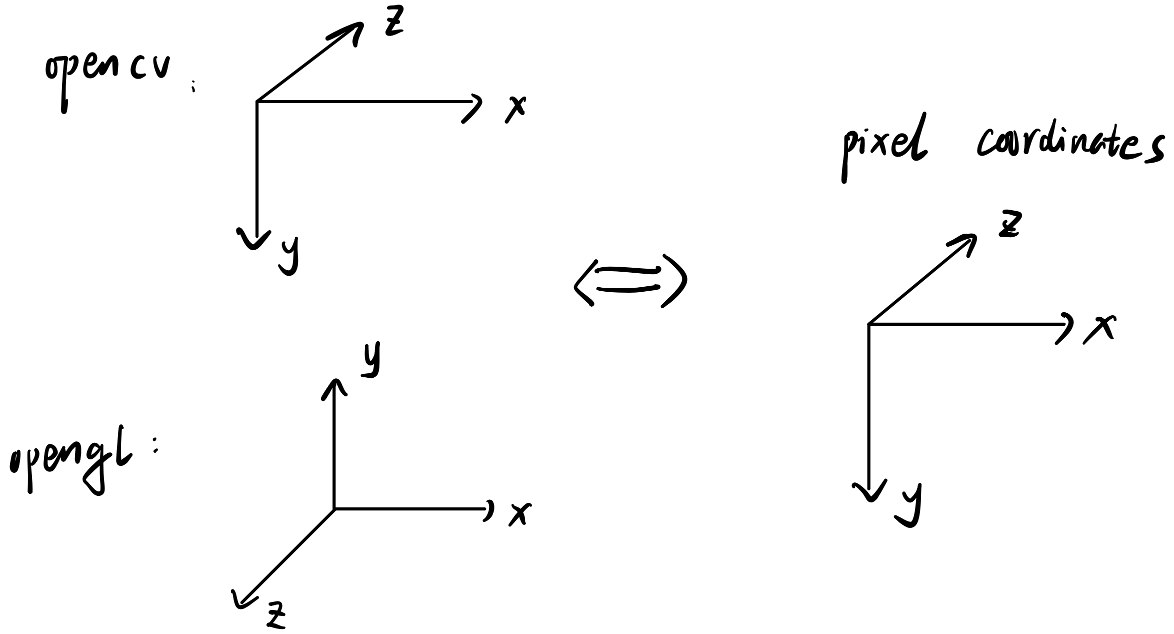
1. pixel plane ⟷ image plane

$$P_p = T_i^p \, P_i$$

$$= T_{(\text{offset}+\text{scaling})} \cdot T_{\text{axis\_rot}} \quad P_i$$

1.1 axis rotation : the definition of the camera coordinate system varies from different systems.



$\Rightarrow \begin{bmatrix} \boxed{\begin{matrix}1\\0\\0\end{matrix}} & \boxed{\begin{matrix}0\\-1\\0\end{matrix}} & \boxed{\begin{matrix}0\\0\\-1\end{matrix}} \end{bmatrix}$ for opengl

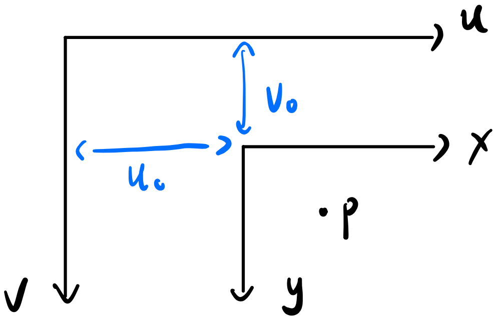
🟧 : the y axis rotates 180°

🟪 : the z axis rotates 180°

🟦 : the x axis is unchanged.

for opengl : the rotation mtx is 3x3 Identity mtx.

