# Ex2

→ Epipolar geometry

a).

$$\tilde{E} = [t]_\times R$$

① $t_1 = (1, 0, 0)^T$:

$$\tilde{E}_1 = [t_1]_\times R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \cdot I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

For point $(a, b, 1)^T$

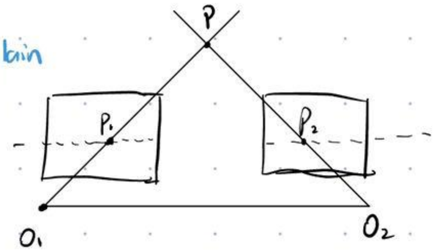
$$\tilde{l}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ b \end{pmatrix} \rightarrow y = b$$

$$\tilde{l}_1 = \tilde{E}_1^T \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -b \end{pmatrix} \rightarrow y = b$$

* epipole:
baseline × Image plain

∴ $\tilde{l}_1$, $\tilde{l}_2$ parallel to x-axis

epipole line:
epipole plain × image plain




∴ epipoles are located at infinity

② $t_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$$\tilde{E}_2 = [t_2]_\times R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

For a point $(a, b, 1)^T$

$\tilde{l}_2 = \tilde{E}_1 \cdot \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -a \end{pmatrix} \Rightarrow x = a$

$\tilde{l}_1 = \tilde{E}_2^T \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ a \end{pmatrix} \Rightarrow x = a$

$\therefore \tilde{l}_1$、$\tilde{l}_2$ parallel to y-axis

$\therefore$ $e_1$、$e_2$ are located at infinity

③ $t_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$\tilde{E}_3 = [t_3]_\times R = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

For a point $\begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$:

$\tilde{l}_2 = \tilde{E}_3 \cdot \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix} \Rightarrow y = \frac{b}{a} x$

$\tilde{l}_1 = \tilde{E}_3^T \cdot \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix} \Rightarrow y = \frac{b}{a} x$

$\because e_2^T \cdot \tilde{E}_3 = 0 = (x, y, 1) \cdot \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = (y, -x, 0) = 0$

$\therefore e_2 = (0, 0, 1)^T$

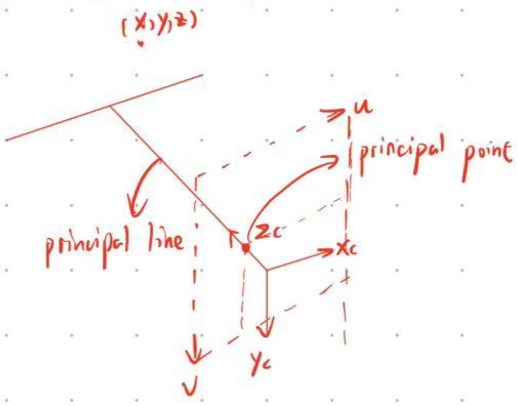
$\because \tilde{E}_3 \cdot e_1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} = 0$

$\therefore e_1 = (0, 0, 1)^T$

b). suppose $O_1 = (a, b, c)^T$

$O_2 = (a, b, c+1)$

so the baseline is the line which crosses $O_1$、$O_2$ and parallel to the ~~z-axis~~. principal axis, epipole : principal point




c). $K_1 = K_2 = I$

local ray directions $\tilde{x}_i = K_i^{-1} \bar{x}_i$

$\therefore \quad \tilde{x}_2^T \cdot \tilde{E} \cdot \tilde{x}_1 = \bar{x}_2^T \cdot K_2^{-T} \cdot \tilde{E} \cdot K_1^{-1} \bar{x}_1 = \bar{x}_2^T \tilde{F} \cdot \bar{x}_1 = 0$

$\therefore$ when $K_1 = K_2 = I$, $\tilde{F} = \tilde{E}$

it means $f_x = f_y = 1$ and principal point is not moved

## → Triangulation

a). suppose the original point $x_0$ is $(a, b, c)^T$

$\tilde{P}_1 = K_1 \cdot [R_1 | t_1] = I \cdot [I \ 0] = [I \ 0]$

$$\tilde{P}_2 = K_2 \cdot [R_2 | t_2] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 0 & -2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 1 & -3 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1^s \cdot \tilde{p}_{13}^T - \tilde{p}_{11}^T \\ y_1^s \tilde{p}_{13}^T - \tilde{p}_{12}^T \\ x_2^s \cdot \tilde{p}_{23}^T - \tilde{p}_{21}^T \\ x_2^s \cdot \tilde{p}_{23}^T - \tilde{p}_{22}^T \end{bmatrix} \cdot \tilde{X}_0 = 0$$

$$\Rightarrow \begin{bmatrix} -1 & 0 & \frac{1}{4} & 0 \\ 0 & -1 & \frac{1}{2} & 0 \\ 2 & 0 & -\frac{6}{5} & \frac{14}{5} \\ 0 & 2 & -\frac{4}{5} & -\frac{4}{5} \end{bmatrix} \cdot \tilde{X}_0 = 0$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \tilde{X}_0 = 0$$

$$\therefore \tilde{X}_0 = \begin{pmatrix} w \\ 2w \\ 4w \\ w \end{pmatrix}$$

$$\therefore \text{when } w = 1, \quad \tilde{x}_0 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 1 \end{pmatrix}$$

$$\therefore X_0 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

## → Stereo Vision

a). suppose the measured disparity is a

$\because Z(d) = \frac{fb}{d}$

$\therefore Z(a) = \frac{fb}{a}$

$Z'(a) = \frac{-fb}{a^2}$

$\therefore$ with Taylor expension,

$Z(a+\Delta d) = Z(a) + Z'(a) \cdot (a+\Delta d - a) + o(\Delta d)$

$= Z(a) - fb \cdot \frac{1}{a^2} \cdot \Delta d + o(\Delta d)$

$\therefore \Delta Z = Z(a+\Delta d) - Z(a) = -fb \frac{\Delta d}{a^2} + o(\Delta d)$

$= -\frac{Z(a)^2 \cdot \Delta d}{fb} + o(\Delta d)$

$\therefore$ the depth measurement error grows quadratically with depth

not the system setup

b). ① make the focal length of the cameras and the length of base line (in other words, the distance of camera) larger, $\Delta Z \downarrow$.

