

What will we learn today?

- Why Geometric Vision Matters
- Geometric Primitives in 2D & 3D
- 2D & 3D Transformations

**Images are
2D projections of
the 3D world**

Simplified Image Formation

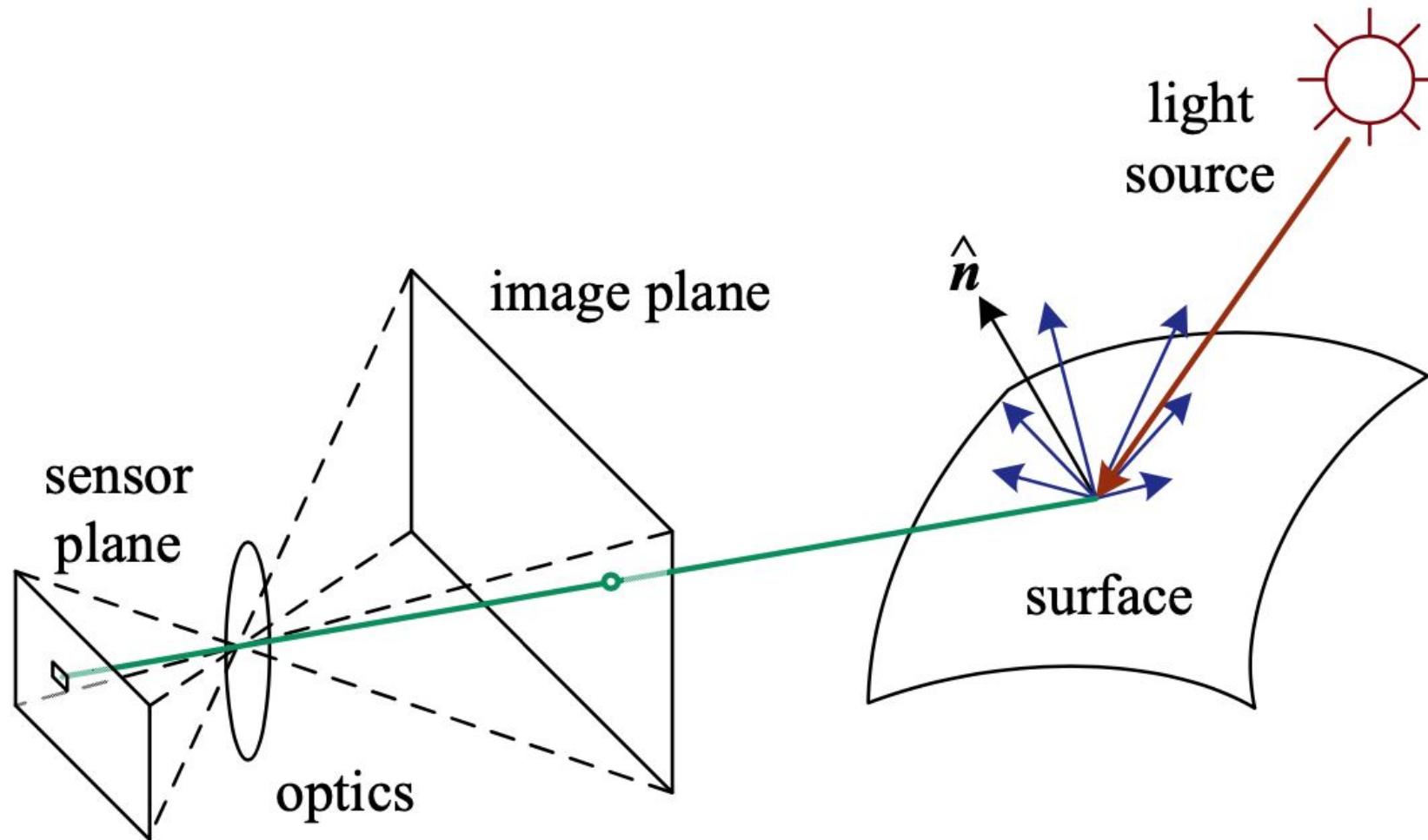


Figure: R. Szeliski

Perspective Projection

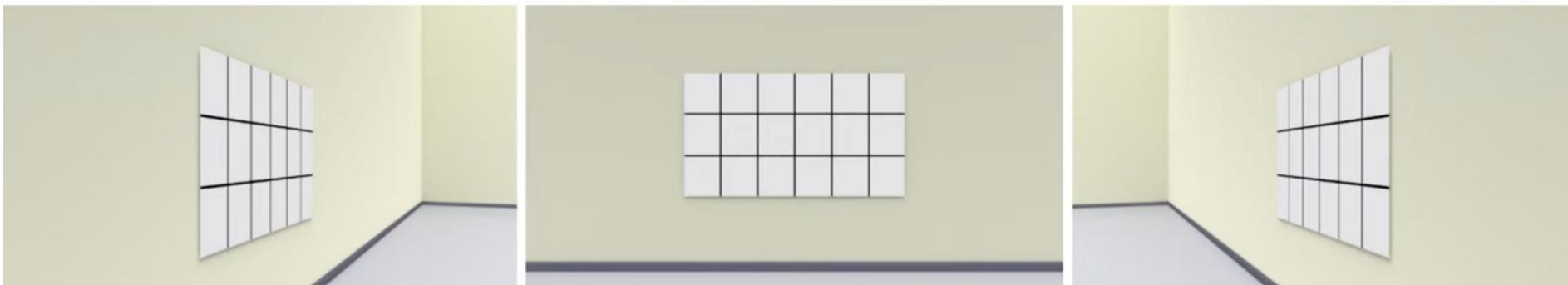
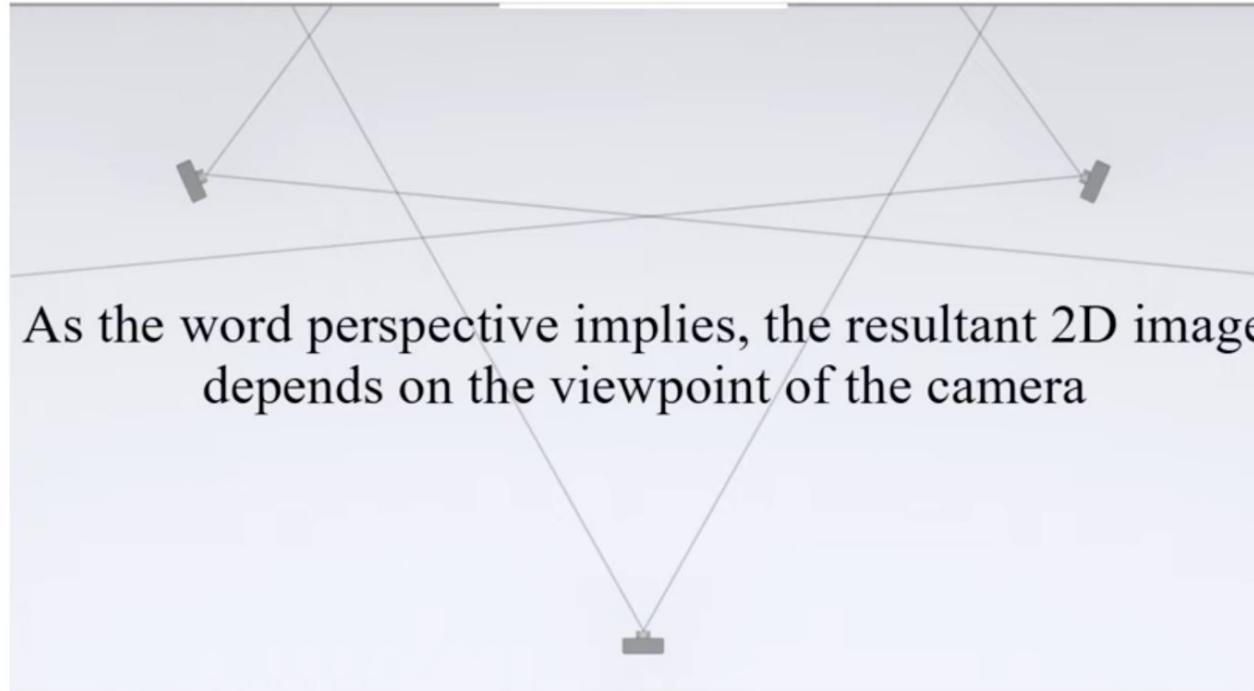
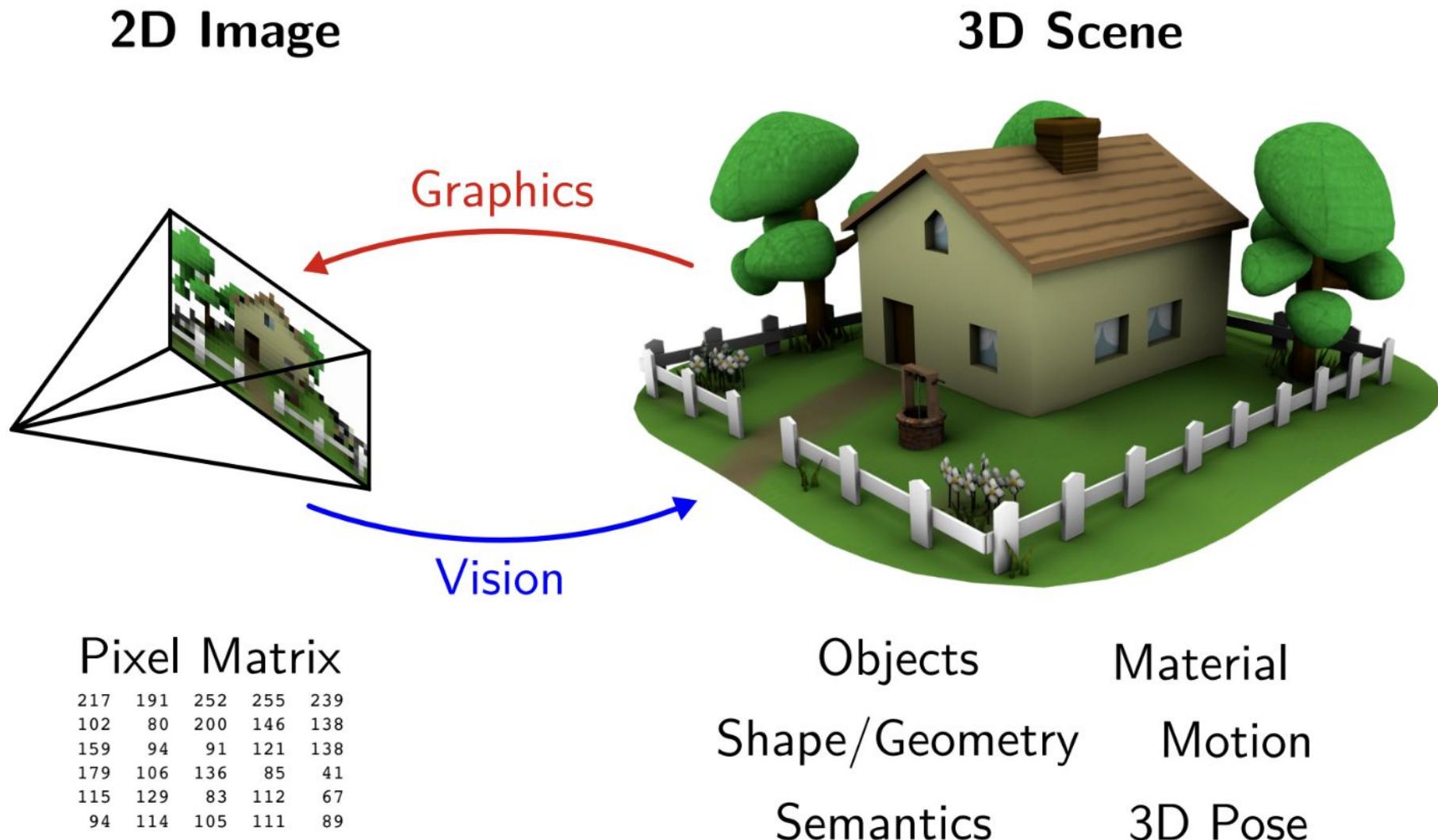


Figure: https://www.youtube.com/@huseyin_ozde...

**Can we understand
the 3D world
from 2D images?**

CV is an **ill-posed** inverse problem



What will we learn today?

Why Geometric Vision Matters

Geometric Primitives in 2D & 3D

2D & 3D Transformations

Points in Cartesian and Homogeneous Coordinates

2D points: $\mathbf{x} = (x, y) \in \mathcal{R}^2$ or column vector $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$

3D points: $\mathbf{x} = (x, y, z) \in \mathcal{R}^3$ (often noted \mathbf{X} or \mathbf{P})

Homogeneous coordinates: append a 1

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

Why?

