ASSIGNMENT→1

NAME :- PUSHPENDRA SINGH CHAUHAN
STUDENT ID :- A20472647
COURSE NO. :- CS512
SEMESTER :- FIRST (FALL 2020)

REVIEW QUESTIONS

Answer the following questions. Make sure that your answers are concise. In questions requiring explanation, make sure your explanation is brief.

1. Geometric image formation:

(a). Let $f=10$ be the focal length of a camera. Let $p=(3,2,1)$ be a world point. Find the co-ordinate of the point p when projecting it onto the image. Assume that the projection is done in camera co-ordinates so there is no need for a transformation between coordinate systems.

Soln:- Let $p'=(u,v)$ be the coordinate of the point $p=(3,2,1)$ when projecting it onto the image.

Here,

$X=3,\ Y=2,\ Z=1,\ f=10$ (focal-length of camera)

Now, $p'=(u,v)$ will be given by : $u = -f\dfrac{X}{Z}$

$$u = -(10) \times \left(\frac{3}{1}\right)$$

$$\boxed{u = -30} \text{ Ans}$$

$$v = -f \frac{y}{z}$$

$$v = -(10) \times \left(\frac{2}{1}\right)$$

$$\boxed{v = -20} \text{ Ans.}$$

Therefore, $p' = (-30, -20)$ are the coordinate of the point $p = (3, 2, 1)$ when projecting it onto the image.

(b). Explain the difference between the pinhole camera model where the image plane is behind the center of projection and the pinhole camera model where the image plane is in front of the center of projection. Which model corresponds better to a physical pinhole camera model? How is the other model justified?

Ans:- • Difference :-

The image formed by the pinhole camera model where the image plane is behind the center of projection is inverted.

While, the image formed by the pinhole camera model where the image plane is in front of the center of projection is not inverted.

- The pinhole camera model where the image plane is behind the center of projection corresponds better to a physical pinhole camera model.

- Though the other model (i.e. the pinhole camera model where the image plane is in front of the center of projection) cannot be implemented in practice, but it provides a theoretical camera which may be simpler to analyse than the real one.

(c). Eexplain what happens to the projection of an object when the focal length gets bigger and what happens to the projection when the distance to the object gets bigger.

Ans:- • The size of the projection of an object (image size) increases when the focal length gets bigger.

- The size of the projection of an object (image size) decreases when the distance to the object gets bigger.

(d). Given the 2D point (1,1) find its coordinates in homogenous coordinates (2DH). Find another 2DH point that corresponds to the same 2D point.

Ans:- We know that,

Homogenous coordinates are given by:

2D $\qquad$ 2DH

- $(x, y) \longrightarrow (x, y, 1)$.

- 2DH point $(x, y, 1)$ multiplied by a constant t (Here, $t \neq 0$) gives another 2DH point ~~tx~~, (tx, ty, t) which corresponds to same 2D point (x, y).

Given: 2D point $= (1, 1)$

Now,

- $\left.\begin{matrix}\text{Homogenous coordinates} \\ \text{(2DH) point}\end{matrix}\right\} = \underline{\underline{(1, 1, 1)}}$. Ans

- $\left.\begin{matrix}\text{Another 2DH point that} \\ \text{corresponds to the same} \\ \text{2D point}\end{matrix}\right\} = \underline{\underline{(2, 2, 2)}}$. Ans

(e). Given the 2DH point $(1, 1, 2)$, find the 2D point corresponding to it.

Ans:- 2D point from a given 2DH point:

2DH $\qquad$ 2D

$(x, y, w) \longrightarrow (x/w, y/w)$

Now,

2DH (given) $\qquad$ 2D

$(1, 1, 2) \longrightarrow \underline{\underline{(1/2, 1/2)}}$ Ans.

(f). Explain the meaning of the 2DH point $(1, 1, 0)$.

Ans:- The 2DH point (1, 1, 0) denotes the point at infinity or direction.

(g). Eexplain what makes it possible to write the non-linear projection equation as a linear equation in homogenous coordinates.

Ans:- Non-linear projection equations are given by:

$$u = \frac{fx}{z} \quad \text{and} \quad v = \frac{fy}{z}$$

Representation of projection equation using linear equation in homogenous coordinates:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \begin{matrix} U = fx \\ V = fy \\ W = z \end{matrix}$$

Here, 2D $\longrightarrow$ 2DH

$$(u, v) \longrightarrow (U, V, W)$$

$$\Rightarrow u = \frac{U}{W}, \quad v = \frac{V}{W}$$

From above, we can clearly see that the extra term 'W' added in 2DH coordinate is acting like a 3rd dimension and due to this on equating, both sides are having same dimension. That is why we are getting linear projection equation in homogenous coordinates.

(h). Given the projection matrix $M = K[I|0]$. Write the dimensions ~~of~~ of M and the sub-matrices K, I, O.

Ans:-

$$\underset{(3\times4)}{M} = \underset{(3\times3)}{K}\overbrace{[\underset{(3\times3)}{I} \mid \underset{(3\times1)}{O}]}^{(3\times4)}$$

- Dimension of $M = (3\times4)$ rows → columns
- Dimension of $K = (3\times3)$
- Dimension of $I = (3\times3)$
- Dimension of $O = (3\times1)$

(i). Given a projection matrix M whose rows are [1,2,3,4], [5,6,7,8], [1,2,1,2] and a 3D point P = [1,2,3], find the coordinate of the 2D point p obtained by projecting P using M.

Ans:- We know that,

$$r = M \times R$$

Here,

r = 2DH point obtained after projection
M = projection matrix
R = 3DH point which will be projected using M.