② problems: larger f results in the decrease of the information contained in a picture, while larger b results in the increase of difficulty to match the block

## → Block Matching

a). $W_i' = \alpha_i W_i + 1\beta_i$

$\therefore \bar{W}_i' = \alpha_i \cdot \bar{W}_i' + 1 \cdot \beta_i$ , $\alpha_i \in R$ , $\beta_i \in \mathbb{R}^{K^2}$

$\therefore W_i' - \bar{W}_i' = \alpha_i ( W_i - \bar{W}_i' )$

$$\| W_i' - \bar{W}_i' \|_2 = \| \alpha_i ( W_i - \bar{w}_i' ) \|_2 = \sqrt{\alpha_i^2} \cdot \| W_i - \bar{w}_i' \|_2$$
$$= \alpha_i \cdot \| W_i - \bar{W}_i' \|_2$$

$$\therefore ZNCC'(x, y, d) = \frac{(W_L'(x,y) - \bar{W}_L'(x,y))^T \cdot (W_R'(x-d,y) - \bar{W}_R'(x-d,y))}{\| W_L'(x,y) - \bar{W}_L'(x,y) \|_2 \cdot \| W_R'(x-d,y) - \bar{W}_R'(x-d,y) \|_2}$$

$$= \frac{\alpha_L ( W_L(x,y) - \bar{W}_L(x,y))^T \cdot (W_R(x-d,y) - \bar{W}_R(x-d,y)) \cdot \alpha_R}{\alpha_L \cdot \| W_L(x,y) - \bar{W}_L(x,y) \|_2 \cdot \alpha_R \| W_R(x-d,y+d) - W_R(x-d,y) \|_2}$$

$$= \frac{( W_L(x,y) - \bar{W}_L(x,y))^T \cdot (W_R(x-d,y) - \bar{W}_R(x-d,y))}{\| W_L(x,y) - \bar{W}_L(x,y) \|_2 \cdot \| W_R(x-d,y+d) - W_R(x-d,y) \|_2}$$

$$= ZNCC(x, y, d)$$

$\therefore$ invarient to changes in brightness of windows b).

$$SSD(3,5,0) = \| W_L(3,5) - W_R(3,5) \|_2^2$$

$$= \left[ (10-5)^2 + (5-6)^2 + (6-7)^2 \right] \times 3$$

$$= 81$$

$$SSD(3,5,1) = \| W_L(3,5) - W_R(3,4) \|_2^2$$

$$= 3 \times \left[ (10-4)^2 + (5-5)^2 + (6-6)^2 \right]$$

$$= 108$$

$$SSD(3,5,2) = \| W_L(3,5) - W_R(3,3) \|_2^2$$

$$= 3 \times \left[ (10-10)^2 + (5-4)^2 + (6-5)^2 \right]$$

$$= 6$$

$\therefore SSD(3,5,2) < SSD(3,5,0) < SSD(3,5,1)$

$\therefore d=2$ is the WTA disparity

c). ① all of the three points pass the test.

② the red block passes the test, but in the left image, the brightness is 3, after moving to the right, the brighness comes to 2, which is not consistent. So it may not be successful.

## → Learned Stereo and End-to-End Models

a).

| Layer | Input Shape | Output Shape | Trainable Parameters | Memory |
|---|---|---|---|---|
| Conv2d(32,64,3) | (32, 128, 128) | (64, 128, 128) | 18496 | 4.07 mb |
| Conv2d(64,128,3) | (64, 128, 128) | (128, 128, 128) | 73856 | 8.28 mb |
| Conv3d(1,64,3) | (1, 32, 128, 128) | (64, 32, 128, 128) | 640 | 128 mb |
| Conv3d(64,128,3) | (64, 32, 128, 128) | (128, 32, 128, 128) | 73856 | 256.28 mb |

b).

(1) ① For $p_1$

$$d^* = E[d] = \sum_{d=0}^{4} \text{softmax}(-C_\theta(d)) \cdot d$$

$$= \frac{e^{-1} \cdot 0 + e^{-3} \cdot 1 + e^{-10} \cdot 2 + e^{-3} \cdot 3 + e^{-1} \cdot 4}{e^{-1} + e^{-3} + e^{-10} + e^{-3} + e^{-1}}$$

$$= \frac{e^{-3} + 2 \cdot e^{-10} + 3 \cdot e^{-3} + 4e^{-1}}{e^{-1} + e^{-3} + e^{-10} + e^{-3} + e^{-1}}$$

$$= 2$$

② For $p_2$

$$d^* = E[d] = \sum_{d=0}^{4} \text{softmax}(-C_\theta(d)) \cdot d$$

$\approx 2$

(2) encourage the disparity to take an appropriate value between the several disparity (actually by the expectation)
which has similarly small cost (especially when there're some disparity shares the same min cost)

solution: encourage the predictions to be unimodally distributed.