Homogeneous coordinates in 2D

2D Projective Space: $\mathcal{P}^2 = \mathcal{R}^3 - (0, 0, 0)$ (same story in 3D with \mathcal{P}^3)

- heterogeneous \rightarrow homogeneous

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

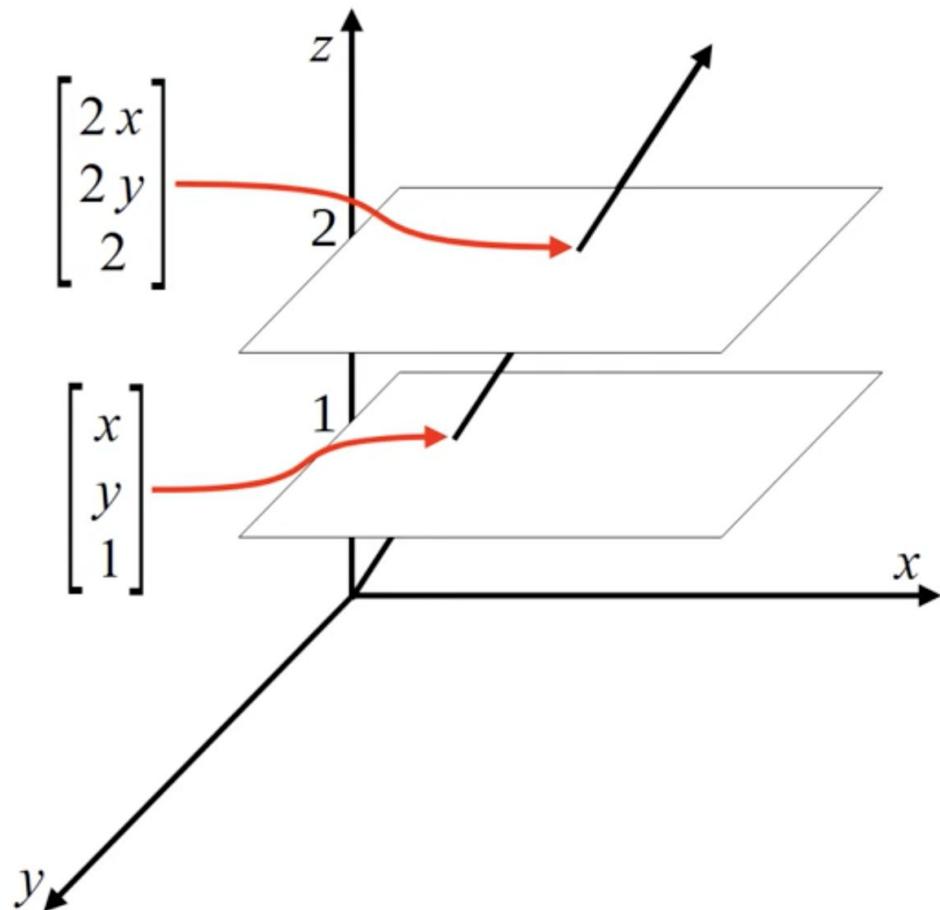
- homogeneous \rightarrow heterogeneous

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

- points differing only by scale are *equivalent*: $(x, y, w) \sim \lambda (x, y, w)$

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

Homogeneous coordinates in 2D



In homogeneous coordinates, a point and its scaled versions are same

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = w \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} wx \\ wy \\ w \end{bmatrix} \quad w \neq 0$$

Figure: https://www.youtube.com/@huseyin_ozde...

Everything is easier in Projective Space

2D Lines:

Representation: $l = (a, b, c)$

Equation: $ax + by + c = 0$

In homogeneous coordinates: $\bar{x}^T l = 0$

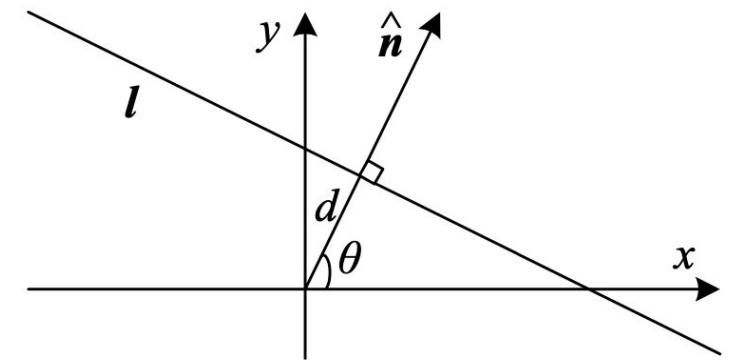
General idea: homogenous coordinates unlock the full power of linear algebra!

Everything is easier in Projective Space

2D Lines:

$$\tilde{\mathbf{x}}^T \mathbf{l} = 0, \forall \tilde{\mathbf{x}} = (x, y, w) \in P^2$$

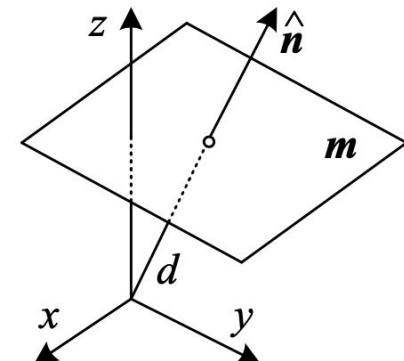
$$\mathbf{l} = (\hat{n}_x, \hat{n}_y, d) = (\hat{\mathbf{n}}, d) \text{ with } \|\hat{\mathbf{n}}\| = 1$$



3D planes: same!

$$\tilde{\mathbf{x}}^T \mathbf{m} = 0, \forall \tilde{\mathbf{x}} = (x, y, z, w) \in P^3$$

$$\mathbf{m} = (\hat{n}_x, \hat{n}_y, \hat{n}_z, d) = (\hat{\mathbf{n}}, d) \text{ with } \|\hat{\mathbf{n}}\| = 1$$



Lines in 3D

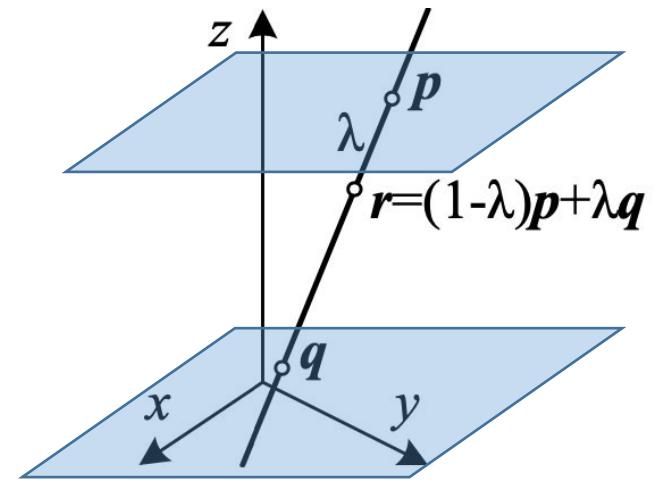
Two-point parametrization:

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

Two-plane parametrization:

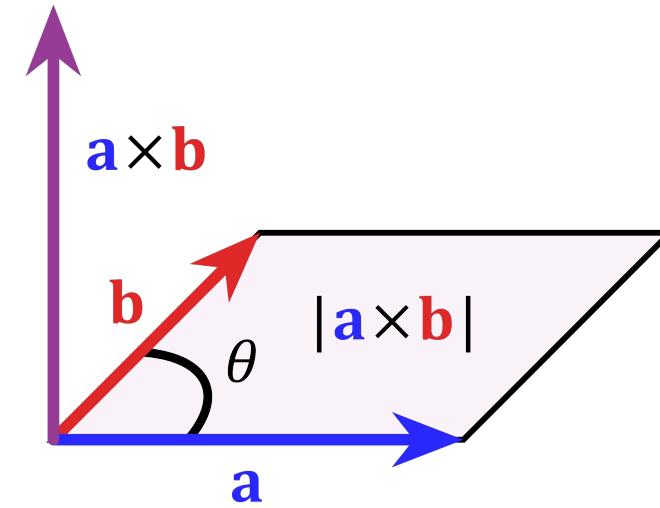
coordinates (x_0, y_0) & (x_1, y_1) of intersection

with planes at $z = 0, 1$ (or other planes)



Cross-product quick reminder

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



$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times} \mathbf{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Benefits of Homogeneous Coordinates

- Line – Point duality:
 - line between two 2D points: $\tilde{l} = \tilde{x}_1 \times \tilde{x}_2$
 - intersection of two 2D lines: $\tilde{x} = \tilde{l}_1 \times \tilde{l}_2$
- Representation of Infinity:
 - points at infinity: $(x, y, 0)$; line at infinity: $(0, 0, 1)$
- Parallel & vertical lines are easy (take-home: intersect //)
- Makes 2D & 3D transformations linear!

Questions?

What will we learn today?

Why Geometric Vision Matters

Geometric Primitives in 2D & 3D

2D & 3D Transformations

The camera as a coordinate transformation

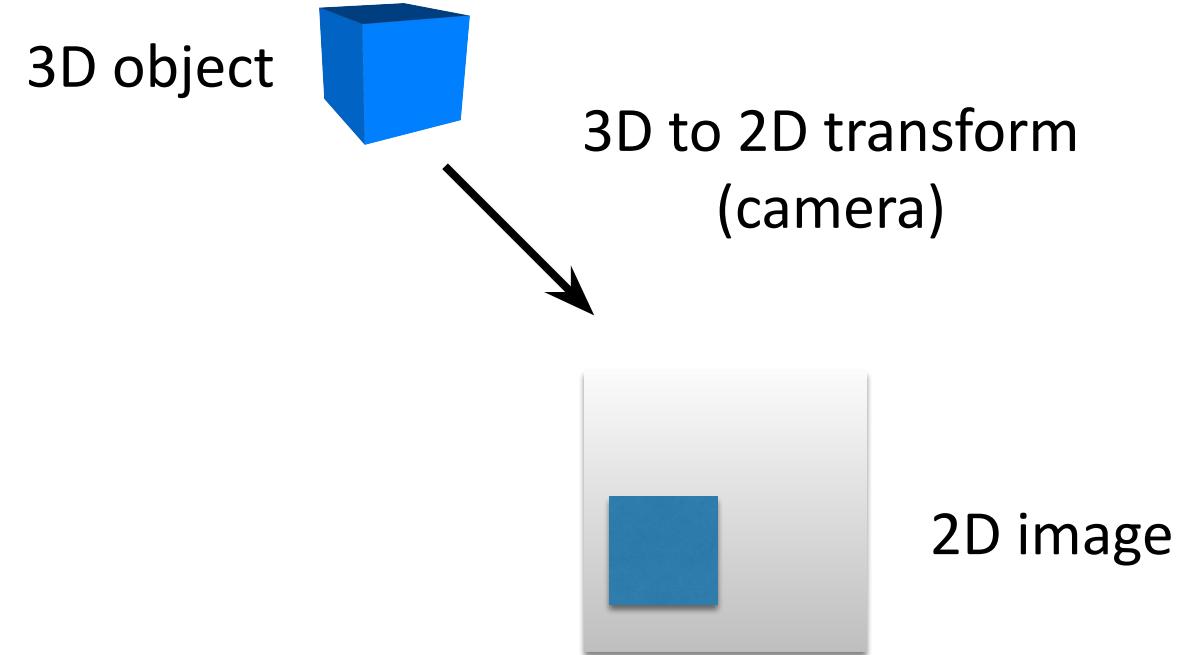
A camera is a mapping

from:

the 3D world

to:

