UNIT-I

DIGITAL IMAGE FUNDAMENTALS

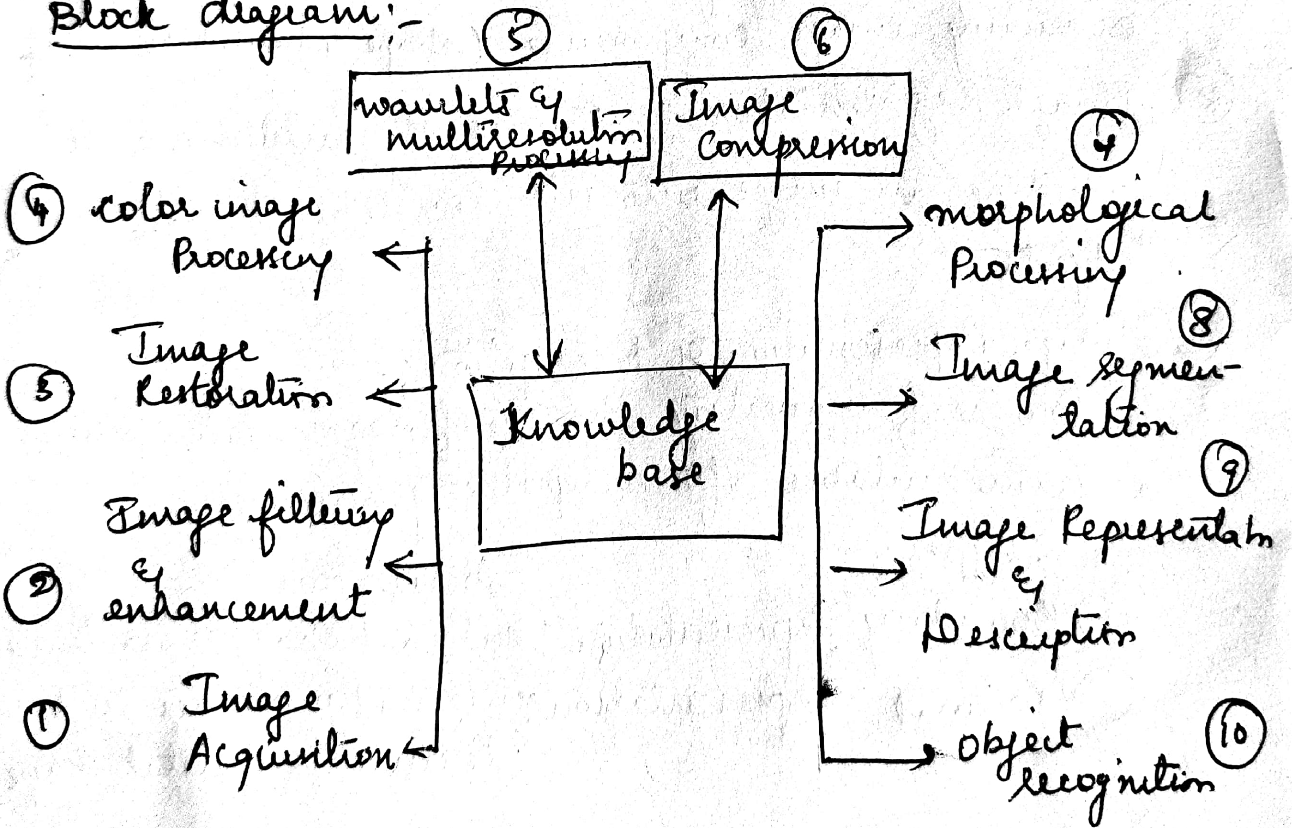
* An image contains descriptive information about the object it represents.
* An image is defined as 2 dimensional fn, $f(x,y)$ that carries some information. where x & y known as spatial a plane co-ordinates.

Image Processing:-
The Process of analyzing & manipulating images using a computer.

I Steps in Digital image Processing:-

* 2 categories
  1) Methods whose i/p & o/p are images
  2) Methods whose i/p are images but o/p's are attributes.