Let,

$p = (u, v)$ {2D point obtained after projection}

$r = (U, V, W)$

Now,

3D (P)		3DH (R)
(1,2,3)	$\longrightarrow$	(1,2,3,1)

So,

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix}$$

So,

$$r = M \times R$$

$$\Rightarrow \begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} (1\times1)+(2\times2)+(3\times3)+(4\times1) \\ (5\times1)+(6\times2)+(7\times3)+(8\times1) \\ (1\times1)+(2\times2)+(1\times3)+(2\times1) \end{bmatrix}$$

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 18 \\ 46 \\ 10 \end{bmatrix}$$

On converting 2DH point to 2D, we get

$$\boxed{u = \frac{U}{W} = \frac{18}{10}} \text{ Ans} \qquad \boxed{v = \frac{V}{W} = \frac{46}{10}} \text{ Ans.}$$

Therefore, the coordinate of the 2D point p obtained by projecting P using M is

$$\boxed{p = \left(\tfrac{18}{10}, \tfrac{46}{10}\right) \text{ or } (1.8, 4.6)} \text{ Ans.}$$

2. Modeling transformations:

(a). Given the point (1,1) find its coordinates after translating it by (2,3). Perform the computation using a transformation matrix.

Ans: Transformation matrix for translation in homogenous co-ordinates for 2D point is given by:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Transformation matrix}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Here,

$$x = 1,\ y = 1;\ t_x = 2,\ t_y = 3$$

So,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} (1\times1)+(0\times1)+(2\times1) \\ (0\times1)+(1\times1)+(3\times1) \\ (0\times1)+(0\times1)+(1\times1) \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} \Rightarrow \boxed{(x', y') = (3,4)} \text{ Ans.}$$

Therefore, the coordinates of the point (1,1) after translating it by (2,3) are (3,4) Ans.

(b). Given the point (1,1) find its coordinates after scaling it by (2,2). Perform the computation using a transformation matrix.

Ans :- Transformation matrix for scaling in homogenous coordinates for 2D point is given by :

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Transformation matrix}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Here,

$$x = 1,\ y = 1\ ;\ S_x = 2\ ;\ S_y = 2$$

So,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} (2\times1)+(0\times1)+(0\times1) \\ (0\times1)+(2\times1)+(0\times1) \\ (0\times1)+(0\times1)+(1\times1) \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \Rightarrow \boxed{(x', y') = (2,2)}\ \text{Ans.}$$

Therefore, the coordinates of the point (1,1) after scaling it by (2,2) are (2,2). Ans

(c). Given the point (1,1) find its coordinates after rotating it by 45 degrees.

Ans:- 2D rotation about origin is given by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Here,

$$x = 1,\ y = 1 \text{ and } \theta = 45^\circ.$$

So,
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(45^\circ) & -\sin(45^\circ) \\ \sin(45^\circ) & \cos(45^\circ) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} (1/\sqrt{2} - 1/\sqrt{2}) \\ (1/\sqrt{2} + 1/\sqrt{2}) \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 0 \\ 2/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix} \Rightarrow \boxed{(x', y') = (0, \sqrt{2})} \text{ Ans}$$

Therefore, the coordinates of the point (1,1) after rotating it by 45° degrees about origin are $(0, \sqrt{2})$ Ans.

(d). Given the point (1,1) find its coordinates after rotating it by 45 degrees about the point (2,2).

Ans:- Steps to be followed :-

(i). Translate the reference point (2,2) by (−2,−2) to the origin (0,0) and also translate the point (1,1) by (−2,−2).

Let T' be the transformation matrix, then,

$$T' = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} t_x = -2 \\ t_y = -2 \end{Bmatrix}$$

(ii). Then rotate by 45° about origin.

Let R be the transformation matrix.
then,

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(45^\circ) & -\sin(45^\circ) & 0 \\ \sin(45^\circ) & \cos(45^\circ) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(iii) Translate back by (2,2)

Let T be the transformation matrix,
then,

$$T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \left\{ \begin{array}{l} t_x = 2 \\ t_y = 2 \end{array} \right\}$$

Now,

$$P' = T * R * T' * P$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -\sqrt{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2-\sqrt{2} \\ 1 \end{bmatrix}$$

$$\Rightarrow \boxed{(x', y') = (2, 2-\sqrt{2})} \text{ Ans.}$$

Therefore, the coordinates of the point (1,1) after rotating it by 45° degrees about the point (2,2) are $(2, 2-\sqrt{2})$ Ans.

(e). Given that I want first to rotate an object using a matrix R and then translate it using a matrix T, what should be the combined matrix (expressed in terms of R and T) that needs to be applied to the object.

Ans: Let combined matrix be M then,

$$\boxed{M = T * R} \text{ Ans.}$$

(f). Let M be a 2D transformation matrix in in homogenous coordinates whose rows are [3,0,0], [0,2,0], [0,0,1]. What is the effect of applying this matrix to transform a point p.

Ans: We know that, 2D transformation matrix in homogenous coordinates for scaling is given by :

$$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = M$$

{Given matrix}

So, $S_x = 3$ and $S_y = 2$.

Therefore, the point p will be scaled by (3,2) if we apply the given transformation matrix M to the point p.

(g). Let M be a 2D transformation matrix in homogenous coordinates whose rows are [1,0,1], [0,1,2], [0 0 1]. What is the effect of applying this matrix to transform a point p.

Ans: We know that, 2D transformation matrix in homogenous coordinates for translation is given by:
$$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} ; \text{ given: } M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

So,
$t_x = 1$ and $t_y = 2$.

Therefore, the point p will be translated by (1,2) if we apply the given transformation matrix M to the point p.

(h). Let M be a 2D transformation matrix in homogenous coordinates whose rows are [3,0,0], [0,2,0], [0,0,1]. What is the transformation matrix which will reverse the effects of this transformation?

