Class 2

Shikhar Saxena

January 06, 2024

We want the aperture to be ideal (small/large both are bad).

Object size vs Object motion vs Camera motion

The position of the image captured depends on the height of the camera (and where it's placed compared to the object).

Problem: Person who is 1.75m tall standing at 7m from camera. Pinhole camera has focal length of 20mm. The sensor is 1cm tall and has a resolution of 4000x3000. Find the height of the person in pixels in the image.

Using similar triangle concepts:

$$\frac{1.75}{x} = \frac{7000}{20}$$

$\implies x = 0.5$ cm

The aperture covered (from center to bottom): bottom half. The image completely fits on the sensor.

Convention: WxH resolution usually. Given, height of the sensor is 1cm (which will be 3000 pixels). So 0.5cm will be 1500 pixels.

If we raise the camera by 1m, how much does the person move in the sensor (in pixels)?

The person in the image will start going from bottom half to top half. This can be calculated. Say we move 1.75/2 m up then the image will be at the center of the sensor for example.

How much does the sun move in the above case. Note: Sun is 150 million kms away (in pixels)?

Zero. This is because the object is so far away that the light coming in is parallel.

Coordinate System

- $[X, Y, Z]$: World coordinate system

    - ⋆ This assumes a frame of reference (origin)
- ∘ $[x, y]$: Image coordinate system
- ∘ Camera model: a function that maps world coordinates to image coordinates
    - ⋆ This function is not one-to-one. But many-to-one mapping.
- ∘ Perspective mapping: How the many-to-one mapping is achieved.

Transformations: 3D Translation

Linear Transformation: $X' = TX$

Having an extra dimension accounts for homogeneous coordinate system. Projective space becomes the 4th dimension. Adding a 4th dimension makes it easier to compute translation essentially. And you can easily get back to cartesian coordinate system.

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X + d_x \\ Y + d_y \\ Z + d_z \end{bmatrix}$$

Scaling

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X\ S_x \\ Y\ S_y \\ Z\ S_z \end{bmatrix}$$

Rotation (around z-axis)

$X = R\cos\phi$ and $Y = R\sin\phi$.

$X' = R\cos(\phi + \theta)$ and $Y' = R\sin(\phi + \theta)$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

We can also rotate arbitrarily. We'll have angles of rotations along each axis then. Order of Rotation is important because they are not permutation-insensitive. Rotate x in one axis then y in another will end up in a different position than say y in other axis then x in the first axis.

Cartesian ↔ Homogeneous coordinates

Add a new dimension:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} kX \\ kY \\ k \end{bmatrix}$$

Getting back to cartesian:

$$\begin{bmatrix} X \\ Y \\ w \end{bmatrix} = \begin{bmatrix} X/w \\ Y/w \end{bmatrix}$$

—

- Perspective Projection
- Basic Camera Equation