Block diagram:-



(1) Image Acquisition: Process of capturing or generating digital image using image sensors. (Preprocessing)

(2) Image enhancement: Process of manipulating an

image so that the result is more suitable than the original image for specific application (subjective)

③ Image Restoration :- Process of improving the appearance of an image. (objective)

④ color image Processing :- Process an image considering its color as one of the important attribute in addition.

⑤ wavelets & multiresolution Processing :

* It is used for

a) Image data compression

b) Pyramidal representation – subdividing image into smaller region

⑥ Image compression :- Process of reducing the storage required to save an image or reducing BW to Transmit an image a) Lossless Compression b) Lossy compression

⑦ Morphological Processing :- Deals with the tools for extracting image components. (shape & boundary)

⑧ Image segmentation :- Process of partitioning or dividing an image into its constituent parts or objects

⑨ Image Representation & Description :-

Process of converting the o/p of segmentation process into a form suitable for computer processing.

2 types.

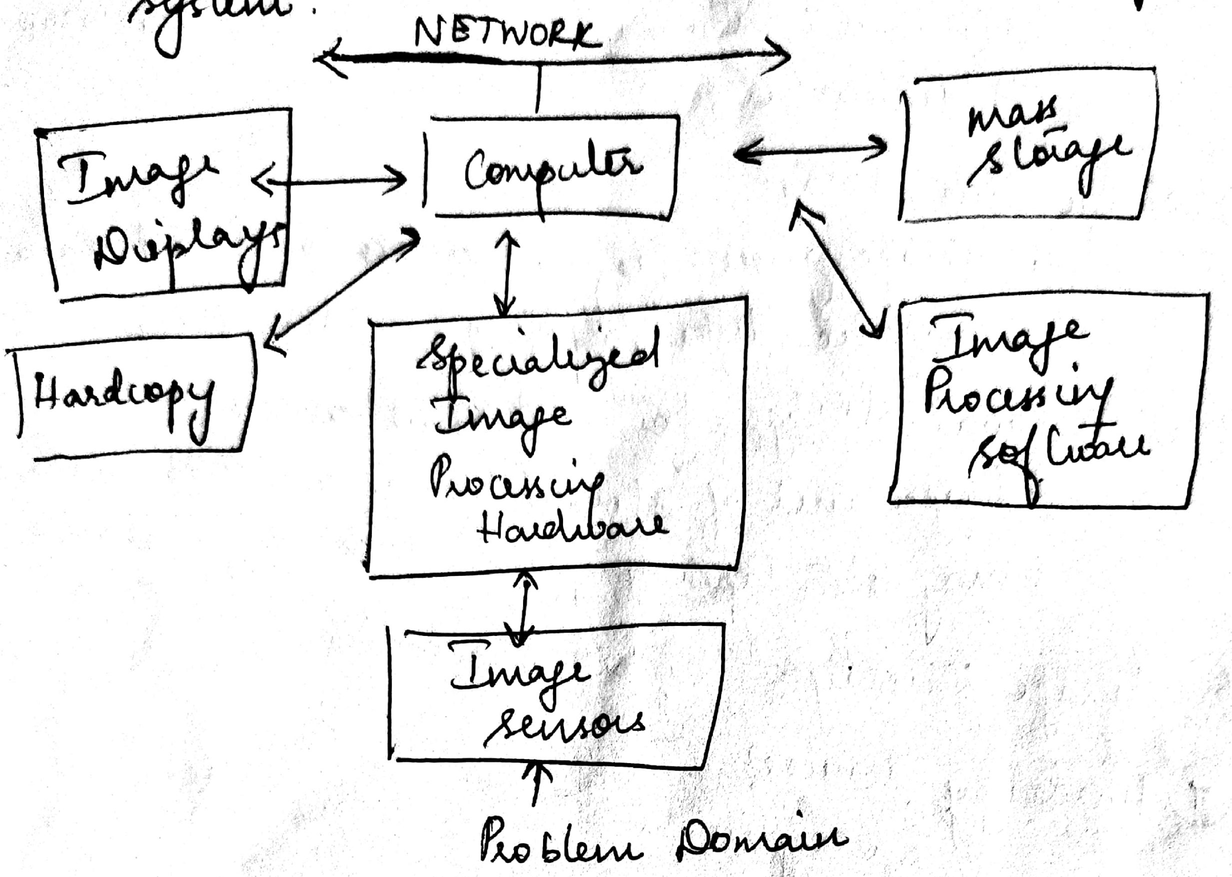
① Boundary Representation : Focuses on external shape characteristics

② Regional Representation : internal properties such as texture or skeletal shape.