a 2D image



The camera as a coordinate transformation

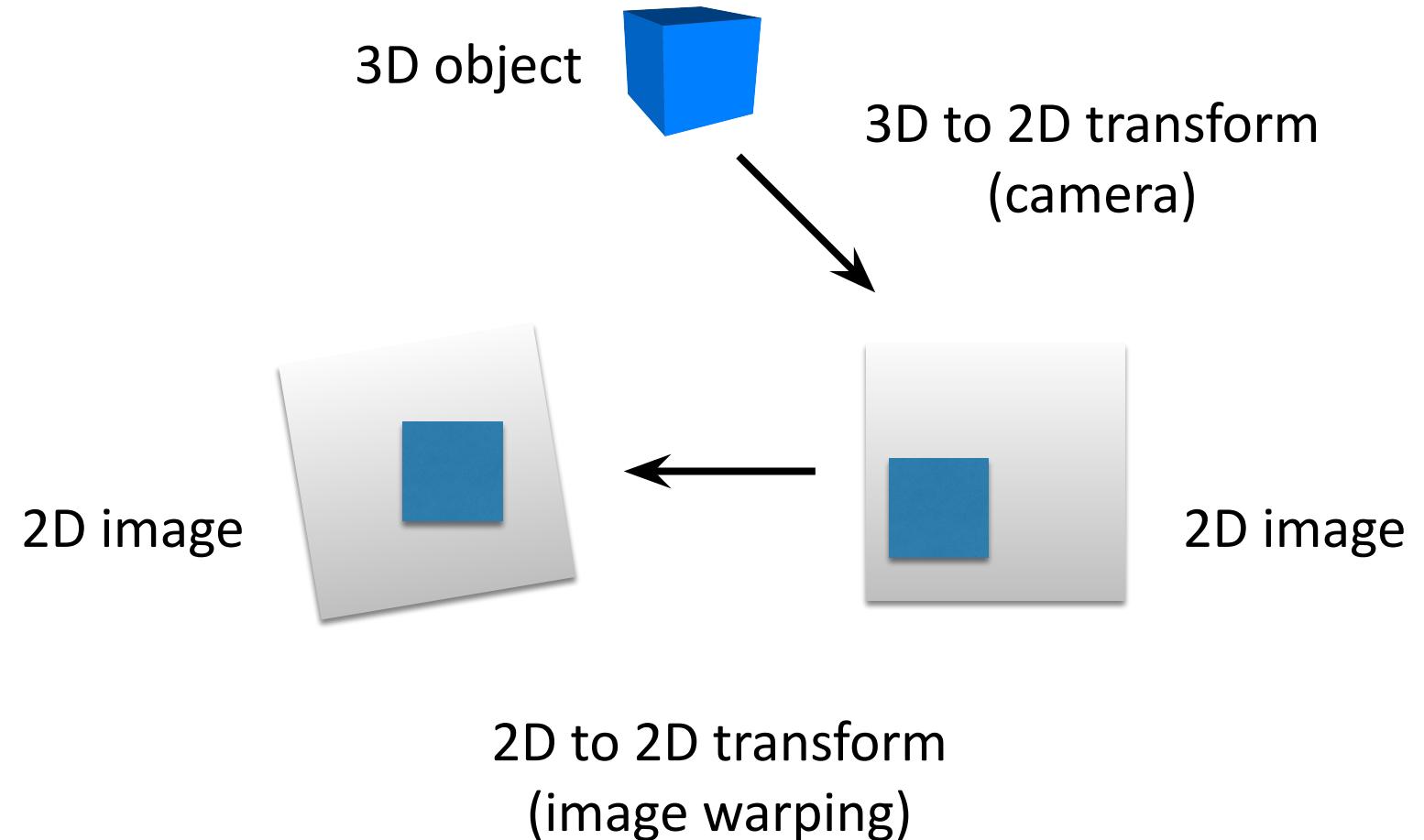
A camera is a mapping

from:

the 3D world

to:

a 2D image



Cameras and objects can move!

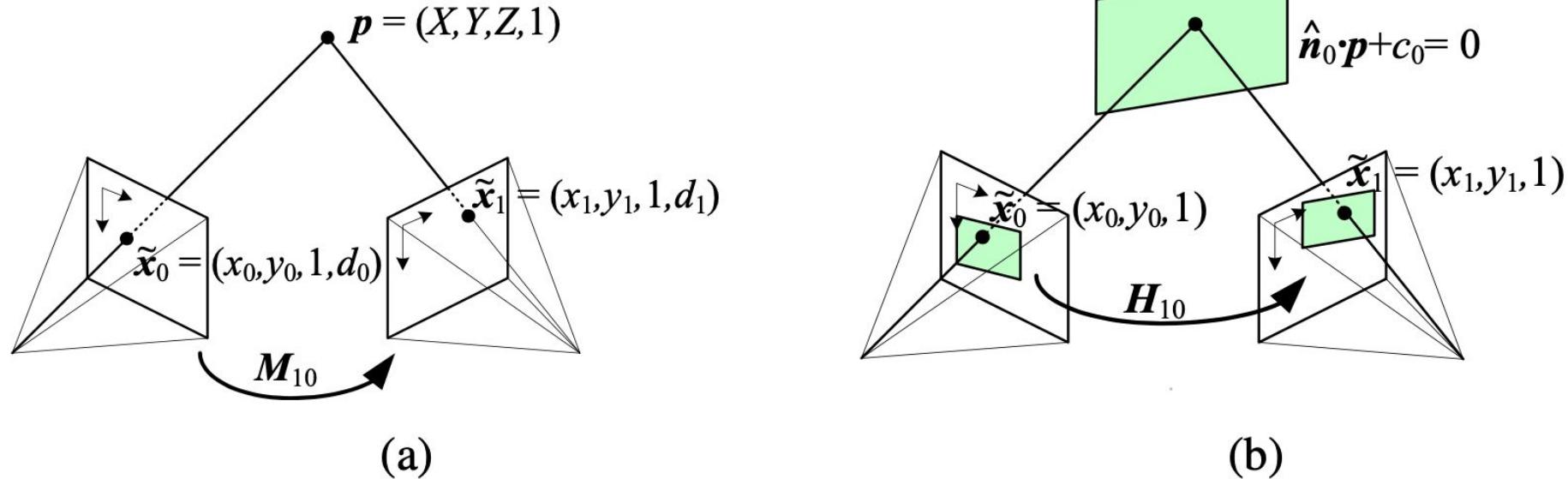


Figure 2.12 A point is projected into two images: (a) relationship between the 3D point coordinate $(X, Y, Z, 1)$ and the 2D projected point $(x, y, 1, d)$; (b) planar homography induced by points all lying on a common plane $\hat{\mathbf{n}}_0 \cdot \mathbf{p} + c_0 = 0$.

2D Transformations Zoo

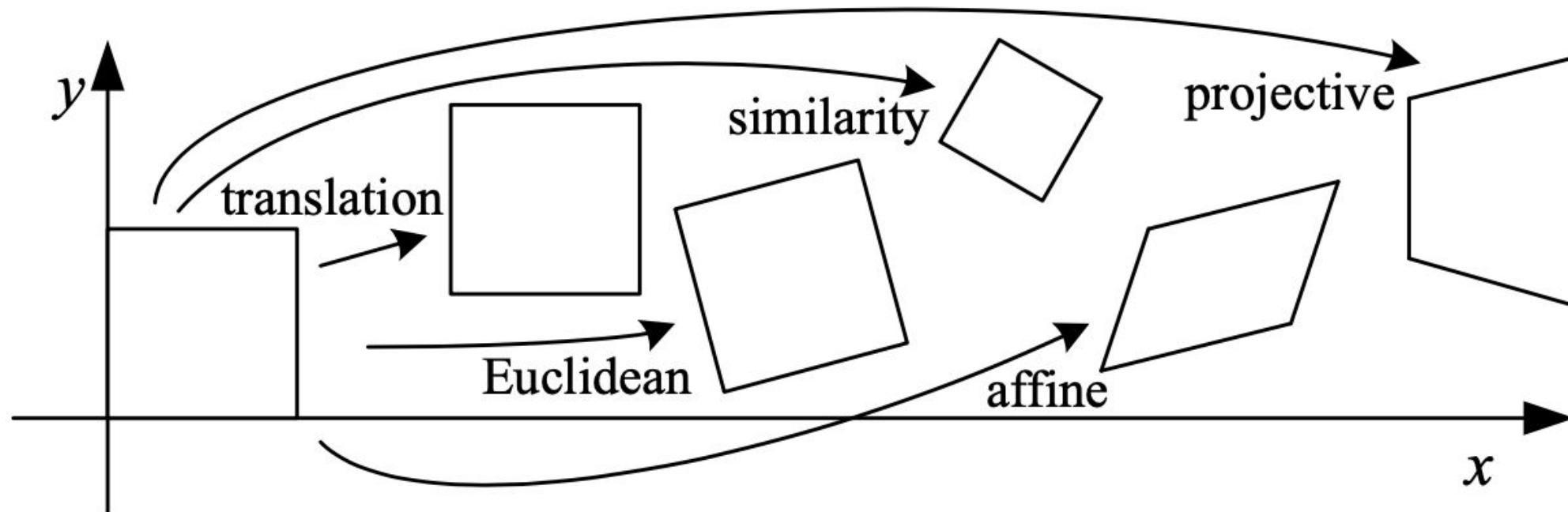


Figure: R. Szeliski

Transformation = Matrix Multiplication

Scale

$$\mathbf{M} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$$

Flip across y

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

Rotate

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

Flip across origin

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

Shear

$$\mathbf{M} = \begin{bmatrix} 1 & s_x \\ s_y & 1 \end{bmatrix}$$

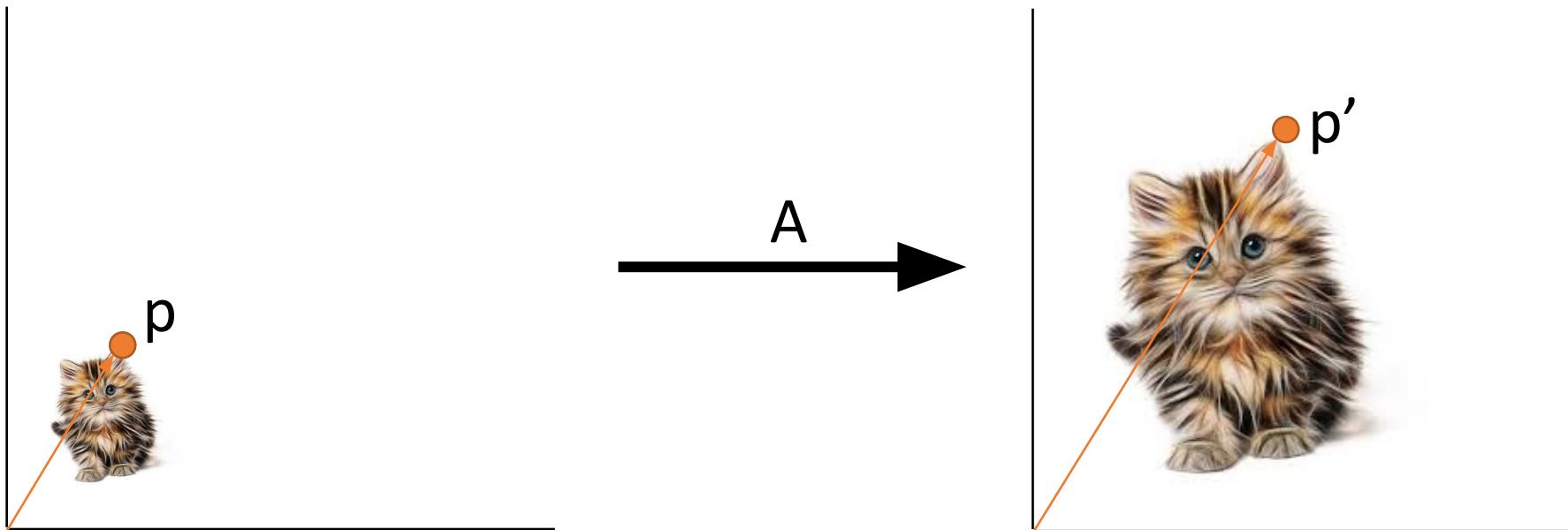
Identity

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

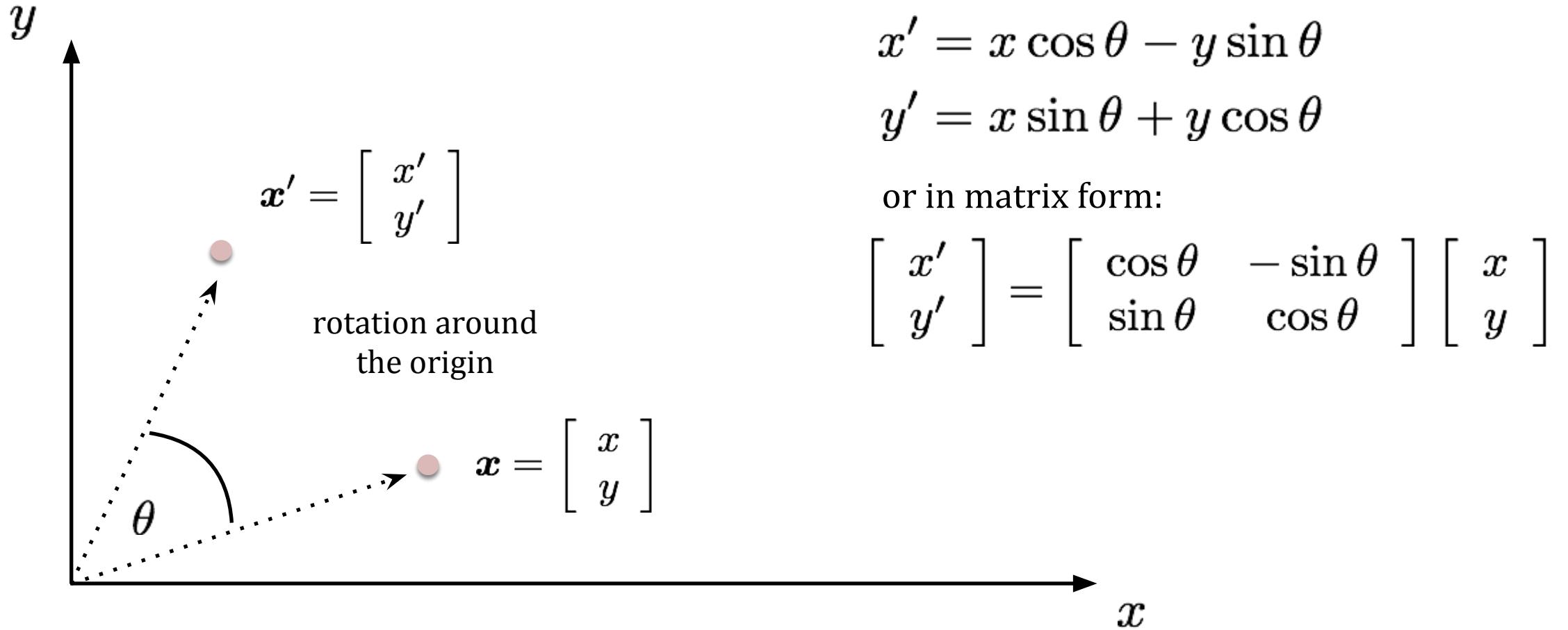
Scaling

$$\begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} s_x x \\ s_y y \end{bmatrix}$$

