$\boxed{\text{LEC 9}}$

1. Starting from Homogeneous Coordinate

$$\begin{cases} \begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow \begin{bmatrix} kx \\ ky \\ k \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} & \text{(point)} \\ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \longrightarrow \text{vector} & \end{cases}$$

Benefit : ① consider intersaction between $\begin{cases} Ax + By + Cw = 0 \\ Ax + By + Dw = 0 \end{cases}$ $C \neq D$

$$\Rightarrow \begin{cases} w = 0 \\ x = -B \\ y = A \end{cases} \Rightarrow \begin{bmatrix} -B \\ A \\ 0 \end{bmatrix} \rightsquigarrow \boxed{\text{infinite point}}$$

② A line: $Ax + By + C = 0$

$$\Leftrightarrow (A, B, C) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

2 points $\rightarrow$ 1 line

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = (x_1, y_1, 1) \times (x_2, y_2, 1)$$

2 lines $\longrightarrow$ 1 point

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (A_1, B_1, C_1) \times (A_2, B_2, C_2)$$

③ Rotation $\begin{pmatrix} x \\ y \end{pmatrix} = R \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}$

Translation $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} I & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}$

④ Linear Transform $A = R(\beta) \begin{pmatrix} w_1 & \\ & w_2 \end{pmatrix} R(\gamma)$

$\boxed{\text{SVD}}$

2. World Coord $\xrightarrow{\text{extrinsic}}$ Camera Coord $\xrightarrow{\text{intrinsic}}$ Image Coord

Camera Coord $\rightarrow \boxed{\text{2D phase}} \rightarrow$ Image Coord

Extrinsic :

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R_{\alpha\beta\gamma} & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Intrinsic :

①
$$\underbrace{\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \equiv \begin{pmatrix} f & & & \\ & f & & 0 \\ & & 1 & \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix}}_{\text{ideal pinhole camera}}$$

②
$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{dx} & s & u_0 \\ & \frac{1}{dy} & v_0 \\ & & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \equiv \begin{pmatrix} \alpha & s & u_0 & 0 \\ & \beta & v_0 & 0 \\ & & 1 & 0 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix}$$

- $(u_0, v_0) \rightarrow$ coordinate of "camera origin" in "image coordinate system"
- $\left\{ \begin{matrix} dx \\ dy \end{matrix} \right. \Rightarrow$ pixel length
- $s \Rightarrow$ skew effect

3. Learning Affine Transform (given point match)

Model: $\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

$\boxed{\begin{pmatrix} x_i \\ y_i \end{pmatrix} \longleftrightarrow \begin{pmatrix} x_i' \\ y_i' \end{pmatrix}}$ $i = 1, 2, \ldots, n$

$\Rightarrow \underbrace{A \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = b}_{\text{LSQ}}$ $\begin{cases} A \in \mathbb{R}^{2n \times 6} \\ b \in \mathbb{R}^{6} \end{cases}$

$n \geq 3$

4. Learning Homography

Model: $\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

$$\Rightarrow \begin{cases} x' = \dfrac{ax + by + c}{gx + hy + i} \\ y' = \dfrac{dx + ey + f}{gx + hy + i} \end{cases}$$

$$\Leftrightarrow \begin{cases} (gx + hy + i)\, x' = ax + by + c \\ (gx + hy + i)\, y' = dx + ey + f \end{cases}$$

$$\Rightarrow A \underbrace{\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{pmatrix}}_{:= t} = 0 \quad \begin{cases} A \in \mathbb{R}^{2n \times 9} \\ 0 \in \mathbb{R}^{2n} \end{cases}$$

s.t $\|t\|_2 = 1$

$n \geq 4$

$\begin{cases} \min \; t^T A^T A t \\ \text{s.t} \; \|t\| = 1 \end{cases}$

$\rightarrow$ smallest eigenvector of $A^T A$

5. Align 2 images

① SIFT

② Feature Match (Mis-match)

