1. Camera model

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{d_x} & 0 & c_x \\ 0 & \frac{1}{d_y} & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Camera Intrinsics}} \underbrace{\begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}}_{\text{Camera Extrinsics}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

origin point is left-up corner; unit is pixel

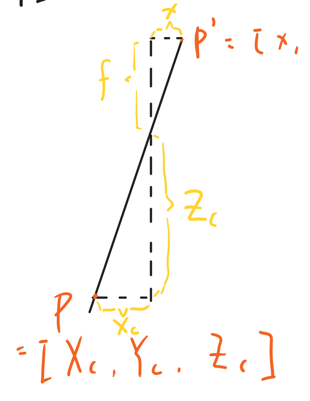
image plane coordinate (origin point is the center of image plane)

camera coordinate

world coordinate

1.1 Camera Extrinsics : easy to understand

1.2 Camera Intrinsics : 3D $P \to$ 2D p'



$\leftarrow$ imagin the 3D situation

$$\frac{Z_c}{f} = \frac{X_c}{x} = \frac{Y_c}{y}$$

$$\Rightarrow \begin{cases} x = f\frac{X_c}{Z_c} \\ y = f\frac{Y_c}{Z_c} \\ z = f \end{cases}$$

$$\Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

now $\begin{bmatrix} x \\ y \end{bmatrix}$'s origin point is image plane center.

we want to ① change the origin point to left-up center
translate (c_x, c_y)

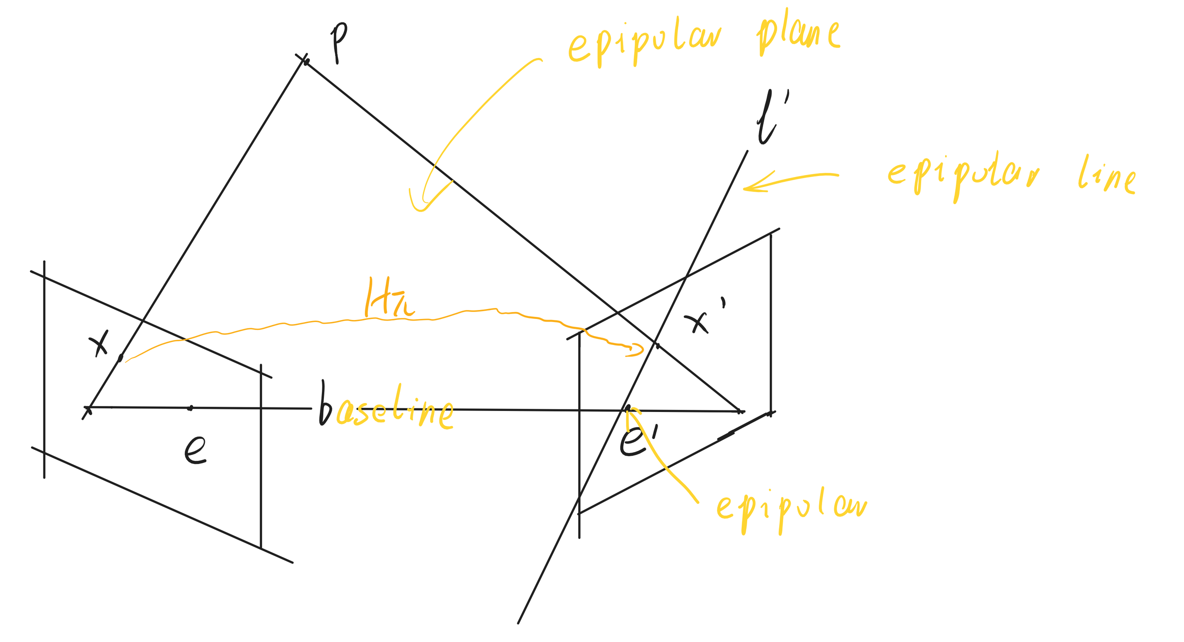
② use pixel size as the unit.
assume pixel size is (α, β) $(\frac{1}{d_x}, \frac{1}{d_y})$

