ASSIGNMENT - 4

AKSHAY R
A20442409

Review Questions:

1. Camera Callibration:

a. Given the equation $p = MP$

Forward Projection: Given a world point P project it onto an image using a projection matrix M.

Callibration: Given image points $\{p_i\}_{i=1}^{m}$ and world points $\{P_i\}_{i=1}^{m}$ find the camera intrensic and extrensic parameters $(\alpha_u, \alpha_v, s, u_0, v_0, R^*, T^*, K)$.

Reconstruction: Given several images with $N \times \{p_i\}_{i=0}^{m}$ and the projection matrix M reconstruct the 3D object co ordinates.

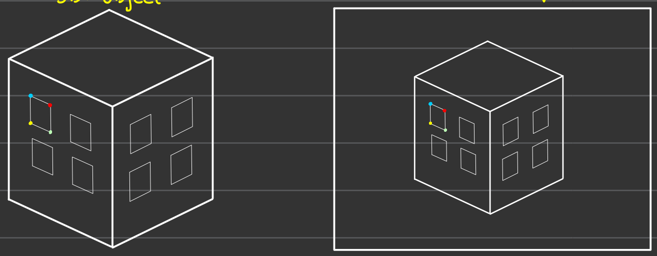
Reconstruction is the most difficult out of the three discussed above. Forward projection is the simplest.

b. The necessary inputs for camera callibration are, the points $\{p_i\}_{i=1}^{m}$ which are the image coordinates and points $\{P_i\}_{i=1}^{m}$ which are the corresponding world co-ordinates.

We need atleast 6 point pairs in order to do a 3D $\{P_i\}_{i=1}^{m}$ camera calibration or at least 4 points to perform planar camera callibration.

c. Given $\{p_i\}_{i=1}^{m} \longleftrightarrow \{P_i\}_{i=1}^{m}$ find camera parameters:

3D- object 2D- Image.



$\{P_i\}$ 3DH $\{p_i\}$ 2DH

Steps:

- Find the projection matrix M. (estimate).
- Find parameters from M.

To estimate the projection matrix:

- Find the point pair matrix in the form of

$$\underset{2m \times 12}{\overset{A}{\begin{bmatrix} P_1^T & 0 & -x_1 P_1^T \\ 0 & P_1^T & -y_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_m^T & 0 & -x_m P_m^T \\ 0 & P_m^T & -y_m P_m^T \end{bmatrix}}} \underset{12\times 1}{\overset{x}{\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}}} \overset{=}{=} \underset{2m\times 1}{\overset{0}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}}$$

Number of corresponding points.

- Then solve $Ax = 0$ using SVD to obtain $\hat{m}$

- Find ρ so that

$$M = \rho \hat{M} = K^* [\, R^* \mid T^* \,]$$

- Find the solution for the below equation

$$\underbrace{\overbrace{\begin{bmatrix} \alpha_u r_1^T + s r_2^T + u_0 r_3^T & \\ \alpha_v r_2^T + v_0 r_3^T & K^* T^* \\ r_3^T & \end{bmatrix}}^{K^* R^*}}_{\text{UNKNOWNS}} = \rho \underbrace{\overbrace{\begin{bmatrix} \overbrace{\begin{matrix} — a_1^T — \\ — a_2^T — \\ — a_3^T — \end{matrix}}^{3\times 3} & \overbrace{b}^{3\times 1} \end{bmatrix}}^{3\times 4}}_{\text{KNOWN}} = \rho \hat{M}$$

d. Given

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad P_i : \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

$$p_i = M P_i$$

$$p_i : \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+4+9+4 \\ 1+0+9+4 \\ 1+2+3+1 \end{bmatrix} = \begin{bmatrix} 18 \\ 14 \\ 7 \end{bmatrix}$$

The 2D point is $\left(\frac{18}{7}, 2 \right)$

e. given $\{p_i\} \longleftrightarrow \{P_i\}$

$(100, 200) \longleftrightarrow (1, 2, 3)$

The first two lines must be of the form:

$$= \underbrace{\begin{bmatrix} P_i^T & 0 & -x_i P_i^T \\ 0 & P_i^T & -y_i P_i^T \end{bmatrix}}_{2 \times 12} \qquad P_i^T = [1, 2, 3, 1]$$

$$= \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 & 0 & 0 & -100 & -200 & -300 & -100 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 & 1 & -200 & -400 & -600 & -200 \end{bmatrix}$$

f. The minimum number of points needed to be able to find a unique solution for M is 6 because we have 11 unknowns. Hence we need 12 equations $\Rightarrow$ 6 points.