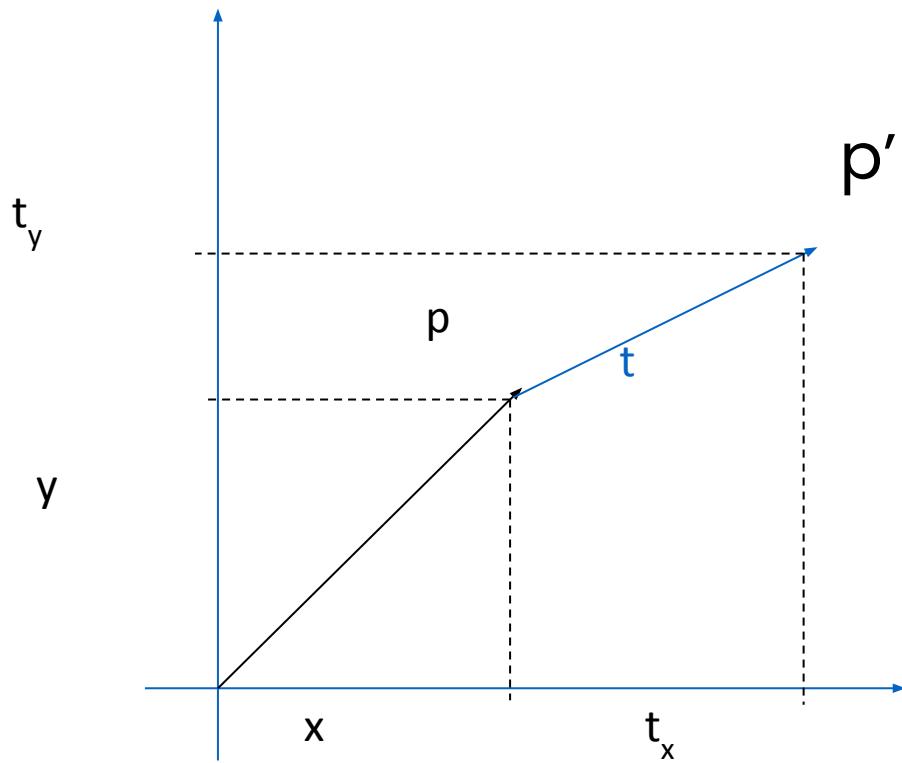
A p p'



Rotation



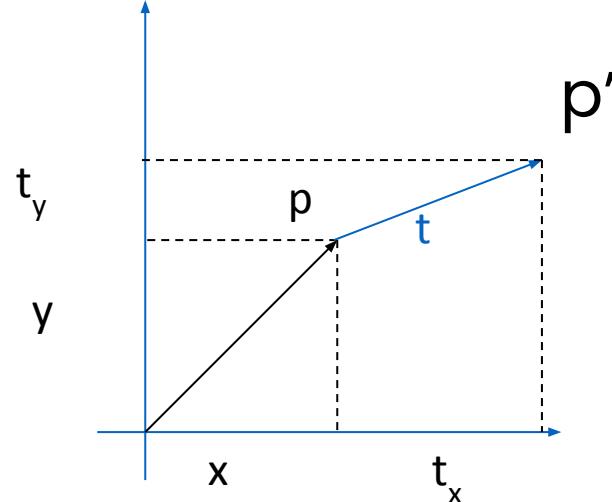
Translation



$$x' = x + t_x$$
$$y' = y + t_y$$

As a matrix?

Translation with homogeneous coordinates



$$p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$t = \begin{bmatrix} t_x \\ t_y \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} t_x \\ t_y \\ 1 \end{bmatrix}$$

$$p' = Tp$$

$$p' \rightarrow \begin{bmatrix} x + tx \\ y + ty \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} I & t \\ 0 & 1 \end{bmatrix} p = Tp$$

2D Transformations with homogeneous coordinates

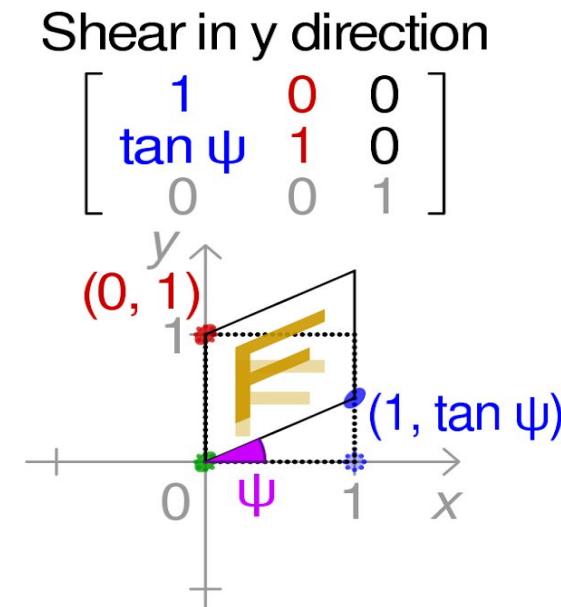
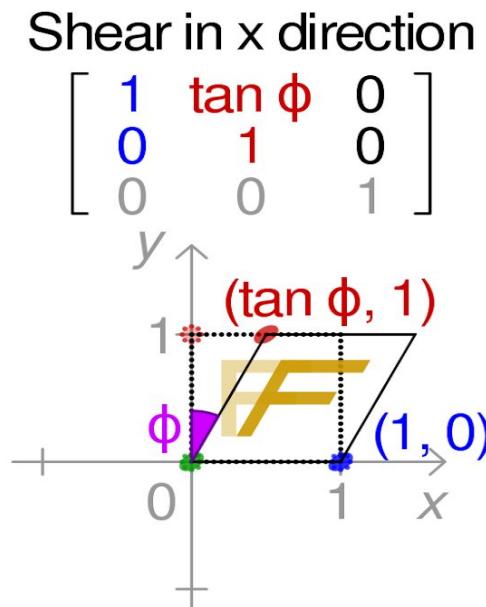
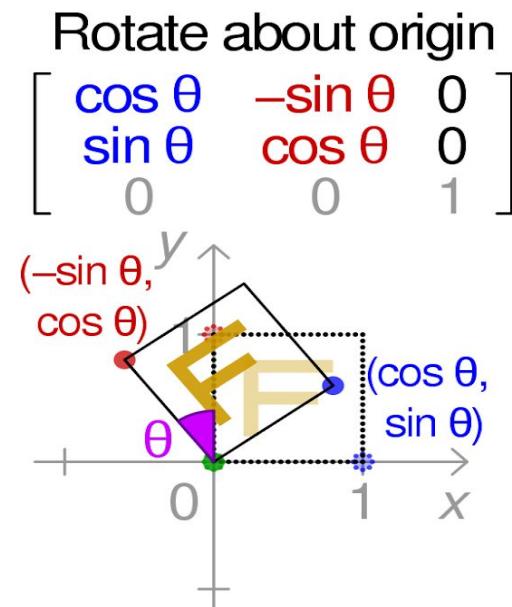
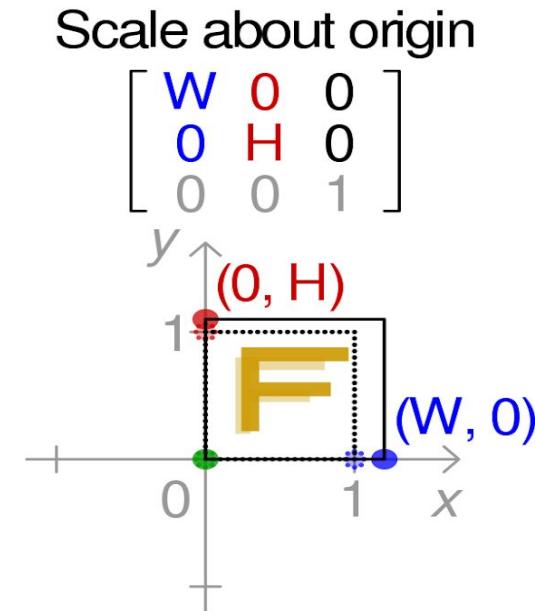
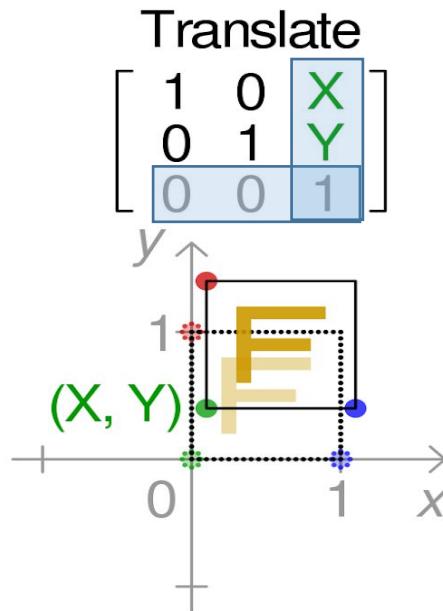
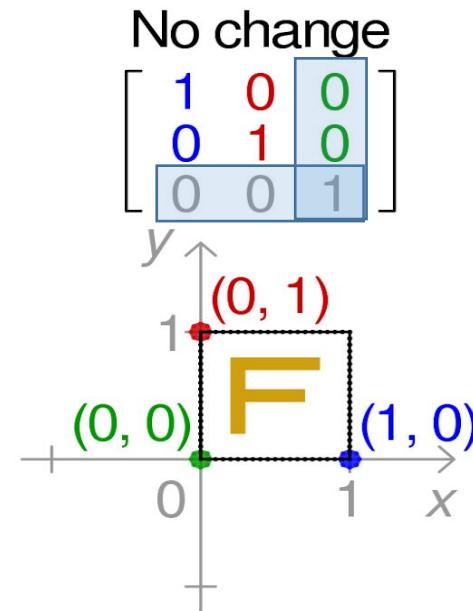


Figure: Wikipedia

Questions?

