Name: Irfan Ahmed.

ID : 113166464.

7.)

a.) Translation - 2 DOF

b.) Euclidean - 3 DOF

c.) Similarity - 4 DOF

d.) Affine - 6 DOF

e.) Perspective - 8 DOF

2.) a.)

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} 0 \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} P_{12} & P_{13} & P_{14} \\ P_{22} & P_{23} & P_{24} \\ P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} Y \\ Z \\ 1 \end{bmatrix}$$

world co-ordinates.

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} P_{12} & P_{13} & P_{14} \\ P_{22} & P_{23} & P_{24} \\ P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} Y \\ Z \\ 1 \end{bmatrix}$$

$S \rightarrow$ scaling.

b.) Degrees of freedom - 8

c.) 4 point correspondences are required.

d) Yes. The accuracy will be increased when ~~c~~ calibrating. We can ~~used~~ use linear least squares method and optimize the objective function to get the best fit.

e.) Invariants of planar projective transform:

- concurrency and collinearity
- order of contact
- tangent discontinuities and cusps
- tangency (2 pt contact)

f.) Parallelism, ratio of areas

3.) a.) First derivative of an image:

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

So, the filter should be [-1 1] with which the image should be correlated. for convolution [1 -1] should be used because we rotate by 180° before correlation.

[56 64 79 98 115 126 132 133]

Kernel
[1 -1]

Convolution: [8 15 19 17 11 6 1]

After performing first derivative, the maximas are the points where there was a huge change in intensity. '19' is the max which happens ~~for~~ between pixels with intensity 79 & 98.

To cross verify, we can do 2nd derivative.

2nd Derivative: [7 4 -2 -6 -5 -5]

Zero crossings are edges for 2nd derivative.

b.) i) To approximate a first derivative, sum of values of kernel should be '0'. [because DC Gain should be '0' when convolved]

ii) To approximate second derivative, sum of kernel values should be '0'.

iii) Approximating a Gaussian:

Sum of kernel values should be 1 after normalization.

(to keep the mean of image intact after convolving]

3.) c.) Detecting corners will be more appropriate when you want to match or identify objects or people. When we want to know ~~how~~ similar object in different images (as in keypoint matching) we require corners more than edges because corner are less in number than edges (i.e. two edges make a corner). So, when we need keypoints / key descriptors corners are preferred.

Ex: Motion tracking.

4.) a.)

Naive: Convolution with Gaussian 5x5 filter : $O(mn(25)) = O(mn)$

No. of multiplications : $25mn$

Using FFT

Using FFT, no. of multiplications : $mn \log(mn)$

We can use separable filters of 2D Gaussian as 2 1D filters.

No. of mult : $2 \times 5 \times m \times n = 10mn$

To approximate gaussian we can use repeated convolutions of box filter $\frac{1}{2}[1 \ 1]$ (To reach 5x5 we need 5 box filter convolutions along x & y)

No. of multiplications when we optimize it further by

- retaining previous sums (using sliding window) : $\underset{x\,direction}{mn} + \underset{y\,direction}{mn} = 2mn$.

b.)

$$\frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Convolving with this filter → we can just add up pixel values of 5x5 size and in end multiply with $1/25$.

So, no. of multiplications : mn.

c.)

$$\frac{1}{A}[1 \quad 4 \quad 6 \quad 4 \quad 1]$$

Gaussian smoothing : Sum of values after normalization is 1.

$$\frac{1}{A}(1+4+6+4+1) = 1$$

$$A = 16$$

d.) Yes. Box filter $\frac{1}{2}[1 \quad 1]$.

Convolve box with box (additional padding) : $\frac{1}{4}[1 \quad 2 \quad 1]$

This actually looks like a pascal pyramid.

$$\begin{array}{ccccccccc} & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1. \end{array}$$

So, 3 ~~4~~ convolutions are required.

$$\frac{1}{2}[1 \quad 1] * \frac{1}{2}[1 \quad 1] = \frac{1}{4}[1 \quad 2 \quad 1]$$

$$\frac{1}{4}[1 \quad 2 \quad 1] * \frac{1}{2}[1 \quad 1] = \frac{1}{8}[1 \quad 3 \quad 3 \quad 1]$$

$$\frac{1}{8}[1 \quad 3 \quad 3 \quad 1] * \frac{1}{2}[1 \quad 1] = \frac{1}{16}[1 \quad 4 \quad 6 \quad 4 \quad 1]$$

Additional 0's are padded.