Ans:- Given transformation matrix:
$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It will scale by (3,2).

Therefore, for reversing this effect we have to scale again by $(1/3, 1/2)$.

So,
$$\left.\begin{array}{l}\text{the required}\\ \text{transformation matrix}\end{array}\right\} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ **Ans.**

(i). Let $M = R(45)\,T(1,2)$ be a transformation matrix in homogenous coordinates composed of rotation by 45 degrees and a translation by (1,2). Express the inverse of this transformation in terms of a rotation and translation matrix.

Ans:- Given : $M = R(45)\,T(1,2)$

Now,
$$M^{-1} = [R(45)\,T(1,2)]^{-1}$$
$$\boxed{M^{-1} = T^{-1}(1,2)\,R^{-1}(45)}$$ **Ans**

By using inverse property: $(ab)^{-1} = b^{-1}a^{-1}$

(j). Find a vector which is perpendicular to the vector (1,3).

Ans: Let $\vec{a} = (x,y)$ be the perpendicular vector to the given vector (1,3).

We know that,
Dot product of two perpendicular vectors is equal to 0 (zero).

So,

$$(x, y) \cdot (1, 3) = 0$$
$$1 \times x + 3 \times y = 0$$
$$x + 3y = 0.$$

if $\boxed{x = 3}$ then,

$$3 + 3y = 0 \Rightarrow \cancel{3}y = -\cancel{3}$$
$$\boxed{\vec{a} = (3, -1)} \text{ Ans} \qquad \boxed{y = -1}$$

Therefore, the vector $\vec{a} = (3, -1)$ is perpendicular to the given vector $(1, 3)$.

(k). Find the projection of the vector $(1,3)$ onto the direction defined by the vector $(2,5)$.

Ans:- We know that,

$$\left.\begin{array}{r}\text{projection of vector } \vec{a} \text{ onto the} \\ \text{direction defined by the vector } \vec{b}\end{array}\right\} = \frac{(\vec{a} \cdot \vec{b})}{|\vec{b}|^2} \vec{b}$$

So,

$$\left.\begin{array}{r}\text{the projection of vector } (1,3) \\ \text{onto the direction defined by} \\ \text{the vector } (2,5)\end{array}\right\} = \frac{\{(1,3) \cdot (2,5)\}}{|(2,5)|^2} (2,5)$$

$$= \frac{\{1 \times 2 + 3 \times 5\}}{\left(\sqrt{(2)^2 + (5)^2}\right)^2} (2,5)$$

$$= \frac{17}{29} (2,5) = \left(\frac{34}{29}, \frac{85}{29}\right) \text{ Ans}$$

Therefore, the ~~proje~~ projection of the vector $(1,3)$ onto the direction defined by the vector $(2,5)$ is $\left(34/29, 85/29\right)$ Ans

3. General Camera model :-

(a). Explain the need for a general projection matrix that uses different coordinate systems for camera and image.

Ans:- The need for a general projection matrix that uses different coordinate systems for cameras and image is to properly relate world points to image points by taking into account different coordinate systems.

(b). Given that the camera is rotated by R and translated by T with respect to the world, write the transformation matrix that will convert world to camera coordinates.

Ans: The transformation matrix that will convert world to camera coordinates is given by :

$$\boxed{M_{C \leftarrow W} = \left[\begin{array}{c|c} R^* & T^* \\ \hline 0 & 1 \end{array}\right]}$$ Ans.

Here,
R^* = Rotation of world with respect to camera.
T^* = Translation of world with respect to camera.

or

$$\boxed{M_{C \leftarrow W} = \begin{bmatrix} R^T & -R^T T \\ 0 & 1 \end{bmatrix} = \left[\begin{array}{c|c} R^* & T^* \\ 0 & 1 \end{array}\right]}$$ Ans

(c). Given three unit vectors $\hat{x}, \hat{y}, \hat{z}$, write the rotation matrix describing the rotation of the camera with respect to the world.

Ans: The rotation matrix describing the rotation of the camera with respect to the world is given by:

$$\boxed{R^T = \begin{bmatrix} \hat{x}_c^T \\ \hat{y}_c^T \\ \hat{z}_c^T \end{bmatrix}}$$

Ans.

Here, c = Camera coordinate

(d). Given the transformation matrix between world and camera coordinates $M = \begin{bmatrix} R^* & T^* \\ 0 & 1 \end{bmatrix}$. Explain the meaning of R^* and T^*.

Ans:- Here,

R^* = Rotation of world-coordinates with respect to camera coordinates.

T^* = Translation of world coordinates with respect to camera coordinates.

(e). Given that ~~the~~ there are k_u pixels per mm in the direction of X, k_v pixels per mm in the Y direction, and the optical center of the camera is translated by $(u_0, v_0) = (512, 512)$ pixels, write the transformation matrix that will convert camera coordinates to image coordinates.

Ans:- Given,

K_u = pixels per mm in the direction of X
K_v = pixels per mm in the direction of Y
u_0 = 512 pixels
v_0 = 512 pixels

The transformation matrix that will convert camera coordinates to image coordinates is given by:

$$\boxed{M_{i \leftarrow c} = \begin{bmatrix} K_u & 0 & u_0 \\ 0 & K_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} K_u & 0 & 512 \\ 0 & K_v & 512 \\ 0 & 0 & 1 \end{bmatrix}}$$ Ans

(f). Let the projection matrix M of a general camera be given by $K^*[R^*|T^*]$. Explain which parts contain the intrinsic and extrinsic parameters of the camera.

Ans:- • K^* contain the intrinsic and $[R^*|T^*]$ contain the extrinsic parameters of the camera.

(g). Eexplain the reason for including a 2D skew parameter in the camera model.