③ Estimate Transform Matrix

↓

Via RANdom SAmple Consensus (RANSAC)

↓

a) sample
b) measure the quality
c) re-sample
d) pick the best one

6. image warping
- forward
- backward

LEC 20 → Recap : in lec 9, from world coordinate system to image coordinate system,

we have

$$\begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha & s & u_0 & 0 \\ 0 & \beta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{\alpha\beta\gamma} & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} \alpha & s & u_0 \\ & \beta & v_0 \\ & & 1 \end{pmatrix} R_{\alpha\beta\gamma} \left(I ; -\tilde{t} \right) \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$:= M_{int} M_{ext} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$= \boxed{M_{int} \cdot R_{\alpha\beta\gamma}} \left(I ; \tilde{t} \right) \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

1. Learning $M_{int} \cdot M_{ext} := M$ [like Homography]

$$= M_{int} R_{\alpha\beta\gamma} (I; -\tilde{t})$$

Assume that we have the match between $\begin{cases} \text{World Coordinate System} \\ \text{Image Coordinate System} \end{cases}$

<u>Model</u>: $\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \equiv M \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}$ $\quad (a \equiv b \Rightarrow a \times b = 0)$

$$\Rightarrow \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \times \left[M \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \right] = \underline{0}$$

$$\Rightarrow \begin{bmatrix} 0 & -1 & y_i \\ 1 & 0 & -x_i \\ -y_i & x_i & 0 \end{bmatrix} M \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = \underline{0}$$

$$\Rightarrow \begin{pmatrix} 0 & -1 & y_i \\ 1 & 0 & -x_i \\ -y_i & x_i & 0 \end{pmatrix} \begin{pmatrix} m_{1\cdot} \, p_w \\ m_{2\cdot} \, p_w \\ m_{3\cdot} \, p_w \end{pmatrix} = \underline{0}$$

$$\Rightarrow \begin{pmatrix} 0 & -1 & y_i \\ 1 & 0 & -x_i \\ -y_i & x_i & 0 \end{pmatrix} \begin{pmatrix} p_w^T \cdot m_{1\cdot}^T \\ p_w^T \, m_{2\cdot}^T \\ p_w^T \, m_{3\cdot}^T \end{pmatrix} = \underline{0}$$

effective

$$\Rightarrow \underbrace{\begin{pmatrix} \boxed{\begin{matrix} 0 & -1 & y_i \\ 1 & 0 & -x_i \end{matrix}} \\ \begin{matrix} -y_i & x_i & 0 \end{matrix} \end{pmatrix}}_{\Downarrow \\ \mathbb{R}^{3\times3}} \underbrace{\begin{pmatrix} p_w^T & & \\ & p_w^T & \\ & & p_w^T \end{pmatrix}}_{\downarrow \\ \mathbb{R}^{3\times12}} \underbrace{\begin{pmatrix} m_{1\cdot}^T \\ m_{2\cdot}^T \\ m_{3\cdot}^T \end{pmatrix}}_{\Downarrow \\ \mathbb{R}^{12\times1}} = \underline{0}$$

$$\Rightarrow A \, m = 0 \qquad \underline{A \in \mathbb{R}^{\boxed{2}\times 12} \quad m \in \mathbb{R}^{12\times1}}$$

$\longrightarrow$ use LSQ to estimate $m \in \mathbb{R}^{12 \times 1}$

$$m \in \mathbb{R}^{12\times 1} \xrightarrow{\text{reshape}} \boxed{M \in \mathbb{R}^{3\times 4}}$$

Here, $M = M_{int} M_{ext}$

$$= M_{int} \left(R \,;\, -R\tilde{t} \right) = \left(M' \,;\, -M'\tilde{t} \right)$$

effect of M_{ext}:

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \left(\begin{array}{c|c} R_{\alpha\beta\gamma} & R_{\alpha\beta\gamma}(-\tilde{t}) \\ \hline 0 & 1 \end{array} \right) \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R_{\alpha\beta\gamma} \left[\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} - \tilde{t} \right]$$