2D Transformations Zoo

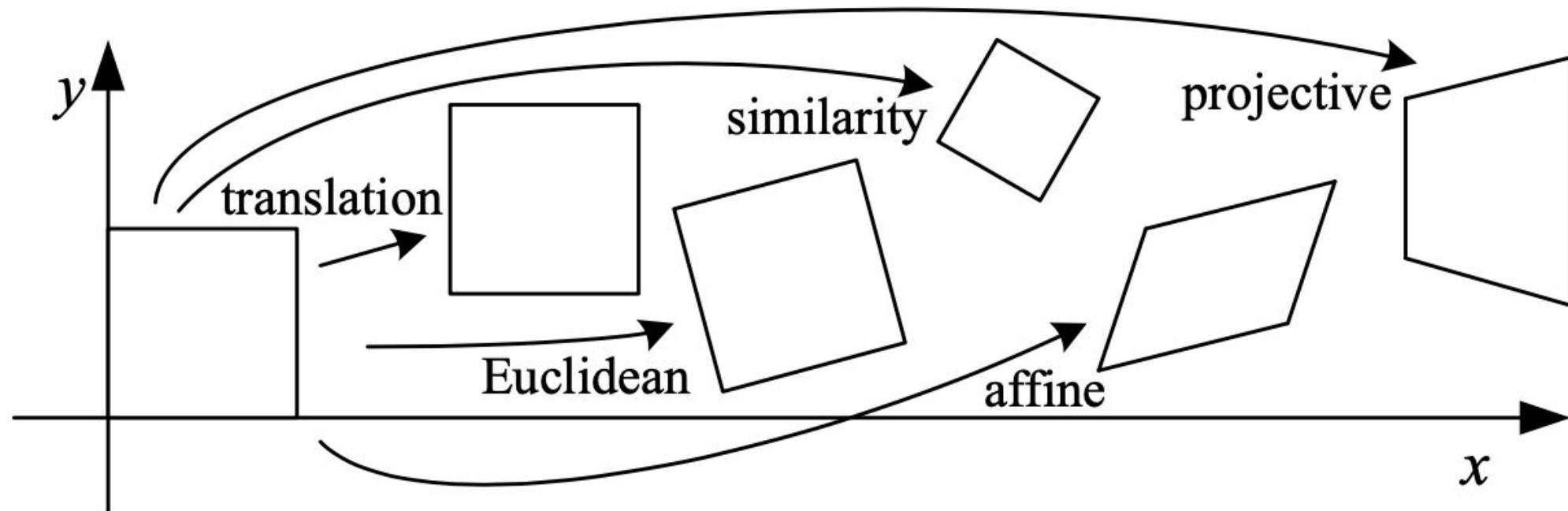


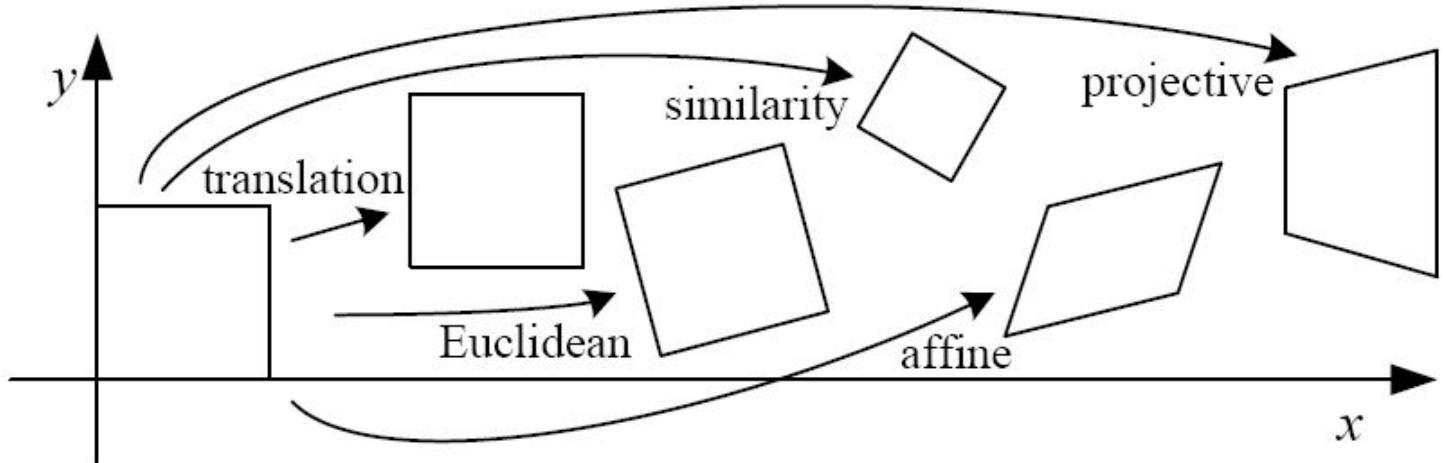
Figure: R. Szeliski

Euclidean / Rigid Transformation

Euclidean (rigid): rotation + translation

$$\begin{bmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

How many degrees of freedom?

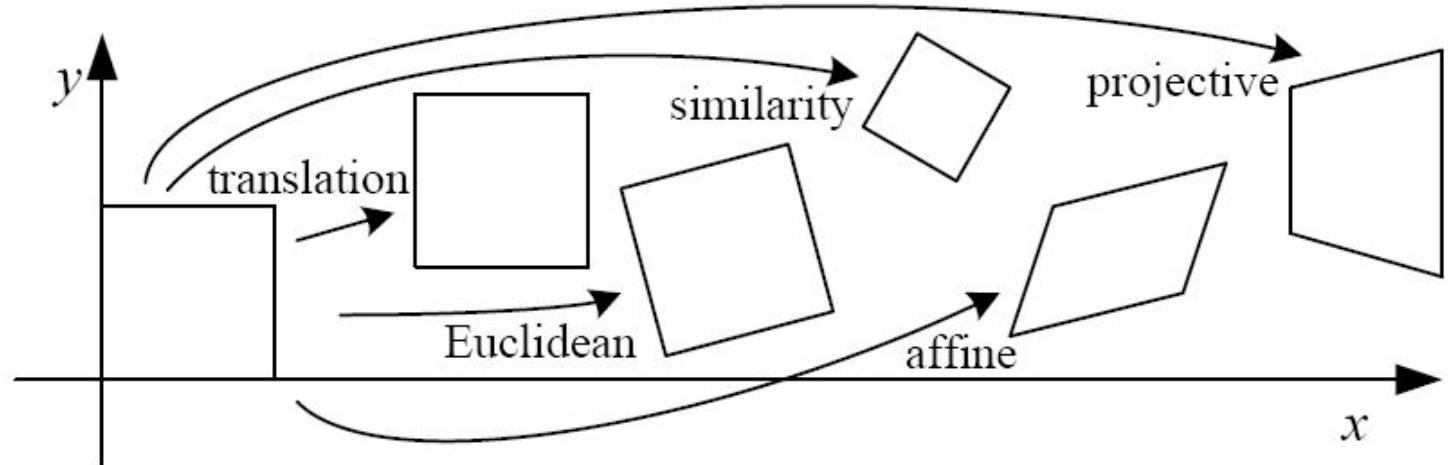


Similarity Transformation

Similarity: Scaling + rotation + translation

$$\begin{bmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

How many degrees of freedom?

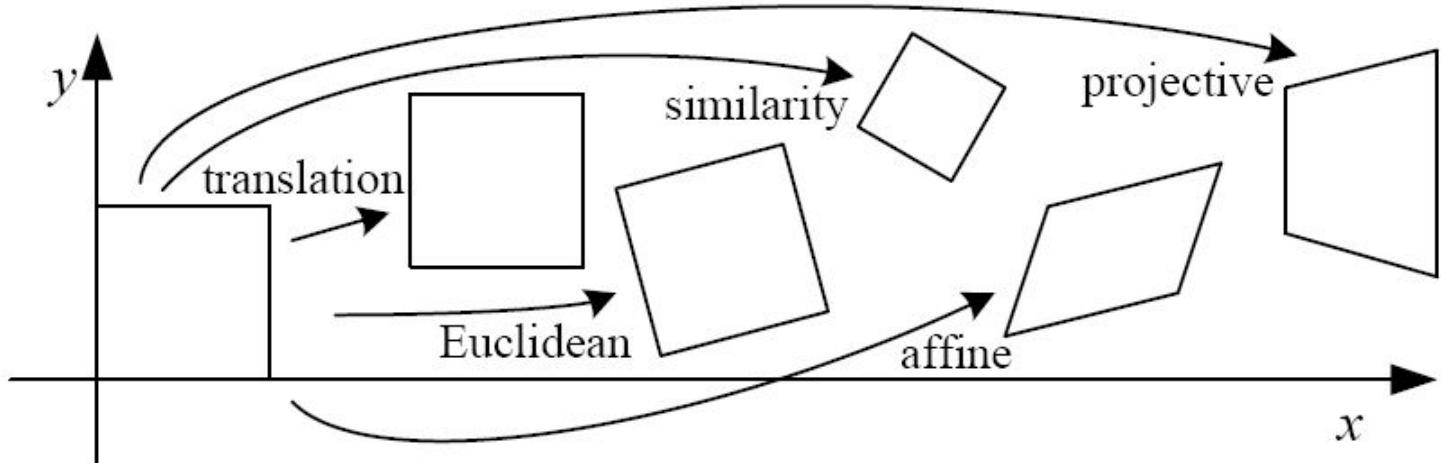


Similarity Transformation

Similarity: Scaling + rotation + translation

$$\begin{bmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha \cos \theta & -\alpha \sin \theta & b_0 \\ \alpha \sin \theta & \alpha \cos \theta & b_1 \\ 0 & 0 & 1 \end{bmatrix}$$

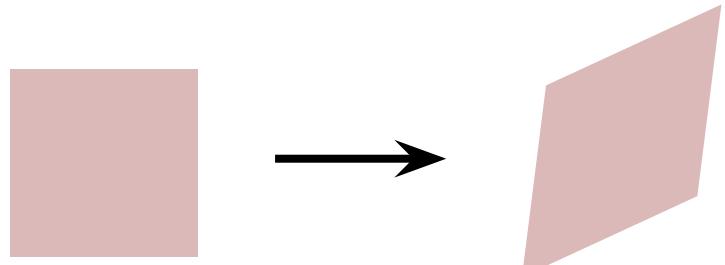
How many degrees of freedom?



Affine Transformation

Affine transformations are combinations of

- Arbitrary (4-DOF) linear transformations + translations



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \quad \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad \longrightarrow$$

Cartesian
coordinates

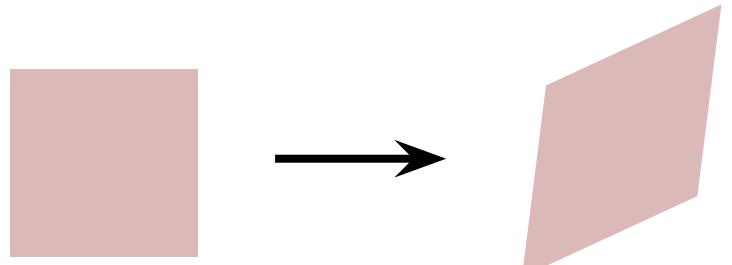
Homogeneous
coordinates

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Affine Transformation

Affine transformations are combinations of

- Arbitrary (4-DOF) linear transformations + translations



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \quad \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

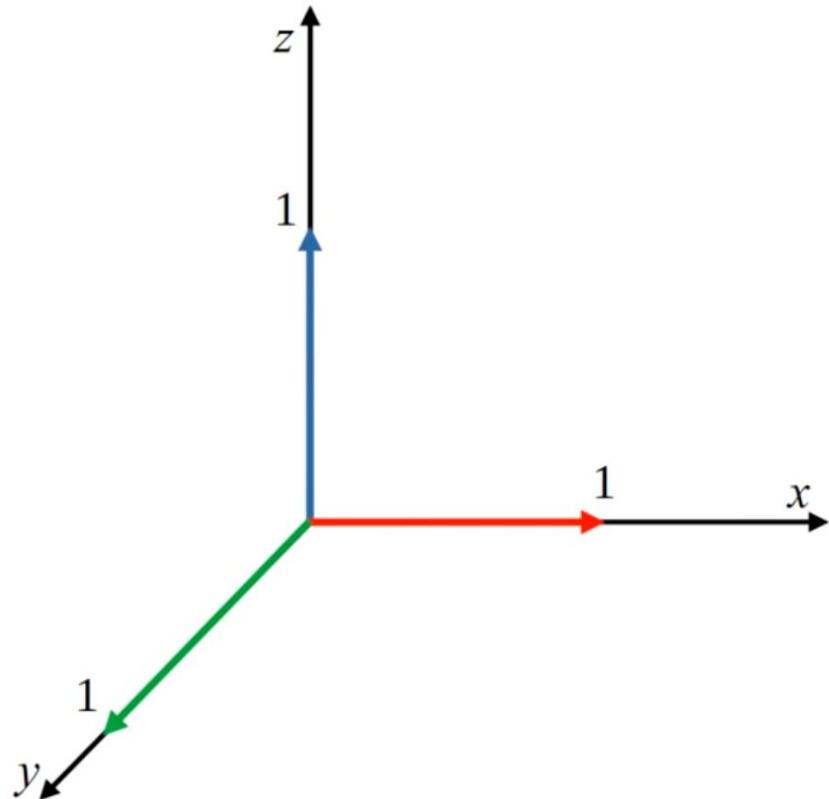
Cartesian
coordinates

Homogeneous
coordinates

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

How many degrees of freedom?

