Computer Vision, 3D Geometry and Machine Learning Notes

Samarth Brahmbhatt

November 18, 2018

Contents

Chapter 1

3D Geometry

1.1 Homogeneous Coordinates

The homogeneous representation of an image point $\mathbf{x} = [u, v]$ is denoted as $\tilde{\mathbf{x}} = [u, v, 1]$. A line in the image passing through $\tilde{\mathbf{x_1}}$ and $\tilde{\mathbf{x_2}}$ is $\tilde{\mathbf{l}} = \tilde{\mathbf{x_1}} \times \tilde{\mathbf{x_2}}$.

1.2 Plücker Coordinates

1.3 Types of Transforms

Table 1.1 shows the types of transforms on 2D homogeneous coordinates. Each transform preserves all the quantities in its row and below it.

Transform	DoF	Preserves
Translation	2	absolute orientation
Rotation	1	distances
Similarity	4	angles
Affine	6	parallel lines
Projective	8	ratios of distances

Table 1.1: Types of transforms on 2D homogeneous coordinates

1.4 Single View Geometry

If the transformation from world to camera coordinates is ${}^{c}T_w = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$, the pinhole camera projection equation is

$$\tilde{\mathbf{x}} = K[R|t]^{\mathbf{w}}\tilde{\mathbf{X}} \tag{1.1}$$

where $\tilde{\mathbf{X}}$ is the homogeneous representation of the 3D point in world coordinates, and

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1.2}$$

is the ideal pinhole camera intrinsics matrix. $P = K[R|t]$ is called the camera projection matrix.

1.4.1 Dissecting the Camera Projection Matrix

Let's denote the columns of P as $P = [P_1|P_2|P_3|P_4]$. Now, $P_1 = P \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Hence P_1 is the image of vanishing point in the world coordinate X direction. Similarly for P_2 and P_3. P_4 is the image of the world origin.

Let's denote the rows of P as $P = \begin{bmatrix} P_1^T \\ P_2^T \\ P_3^T \end{bmatrix}$. Since $P_1^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} P$, P_1^T is the plane passing through the camera center and the image line $[1, 0, 0]$. To see this, consider a projected point $\tilde{\mathbf{x}} = P\tilde{\mathbf{X}}$. $\tilde{\mathbf{x}}$ lies on the image line $\tilde{\mathbf{l}}$ iff

$$\tilde{\mathbf{l}}^T\tilde{\mathbf{x}} = 0 \tag{1.3}$$

$$\iff \tilde{\mathbf{l}}^T P\tilde{\mathbf{X}} = 0 \tag{1.4}$$

$$\iff (P^T\tilde{\mathbf{l}})^T\tilde{\mathbf{X}} = 0 \tag{1.5}$$

Thus the plane corresponding to image line $\tilde{\mathbf{l}}$ is $P^T\tilde{\mathbf{l}}$.

Representing lines with Plücker coordinates