ECE 661
Purdue Univ.
Avi Kak
Lecture 3

World 2D: Planar Projective Transformations and Transformation Groups

Reference: "Multiple View Geometry in Computer Vision" by Hartley and Zisserman

- From now on, we will call an algebraic structure homogeneous if it conveys real-world information through ratios. Homogeneous structures you have seen so far: ① 3-vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ to represent points (x,y) in $\mathbb{R}^2$. ② 3-vectors $l = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$ to represent lines in $\mathbb{R}^2$. ③ 3x3 matrices $C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ c_{13} & c_{23} & c_{33} \end{bmatrix}$ to represent conics in $\mathbb{R}^2$.
- Let's now talk about constructing mappings from one (x,y)-plane to another (x,y)-plane. If we denote a mapping by H, we are obviously talking about $H: \mathbb{R}^2 \rightarrow \mathbb{R}^2$.
- Representing the points in the planes by homogeneous 3-vectors, we have $H: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.
- A Planar Projective Transformation is a linear transformation on homogeneous 3-vectors, the transformation being represented by a non-singular 3x3 matrix H, as in

Therefore, H^{-1} always exists

$$\boxed{\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = H \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}} \rightarrow x' = Hx$$

Since only the ratios of the 3-vectors are important to us, it is only the ratios of the elements of H that matter. In other words, H multiplied by an arbitrary non-zero scalar does not alter the transformation. So we say H is homogeneous.

- A planar projective transformation is more popularly known as Homography

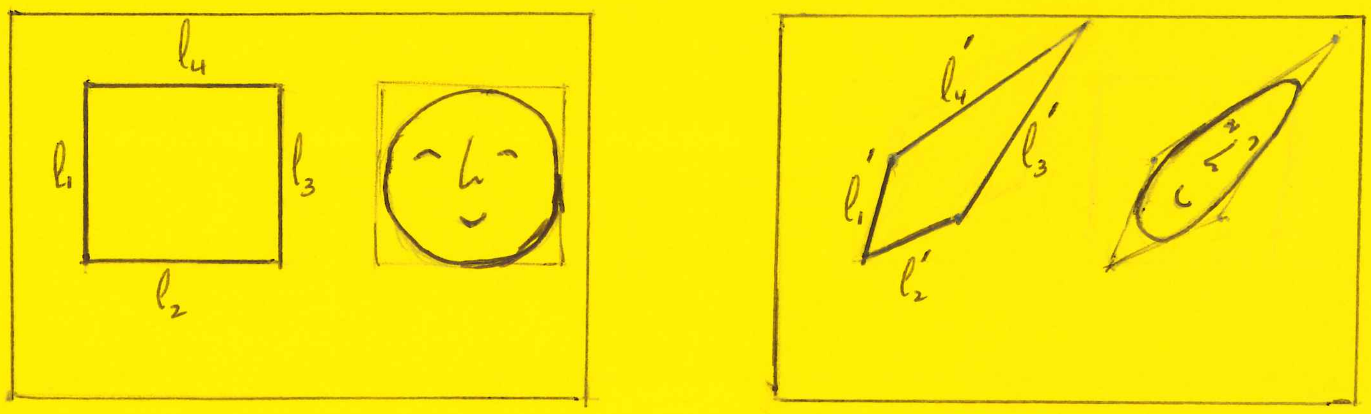
---

- A fundamental property of a homography is that it always maps a straight line to a straight line.
- To prove the above, let x_1, x_2, and x_3 be three points on a line l in the domain of a homography H. We have $\boxed{l^T x_i = 0,\ i=1,2,3}$. We are also given $\boxed{x_i' = H x_i}$, implying $\boxed{x_i = H^{-1} x_i'}$. Therefore, $l^T H^{-1} x_i' = 0$, which means that the mapped points x_i' are all on a line $\boxed{l' = H^{-T} l}$.

$H^{-T} = (H^{-1})^T$

- The proof above also tells us that as points transform according to H, lines transform according to H^{-T}.
- Let's now see how conics are mapped. A conic C in the domain of H is the set of points that obey $\boxed{x^T C x = 0}$. With $\boxed{x = H^{-1} x'}$, we can write $\boxed{x'^T H^{-T} C H^{-1} x' = 0}$, implying that the mapped conic is given by $C' = H^{-T} C H^{-1}$

- Considering that straight lines go into straight lines, what sort of visual distortion might we expect from a general homography? You could see a scene such as the one shown at left below turn into a scene that is shown at right:



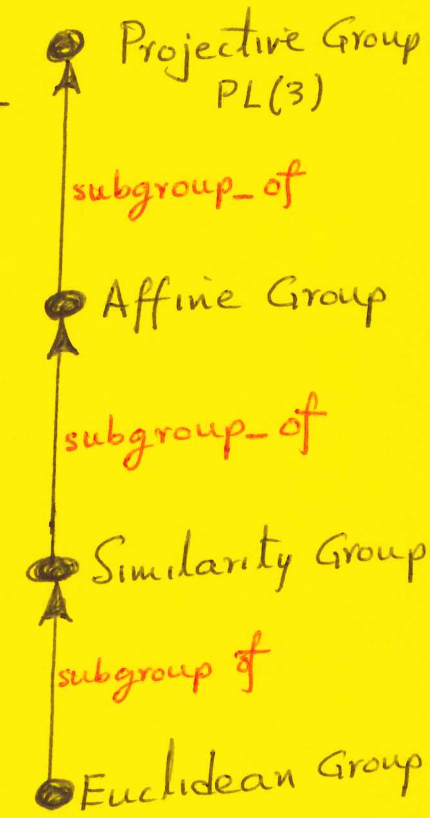
This is what you'd get if you took a photograph of windows and picture frames mounted on a wall when the objects you are photographing are high on the wall and off to a side in relation to where you are standing

A Hierarchy of Transformations

- Now that you understand what we mean by a general homography (meaning, a Planar Projective Transform), it is time to consider its various special cases. But first we must define a group, since the set of all Planar Projective Transforms forms a group, as do its various special cases.
- A group is a set, along with a binary operator defined for the elements of the set, provided the following four conditions are satisfied: (i) The set must be closed with respect to the operator ($a \circ b$ is in the set if a and b are); (ii) The operator must be associative $\left(a \circ (b \circ c) = (a \circ b) \circ c\right)$; (iii) The set must contain an identity element (that is, an element i such that $a \circ i = i \circ a = a$ for all a in the set); and (iv) the inverse of every element must be in the set (for every a we can identify an element that we denote a^{-1} such that $a \circ a^{-1} = i$).
- The set of all 3x3 nonsingular matrices, along with matrix multiplication as the operator, forms a group. We will denote this group $GL(3)$ where 'GL' stands for 'General Linear'.
- While every element of $GL(3)$ can be used as a homography, we know that any two elements that are related by a scalar multiplier stand for the same homography. So, in order to recognize distinct homographies in $GL(3)$, we partition the set into equivalence classes where all the 3x3 matrices in the same equivalence call are related by scalar multipliers.
- We refer to the set of equivalence classes as the $\boxed{\text{quotient group } PL(3)}$, where 'PL' stands for Projective Linear.

- The "Hierarchy of Transformations" refers to the group PL(3) and its various subgroups. PL(3) is at the root of the hierarchy. (For a subset of a group to constitute a subgroup, the subset must satisfy all four conditions that were mentioned earlier.)

- The specific subgroups of PL(3) that we are interested in are the Affine, the Similarity, and the Euclidean, as shown in the figure at right:



- What are the engineering benefits of recognizing the group structures shown at right? The group structures guarantee us that, say, two successive applications of the affine transform (with two different matrices) will be affine. For another example, applying first an affine transform followed by a similarity transform will result in an overall application of an affine transform (because a similarity transform ISAn affine transform).

- Before providing definitions for the three subgroups of PL(3), let's first express the general planar projective transform in the following manner:

$$H = \begin{array}{c} \begin{bmatrix} A & \vec{t} \\ \vec{v}^T & v \end{bmatrix} \updownarrow 3 \\ \xleftrightarrow{\;3\;} \end{array} \qquad \text{with} \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \vec{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

The reason for the choice of symbols will become clear shortly.

- Since H is homogeneous — meaning that only the ratios of the elements of H are important — we will frequently set the element $v = 1$. [However, beware, there will be situations when we will have to allow v to become zero. More on that later.]

- We are now ready to define the three subgroups of PL(3).

Affine Transformations

- The affine group is obtained from the projective group by restricting the last row of the 3x3 transformational matrix to (0 0 1). You can easily show yourself that if we multiply two 3x3 matrices, each with its last row restricted to (0, 0, 1), the result is a matrix with its last row as (0, 0, 1).

