ASSIGNMENT 6

1)

a) The sparse approach is one which finds the correspondance for special featureds whereas the dense approach is one which finds the correspondance for every pixel. The advantage of this approach is that it can be used in frames with very high disparities between each. But this approach would not work on frames with little or no difference between them.
The advantage of the dense approach is that it would work for frames with very less disparities but would fail for frames with high disparity.

b) Using NCC, $\Psi(\omega_1, \omega_2) = \sum_i \omega_1(x_i, y_i) \cdot \omega_2(x_i, y_i)$

Using SSD, $\Psi(\omega_1, \omega_2) = -\sum_i \left(\omega_1(x_i, y_i) - \omega_2(x_i, y_i)\right)^2$

NCC computes normalized cross corelation of two windows whereas SSD computes the sum of square distance between the windows.
If the search space is allowed to be the entire image, there might be mismatching points since the chances of error are more.
We can use the epipole line for computation.

c) $z = f \frac{T}{d}$

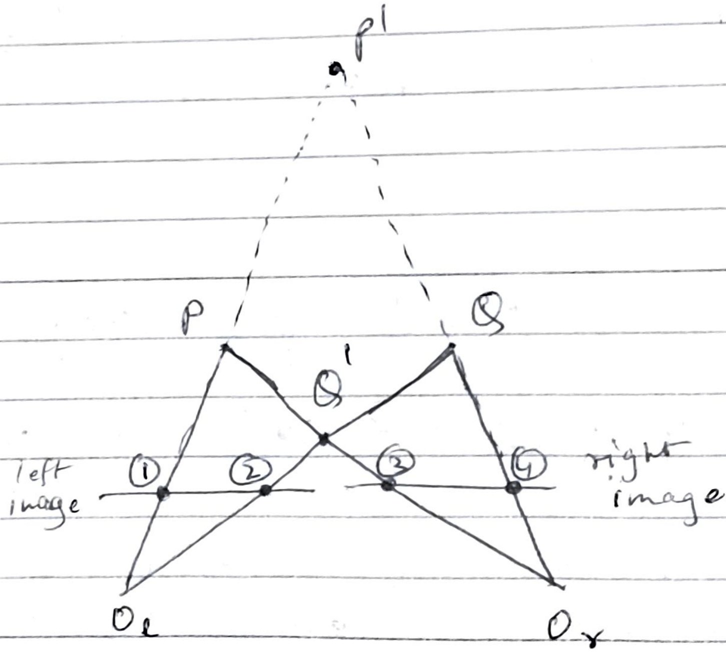
$= f \frac{T}{x_r - x_l}$

$= 10 \times \frac{100}{(103-100)}$

$= 1000/3$

$= 333.33$

d)



Let P and Q be two 3D points. Projecting the points P and Q to the optical centre we get points ① and ③ for P and ② and ④ for Q on the left and right images.

During recontruction, if the correspondance is done correctly we match points 1 and 3 and get P. Similarly we match points 2 and 4 and get point Q.

In case we make a mistake and somehow we match points 1 ~~and~~ and 4 and during ~~recons~~ reconstruction we get the point P'. Similarly we match 2 and 3 and get point Q' ~~a~~ ~~aft~~ after reconstruction.

We see that even with ~~t~~ two points if we make a mistake the incorrect points are very far from the original points.

Some post processing is required to detect the incorrect matches.

e) $M_{left \leftarrow right} = R_l^T T_l^{-1} T_r R_r$

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

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

$R = R_l^T R_r \qquad T = R_l^T (T_r - T_l)$

2)

a) The intersection of the baseline, which is the line connecting the optical centers of the cameras, and the images is ~~the~~ the epipole.
Epipoles do not change and all epipolar lines must pass through the epipole. Epipoles may be inside or outside an image.
Given a point P_l in the left image P_l, O_l, O_r determine an epipolar plane. the intersection of this epipolar plane with the right image determines an epipolar line on it.

b) Essential matrix

$$E = R^T [T]_\times$$

Epipolar constraint equation

$$P_r^T E P_l = 0$$

c) Fundamental matrix

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

Epipolar constraint equation

$$P_r^T F P_l = 0$$

d) The essential (E) and fundamental (F) matrices are of rank 2.