Once we have the necessary points form a matrix in the form of

$$\underbrace{\begin{bmatrix} P_1^T & 0 & -x_1 P_1^T \\ 0 & P_1^T & -y_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_m^T & 0 & -x_m P_m^T \\ 0 & P_m^T & -y_m P_m^T \end{bmatrix}}_{\substack{A \\ 2m \times 12}} \underbrace{\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}}_{\substack{x \\ 12 \times 1}} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}_{\substack{0 \\ 2m \times 1}}$$

Number of corresponding points.

And solve for $Ax = 0$ using SVD.

g. Once we have an estimate of the projection matrix M, we use the orthogonality of r_1, r_2, r_3 and the orthogonality of a_1, a_2 and a_3 in the below equations

$$\text{①} \quad \alpha_u r_1^T + s r_2^T + u_0 r_3^T = \rho a_1^T$$
$$\text{②} \quad \alpha_v r_2^T + v_0 r_3^T = \rho a_2^T$$
$$\text{③} \quad r_3^T = \rho a_3^T$$
$$\text{④} \quad K^* T^* = \rho b$$

Some of the properties of orthogonality is stated below.

$r_1 \cdot r_2 = 0 \qquad r_2 \cdot r_3 = 0 \qquad r_1 \cdot r_3 = 0 \qquad r_2 \cdot r_2 = 1 \qquad r_1 \cdot r_1 = 1$

$r_1 \times r_2 = r_3 \qquad r_2 \times r_3 = r_1 \qquad r_3 \times r_1 = r_2 \qquad r_3 \cdot r_3 = 1$

h.

Given $\{p_i\}_{i=1}^{m} \longleftrightarrow \{P_i\}$ and estimated $M = \begin{bmatrix} - m_1^T - \\ - m_2^T - \\ - m_3^T - \end{bmatrix}$ assess the quality of fit using :

$$E = \frac{1}{m} \sum_i \left(\left\| x_i - \frac{m_1^T P_i}{m_3^T P_i} \right\|^2 + \left\| y_i - \frac{m_2^T P_i}{m_3^T P_i} \right\|^2 \right)$$

The above equation provides distance between known and predicted positions

Eg : $(7.3, 12.1, 15.3) \longleftrightarrow (5, 3)$ KNOWN $\rightarrow L_2 \rightarrow$ ERROR

$\longrightarrow (6\ \ 2)$ PREDICTED USING M

i. The below points show the principles and the approach for planar camera callibration.

① Estimate 2D homography (projective map) between calibration plane and image (for several images). → obtain $\hat{H}_1, \hat{H}_2, \hat{H}_3$

② Estimate intrensic parameters. → using $\hat{H}_1, \hat{H}_2, \hat{H}_3$

③ Compute extrensic parameters for view of interest. → (R, T)

For planar camera cellibration we project using a 2D projective map as shown below.

$$\underset{2DH}{p_i'} = \underbrace{K^* [\overset{3\times 3}{R^*} \mid \overset{3\times 1}{T^*}]}_{3\times 4} \underset{3DH}{P_i}$$

$$= K^* \begin{bmatrix} | & | & | & \\ r_1 & r_2 & r_3 & T^* \\ | & | & | & \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{bmatrix}$$

NON CO-PLANAR MODEL

$$\underset{2DH}{P_i'} = \overbrace{\underbrace{K^* [r_1 \quad r_2 \quad T^*]}_{\text{2D projective map}}}^{3\times 3} \underset{2DH}{\begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}}$$

PLANAR MODEL

j. We make a basic assumption the there are no coordinates present in the z plane ⇒ we assume that $z = 0$ We project the world planar point $\overset{2DH}{(X_i , Y_i , 1)}$ to image coordinates $\overset{2DH}{(x_i , y_i , 1)}$

with homography (2D projective map) $= K^* [r_1 \; r_2 \; T^*]$
instead of the projection matrix $M = K^* [r_1 \; r_2 \; r_3 \; T^*]$

2. Camera Callibration 2.

a. given $\{p_i\} \longleftrightarrow \{P_i\}$

$(1, 2) \longleftrightarrow (3, 4, 5)$

The first two lines must be of the form:

$$= \underbrace{\begin{bmatrix} P_i^T & 0 & -x_i P_i^T \\ 0 & P_i^T & -y_i P_i^T \end{bmatrix}}_{2 \times 12} \qquad P_i^T = [3 \;\; 4 \;\; 5 \;\; 1]$$

$$= \begin{bmatrix} 3 & 4 & 5 & 1 & 0 & 0 & 0 & 0 & -3 & -4 & -5 & -1 \\ 0 & 0 & 0 & 0 & 3 & 4 & 5 & 1 & -6 & -8 & -10 & -2 \end{bmatrix}$$

b.