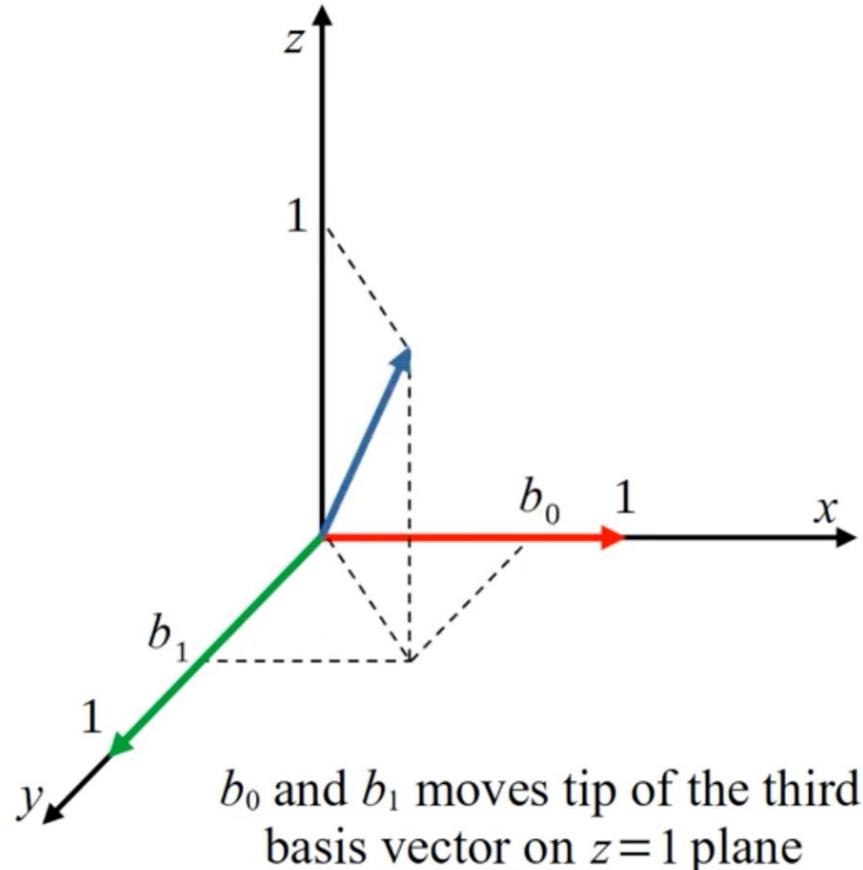
Affine Transformation



This matrix is a linear transformation matrix in 3D

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Affine Transformation



This matrix is a linear transformation matrix in 3D

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Then this column is the third basis vector of transformed vector space

And what b_0 and b_1 do is to change the orientation of that basis vector

Affine Transformation

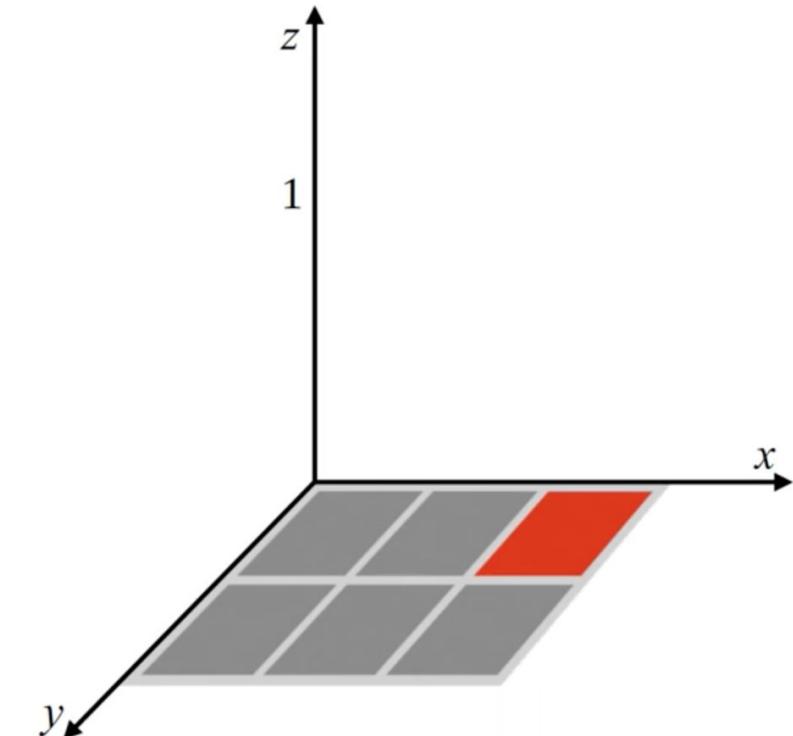


Figure: https://www.youtube.com/@huseyin_ozde...

Affine Transformation

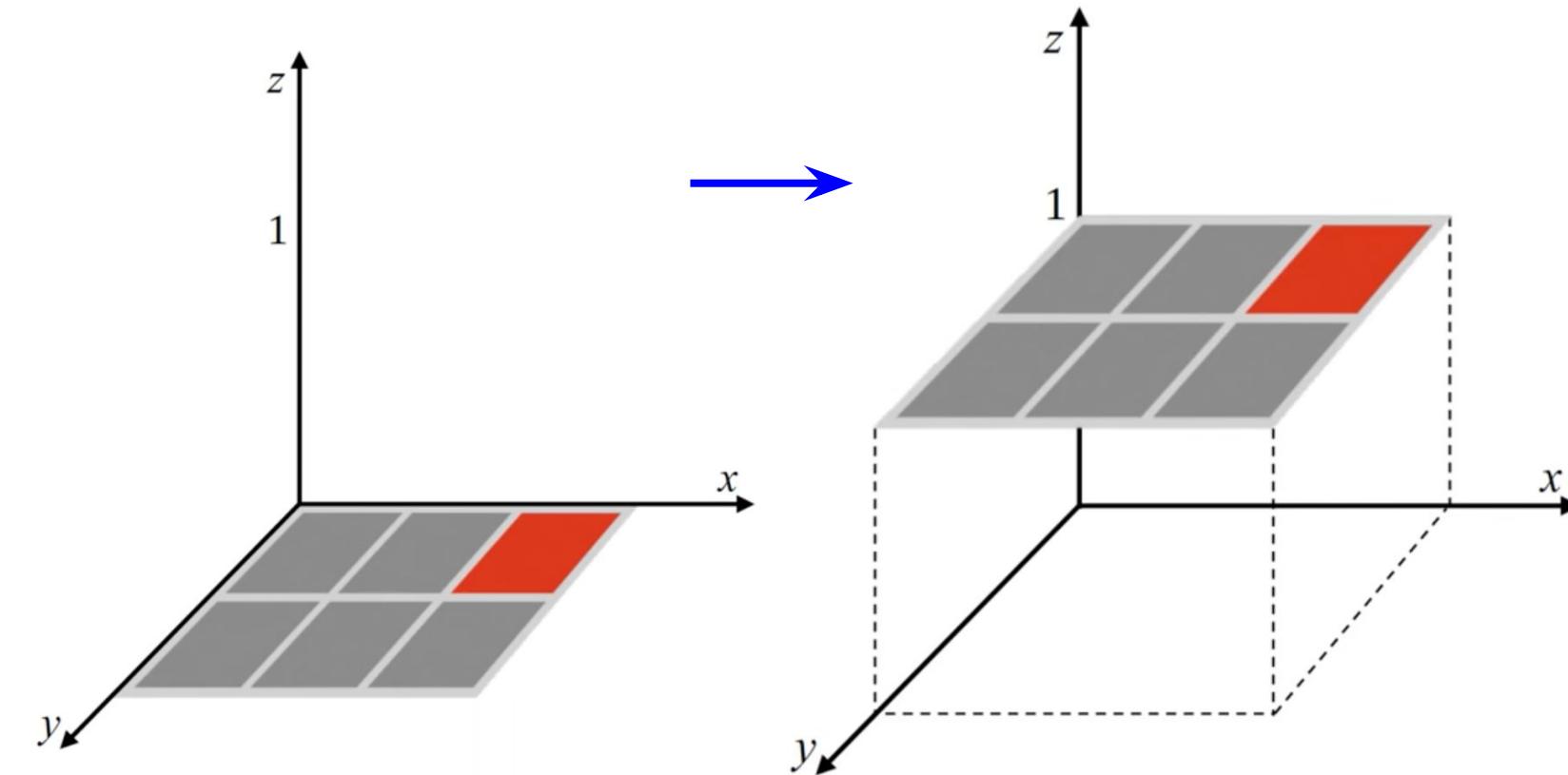


Figure: https://www.youtube.com/@huseyin_ozde...

Affine Transformation

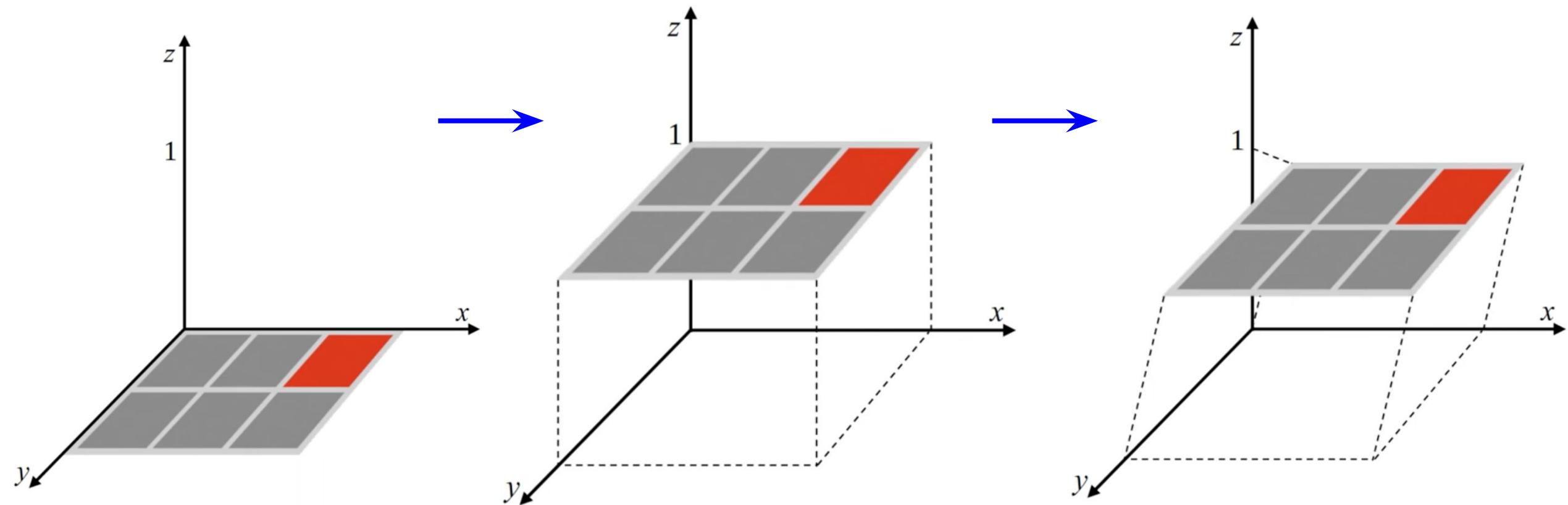


Figure: https://www.youtube.com/@huseyin_ozde...

Affine Transformation

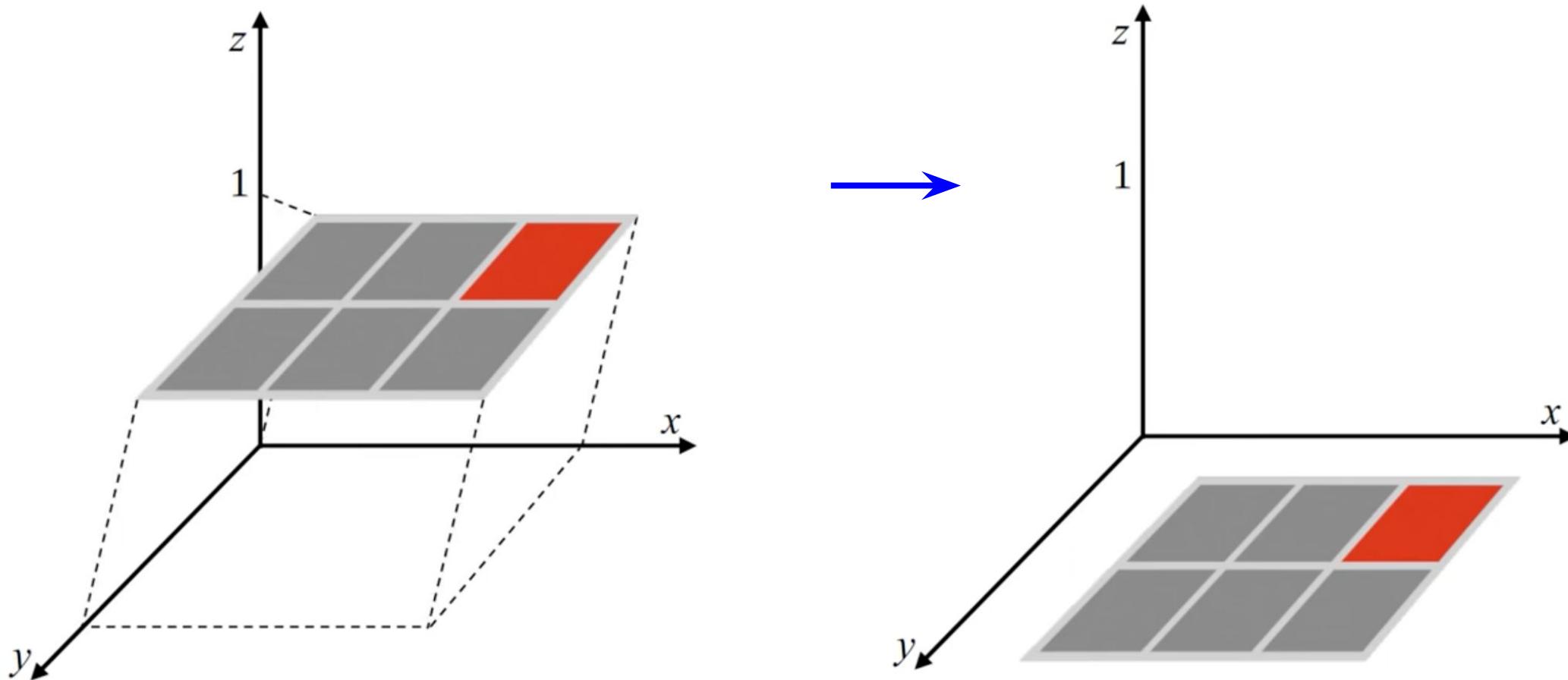
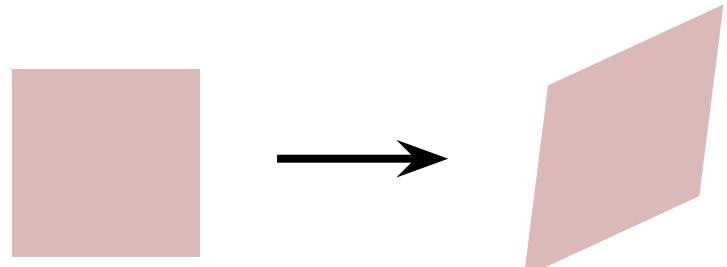


Figure: https://www.youtube.com/@huseyin_ozde...