- Affine transformations are commonly expressed as

$$\begin{pmatrix} x_1' \\ x_2' \\ 1 \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

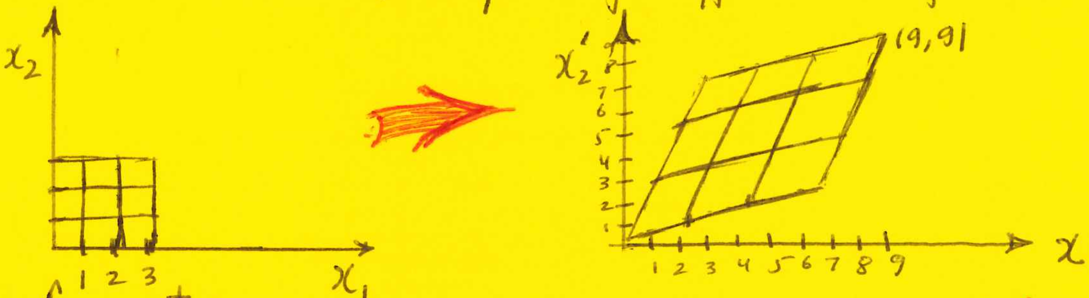
(x_1, x_2) and (x_1', x_2') are physical coordinates in their respective planes

$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is non-singular → $\det(A) \neq 0$

- To better understand the roles played by the different elements of an affine transformation matrix, we'll first consider the case when $t_x = t_y = 0$.

- With t_x and t_y set to zero, consider an example of affine transformation:

$$\boxed{\begin{aligned} x_1' &= 2x_1 + x_2 \\ x_2' &= x_1 + 2x_2 \end{aligned}}$$



- As you can see, an affine transformation not only takes straight lines into straight lines, it also keeps parallel lines parallel. While stretching the plane in one direction, it shrinks it in the orthogonal direction, but in such a way that parallel lines stay parallel.
- It is easy to prove that the affine transform keeps parallel lines parallel. You do so by showing that it always maps a point an infinity into a point at infinity. That is, ideal points always stay ideal. Proof consists of showing that the matrix-vector product $\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$ always gives a vector whose last coordinate is zero — an ideal point.
- For an algebraic understanding of the differential scaling of the domain plane along two orthogonal axes, as you can see in the figure at the top, it is best to express the matrix A through its singular value decomposition:

$$A = UDV^T$$

U : 2x2 orthogonal
V : 2x2 orthonormal
$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$

cols. of V are eigenvectors of A^TA

Being orthogonal, both U and V are rotational matrices. A being nonsingular guarantees that both λ_1 and λ_2 $(\lambda_1 \geqslant \lambda_2)$ exist. Let's rewrite this SVD as

$$A = (UV^T)(VDV^T)$$

since V is orthonormal $V^TV = I$

rotation through V^T aligns principal eigenvector with x-axis

Let V stand for rotation through φ and let UV^T stand for rotation through θ. (Means the largest eigenvector is at ange φ w.r.t x-axis.) The above equation says to rotate the domain plane through $-\varphi$, scale it by λ_1 along x and by λ_2 along y, and then restore the original orientation of the plane. Finally, we must rotate the plane through θ.

- With regard to the role played by t_x and t_y, it is easy to show that these simply translate the pattern after it has been rotated and scaled by A. Just try using $t_x = 4$ and $t_y = 2$ in the example at the top of this page.

Similarity Transformations

- The similarity group is obtained from the affine group by requiring that A be orthogonal: $A^TA = \lambda^2 I$
- Similarity transform preserves angles.
- To prove the above, first note if $X' = AX$ then $\|X'\|^2 = X^TA^TAX = \lambda^2\|X\|^2$
- If X and Y are two different points in the domain plane and X' and Y' the corresp. points in the range plane, then

$$\frac{\cos(X', Y')}{\cos(X, Y)} = \frac{X'^TY'}{X^TY} \cdot \frac{\|X\|\,\|Y\|}{\|X'\|\,\|Y'\|} = \frac{X^TA^TAY}{X^TY} \cdot \frac{1}{\lambda^2} = 1$$

shape is preserved

Try: $\begin{pmatrix} x_1' \\ x_2' \\ 1 \end{pmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$

Euclidean Transformations

- The euclidean group is obtained from the similarity group by further requiring that A be orthonormal. That is $\boxed{A^TA = I}$ (2x2 identity matrix)
- This group represents transformations corresponding to rigid body motions.
- Preserves Euclidean distance between any two points in the domain plane.
- Let X and Y be two different points in the domain plane and X' and Y' the corresp. points in the range plane. Let d denote the Euclidean distance:

$$d(X', Y') = \sqrt{(X'-Y')^T(X'-Y')} = \sqrt{(X-Y)^TA^TA(X-Y)} = d(X, Y)$$