5.)

a) (ii) Gaussian

(iii) Laplacian of Gaussian.

(iv) Derivative of Gaussian along X

b.) Aliasing can be prevented by using low pass filters (such as gaussian blur).

Causes aliasing → Box filter

↳ Laplacian of Gaussian

↳ Derivative of Gaussian along X & Y

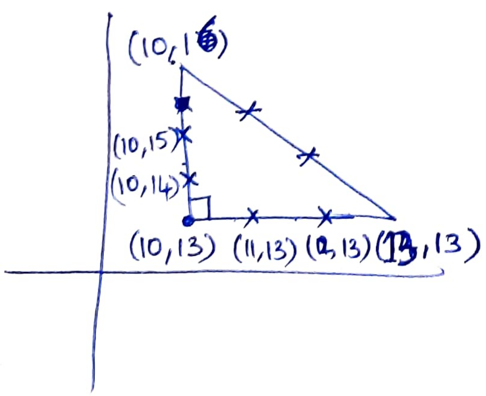
6.) a) There are 2 dimensions in Hough parameter space.

i.e. 'ρ' & 'θ' ← orientation. (In $y = mx + c$)

↳ dist from origin — slope (m) — intercept (c).

(rho, theta)

b.)



A '1' in input image at (10,13) will vote for

(10,16),	(11,13)
(10,15)	(12,13)
(10,14)	(13,13)

c.) To generalize solution for range, my 'ρ' value will be crossed as '1' for all cells from (ρ+3, . . . ρ+20) → To include all lengths.

7.)

a.) Co-ordinates of A w.r.t camera $(3, 2, 10)$

$f = 0.05\ m$

$x' = \frac{fX}{Z}$, $y' = \frac{fY}{Z}$

$x' = \frac{0.05 \times 3}{10}$, $y' = \frac{0.05 \times 2}{10}$

$x' = \frac{0.15}{10}$, $y' = \frac{0.1}{10}$ $\Rightarrow (x', y') = (0.015, 0.01)$

↳ Co-ordinates of A in img plane.

b.) The person's head co-ordinates are $(0, 0, 8)$ w.r.t camera

$\therefore (x', y') = (0, 0) \rightarrow \left(\frac{0.05 \times 0}{8}, \frac{0.05 \times 0}{8}\right)$

the person is lying in front of origin with depth of 8m (which does not reflect on image plane).

c.) $(24 \times 32) \rightarrow (480 \times 640)$

$(u_0, v_0) \rightarrow (300, 250)$.

$k_u = \frac{\cancel{480}^{20}}{\cancel{24} \times 10^{-3}}$ (mm to m) $= 2 \times 10^4$

$k_v = \frac{640}{32 \times 10^{-3}} = 2 \times 10^4$

for) a) $x' = u_0 + \frac{k_u f X}{Z} = 300 + \frac{\cancel{2} \times 10^4 \times \cancel{5} \times 10^{-2} \times 3}{\cancel{10}} = 300 + 300 = 600.$

$y' = v_0 + \frac{k_v f Y}{Z} = 250 + \frac{\cancel{2} \times 10^4 \times \cancel{5} \times 10^{-2} \times 2}{\cancel{10}} = 250 + 200 = 450$

b.) $x' = u_0 + \frac{k_u f(0)}{Z} = u_0 = 300$

$y' = v_0 + \frac{k_v f(0)}{Z} = v_0 = 250.$

d.) First the world co-ordinate ~~plane~~ axes were rotated and then translated to camera reference.

After that it has been translated & scaled w.r.t to CCD array.

i.e., ~~to~~ wrt principal points and scaling (K_u, K_v)

$$P_c = C[R|T]P_w.$$

e.)

(u,v) $0.2 m/sec.$

Faster shutter speeds are better for fast moving objects to avoid motion blur.

$$\frac{du}{dt} = \frac{2\times10^{+4} \times 5\times10^{-2} \times 0.2}{10} , \quad \cancel{8\times10^{4}\times5\times10^{-2}\times}$$

$$\frac{du}{dt} = 20 \text{ pixel/sec.}$$

Min Shutter speed should be $1/20$ th of second.