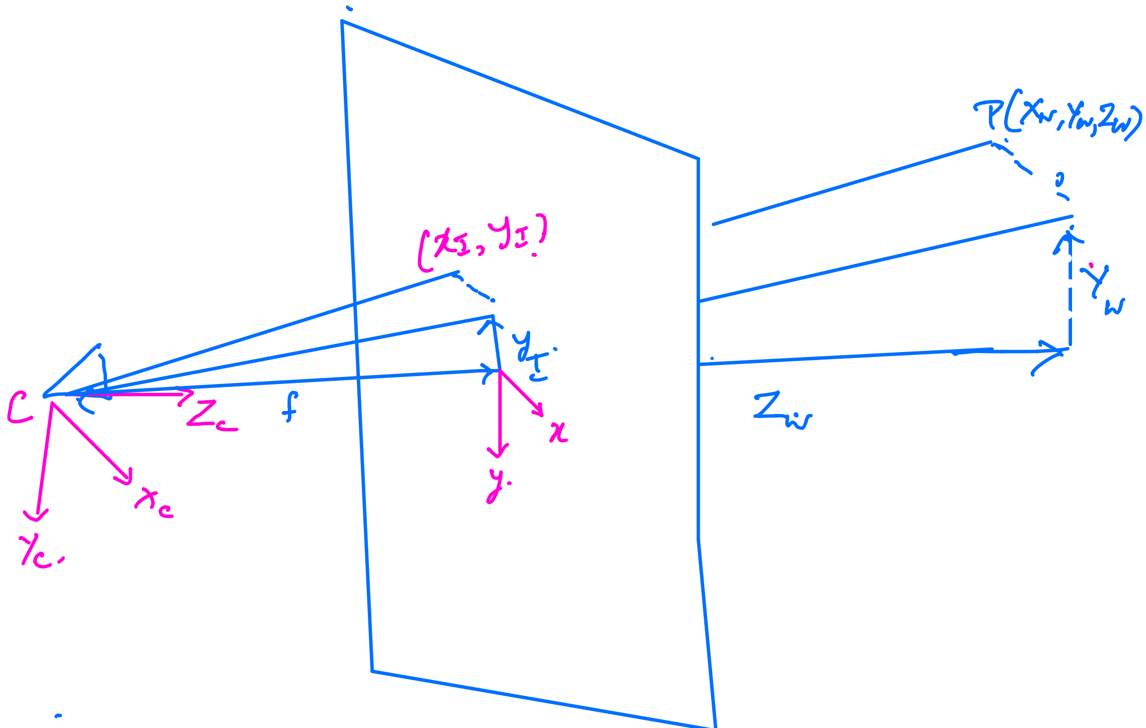


$$\frac{Y_w}{y_I} = \frac{Z_w}{f} \longrightarrow (1)$$

or $$\frac{-Y_w}{-y_I} = \frac{Z_w}{f} \longrightarrow (2)$$ (Since y or Y_c points downward)

Similarly

$$\frac{X_w}{x_I} = \frac{Z_w}{f} \longrightarrow (3)$$

Dropping suffices w, I we get.

$$y = \frac{fY}{Z} \longrightarrow (4)$$

$$x = \frac{fX}{Z} \longrightarrow (5)$$

or $$\begin{bmatrix} Zx \\ Zy \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \longrightarrow (6)$$

or $$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \longrightarrow (7)$$

or $$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K \vec{X} \longrightarrow (8)$$

$(\vec{X} = [X \ Y \ Z]^T)$

or $$\boxed{\lambda \vec{x} = K \vec{X} \longrightarrow (9)}$$

CENTRAL PROJECTION

$$\boxed{\vec{x} \approx K\vec{X}}$$

$\vec{x}$ is in homognous coordinate and is the analog coordinate or the coordinate of an image point formed on the image plane (film) prior to digitization

$\vec{x}$ is in the same metric scale as $\vec{X}$.

If $\vec{X}$ is measured in metres so is $\vec{x}$

The digital camera has CCD sensor array that senses the image formed on the image plane; which maps the image location to a pixel location with a corresponding intensity value

→ Let k be the scaling factor that takes from image to pixel coordinate

$$u = \alpha x \quad \longrightarrow (10). \qquad v = \alpha y \longrightarrow (11).$$

$$u = \frac{\alpha f X}{Z} \longrightarrow (11) \qquad v = \frac{\alpha f Y}{Z} \longrightarrow (12)$$

$$uZ = \alpha f X \longrightarrow (13), \qquad vZ = \alpha f Y \longrightarrow (14).$$

$$\text{or} \quad Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha f & 0 & 0 \\ 0 & \alpha f & 0 \\ 0 & 0 & 1 \end{bmatrix} X$$

$$\text{or} \quad \lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K\vec{X} \longrightarrow (13)$$

($\lambda = Z$ is the depth to the 3D point)

The above assumes that the image coordinates are in the center of the image.

However they are typically in the top left corner and the image center is (u_0, v_0) then, we have

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \vec{X} \qquad (16)$$

Note that any $\gamma \vec{X}$ also project to $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$ as.

$$\gamma \vec{X} = \lambda \begin{bmatrix} \gamma u \\ \gamma v \\ \gamma \end{bmatrix} = \lambda \gamma \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \qquad (17)$$

(In homogenous coordinate system).