$\Rightarrow$ $\tilde{t}$ represents the "origin" of "Camera Coordinate System" in "World Coordinate System"

Therefore, when we achieve the estimate $M \in \mathbb{R}^{3\times 4} = \left(M' \,;\, t \right)$

we can further estimate the "camera location" $\hat{t} = -(M')^{-1} t$

2. Furthermore, when we learn $M = M_{int} M_{ext}$

$$= \underline{M_{int} R_{\alpha\beta\gamma} [\, I_3 ; -\tilde{t} \,]}$$

$$M_{int} = \begin{pmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \boxed{\text{upper triangle}}$$

We can know MORE things.

① $\tilde{t}$

a) $M = M' [\, I ; -\tilde{t} \,] = (\, M' ; t \,)$

$$\Rightarrow \underline{\underline{\tilde{t} = -(M')^{-1} t}}$$

b) observation:

$R(p_w - \tilde{t}) = 0$ for $p_w = \tilde{t}$

$$\Leftrightarrow \underbrace{\begin{pmatrix} \boxed{R \quad -R\tilde{t}} \\ 0 \quad\quad 1 \end{pmatrix}_{4\times4} \begin{pmatrix} \tilde{t} \\ 1 \end{pmatrix}_{4\times1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}}$$

M_{ext}

$$\Leftrightarrow \underbrace{M_{ext}}_{3\times4} \underbrace{\begin{pmatrix} \tilde{t} \\ 1 \end{pmatrix}}_{4\times1} = \underbrace{0}_{3\times1}$$

$$\Rightarrow M \begin{pmatrix} \tilde{t} \\ 1 \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} \tilde{t} \\ 1 \end{pmatrix} \longleftrightarrow \text{smallest eigenvector w.r.t } M^T M$$

SVD of M:

$$M_{3\times4} = U_{3\times3} D \underline{\underline{V^T_{4\times4}}}$$

$$\Uparrow$$

$$\underline{MV = UD}$$

$$\Uparrow$$

$$\boxed{\begin{array}{l} M v_4 = 0 \\ M v_i = \sigma_i u_i \quad i=1,2,3 \end{array}}$$

$M(v_1, v_2, v_3, v_4) = (\sigma_1 u_1, \sigma_2 u_2, \sigma_3 u_3, 0)$

② $M = M_{int} M_{ext}$

$$= M_{int} R_{\alpha\beta\gamma} [\, I_3 ; -\tilde{t} \,]$$

$$= M' [\, I_3 ; -\tilde{t} \,] = [\, M' ; -M'\tilde{t} \,]$$

$\Rightarrow$ $\underbrace{M_{int}}_{\text{upper-triangle}}$, $\underbrace{R_{\alpha\beta\gamma}}_{\text{orthogonal}}$ can be achieved by RQ Decomposition of $\underline{M' = M[0:3, 0:3]} \in \mathbb{R}^{3\times 3}$