⑩ Image Recognition :-

Process assigns a label or name of an object identified from an image.

II Components (or) Elements of image Processing system.



1. Image sensors:- 2 elements:
   i) Physical device – sensitive to the energy radiated by energy eg: digital video camera.
   ii) Digitizer: convert o/p of physical sensing device into digital form.

2. Specialized Image processing hardware:-
   * H/w consist of digitizer & some H/w to perform other basic operations.

3. Computer:- It is general purpose system
4. Image Processing software:- write code
5. Mass storage:- Large amount of storage space, mass storage capability is important.

→ Types:-

* 3 categories of digital storage.

(a) (1) Short-term storage:
a) Computer memory
b) Frame buffer

1024 X 1024 pixels

(b) (2) On-line storage
* storage gives frequent access to stored data
* magnetic disks & optical-media storage

(c) (3) Archival storage for infrequent access.
* Large amt of storage space & stored data
* magnetic tapes & optical disk → jukeboxes

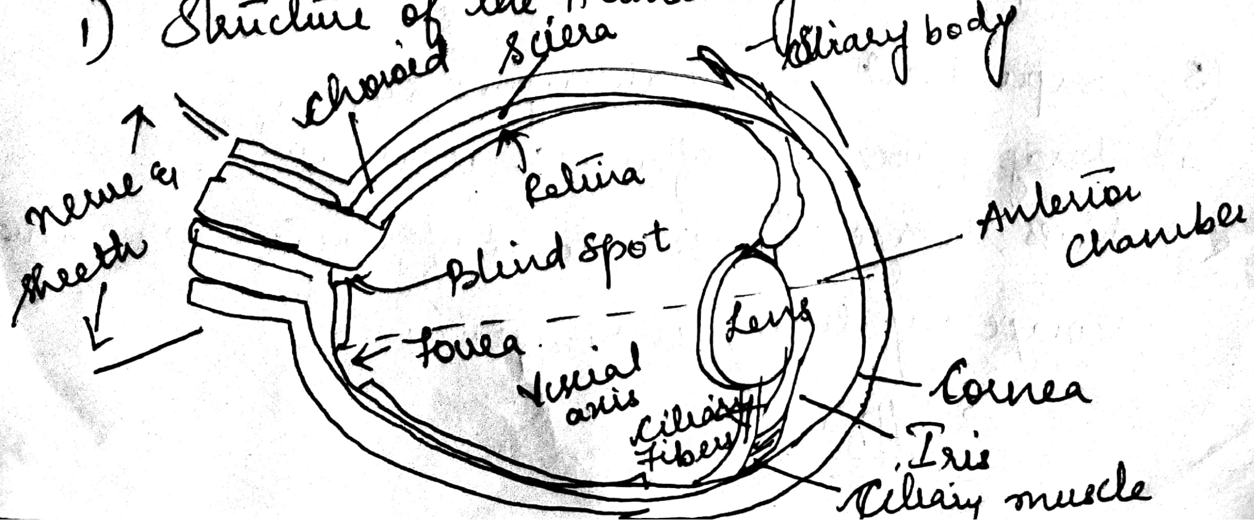
(6) Image Displays:- color TV monitors

(7) Hardcopy devices:
i) Laser printers
ii) Film cameras
iii) Heat-sensitive devices
iv) Inkjet units
v) Digital unit like optical & CD-ROM disk.

(8) Network:
Image processing application need large amount of data.

III Elements of Visual Perception :-

1) Structure of the Human eye



① a) The cornea & sclera outer cover:-

* Tough, Transparent tissue
* Sclera is an opague

② b) The choroid :-

* N/w of blood vessels which are major nutrition sources.

divided into 2 anterior extreme,

i) The ciliary Body

ii) The Iris Diaphragm.