1.2 offset + scaling



$P_i : [x, y, 1]^T$

$P_p : [u, v, 1]^T$

After rotating the image plane to making its axis directions being aligned with the pixel plane, the coords of the point on the image plane need to be further scaled from meter metric to pixel metric, and then an offset (u_0, v_0) is added.

$$\begin{cases} u = \dfrac{x}{dx} + u_0 \\ v = \dfrac{y}{dy} + v_0 \end{cases} \Rightarrow$$

dx, dy are always set to 1

u_0, v_0 are always set to $W/2$, $H/2$

$$\Rightarrow T_{(\text{offset+scaling})} = \begin{bmatrix} \frac{1}{dx} & & u_0 \\ & \frac{1}{dy} & v_0 \\ & & 1 \end{bmatrix}$$

$$\Rightarrow \quad P_p = T_{(o+s)} \, T_{rot} \, P_i$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & & u_0 \\ & \frac{1}{dy} & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} s_1 & & \\ & s_2 & \\ & & s_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

for $P_i = T_{rot}^{-1} \, T_{(o+s)}^{-1} \, P_p$

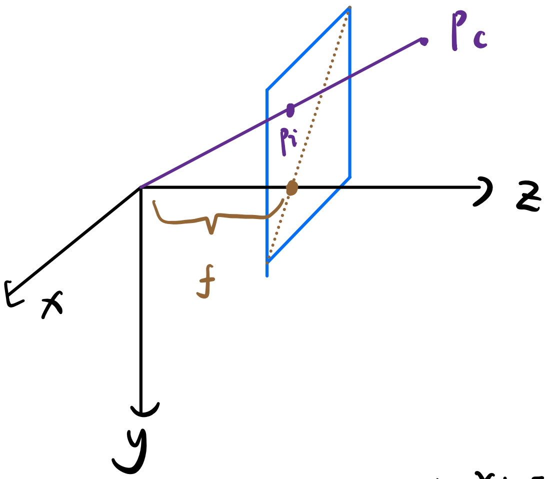
$T_{(o+s)}^{-1}$: de-offset + de-scaling

$$\Rightarrow \begin{cases} x = dx\,(u - v_0) \\ y = dy\,(v - v_0) \end{cases} \Rightarrow T_{(o+s)}^{-1} = \begin{bmatrix} dx & & -dx \cdot u_0 \\ & dy & -dy \cdot v_0 \\ & & 1 \end{bmatrix}$$

$T_{rot}^{-1} = T_{rot}^{T} = T_{rot}$ if each axis is rotated by either 0° or 180°.

$$\Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_1 & & \\ & s_2 & \\ & & s_3 \end{bmatrix}^T \begin{bmatrix} dx & & -dx \cdot u_0 \\ & dy & -dy \cdot v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

2. image plane $\Longleftrightarrow$ camera coordinates.



$P_c : [X_c, Y_c, Z_c, 1]^T$

$P_i : [X_i, Y_i, Z_i, 1]^T$

$$\begin{cases} \frac{X_c}{X_i} = \frac{Y_c}{Y_i} = \frac{Z_c}{Z_i} = \text{constant} \\ Z_i = f. \end{cases}$$

$$\Rightarrow X_i = f \cdot \frac{X_c}{Z_c}, \quad Y_i = f \cdot \frac{Y_c}{Z_c}$$

$$\Rightarrow P_i' = f\left[\frac{X_c}{Z_c}, \frac{Y_c}{Z_c}, 1, \frac{1}{f}\right]^T$$

take x_i, y_i only, set $z=1$

$$\Rightarrow P_i = [x, y, 1]^T$$

$$= \frac{1}{Z_c}\begin{bmatrix} fX_c \\ fY_c \\ Z_c \end{bmatrix} = \frac{1}{Z_c}\begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix}\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

$$\Rightarrow T_c^i = \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix},$$

and the result should be normalized with the third elem = 1.

$$T_i^c = \begin{bmatrix} 1/f & & \\ & 1/f & \\ & & 1 \end{bmatrix}.$$

and the depth should be multiplied as a scalar to the image coords $[x, y, 1]^T$ before $* \; T_c^i$.

$$P_c = T_i^c \cdot z_c \cdot P_i$$

3. Camera coordinate $\Longleftrightarrow$ World coordinate.

$$P_w = T_c^w P_c = \begin{bmatrix} R_{cam} & C_{cam} \\ & 1 \end{bmatrix} P_c = \begin{bmatrix} R^T & -R^T t \\ & 1 \end{bmatrix} P_c$$

$$P_c = T_w^c P_w = \begin{bmatrix} R & t \\ & 1 \end{bmatrix} P_w = \begin{bmatrix} R_{cam}^T & -R_{cam}^T C_{cam} \\ & 1 \end{bmatrix} P_w$$