Affine Transformation

Affine transformations are combinations of

- Arbitrary (4-DOF) linear transformations + translations



Properties of affine transformations:

- origin does not necessarily map to origin
- lines map to lines
- parallel lines map to parallel lines
- ratios are preserved

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

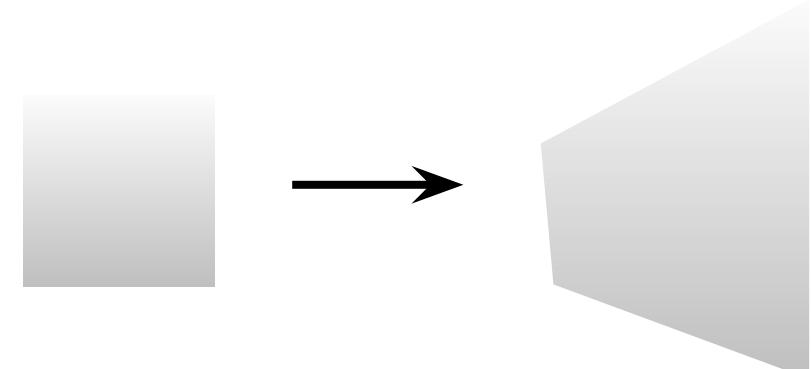
Projective Transformation (homography)

Projective transformations are combinations of

- Affine transformations + projective warps

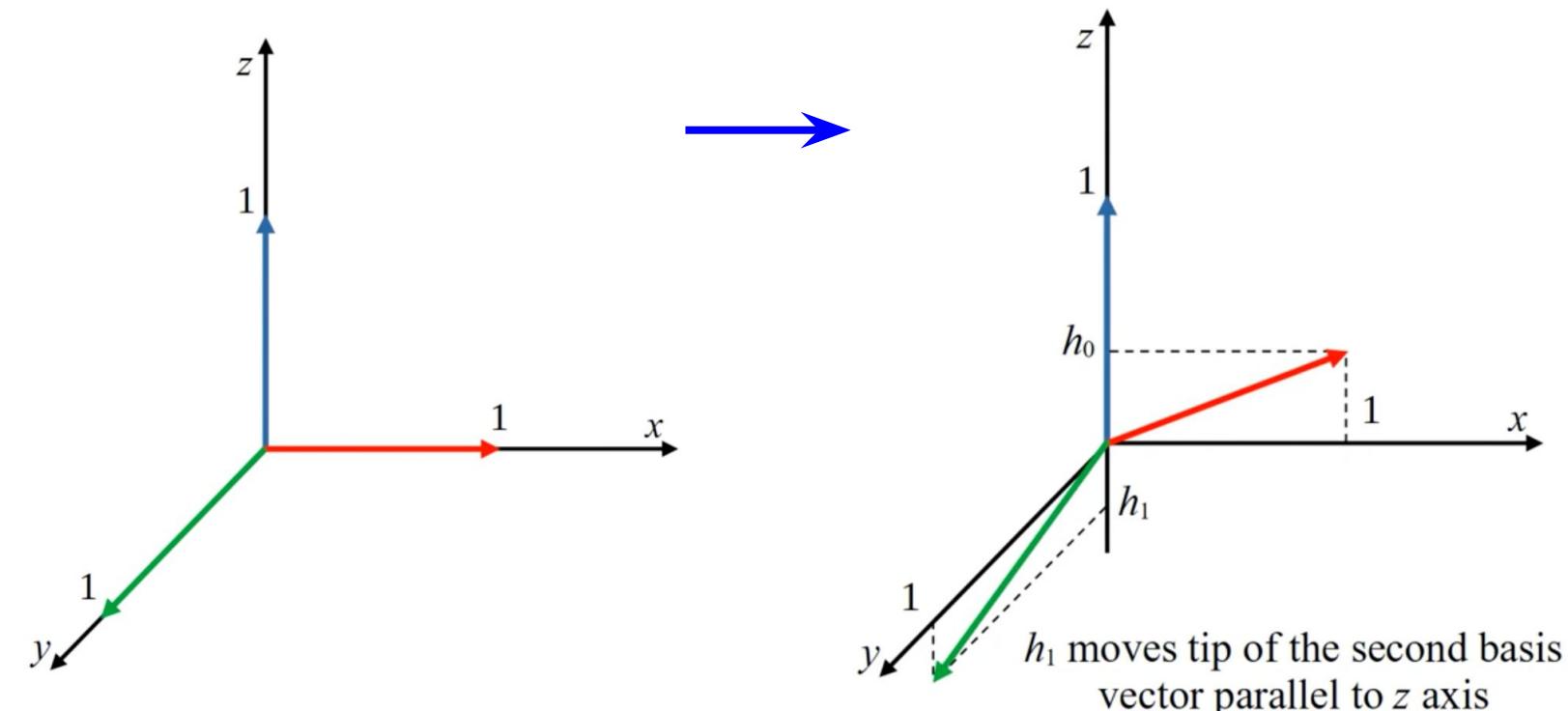
$$w \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ h_0 & h_1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

How many degrees of freedom?



Source: K. Kitani

Projective Transformation (homography)



This matrix is a linear transformation matrix in 3D

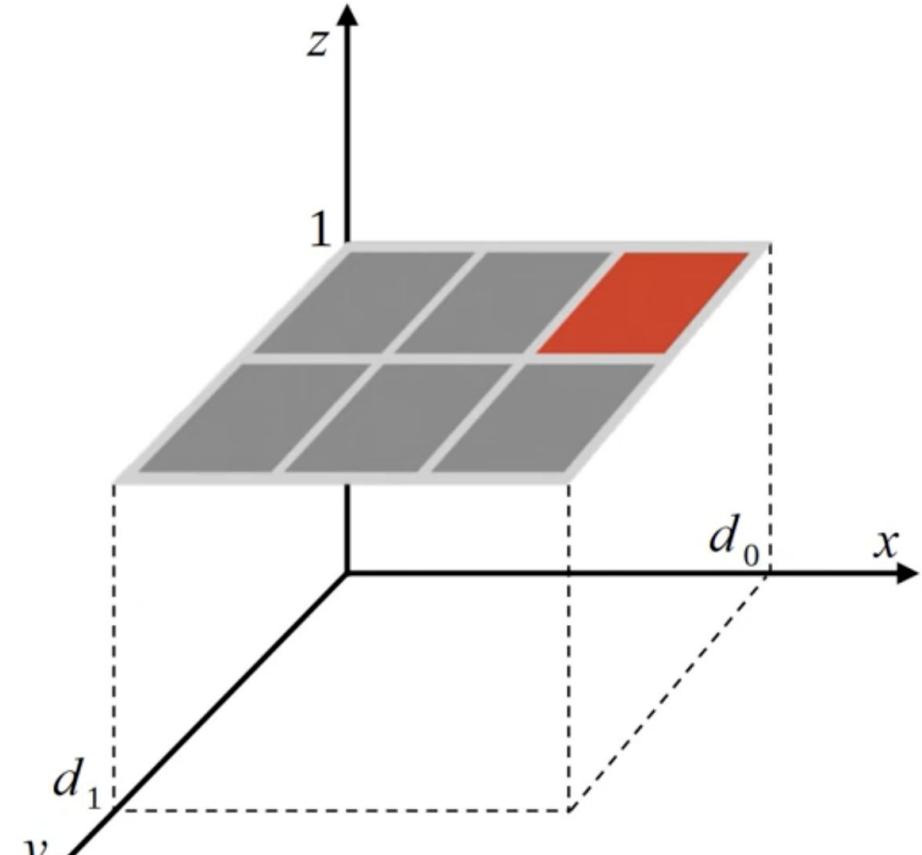
$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_0 & h_1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Then these two columns are the first and the second basis vectors of transformed vector space

And what h_0 and h_1 do is to change the orientation of those basis vectors

Figure: https://www.youtube.com/@huseyin_ozde...

Projective Transformation (homography)



As an example let
 $h_0 > 0$, $h_1 < 0$ and $|h_0| > |h_1|$

Figure: https://www.youtube.com/@huseyin_ozde...

Projective Transformation (homography)

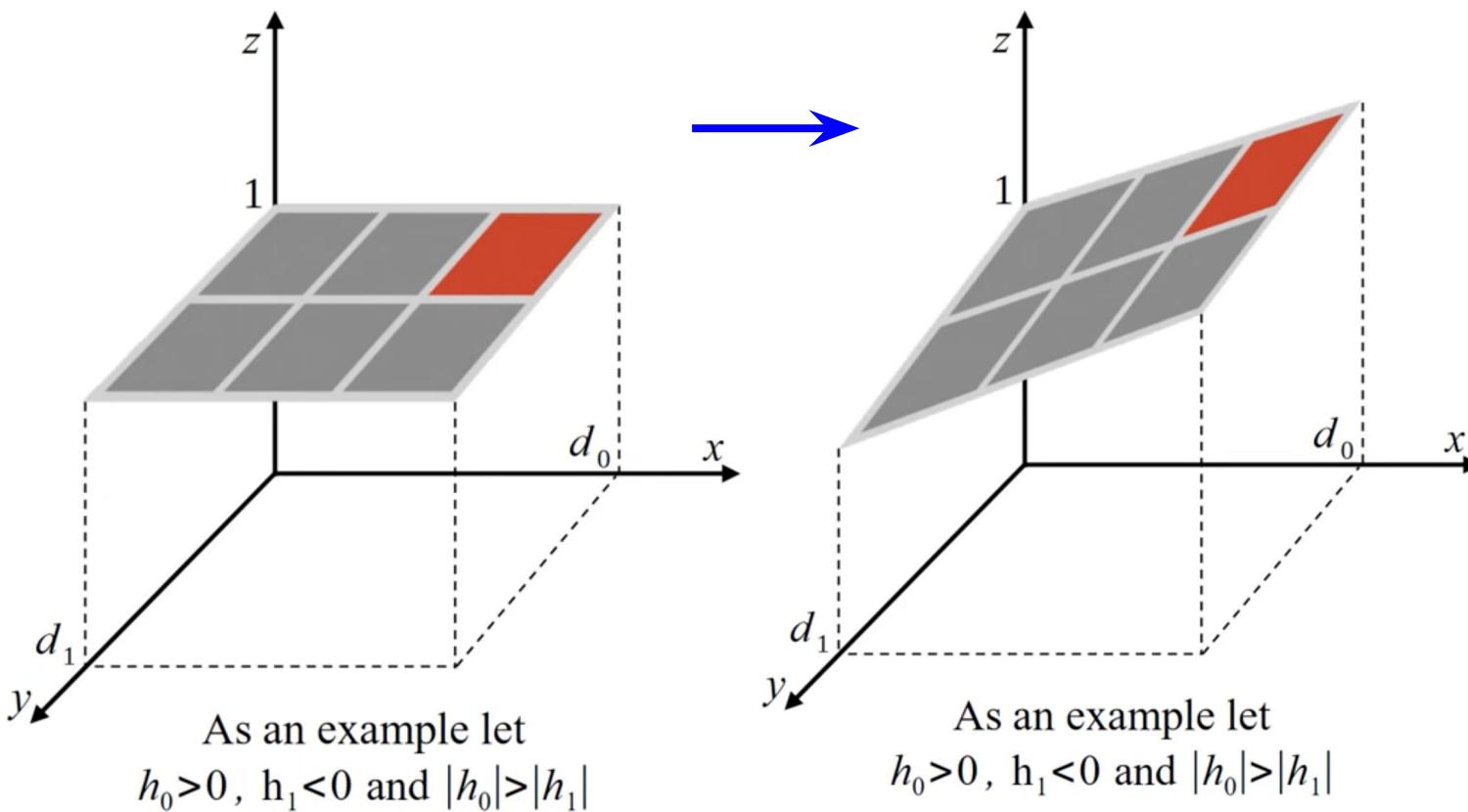


Figure: https://www.youtube.com/@huseyin_ozde...

