1a) Sparse stereo matching and dense stereo matching.

In both the cases we try to find corresponding points, but the difference is that,

in sparse ~~we~~ we compute corresponding <u>points</u> of one image and another.

in dense we compute corresponding <u>pixel</u> of one image and another

- When there are large disparity in the images we use sparse, where images are far apart, and we use dense when there is low disparity or images are close.

Matching cannot be done in
→ uniform regions
→ not visible in one image but in another
→ ambiguity

Scanned by CamScanner

1 b) Normalised cross - relation (NCC)

we will take window w_1 & w_2 → for both images & multiply the corresponding elements, but in normalised version we subtract each by mean & divide by standard deviation.

$$(ZNCC)\ \varphi(w_1, w_2) = \sum_i \left(\frac{w_1(x_i, y_i) - \mu_1}{\sigma_1} \right) \cdot w_2\left(\frac{(x_i, y_i) - \mu_2}{\sigma_2} \right).$$

higher values → similar window.

Squared Distance (SSD)

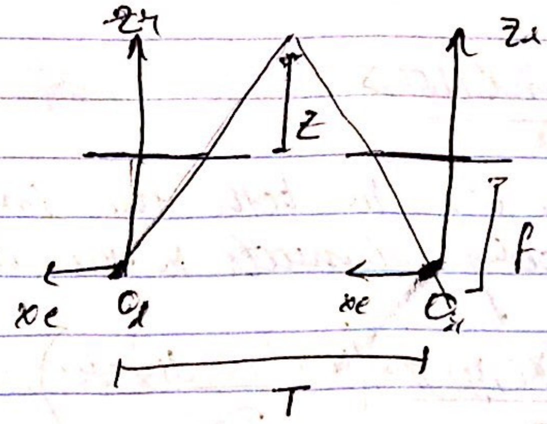
Here we take difference of product of corresponding elements, but for normalised version we again subtract by mean & divide by std. deviation.

$$(ZSSD)\ \varphi(w_1, w_2) = \sum_i \left[w_1\left(\frac{(x_i, y_i) - \mu_1}{\sigma_1} \right) - w_2\left(\frac{(x_i, y_i) - \mu_2}{\sigma_2} \right) \right]^2$$

The risk is that there will be a lot of options corresponding to the point in a image. so chances of error making increases.

Here Search space is reduced by only considering epipolar line instead to whole image.

c)



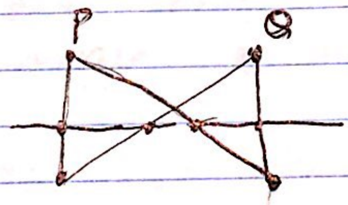
Given:–
$X_R = (100, 200)$
$X_l = (103, 200)$
$f = 10$
$T = 100$

$$\text{depth}, Z = f\frac{T}{d}$$

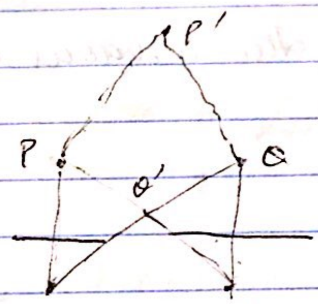
$d = X_r - X_l$
$\Rightarrow d = 103 - 100$
$= 3$

$$\text{depth}, Z = \cancel{\frac{10 \times 13}{100}} \quad 10 \times \frac{100}{3}$$
$$= \cancel{0.13} \quad = 33.3\overline{3}$$

d) <u>Ambiguity Problem</u>



perfect



reconstruction
wrong
matching

we see that if we reconstruct with wrong points, it will make mistake.
thus a small change/error in correspondance can make big error in reconstruction.

e) R_l, T_l
R_r, T_r

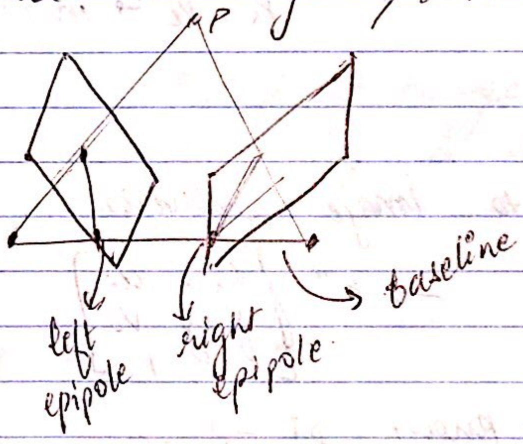
$$M_{left \leftarrow right} = R_l^T T(-T_l)\ T(T_r) R_r$$

$$= \begin{bmatrix} R_l^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -T_l \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & T_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_r & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} R_l^T R_r & R_l^T (T_r - T_l) \\ 0 & 1 \end{bmatrix}$$

$$\boxed{R = R_l^T R_r} \ \& \ \boxed{T = R_l^T (T_r - T_l)}$$

2 a) → Epipoles are the points intersecting by the baseline connecting two ~~two~~ image planes with the two image planes.



