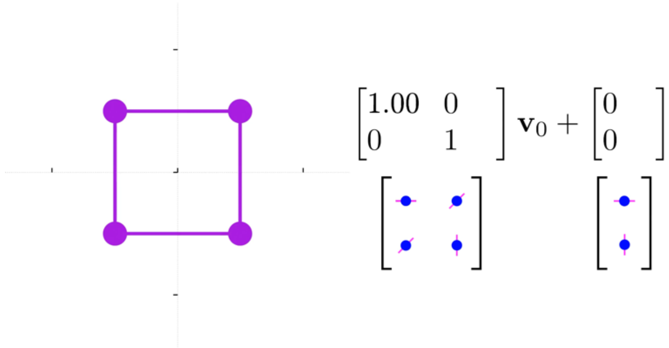
$$\mathbf{x}' = \mathbf{A}\bar{\mathbf{x}} \qquad \bar{\mathbf{x}} = (x, y, 1)$$

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

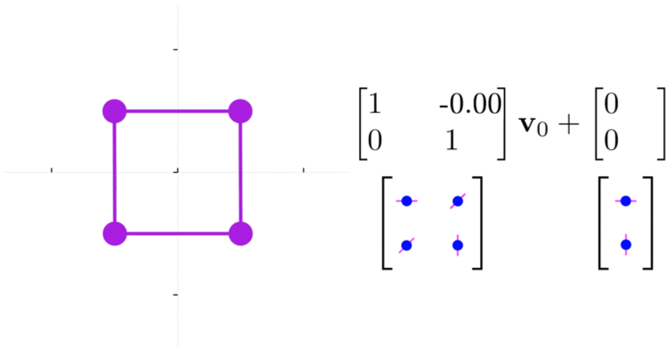


Parallel lines remain parallel under affine transformations.