E is of rank 2 because there is only two lines or columns that are linearly independant and the third depends on the first two.
Since ~~E~~ F is generated based on E and E is of rank 2, so F is also of rank 2.

e) Right epipolar line

$$\ell = F P_l$$

f) Left epipolar line

$$\ell = F^T P_r$$

g) Computing E and F directly without the parameters $R, T, K_l^{int}, K_r^{int}$ is known as weak calibration

h) Given $(x'_1, y'_1) = (100, 200)$ and $(x_1^{r}, y_1^{r}) = (50, 100)$

The first line of the matrix is written as:

$$\left[x_i x'_i \quad x_i y'_i \quad x_i \quad y_i x'_i \quad y_i y'_i \quad y_i \quad x'_i \quad y'_i \quad 1 \right]$$

$$= \left[5000 \quad 10000 \quad \cancel{100}\ 50 \quad 10000 \quad 20000 \quad 100 \quad 100 \quad 200 \quad 1 \right]$$

i)

1) Let P_i and P_i' be ~~a~~ a point in the left and right image and q_i and q_i' be the ~~norma~~ ~~normalized~~ corresponding normalized points.

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

where μ is the mean and σ is the standard deviation

In matrix form it can be written as

$$q_i = \underbrace{\begin{bmatrix} 1/\sigma_p & 0 & 0 \\ 0 & 1/\sigma_p & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -\mu_{p_x} \\ 0 & 1 & -\mu_{p_y} \\ 0 & 0 & 1 \end{bmatrix}}_{\equiv M_p} P_i$$

$$q_i = M_p P_i$$
$$q_i' = M_{p'} P_i'$$

Normalization of the point sets is necessary to get a stable solution.

Recovering fundamental matrix of original points:

Convert the points P_i and P_i' to q_i and q_i' using M_p and $M_{p'}$

$$q_i = M_p P_i$$
$$q_i' = M_{p'} P_i'$$

$$q_i^T F' q_i' = 0$$

where F' is the estimated fundamental matrix based on the points q_i and q_i'

$$(M_p P_i)^T F' (M_{p'} P_i') = 0$$

$$P_i^T \underbrace{M_p^T F' M_{p'}}_{=F} P_i' = 0$$

$$F = M_p^T F' M_{p'}$$

where F is the fundamental matrix of the original points.

j) $P_r^T F P_l = 0$

Given P_l we have epipolar line FP_l in the right image.
Every point P on this line satisfies $P_r^T F P_l = 0$
specifically $e_r^T F P_l = 0$

Since this is for all P_l it must be that
$e_r^T F = 0$

$\Rightarrow F^T e_r = 0$

e_r is the left null space of F

Similarly for left epipole:

$P_r^T F e_l = 0 \quad \forall P_r \Rightarrow F e_l = 0$

e_l is the right null space of F

3)

a) ~~Re~~ Steps in rectification:

- → move from image to camera coordinates
- → Align the right image with the left
- → Align both the images with baseline
- → ~~Make~~ make both of them coplanar.

After ~~rec~~ rectification, we have two points p_i and p_i' which can be used to find the 3D point $\underline{P_i}$

b) Different approaches for reconstruction:

(i) When ~~both~~ intrinsic and extrinsic parameters are known → Everything is fully calibrated and we can do an absolute reconstruction

(ii) When only intrinsic parameters are known → We can get an Euclidean reconstruction which means that we can find out the shape of the object but the scale

(iii) When none of the parameters are known → In this case the reconstruction is possible but only upto unknown 3D projective map.

c) The matrix required to solve for a, b, c

$$\left[P_l \quad -RP_r \quad P_l \, R P_r \right]$$

d)

$$P = aP_l + \frac{1}{2} cw$$

or

$$P = \frac{1}{2}\left(aP_l + bRP_r + T\right)$$

e) The scale in Euclidean reconstruction is unknown because the baseline is unknown. This can be removed by ~~measuring the~~ making sure that the baseline is constant i.e equal to one.

f) $$E^T E = \left(R^T [T]_x\right)^T R^T [T]_x$$

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

$$tr\left(E^T E\right) = 2\|T\|^2$$

$$\hat{E} = \frac{2E}{tr\left(E^T E\right)}$$

9) There are 4 possibilities of the signs for matrix R and T
(+, +), (+, -), (-, +), (-, -)

We reconstruct using all the above possibilities and choose the one where the ~~zeros~~ z coordinates are positive.