Ans:- The pixels in a CCD sensor may not be perfectly square, resulting in a small distortion in the X or Y directions.
In order to take such skew effect into account we include a 2D skew parameter in the

camera model. (It takes into account shear or skew transformation).

(h). Eexplain what happens to the camera model when taking into account radial lens distortion. What is the complication introduced by the radial lens distortion?

Ans: Radial distortion occurs when light rays bend more near the edges of a lens than they do at its optical sensor. The smaller the lens, the greater the distortion. It is more frequent in wide angle lenses.

- The camera model when taking into account radial lens distortion is given by:

Scale transformation.

$$p^{(i)} = \begin{bmatrix} 1/\lambda & 0 & 0 \\ 0 & 1/\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} K^* [R^* | T^*] \, P^{(w)}$$

Here,

linear distortion coefficient

$$\lambda = 1 + k_1 d + k_2 d^2$$

Quadratic distortion coefficient

d = distance from the center

- The complication introduced by the radial lens ~~distortat~~ distortion is that the scale is not uniform while applying scale transformation.
  Scale transformation matrix is not fixed, it will change at each location.

The more we move away from the center, the larger shrink will happen.

(i). Explain the meaning of a weak-perspective camera and of an affine camera.

Ans: • Weak-perspective cameras are cameras where we don't see ~~so~~ much perspective-effects like fore-shortening (Distant parts shorter than closer parts).

Projection matrix of a weak-perspective camera is given by:

$$\boxed{M_\infty = \begin{bmatrix} \text{———————} \\ \text{———————} \\ 0 \quad 0 \quad 0 \quad 1 \end{bmatrix}}$$

• Affine camera is made of affine matrix.

$$M_{affine} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$(a, b, c, d, e, f, g, h) \rightarrow$ These are just arbitrary 8 numbers.

It is a computational model so there is no guarantee that affine camera will produce results like real camera.

4. Color and photometric image formation:

(a). Explain the difference between surface radiance and image irradiance.

Ans: Surface radiance is defined as power of light per unit area reflected from the surface.

Image irradiance is defined as power of light per unit area received at the image.

(b). Write the radiosity equations relating surface radiance and image irradiance.

Ans: Radiosity equation relating surface radiance and image irradiance is given by:

$$\boxed{E(p) = L(p)\frac{\pi}{4}\left(\frac{d}{f}\right)^2 (\cos\alpha)^4} \text{ Ans.}$$

Here,

$E(p)$ = Image irradiance
$L(p)$ = Surface radiance
f = focal length
d = diameter of lens
'α' = Angle between principal axis and surface normal

(C). Define the albedo of a surface

Ans: Surface albedo is a reflection coefficient. It lies between 0 and 1 i.e. surface albedo $\in [0,1]$.

It tells us how good reflector the surface is.

→ If the surface is a good reflector then surface albedo ≈ 1.

→ If the surface is not a good reflector then surface albedo ≈ 0.

(d). Explain what is the reason for using the RGB color model to represent colors.

Ans: Human vision use R, G, B receptors. because of this we can imitate real colors with R, G, B combinations.

This is the reason why we use the RGB color model to represent colors.

(e). Given the RGB color cube, what are the colors along the line that connects (0,0,0) with (1,1,1).

Ans: We will have different shades of gray colors along the line that connects (0,0,0) with (1,1,1) in the RGB color cube.

(f). Explain the way by which RGB colors are mapped to real-world colors.

Ans: The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors.

In the real world, light consists of all visible colors, not just red, green, and blue wavelengths. We have structures in our eyes called "cones" that respond to red, green, and blue light sources. A monochromatic yellow light excites both the red and green cones in our eyes, and we see it as yellow.

(g). Given the CIE RGB color model and its conversion to the XYZ model, explain what is the ~~...~~ use for the luminance component Y.

Ans: The luminance component Y is used to measure the brightness or intensity of a color.

(h). Eexplain the advantage of the LAB color space.

Ans: The advantage of the LAB color space is that it measures euclidean distance in perception space.

Eeuclidean distance in LAB color space does correspond to human perception.

Color 1 Color 2

Eg:- $|C_1 - R| < |C_2 - R| \Rightarrow C_1$ is more similar to R

Reference point

5. Noise and filtering:

(a). Explain how to estimate the signal to noise ratio (SNR) in an image.

Ans: Signal to noise ratio (SNR) in an image is defined as:

$$\boxed{SNR = \frac{E_s}{E_n} = \frac{\sigma_s^2}{\sigma_n^2} = \frac{\frac{1}{n}\sum_{i,j}(I(i,j)-\bar{I})^2}{\sigma_n^2}}$$

Here,

E_s = Energy of signal
E_n = Energy of noise

σ_s^2 = Variance of signal
σ_n^2 = Variance of noise

n = total no. of pixels in the image
$I(i,j)$ = Intensity of the pixel at location (i,j)
$\bar{I}$ = Average intensity of all pixels.

σ_n^2 = Variance for multiple frames of a static scene
or
Variance in a uniform image region

- SNR is measured in decibels (db).

$$\boxed{SNR[db] = 10 \log_{10}\left(\frac{E_s}{E_n}\right)}$$

(b). Eexplain the difference between Gaussian and impulsive noise. Which filter handles better impulsive noise: an averaging filter or a median filter.

Ans: • The difference between Gaussian and impulsive noise is that:
Gaussian noise is in the form of Gaussian distribution but impulsive noise can have random values from different distribution.

• A median filter handles impulsive noise better than an averaging filter.

(c). Given an image having the value of 2 in each cell, write the value of the pixels in this image after applying a 3X3 convolution filter having all 1-s in its entries.