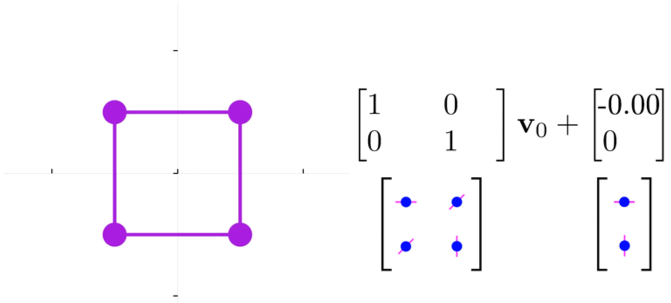
2D Affine Transformation Examples



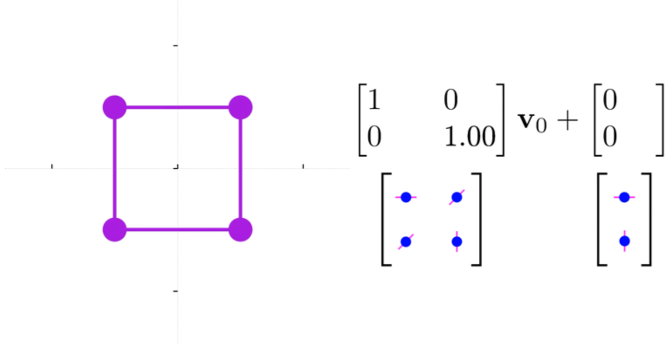
Scaling along x



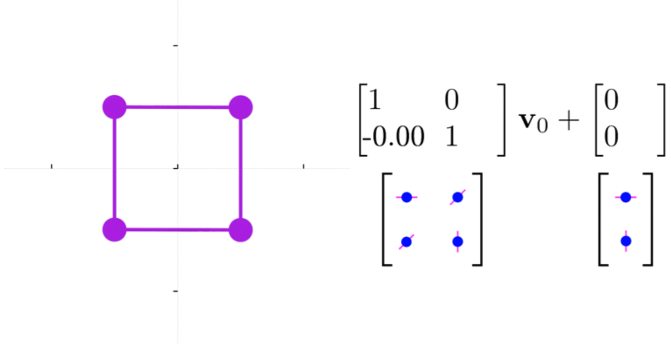
Shearing along x



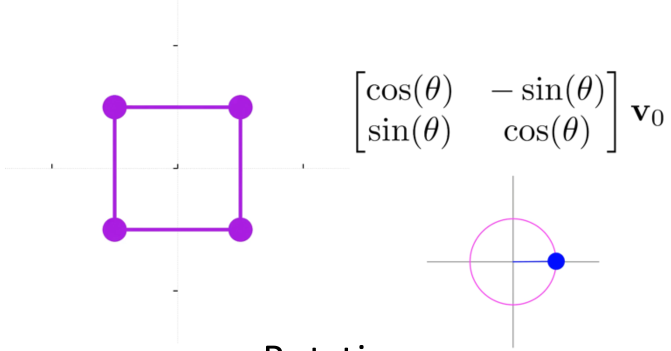
Translation along x



Scaling along y



Shearing along y



Rotation

https://www.algorithm-archive.org/contents/affine_transformations/affine_transformations.html

2D Projective Transformation

Also called perspective transform or homography

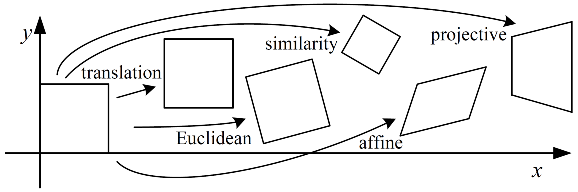
$$\tilde{\mathbf{x}}' = \tilde{\mathbf{H}}\tilde{\mathbf{x}}$$

homogeneous coordinates

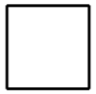
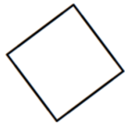
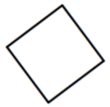
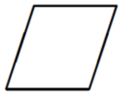
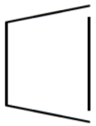
3×3 $\tilde{\mathbf{H}}$ is only defined up to a scale

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

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

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

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

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

$$3 \times 4$$

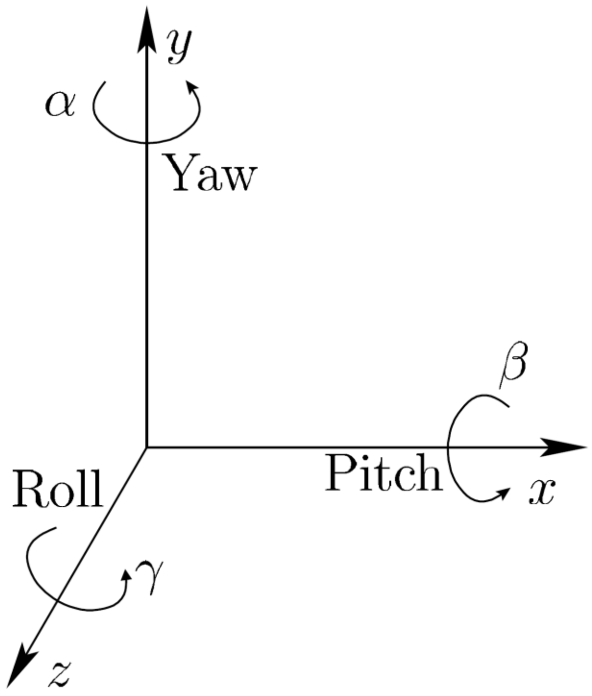
augmented vector $\bar{\mathbf{x}} = (x, y, z, 1)$

3D Rotation

The yaw, pitch, and roll rotations can be combined sequentially to attain any possible 3D rotation.

$$R(\alpha, \beta, \gamma) = R_y(\alpha) R_x(\beta) R_z(\gamma)$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad R_x(\beta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \quad R_y(\alpha) = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}$$



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

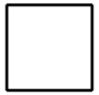
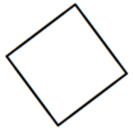
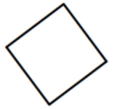
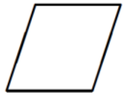
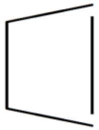
Also called 3D perspective transform or homography

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