$$\underbrace{\left[\begin{array}{c|c} \overbrace{\begin{matrix} \alpha_u r_1^T + s r_2^T + u_0 r_3^T \\ \alpha_v r_2^T + v_0 r_3^T \\ r_3^T \end{matrix}}^{K^* R^*} & K^* T^* \end{array}\right]}_{\text{UNKNOWNS}} \equiv \rho \underbrace{\overbrace{\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{array}\right]}^{3\times 4 \;(3\times 3 \,|\, 3\times 1)}}_{\text{KNOWN}} = \rho \hat{M}$$

UNKNOWNS

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} = \left[\begin{array}{c|c} — a_1^T — & \\ — a_2^T — & b \\ — a_3^T — & \end{array}\right]$$

$$|\rho| = \frac{1}{|a_3|}$$

$$u_0 = |\rho|^2 a_1 . a_3$$

$$v_0 = |\rho|^2 a_2 . a_3$$

$$|\rho|^2 = \frac{1}{3^2+4^2+5^2} = 0.02$$

$$u_0 = 0.02 \left(3+8+15 \right) = 0.02 \times 26 = 0.52$$

$$v_0 = 0.02 \left(6+12+20 \right) = 0.02 \times 38 = 0.76$$

c. $$E = \frac{1}{m} \sum_i \left(\left\| x_i - \frac{m_1^T P_i}{m_3^T P_i} \right\|^2 + \left\| y_i - \frac{m_2^T P_i}{m_3^T P_i} \right\|^2 \right)$$

$m \leftarrow$ Number of corresponding points.

here $m = 1$ $\quad x_i = 1$ $\quad y_i = 2$

$$m_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad m_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \quad m_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix} \quad P_i = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 1 \end{bmatrix}$$

$3+8+15+4$ $\quad 6+12+20+5$ $\quad 9+16+25+6$

$$E = \frac{1}{1} \left(\left(1 - \frac{30}{56}\right)^2 + \left(2 - \frac{43}{56}\right)^2 \right)$$

$$= 0.2155 + 1.5181 = 1.7336.$$

Validation:

$$p_i = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 + 8 + 15 + 4 \\ 6 + 12 + 20 + 5 \\ 9 + 16 + 25 + 6 \end{bmatrix} = \begin{bmatrix} 30 \\ 43 \\ 56 \end{bmatrix}$$

actual point = $(0.53, 0.76)$

Yes the points are off.

d. Given $R^* = I + Q$

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Q = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R^* = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad T^* = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

$$R = (R^*)^T$$

$$T = -(R^*)^T T^*$$

$$R = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad T = -\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -6 \\ -2 \\ -3 \\ -1 \end{bmatrix}$$

3DM → 31

e. A

$$\begin{bmatrix} p_1^{*T} & 0 & -x_1 p_1^{*T} \\ 0 & p_1^{*T} & -y_1 p_1^{*T} \\ \vdots & \vdots & \vdots \\ p_m^{*T} & 0 & -x_m p_m^{*T} \\ 0 & p_m^{*T} & -y_m p_m^{*T} \end{bmatrix}$$

$2m \times 9$

$p^{*T} = [3 \ 4 \ 1]$

$x_1 = 1 \quad y_1 = 2.$

$$\begin{bmatrix} 3 & 4 & 1 & 0 & 0 & 0 & -3 & -4 & -1 \\ 0 & 0 & 0 & 3 & 4 & 1 & -6 & -8 & -2 \\ & & & & \vdots & & & & \end{bmatrix}$$

The above matrix shows the first two rows.

3. Multiple view geomeatry :

a. The major difference between sparse and dense is that sparse is feature based and dense is correletion and SSD based.

The advantages of sparse sterio metching is that it can handle large disparities.
The advantages of dense sterio netching is that it produces more points. Disadvantage is that occlusion or ambiguous points can arise where it becomes difficult to correlete.

b.

$$\begin{bmatrix} 10 & 11 & 10 \\ 13 & 11 & 10 \\ \cdots & \cdots & \cdots \end{bmatrix} \quad \begin{bmatrix} 11 & 13 & 10 \\ 10 & 13 & 11 \\ \cdots & \cdots & \cdots \end{bmatrix} \qquad \begin{bmatrix} 10 & 11 & 10 \\ 13 & 11 & 10 \\ \cdots & \cdots & \cdots \end{bmatrix} \quad \begin{bmatrix} 100 & 50 & 10 \\ 70 & 100 & 10 \\ \cdots & \cdots & \cdots \end{bmatrix}$$

lower similarity with higher correlation.
only because of higher value #s.