Ans: 18 Ans.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$ ← Convolution filter (3X3)

Convolution operation

$\boxed{9\,(1\times 2) = 18}$ Ans

Image ←

$$\begin{bmatrix} 2 & 2 & 2 & 2 & \cdots \\ 2 & 2 & 2 & 2 & \cdots \\ 2 & 2 & 2 & 2 & \cdots \\ 2 & 2 & 2 & 2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

Therefore, the value of the pixels in this image after applying a 3X3 convolution filter having all 1-s in its entries is 18 Ans.

(d). Given that we need the derivative of an image convolved with a filter. Explain how the operation can be applied more efficiently.

<u>Ans</u>: Let the image be 'I' and convolution filter be 'f'.

$$\left.\text{Image after applying convolution filter}\right\} = (I * f)$$

$$\left.\begin{array}{l}\text{Derivative of an image}\\ \text{convolved with a filter}\end{array}\right\} = \frac{d}{dx}(I * f)$$

$$= \frac{d}{dx} I * f \quad \text{Ans}$$

or

$$= I * \frac{d}{dx} f \quad \text{Ans.}$$

(e). Explain the three different ways to handle boundaries during convolution.

<u>Ans</u>:- The three different ways to handle boundaries during convolution are :-

(i) Zero padding

0	0	0	0	0
0	1	2	3	0
0	4	5	6	0
0	7	8	9	0
0	0	0	0	0

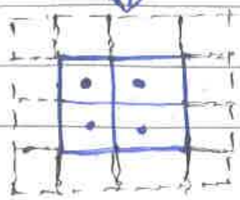
(ii) Mirror / Replicate

1	1	2	3	3
1	1	2	3	3
4	4	5	6	6
7	7	8	9	9
7	7	8	9	9

(iii) Ignore

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

⇓



(f). Write a basic 3x3 smoothing filter. What is the sum of all entries in this filter? Explain the reason for the sum to be selected as it is.

Ans: Basic 3x3 smoothing filter :

1	1	1
1	1	1
1	1	1

Sum of all entries in this filter $= 9 \times 1 = \underline{\underline{9}}$.

The reason for the sum to be selected as it is
→ It is a linear and basic filter which contain 9 cells with 1 in each ~~entry~~ cell.

(g). Explain how to implement a 2D convolution with a Gaussian using two 1D convolution filters. Which option is more efficient? Is it possible to implement any 2D filter in this way?

Ans: • 2D Gaussian is given by : $G_\sigma(x,y) = e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)}$

Procedure to implement a 2D convolution with a Gaussian using two 1D convolution filters is :

$$I_G = I * G = \sum_i \sum_j I(i,j)\, e^{-\left(\frac{i^2+j^2}{2\sigma^2}\right)}$$

$$= \sum_i e^{-\left(i^2/2\sigma^2\right)} \sum_j I(i,j)\, e^{-\left(j^2/2\sigma^2\right)}$$

$$= (I * G_y) * G_x = \underline{\underline{I * G_x * G_y}}$$

(i) Convolve with 1D Gaussian filter along rows.
(ii). Then, convolve with another 1D Gaussian filter along columns.

• Implementing a 2D convolution with a Gaussian using two 1D convolution filter is computationally more efficient than a 2D convolution filter.

• No, it is not possible to implement any 2D filter in this way. To implement a 2D filter in this way, it must have a 'separable' property.

(h). Given a 1D Gaussian filter with $\sigma = 2$, what should be the size of this filter?

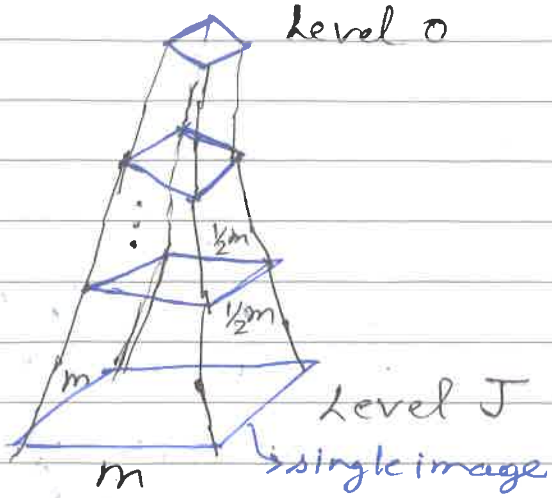
Ans: Size of a 1D Gaussian filter with $\sigma = 2$ is $\left.\right\} = 5\sigma$
$= 5 \times 2$
$= 10$ Ans.

(i). Explain how a Gaussian image pyramid is produced. What is the reason for producing such pyramids? What is the amount of additional processing done in a pyramid compared with a single image?

Ans: When we analyse an ~~Gaussian~~ image at different levels then it produces a Gaussian image pyramid.

$m = 2^J \Rightarrow J = \log_2 m$

$$\left.\begin{array}{l}\text{total no. of}\\ \text{pixels}\end{array}\right\} = m^2 + \frac{1}{4}m^2 + \frac{1}{16}m^2 + \dots < \frac{4}{3}m^2.$$