→ when a world point is joined with the optical centres, it creates a plane, when this plane intersects the images, it creates a lines on both images ; which is called epipolar lines on corresponding left & right image.

2 b) Essential Matrix → It captures essential relationship between two images.

$$\boxed{E = R^T [T]_x}$$

$\boxed{P_r^T E P_l = 0}$ → epipolar constraint equation where P_r, P_l → in camera coordinates.

2c) While moving to image coordinates

$\bar{P}_l = K_l^* P_l$ → $K^* = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$

$\bar{P}_r = K_r^* P_r$

as we know, $P_l^T E P_l = 0$

$$P_r^T K_r^{*-T} E K_l^{*-1} P_l = 0$$

so $$\boxed{F = K_r^{*T} E K_l^{*-1}}$$

epipolar constraint equation is

$$\boxed{P_r^T F P_l = 0}$$

d) The rank of Essential matrix is 2 because $E = R^T [T_x]$ → where T is a rank 2 matrix and has only 2 rows independent of each other.

The rank of Fundamental matrix is 2 - because $F = K_r^{*T} E K_l^{*-1}$ → where E is a rank 2. as described above, thus making F a rank 2 matrix.

e) point = P_l
Fundamental matrix = F

The epipolar line in the right image is $\boxed{F P_l}$ where the corresponding P_r will be found

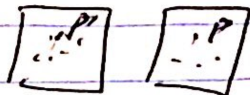
f) point = P_r
Fundamental matrix = F.

$$P_r^T F P_l = 0$$
Taking transpose.
$$P_l^T F^T P_r = 0$$

left ~~right~~ epipolar line → $\boxed{l = F^T P_r}$ → so: $\boxed{P_l^T l = 0}$
↳ point P_l → corresponding to P_r in right image

2g) Weak Calibration $\longrightarrow$

Here we find 'F' directly from 8-point correspondance and use 8-point algorithm.



Given $\{P_i\}_{i=1}^{n} \rightarrow \{P'_i\}_{i=1}^{n}$ $\quad n \geq 8$

To find : F,

$$= \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

$\rightarrow$ there are 8 unknowns. as one we get from

$\underline{P_r^T\, F\, P_l = 0}$.

$P_1 = (x_1, y_1, 1)$ $\quad P'_1 = (x'_1, y'_1, 1)$

$\downarrow$ right $\qquad \searrow$ left

From epipolar constraint we get

$$\Rightarrow (x, y, 1) \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{23} & f_{33} \end{bmatrix} \begin{bmatrix} x'_1 \\ y'_1 \\ 1 \end{bmatrix} = 0$$

$$\Rightarrow x_1 x'_1 f_{11} + x_1 y'_1 f_{12} + x_1 f_{13} + y_1 x'_1 f_{21} + y_1 y'_1 f_{22} + y_1 f_{23} + x'_1 f_{31} + y'_1 f_{32} + f_{33} = 0$$

$$\Rightarrow \underbrace{\begin{bmatrix} x_1 x'_1 & x_1 y'_1 & x_1 & y_1 x'_1 & y_1 y'_1 & y_1 & x'_1 & y'_1 & 1 \\ & & \vdots & & & & & & \\ & & \vdots & & & & & & \end{bmatrix}}_{n \times 9}^{(A)} \underbrace{\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix}}_{9 \times 1}^{X} = \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}}_{9 \times 1}^{0}$$

we have to solve

$$A x = 0$$

$$A = U D V^T$$

take last column of V → called right null space.
which will give 'F' fundamental matrix.
& make it rank 2 matrix.

2(h) left points ⇒ (100, 200) → (x_1, y_1)
right points ⇒ (50, 100) → (x_1', y_1')
~~To so~~ so A matrix is:-

$$\begin{bmatrix} x_1 x_1' & x_1 y_1' & x_1 & y_1 x_1' & y_1 y_1' & y_1 & x_1' & y_1' & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5000 & 10000 & 100 & 10000 & 20000 & 200 & 50 & 100 & 1 \end{bmatrix}$$

2(i) For 8-point algorithm to work well, we normalise the points first, because otherwise it was not working.