We use $x = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ interchangeably with $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$ to denote the pixel coordinate

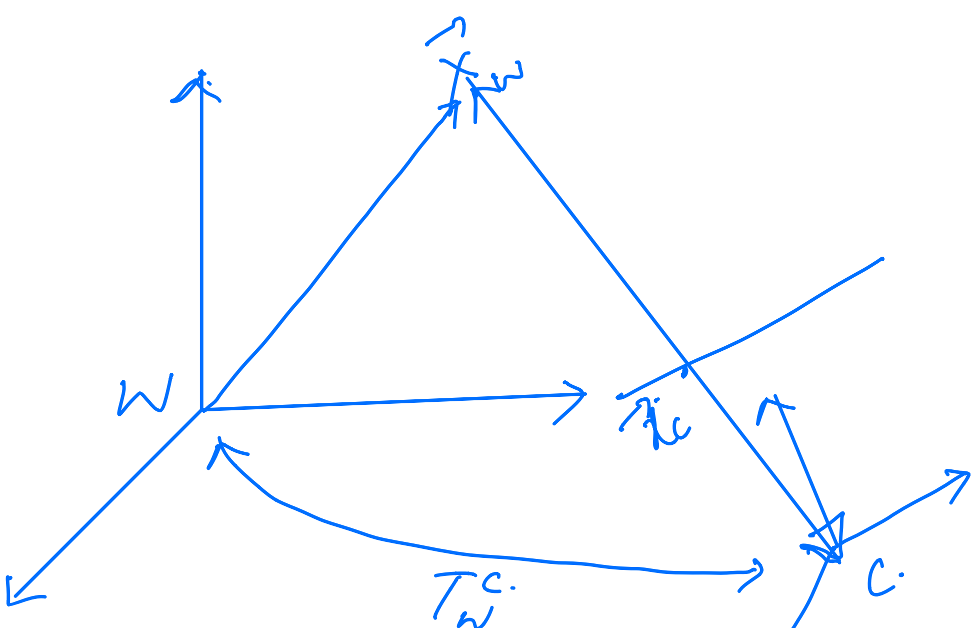
$\lambda \vec{x} = K\vec{X}$

$\vec{x} \approxeq K\vec{X}$ (projective equivalence).

or $K^{-1}\lambda\vec{x} = \vec{X}$.

or $\boxed{\lambda K^{-1}\vec{x} = \vec{X}} \rightarrow (18)$

or $K^{-1}\vec{x}$ is parallel to or points to the direction of the 3D vector $\vec{X}$



$$\lambda \vec{x}_c = K T_w^c \vec{X}_w \quad (4\times 1)$$

$$= \underbrace{K}_{\text{camera intrinsics}} \underbrace{[R \quad t]}_{\text{extrinsics.}} \vec{X}_w \quad (3\times 1).$$

or $\vec{x}_c \simeq K [R \quad t] \vec{X}_w \quad 3\times 1.$

or $\vec{x}_c = K [R \quad t] \vec{X}_w$

(writing losely).

or $\vec{x}_c \simeq K \vec{X}_c$

or $\vec{x}_c$ ||el to $K \vec{X}_c$

or $\text{cross}(\vec{x}_c, K \vec{X}_c) = \vec{0}.$

Suppose $X^c = [4\ 4\ 2]^T$ in cms.

$$K = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

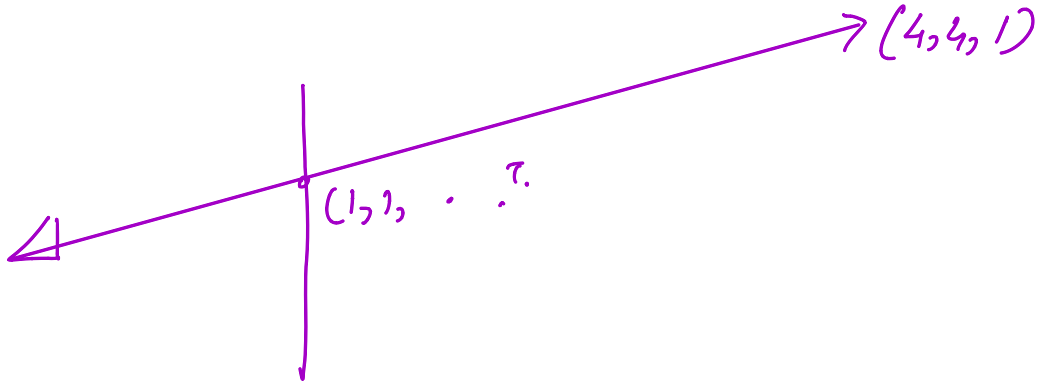
Then $x = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}$

$$= \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

or X^c projects to the image location $(1,1)$ due to the action of K

→ Projection is a non-linear operation.

For the ray or vector (4,4,2). projects to (1,1,1) instead of (1,1,0.5).



It is non linear evidently as

$$\left.\begin{aligned} x &= \frac{fX}{Z} \\ y &= \frac{fY}{Z} \end{aligned}\right\} \text{Non Linear Map}$$

α = scaling factor that takes f in cms to pixels.

$$\alpha = \frac{\text{image-width (pix)}}{\text{sensor-width (cms)}}$$

$f(\mu x) = \alpha f(cms)$

Let $\alpha = 30$, then $f_x = f_y = 30(5.-)$

$= 15.0$

Then $x = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$ or $\begin{bmatrix} u \\ v \\ w \end{bmatrix}$

$$= \begin{bmatrix} 150 & 0 & 0 \\ 0 & 150 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 600 \\ 600 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 300 \\ 300 \\ 1 \end{bmatrix}$$

If the origin is the top left corner of the image and the image center is at $c_x, c_y = (320, 240)$.

Then $x = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 150 & 0 & 320 \\ 0 & 150 & 240 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}$

$= \begin{bmatrix} 1240 \\ 1080 \\ 2 \end{bmatrix} = \begin{bmatrix} 620 \\ 540 \\ 1 \end{bmatrix}$