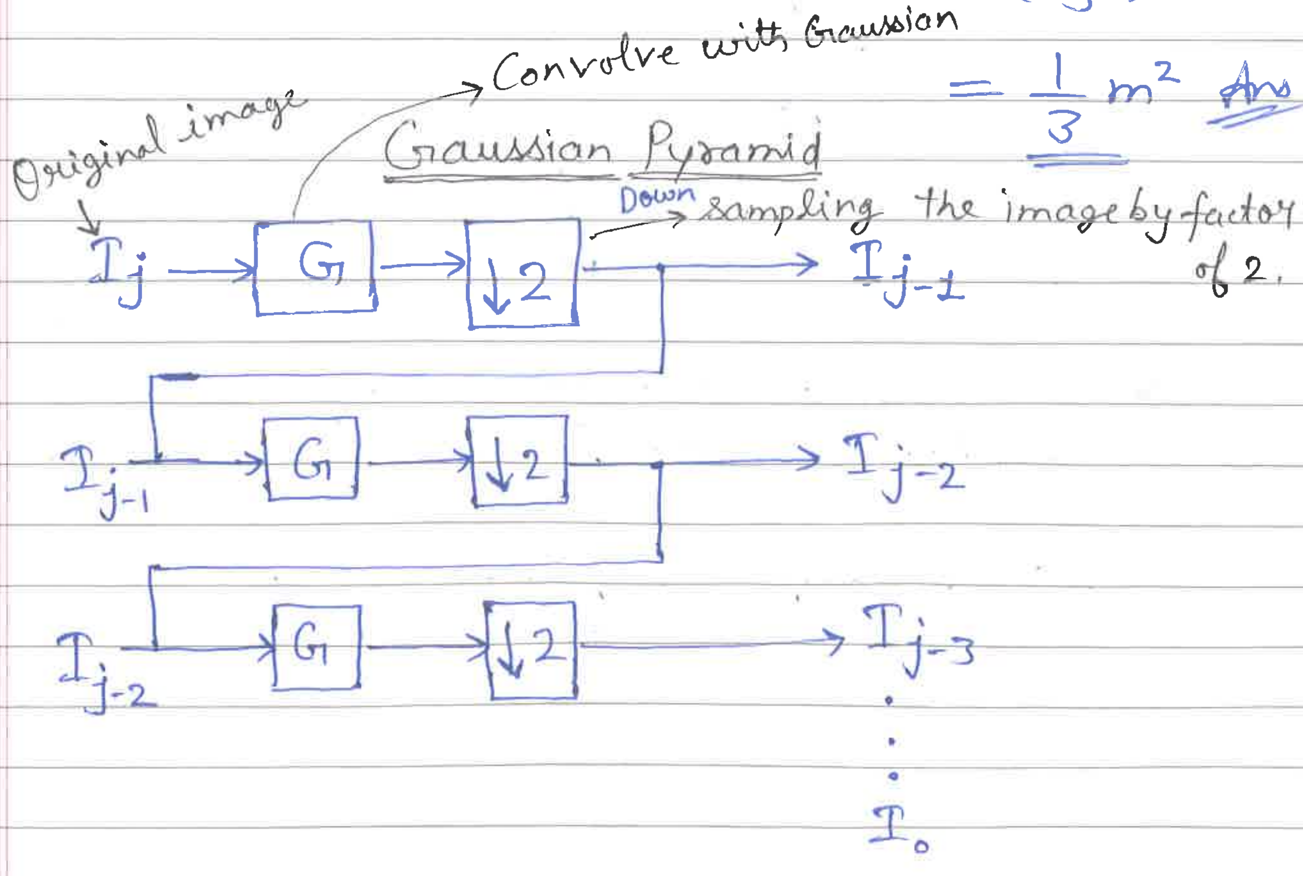
- The reason for producing such pyramids is to do multiple scale ~~image~~ analysis of an image.

- $$\left.\begin{array}{l}\text{The amount of additional}\\ \text{processing done in a pyramid}\\ \text{compared with a single image}\end{array}\right\} = \frac{4}{3}m^2 - m^2$$