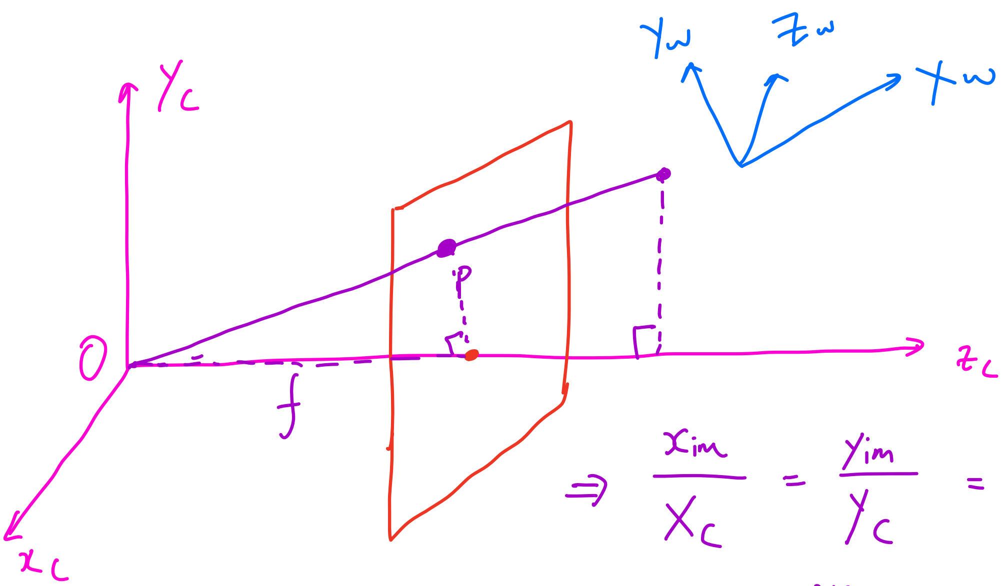
Up to now, we have figured out: [Camera System]

$$\begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} \equiv M \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \qquad \underline{M \in \mathbb{R}^{3\times 4}}$$

$$\boxed{\begin{aligned} M &= M_{int} M_{ext} \\ &= \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} R_{\alpha\beta\gamma} \left(I_3 ; -\tilde{t} \right) \end{aligned}}$$

$\rightarrow$ all can be estimated!

pinhole geometry



$$\Rightarrow \frac{x_{im}}{X_c} = \frac{y_{im}}{Y_c} = \frac{f}{Z_c}$$

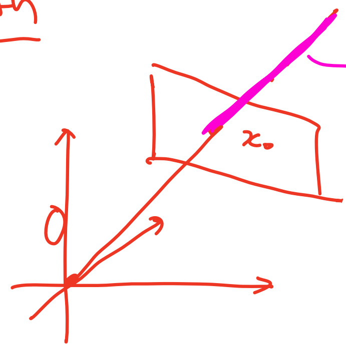
$$\Rightarrow \begin{cases} x_{im} = \dfrac{f X_c}{Z_c} \\ y_{im} = \dfrac{f Y_c}{Z_c} \end{cases} \Rightarrow \begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} \equiv \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix}$$

3D stereosis → Epipolar Geometry

↓

explore dense pixel correspondence

{ single camera → no deepth
{ 2 cameras → depth



all this line is possible for x_0

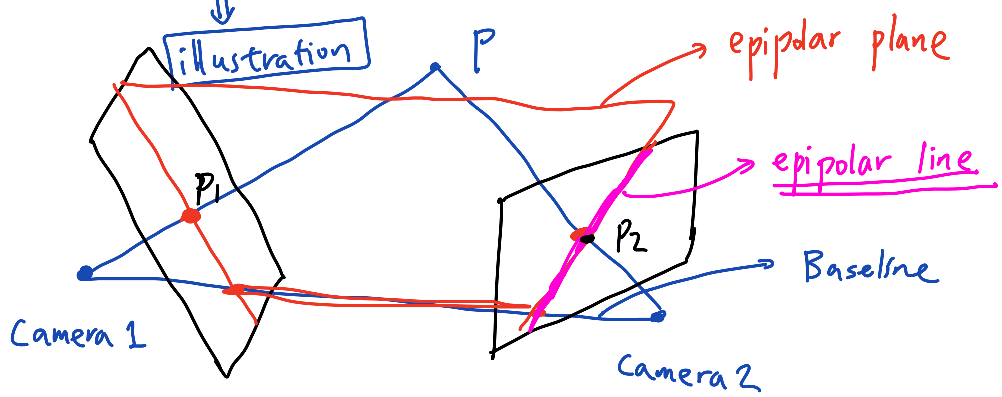
Terminology

{
① baseline : camera 1 ⟷ camera 2
② epipole : e_1, e_2 → intersection between { baseline, image plane }
③ epipolar plane : plane containing baseline
④ epipolar line : intersection between { epipolar plane, image plane }
}