Projective Transformation (homography)

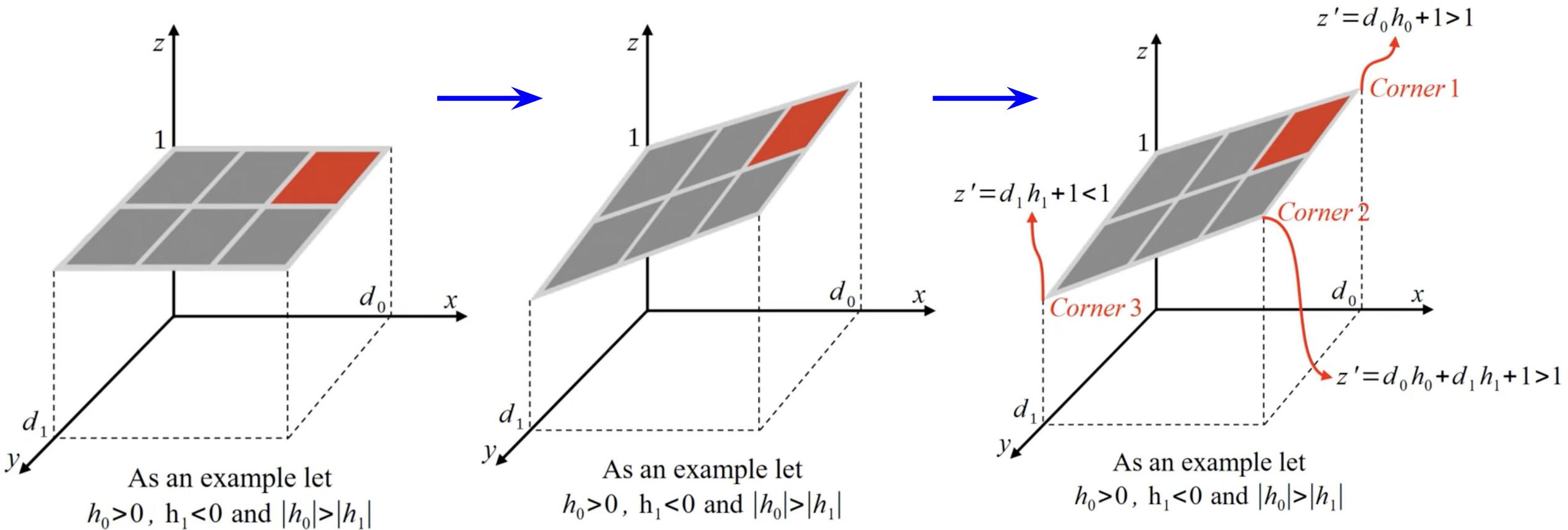


Figure: https://www.youtube.com/@huseyin_ozde...

Projective Transformation (homography)

When going back to
Cartesian coordinates

$$\frac{x'}{z'}$$

$$\frac{y'}{z'}$$

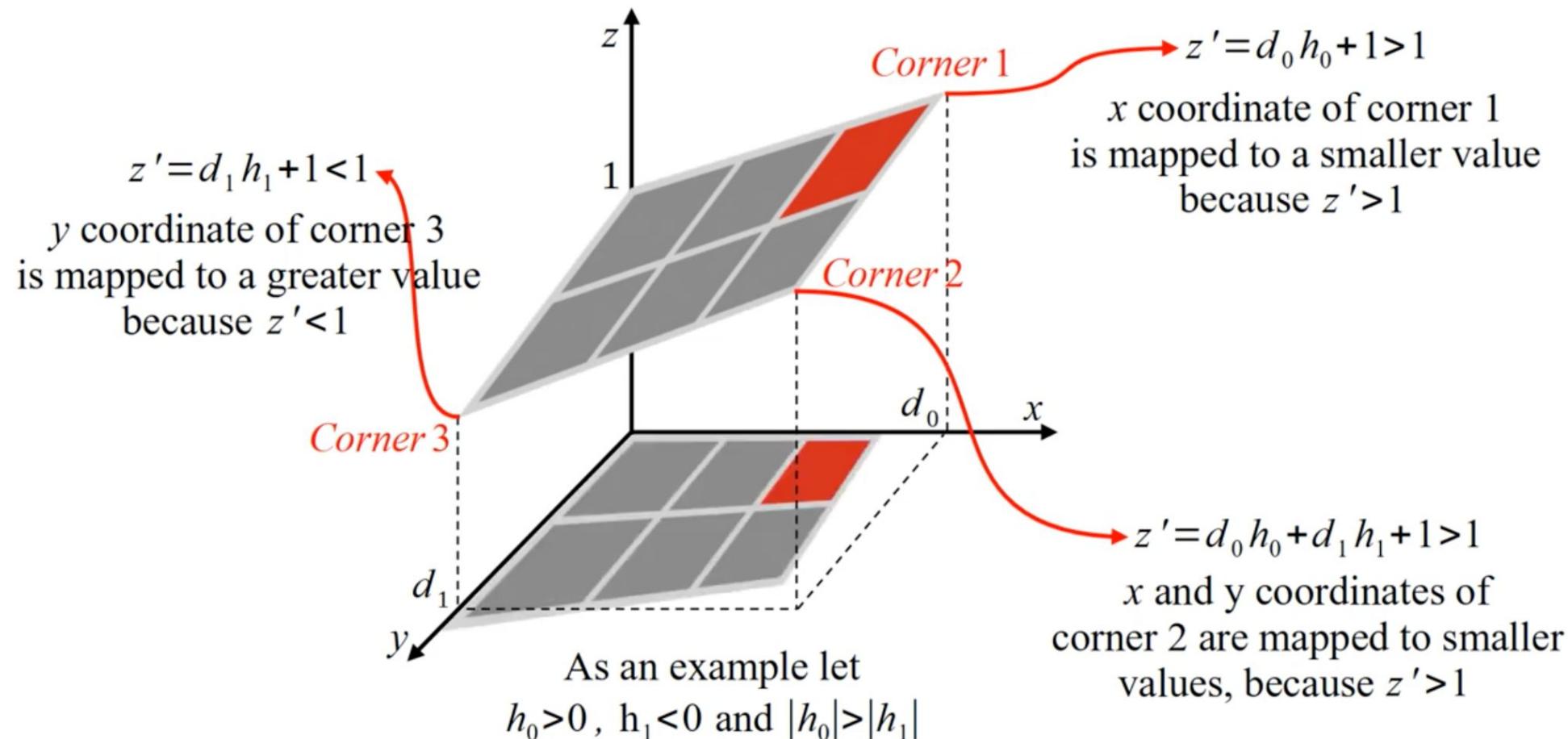
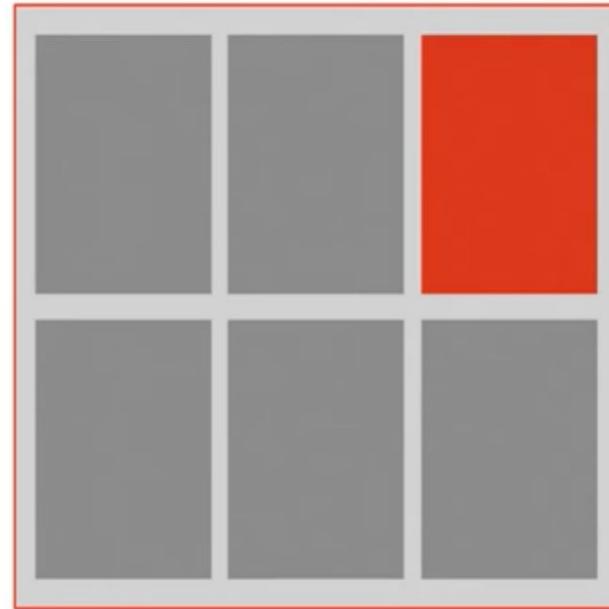


Figure: https://www.youtube.com/@huseyin_ozde...

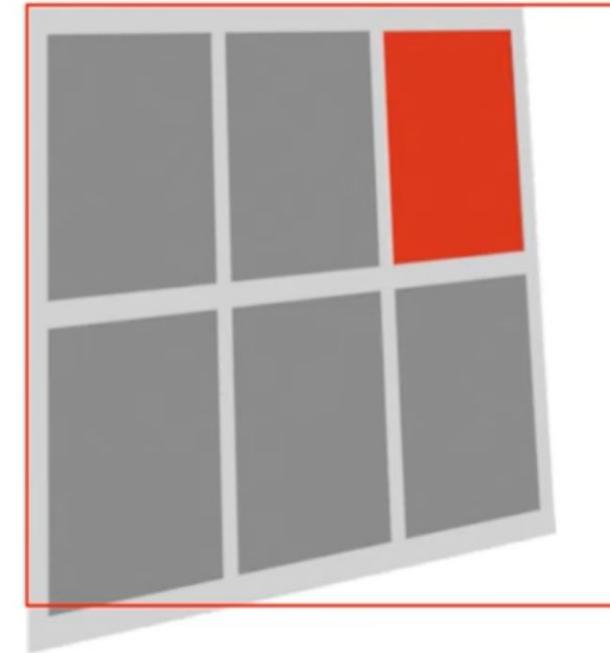
Projective Transformation (homography)



Original Image

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_0 & h_1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$h_0 > 0$, $h_1 < 0$ and $|h_0| > |h_1|$



Warped Image

Figure: https://www.youtube.com/@huseyin_ozde...

Projective Transformation (homography)

Projective transformations are combinations of

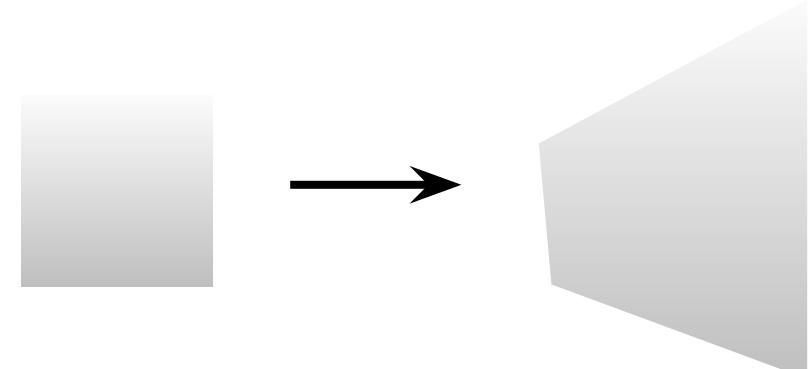
- Affine transformations + projective warps

$$w \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_{00} & A_{01} & b_0 \\ A_{10} & A_{11} & b_1 \\ h_0 & h_1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

How many degrees of freedom?

Properties of projective transformations:

- origin does not necessarily map to origin
- lines map to lines
- parallel lines do not necessarily map to parallel lines
- ratios are not necessarily preserved



Questions?

Composing Transformations

Transformations = Matrices => Composition by Multiplication!

$$p' = R_2 R_1 S p$$

In the example above, the result is equivalent to

$$p' = R_2(R_1(Sp))$$

Equivalent to multiply the matrices into single transformation matrix:

$$p' = (R_2 R_1 S)p$$

Order Matters! Transformations from *right to left*.

Scaling & Translating != Translating & Scaling

$$p'' = TSp = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & t_x \\ 0 & s_y & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x x + t_x \\ s_y y + t_y \\ 1 \end{bmatrix}$$

$$p''' = STp = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & s_x t_x \\ 0 & s_y & s_y t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x x + s_x t_x \\ s_y y + s_y t_y \\ 1 \end{bmatrix}$$

Similarity: Translation + Rotation + Scaling

$$p' = (T R S) p$$

$$p' = TRSp = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= [R \quad t] [S \quad 0] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \boxed{\begin{bmatrix} RS & t \\ 0 & 1 \end{bmatrix}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

This is the form of the general-purpose transformation matrix

2D Transforms = Matrix Multiplication

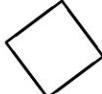
Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	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	

Table 2.1 *Hierarchy of 2D coordinate transformations, listing the transformation name, its matrix form, the number of degrees of freedom, what geometric properties it preserves, and a mnemonic icon. Each transformation also preserves the properties listed in the rows below it, i.e., similarity preserves not only angles but also parallelism and straight lines. The 2×3 matrices are extended with a third $[0^T \ 1]$ row to form a full 3×3 matrix for homogeneous coordinate transformations.*

Figure: R. Szeliski

3D Transforms = Matrix Multiplication

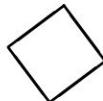
Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix}_{3 \times 4}$	3	orientation	
rigid (Euclidean)	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}_{3 \times 4}$	6	lengths	
similarity	$\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}_{3 \times 4}$	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	

Table 2.2 *Hierarchy of 3D coordinate transformations. Each transformation also preserves the properties listed in the rows below it, i.e., similarity preserves not only angles but also parallelism and straight lines. The 3×4 matrices are extended with a fourth $[0^T \ 1]$ row to form a full 4×4 matrix for homogeneous coordinate transformations. The mnemonic icons are drawn in 2D but are meant to suggest transformations occurring in a full 3D cube.*

Figure: R. Szeliski