$$= \left(\frac{4-3}{3}\right)m^2$$

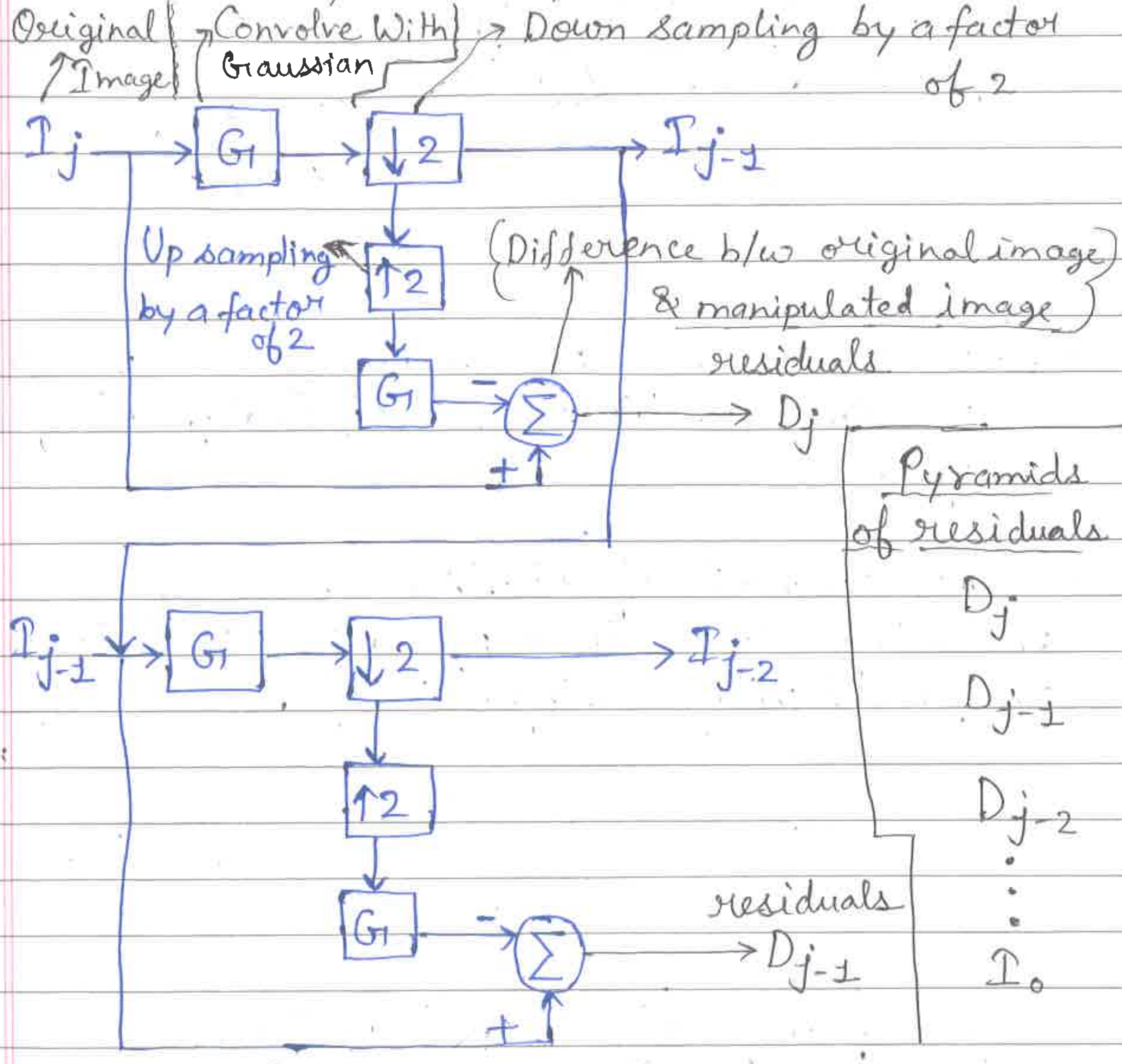
$$= \frac{1}{3}m^2 \text{ Ans.}$$

<u>Gaussian Pyramid</u>



(j) Explain how the Laplacian pyramid is produced and its use.

Ans: Laplacian pyramids



- We can make a pyramid of residuals and this is called Laplacian pyramids.

- It is ~~usefo~~ useful for compression.

6. Edge detection :

(a). Why is edge detection useful? What are the desired properities of edge detection?

Ans: • Edge detection is useful because edges provide good description of the content of the image.

• Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. It is used for image segmentation and data extraction in areas such as image processing, computer vision and machine vision.

• The desirable properities of edge detection are :-

→ Correspond to scene elements.
→ Invariant [illumination, pose, viewpoint, scale].
→ Reliable detection.

(b). Explain the basic steps of edge detection and the need for them: smoothing, enhancement, localization.

Ans: The basic steps of edge detection are:
(i) Smooth to reduce noise (Without affecting edges)
(ii) Enhance edges

(iii) Detect edges
(iv) Localize edges

- The need for smoothing is to reduce noise from the image.
- The need for enhancement is to amplify the edges
- The need for localization is to precisely say where the edge is located.

(c). Describe two filters for computing the image gradient. What is the meaning of the image gradient? What is it used for?

Ans: • Central differences and sobel are two filters used for computing the image gradient.

Central differences :

$$\Delta_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \qquad \Delta_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Sobel filter:

$$\Delta_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad \Delta_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

- An image gradient is a directional change in the intensity or color in an image.
- It is used to extract information from images.

(d). Explain how the sobel filter can be produced from a smoothing and derivative filters.

Ans: In sobel filter, first we apply smoothing and then take derivative.

So,

$$\Delta_x = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}}_{\text{Smoothing filter}} * \underbrace{\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}}_{\text{Derivative filter}} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

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

Note :- Sobel filter is a convolution of smoothing and derivative filter.

(e). Explain how to generate a more accurate derivative filter with an arbitrary σ. Write the elements of a filter for more accurate derivative computation with $\sigma = 2$.

Ans: Accurate derivatives are given by:

$$I_x = I * G'[x] * G[y]$$

Image

1D convolution with horizontal Gaussian derivative

1D convolution with vertical Gaussian

$$I_y = I * G'[y] * G[x]$$

1D convolution with vertical Gaussian derivative

1D convolution with horizontal Gaussian

Here,

$$G(x) = e^{-\frac{x^2}{2\sigma^2}}$$

$$G'(x) = -\frac{\cancel{2}x}{\cancel{2}\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

For $\sigma = 2$,

$$G'(x) = -\frac{x}{4} e^{-\frac{x^2}{8}}$$

$G'(x) \rightarrow$

$x=-2$	$x=-1$	$x=0$	$x=1$	$x=2$
0.3	0.22	0	-0.22	-0.3

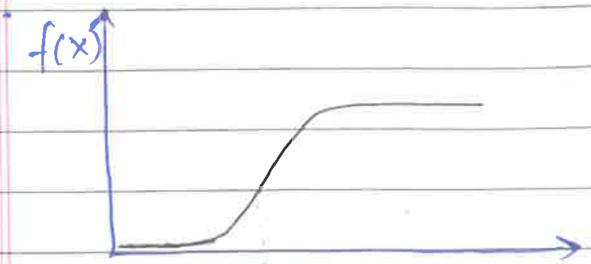
$\rightarrow$ 1D horizontal Gaussian derivative.