Analysis : Interest

{ Camera 1 $\xrightarrow{R,\ t}$ Camera 2
{ shrink searching area for correspondence

⇓



suppose that, we have { P_1, camera parameters } → find corresponding P_2

Naive way : all pixel in image 2

Clever way : epipolar line

$\downarrow$

can be detected through $\begin{cases} P_1 \text{ coordinate} \\ \text{camera parameter} \end{cases}$

Calculation : focus on Camera Coordinate $\begin{cases} P_c \to (1) \\ P_c' \to (2) \end{cases}$

$$P_c = R P_c' + t$$

$$\Rightarrow P_c \perp (R P_c' \times t)$$

$$\Rightarrow P_c^T [t_x] \cdot R P_c' = 0$$

5 degree of freedom

$$\Rightarrow \boxed{P_c^T E P_c' = 0} \to \text{essential matrix}$$

Moreover : $\begin{cases} P_{im} = \begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} \equiv M_{int} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = M_{int} P_c \\ P_{im}' \equiv M_{int}' P_c' \end{cases}$

$$\Rightarrow P_{im}^T (M_{int})^{-T} E (M_{int}')^{-1} P_{im}' = 0$$

$$\Leftrightarrow \boxed{P_{im}^T F P_{im}' = 0} \to \text{fundamental matrix}$$

$$F = (M_{int})^{-T} [t_x] R (M_{int}')^{-1}$$

$\downarrow$

7 degree of freedom

Application :

① estimate for E via $\underline{P_c^T E P_c' = 0} \rightsquigarrow \boxed{Ae = 0} \to \boxed{\text{SVD of } A}$

② estimate for F via $\underline{P_{im}^T F P_{im}' = 0} \rightsquigarrow \boxed{Af = 0}$

③ $\underline{E = [t_\times] \cdot R} \longrightarrow$ decompose to achieve $\begin{cases} R \\ t \end{cases}$

$\Downarrow$

connection between 2 cameras

④ notice that $P_{im}^T F P_{im}' = 0$

$\Rightarrow (P_{im}^T F) \cdot P_{im}' = 0$

$(P_{im}^T F)$ → $\boxed{\text{normal vector}}$ for epipolar line in image 2

↓

$\boxed{\text{our interest}}$

Pf of ⑤:

e is epipole in image 1

$\Leftrightarrow$ $\forall$ epipolar line in image 1 (ℓ_1)

$\ell_1^T e = 0$

$\Leftrightarrow$ $\forall$ point in image 2 p_2

$F p_2 \longrightarrow \forall \ell_1$

$\Rightarrow p_2^T F e = 0 \quad \forall p_2$

$\Leftrightarrow F e = 0$

⑤ characterize epipole e from $x^T F x' = 0$

$\Rightarrow \underline{Fe = 0} \rightsquigarrow$ e is epipole (homogenous coordinate)

⑥ $\boxed{\text{Rectification}}$ → we want make all epipolar line parallel

observation: when 2 image plane is parallel.

then $R = I$, $T = \begin{pmatrix} d \\ 0 \\ 0 \end{pmatrix}$

Approach 2:

$e_1 \longrightarrow$ epipole of image 1

$\downarrow$

$$R_{rect}\, e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$R_{rect} = \begin{pmatrix} \vec{e}_1^{\,T} \\ \vec{e}_2^{\,T} \\ \vec{e}_3^{\,T} \end{pmatrix} \to [illegible]$$

$$F = (K')^{-T} [t]_\times R\, (K)^{-1}$$

$$= \begin{pmatrix} \frac{1}{f} & & \\ & \frac{1}{f} & \\ & & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -d \\ 0 & d & 0 \end{pmatrix} I \begin{pmatrix} \frac{1}{f} & & \\ & \frac{1}{f} & \\ & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -k \\ 0 & k & 0 \end{pmatrix} \Rightarrow \boxed{e = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} \to \boxed{\text{parallel epipolar lines}}$$