Normalize window values (Subtract mean and divide by standard deviation).

Zero mean normalized cross-correlation (ZNCC):

$$\psi(w_1, w_2) = \sum_i \frac{(w_1(x_i, y_i) - \mu_{w_1})(w_2(x_i, y_i) - \mu_{w_2})}{\sigma_{w_1} \sigma_{w_2}}$$

Zero mean normalized SSD (ZNSSD)

$$\psi(w_1, w_2) = \sum_i \left(\frac{w_1(x_i, y_i) - \mu_{w_1}}{\sigma_{w_1}} - \frac{w_2(x_i, y_i) - \mu_{w_2}}{\sigma_{w_2}} \right)^2$$

If we use the entire image as search space we may frequently observe:

- Object points not visible in both views (occlusion)
- Ambiguous points (multiple coordinates)
- Uniform regions (inside points cannot be distinguished).

The search space can be reduced to a line by using axis aligned stereo where we will only need to search along the epipolar lines.

c. Given corresponding points

A (100 , 200) B (103 , 200) $Z = ?$

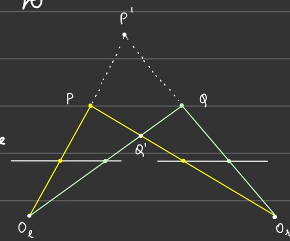
given $f = 10$ $T = 100$

we have: $Z = f\dfrac{T}{d}$ $d = -x_l + x_r = -100 + 103 = 3$

$$Z = 10 \times \frac{100}{3} = \boxed{333.34}$$

d. Consider point P and Q projected on to two images as shown in the figure. O_l and O_r are two points produced by tracing the image points to the world point. (origins)

While reconstruction if if O_l and O_r are projected incorrectly P' and Q' will be the reconstructed points which are severely incorrect in correspondance to the original points.



e. callibration: $\left\{\begin{matrix} R_l, & T_l \\ R_r, & T_r \end{matrix}\right\}$ Rotation / Translation of Left / Right camera with respect to world

$$M_{Right \leftarrow Left} = M_{Right \leftarrow world} \, M_{world \leftarrow Left}$$

$$= (R_r^T T_r^{-1})(R_l^T T_l^{-1})^{-1} = TR$$

$$M_{Right \leftarrow Left} = R_r^T T_r^{-1} T_l R_l = TR$$

4. Multiple view geomeatry 2 :

a. Given = [f11: 10 mm23] [T = 20 mm . parity = 30 mm
2 = 11 + 34 + 23 = 68]

we have $Z = \left[{}^{1}f \right] \dfrac{T}{d} = 10 \times \dfrac{20}{30} = 6.67$ mm.

b. $A \times B = \underset{\text{matrix}}{[A]_\times} \underset{\text{vector}}{B} \equiv \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} B$

Rank 2 skew symetric matrix

given $a_x = 1$ $a_y = 2$ $a_z = 3$

$\Rightarrow$ vector $A_x \equiv \begin{bmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix}$

c. Given $f = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ $p_r^T = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix}$ $p_e^T = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

$$p_r^T F p_e = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2+6+3 & 4+9+4 & 6+12+5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 11 & 17 & 23 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 11 + 34 + 23 = 68$$

Theoretically $p_r^T F p_l = 0$

d. Given $p_1 = (\overset{x_1}{1}, \overset{y_1}{2})$ $\quad p_1' = (\overset{x_1'}{2}, \overset{y_1'}{3})$

$$\underset{\substack{\text{KNOWN} \\ m \times 9}}{\begin{bmatrix} x_1 x_1' & x_1 y_1' & x_1 & y_1 x_1' & y_1 y_1' & y_1 & x_1' & y_1' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m x_m' & x_m y_m' & x_m & y_m x_m' & y_m y_m' & y_m & x_m' & y_m' & 1 \end{bmatrix}} \underset{\substack{\uparrow \\ \text{UNKNOWNS} \\ 9 \times 1}}{\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ \vdots \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix}} = \underset{m \times 1}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

Required row

$$\underset{\substack{\text{KNOWN} \\ m \times 9}}{\begin{bmatrix} \boxed{2 \quad 3 \quad 1 \quad 4 \quad 6 \quad 2 \quad 2 \quad 3 \quad 1} \\ \vdots \\ x_m x_m' \quad x_m y_m' \quad x_m \quad y_m x_m' \quad y_m y_m' \quad y_m \quad x_m' \quad y_m' \quad 1 \end{bmatrix}} \underset{\substack{\uparrow \\ \text{UNKNOWNS} \\ 9 \times 1}}{\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ \vdots \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix}} = \underset{m \times 1}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

END