$$\{p_i\}_{i=1}^{n} \rightarrow \{p_i'\}_{i=1}^{n} \Rightarrow \{q_i\}_{i=1}^{n} \rightarrow \{q_i'\}_{i=1}^{n}$$

original normalised

$$q_i = \frac{p_i - \mu_p}{\sigma_p} \qquad q_i' = \frac{p_i' - \mu_{p'}}{\sigma_{p'}}$$

$$p_i = \underbrace{\begin{bmatrix} \frac{1}{\sigma} & & \\ & \frac{1}{\sigma} & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & -\mu_x \\ & 1 & -\mu_y \\ & & 1 \end{bmatrix}}_{M_p} p_i$$

so $q_i = Mp_i$, $q_i' = Mp_i'$

To recover F for original points from F' of normalised points

we know: - $p_i^T F p_i' = 0$ & $q_i^T F' q_i' = 0$.

so,

$q_i^T F' q_i = 0$

$\Rightarrow (M' p_i') F' (M p_i) = 0$ Transpose.

$\Rightarrow p_i^T \underbrace{M'^T F' M}_{F} p_i = 0$

so $\longrightarrow$ $\boxed{F = M'^T F' M}$

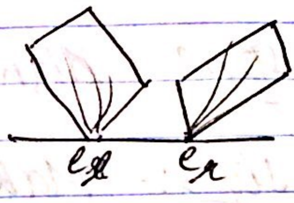
Scanned by CamScanner

Q j) To recover epipole from Fundamental matrix

Epipolar constraint : $P_r^T F P_r = 0$

Given P_l all points on right epipolar line should statisfy $P_r^T F P_l = 0$

in particular,



$$e_r^T F P_l = 0 \quad \forall P_l$$

$$\Downarrow$$

$$e_r^T F = 0$$

$$\Rightarrow \boxed{F^T e_r = 0}$$ we do SVD of F^T

$$F^T = UDV^T$$

$$F = (UDV^T)^T$$

so right epipole is last column of U $(F = UDV^T)$

e_r is left null space of F

for the left epipole

$$P_r^T F P_l = 0 \quad \forall P_l$$

$$\Downarrow$$

$F e_l = 0 \Rightarrow$ left epipole is the right null space of F

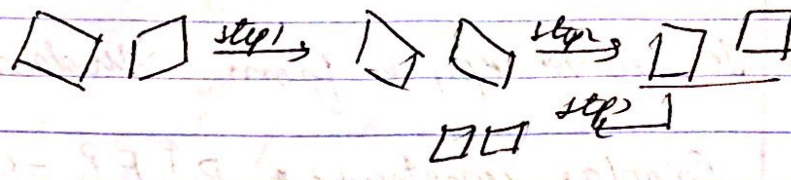
so, SVD of F^T

$= UDV^T$

$\rightarrow$ last column of V will give left epipole.

Scanned by CamScanner

3. a) Rectification.



step1:- align right image with left, so that they have same orientation, rel to each other.

Step2:- Align both with baseline.

Step3:- Make them coplanar

for step1:- → move from pixels to camera coordinates.

$$\bar{P}_l = K_l^{*} P_l \longrightarrow P_l = K_l^{*-1} \bar{P}_l$$

Image (Pixel) camera

similarly, $P_r = K_r^{*-1} \bar{P}_r$

step1:- Apply R^T to right img

step2:- Find coordinate sys aligned with baseline

$$e_1 = \frac{T}{\|T\|}$$

$$\boxed{e_2 = (0,0,1) \times e_1} \quad \boxed{e_3 = e_1 \times e_2}$$

$$R_{rect} = \begin{bmatrix} e_1^T \\ e_2^T \\ e_3^T \end{bmatrix}$$

step3:- Scale using K_l^{*}

$$R_{left-rect} = K_l^* R_{rect} (K_l^*)^{-1}$$

$$R_{right-rect} = K_l^* R_{rect} R^T (K_r^*)^{-1}$$

After the pair is rectified, the corresponding points on both the images will be aligned in a same line (same row) and no disparity.

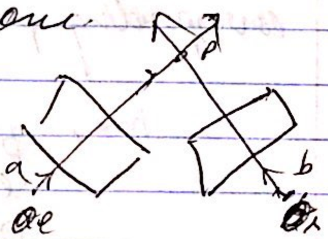
3 b) Approaches for reconstruction :-

— Absolute Reconstruction (Best / complete)

a) complete reconstruction is done

b) know intrinsic & extrinsic parameters

c) In triangulation, ~~send~~ ray is send ~~by~~ through both left & right point to be intersect at some point → our final reconstructed one.



left ray = $a P_l$ =

right ray = $b P_r$

so

$a P_l = b P_r^* \rightarrow$ where $P_r^* = R P_r + T$

(in left coordinate)