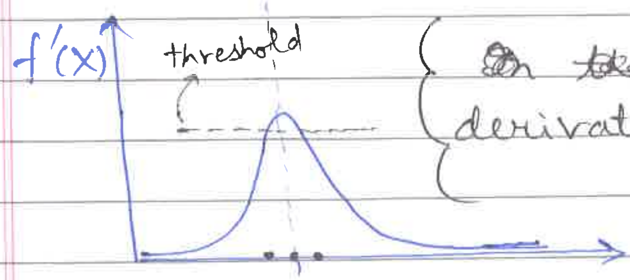
$$G'(y) = \frac{-y}{4} e^{-\left(\frac{y^2}{8}\right)}$$

	$G'(y)$
$y=2$	-0.3
$y=1$	-0.22
$y=0$	0
$y=-1$	0.22
$y=-2$	0.3

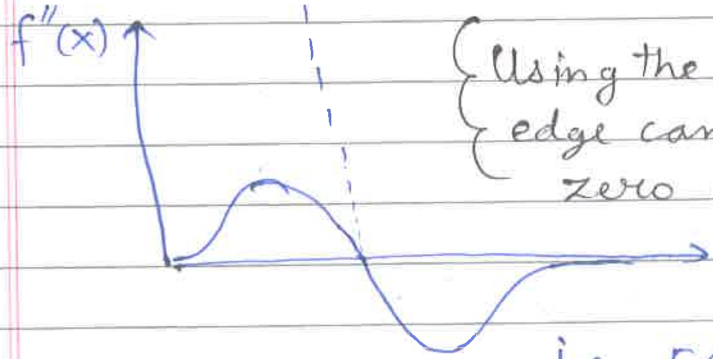
$\rightarrow$ 1D vertical Gaussian derivative (f). Explain how an edge can be localized using the first or second order derivative of the image.

Ans:





~~On~~ ~~the~~ Using the first order derivative edge can be localized by using threshold



Using the second order derivative edge can be localized by using zero crossing

i.e. $E(i,j) = \begin{cases} 0 & \text{if } H(i,j) < 0 \\ 1 & \text{if } H(i,j) \geq 0 \end{cases}$

Edge location

2nd order derivative

(g). Let $\sigma = 1$. Write the Laplacian of Gaussian (LOG) filter using this σ. Explain how to use LOG to detect edges.

Ans: Here,

$$G = e^{-\frac{r^2}{2\sigma^2}} \quad (\text{Here, } r^2 = x^2 + y^2).$$

$$\nabla^2 G = \frac{r^2 - 2\sigma^2}{\sigma^4} \times e^{-\frac{r^2}{2\sigma^2}}$$

$$\nabla^2 G_x = \left(\frac{x^2 - 2\sigma^2}{\sigma^4}\right) e^{-\frac{x^2}{2\sigma^2}} = (x^2-2)\, e^{-x^2/2} \quad \left(\begin{array}{l}\text{Here,}\\ \sigma=1\end{array}\right)$$

$\nabla^2 G_x \longrightarrow$

$x=-2$	$x=-1$	$x=0$	$x=1$	$x=2$
0.27	-0.6	-2	-0.6	0.27

$\rightarrow$ 1D horizontal LOG filter

$$\nabla^2 G_y = \left(\frac{y^2 - 2\sigma^2}{\sigma^4}\right) e^{-\frac{y^2}{2\sigma^2}} = (y^2-2)\, e^{-y^2/2} \quad \{\text{Here, } \sigma=1\}$$

$\rightarrow \nabla^2 G_y$

$y=2$	0.27
$y=1$	-0.6
$y=0$	-2
$y=-1$	-0.6
$y=-2$	0.27

$\rightarrow$ 1D-vertical LOG filter

Edge detection using LOG:

(i) Compute LOG : $H = (\nabla^2 G) * I$

(ii) Threshold : $E(i,j) = \begin{cases} 0 & \text{if } H(i,j) < 0 \\ 1 & \text{if } H(i,j) \geq 0 \end{cases}$

(iii) Mark edges at transitions: $0 \rightarrow 1$
$1 \rightarrow 0$

(Scan left-to-right & top-to-bottom)

(h). Explain the main difference between the canny edge detection algorithm and a standard edge detection that does not use directional derivatives. What is the condition for detecting an edge candidate in Canny?

Ans: Canny edge detection algorithm detects edges at zero crossing of second order directional derivative taken along the gradient.

While a standard edge detection algorithm that does not use directional derivative detects edges at zero crossing of second order derivative or some threshold in first order derivative taken along x and y.

- The condition for detecting an edge candidate in Canny is:

If $|\nabla(I * G)| > \tau$, set edge at maximum of $|\nabla(I * G)|$ along the direction of $\nabla(I * G)$.

or

If $|gradient| > \tau$, detect edges at zero crossing of $\frac{\partial^2}{\partial n^2}(I * G)$.

(i). Eexplain the non-maximum suppression and hysteresis thresholding parts of the Canny algorithm.

Ans: Non-Maximum Suppression

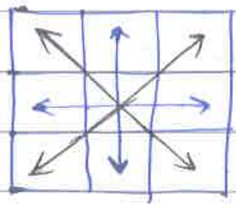
Here, we want local maximum of gradient magnitude in direction of gradient.

$$\nabla(I * G) = (I_x, I_y)$$

$$\theta = \tan^{-1}\left(\frac{I_y}{I_x}\right)$$

$$\theta^* = \text{round}\left(\frac{\theta}{45}\right) * 45 \quad \left\{\begin{array}{l}\text{Angle discretization} \\ \text{Eg:- } 0, 45, 90, 135 \text{ etc.}\end{array}\right\}$$

$$E(i,j) = \begin{cases} 1 & \text{if } \nabla(I * G) \text{ is a local maximum along } \theta^* \\ 0 & \text{otherwise} \end{cases}$$



Compare neighbors after discretization

<u>Hysteresis Thresholding</u>

Use T_H to start tracking and T_L to continue ($T_H > T_L$)

(i). Initialize array at visited pixels
$V(i,j) = 0$.

(ii). Scan image top-to-bottom and left-to-right:
if $(!V(i,j))$ && $(|\nabla I| > T_H)$ then, start tracking an edge.

(iii). Search for additional neighbors in direction orthogonal to ∇I such that

$$|\nabla I| > T_L$$