$$\Rightarrow x'^T F x = 0 \iff \boxed{y' = y}$$

$$\Rightarrow \text{we want } F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -k \\ 0 & k & 0 \end{pmatrix}$$

$$\iff \text{Find } H', H \text{ s.t. } H'^{-T} F H^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Moreover, consider $R = I$ $\quad T = \begin{pmatrix} 0 \\ 0 \\ t_z \end{pmatrix}$

$$F = K'^{-T} [t_\times] R K^{-1}$$

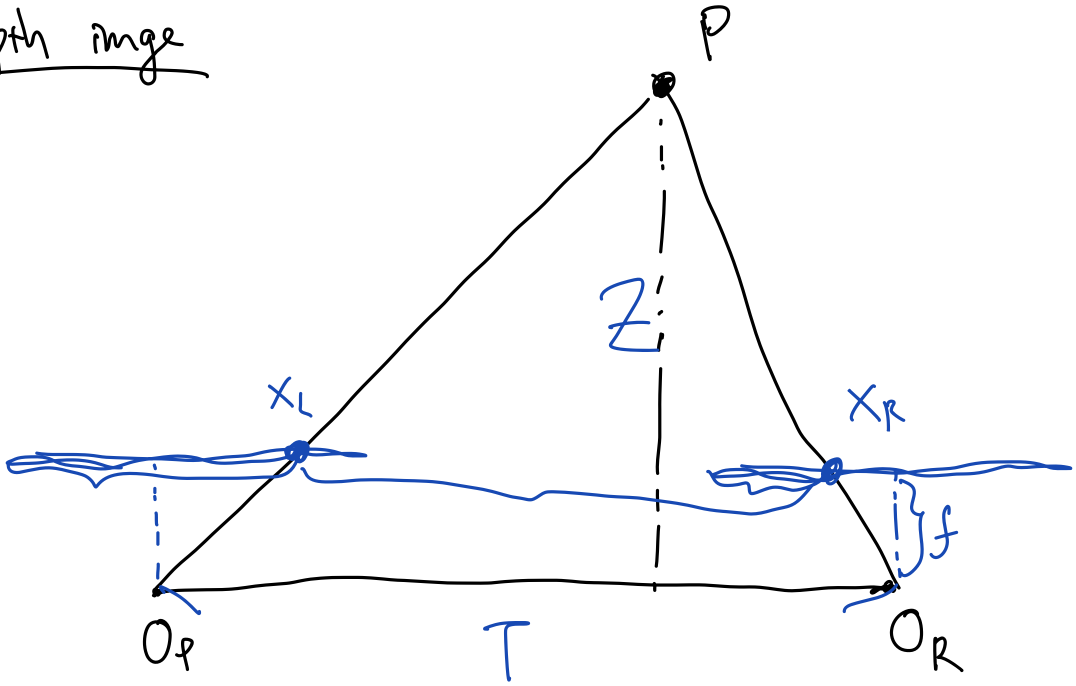
$$= \begin{pmatrix} \frac{1}{f} & & \\ & \frac{1}{f} & \\ & & 1 \end{pmatrix} \begin{pmatrix} 0 & -t_z & 0 \\ t_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{f} & & \\ & \frac{1}{f} & \\ & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -k & 0 \\ k & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \boxed{e = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}} \to \text{epipole}$$

in image 1, we have $(x, y, 1)$

$$x'^T F^T x = 0 \iff F^T x = \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix} \Longrightarrow \text{epipolar line in image 2}$$

Depth image



disparity $d = X_L - X_R$

distance between $X_L, X_R = T - (X_L - X_R)$

$$= T - d$$

$$\Rightarrow \frac{T}{T-d} = \frac{Z}{Z-f}$$

$$\Rightarrow \boxed{Z = f \cdot \frac{T}{d}} \Rightarrow \text{image depth estimation}$$