then $u = \alpha \cdot x + c_x$
$v = \beta \cdot y + c_y$

$$\Rightarrow u = \alpha f \frac{X_c}{Z_c} + c_x$$
$$v = \beta \cdot f \frac{Y_c}{Z_c} + c_y$$

$$\Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

2. Fundamental matrix & Essential matrix.



△ $l' = e' \times x'$

$= [e']_\times x'$

$= \underbrace{[e']_\times H_\pi}_{F \text{ (fundamental matrix)}} x$

$= F x$

Epipolar constraint △ $(x')^T l' = (x')^T F x = 0.$ $\leftarrow$ at least 7 pairs to find F

use camera model describe epipolar constraint. △ For $P = [X, Y, Z]$

$p_1 = KP$, $p_2 = K(RP + t)$

$p_2 = K(R K^{-1} p_1 + t)$

$K^{-1} p_2 = R K^{-1} p_1 + t$

$t_\times K^{-1} p_2 = t_\times R K^{-1} p_1$

$p_2^T K^{-T} t_\times K^{-1} p_2 = p_2^T K^{-1} t_\times R K^{-1} p_1$

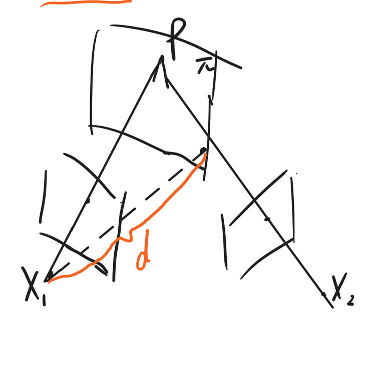
$0 = p_2^T \underbrace{K^{-1} \boxed{t_\times R} K^{-1}}_{F} p_1$

E : essential matrix

3. Homography.

△ Difference between homography and epipolar constraint.
homography: point-2-point ; points must be in same plane
epipolar : point-2-line ; points can be in different plane

△ use camera model to describe homography



P : in camera 1 coordinate X_1
in camera 2 coordinate X_2

$X_2 = RX_1 + T$ ①

Let plane π's normal vector be N, distance be d
Then $N^T X_1 = d$

$\frac{1}{d} N^T X_1 = 1$ $\forall X_1 \in \pi$ Note: have to be in same plane π

① can be writen as

$X_2 = RX_1 + T\frac{1}{d} N^T X_1$

$X_2 = \underbrace{(R + T\frac{1}{d} N^T)}_{H \text{ (homography)}} X_1$ $\leftarrow$ 3D

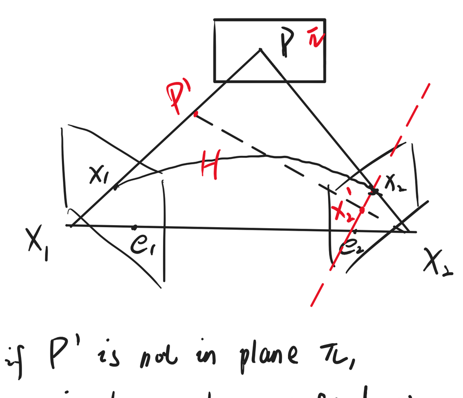
Assume same camera

$x_1 = KX_1$ $x_2 = KX_2$

$K^{-1} x_2 = (R + T\frac{1}{d} N^T) K^{-1} x_1$

$\boxed{x_2 = K(R + T\frac{1}{d} N^T) K^{-1} x_1}$ $\leftarrow$ 2D

△ what if two points not is same plane?



if P' is not in plane π,
using homography, we find wrong x_2 instead of correct x_2'

△ Camera only Rotation, no Translation ?
Then ① homography works even for points in different planes
② can not recover fundamental matrix F,
so that can not find 3D coordinates of points

① because no Translation.

$x_1 = KX_1$ $x_2 = KX_2 = KRX_1$

$\boxed{x_2 = \underbrace{KRK^{-1}}_{H'} x_1}$

compare $H = K(R + T\frac{1}{d} N^T) K^{-1}$

and $H' = KRK^{-1}$

no relation with plane normal vector N

② can not recover 3D coordinate

recall $F = K^{-T} t_\times R K^{-1}$

because $t = (0,0,0)^T$, $F = [0]$