— Euclidean Reconstruction

a) This reconstruction is done upto an unknown ~~status~~ scale.

b) know only intrinsic parameter.

c) As extrinsic parameters are not known here, we cannot tell how big / small a object is.

d) we use 8 corresponding points, & do weak calibration get 'F'

e) get essential matrix from F

$$E = K_r^{*T} F K_l^*$$

e) Normalise E is computed $\hat{E} = \frac{1}{2\|T\|^2} E = \frac{2}{Tr[E^T E]} E$

~~f) To find T~~

f) compute R & T upto unknown sign

— Reconstruction upto unknown 3D projective map:

a) worst kind of reconstruction.

b) No parameters are known.

c) Right image rotated by R
Translated by T.

corresponding points → P_l, P_r

here,

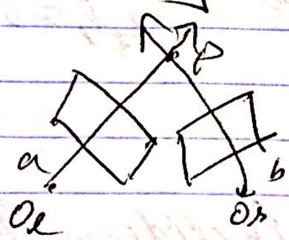
$$P_l = K_l^{*-1} \bar{P}_l$$
$$P_r = K_r^{*-1} P_r$$

[move to camera coordinates]

we give ray from left & right,

left → aP_l

right → bP_r.



to get P → $aP_l - bP_r^*$ (all in left coordinates)

where $P_r^* = RP_r + T$

$\Rightarrow$ $aP_l = b(RP_r + T)$

$aP_l = bRP_r + T \longrightarrow (1)$

To find a, b.

$\Rightarrow$ reconstruction:

$\Rightarrow a P_l + c w = b P_r^*$ where $w \perp$ to both.

so

$w = P_l \times P_r^*$

$= P_l \times R P_r^*$

$\Rightarrow a P_l + c(P_l \times R P_r) - b R P_r = T$

$$\underbrace{\begin{bmatrix} P_l & -RP_r & P_l \times RP_r \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{(X)} = T$$

$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ solve (A, T) or $\begin{pmatrix} a \\ b \\ c \end{pmatrix} = A^{-1} T$

& get results as

reconstructed points $\begin{cases} P = a P_l + \frac{1}{2} c w \\ \underline{P} = \left(\frac{1}{2} a P_l + b R P_r + T\right) \end{cases}$

d) Given points (a, b, c)
equation is

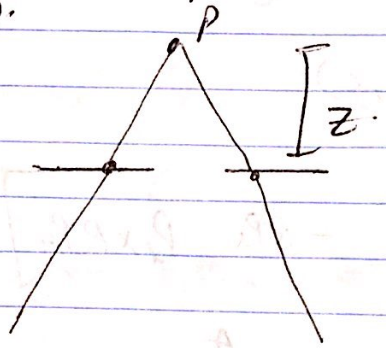
$$\underbrace{\begin{bmatrix} P_l & -RP_r & P_l \times RP_r \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{x} = T$$

solution is

$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = A^{-1} T$ $\rightarrow$ one way to solve.

$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \text{solve}(A, T)$ $\rightarrow$ other way to solve.

3e) In Euclidean reconstruction method, we only
1 know the intrinsic parameter, so to measure distance to
point from image (z).



we can measure the disparity between the points & can measure
Z' • As we don't know the extrinsic parameter
so we don't know distance betwen the optical centres.

3f) we know,

$$F = K_r^{*-T} \cdot E K_l^{*-1}$$

$$E = R^T [T]_x.$$

To normalize E.

$$E^T E = [T]_x^T \underbrace{R R^T}_{I} [T]_x$$

$$\Rightarrow E^T E = [T]_x^T [T]_x \quad \because [T]_x = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$

$$\Rightarrow E^T E = \begin{bmatrix} T_y^2 + T_z^2 & -T_x T_y & +T_x T_z \\ T_x T_y & T_z^2 + T_x^2 & T_y T_z \\ T_x T_z & -T_y T_z & T_x^2 + T_y^2 \end{bmatrix}$$

$$\text{Trace of } E^T E = 2(T_x^2 + T_y^2 + T_z^2) = T_y^2 + T_z^2 + T_z^2 + T_x^2$$

$$= 2 |T|^2$$

$$\text{Normalized } \hat{E} = \frac{2}{Tr(E^T E)} \cdot E$$

The baseline in $\hat{E}$ has length of 1

$\Rightarrow \|T\| = 1$, sign = ? → unknown.

3g) Sign of normalized $\hat{E}$ is unknown, since E is used for computing the translation & rotation of Euclidean construction. → they are upto unknown sign

We can use four possible signs for T & R:-

(+,+) (+,-) (-,+), (-,-)

& choose such that z-coordinate are positive