Iris Diaphragm :

* central opening of iris known as pupil, diameter varies 2 to 8 mm.

Lens :-

* layers of fibrous cells.
* Contains 60% to 70% fat & more protein.

Cataracts :

* Excessive clouding of lens happen in extreme cases known as cataracts.

③ The Retina :-

* Inner most membrane of eye.

a) Fovea; central portion of retina.

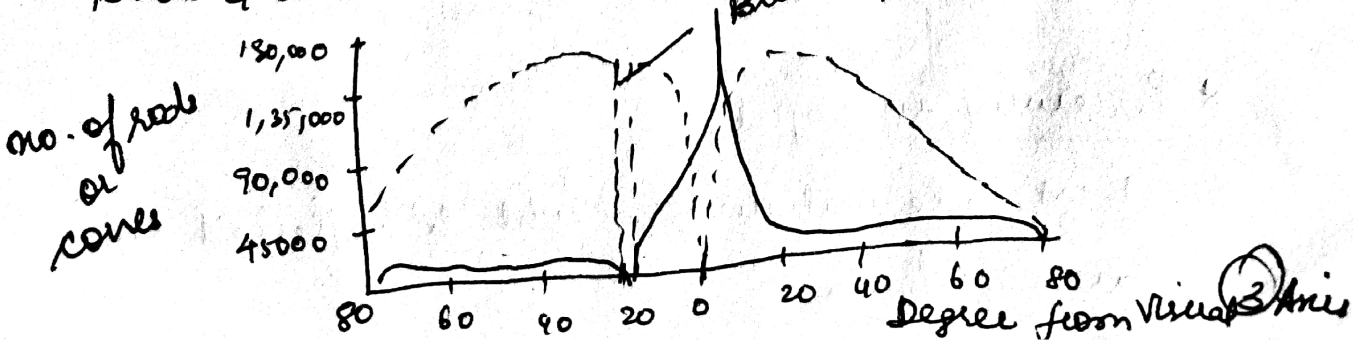
b) light receptors :

* provide pattern vision

i) Cons — 6 to 7 million (Photopic)

ii) Rods — 75 to 159 million

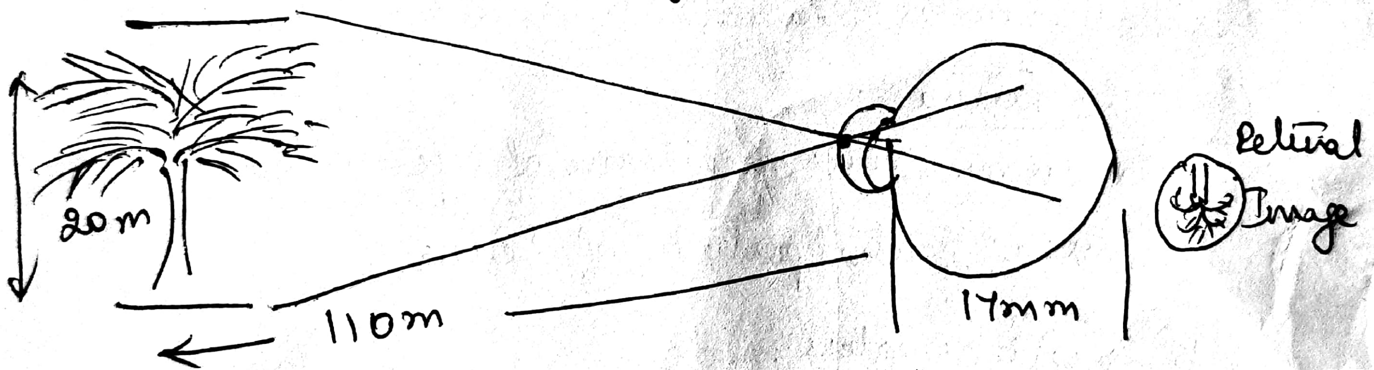
Rods & cons distribution



Blind spot: The area in which is no presence of light receptors called blind spot.

2) Image Formation in the Eye:-

* <u>Role of length</u>: The lens of the eye is flexible, whereas ordinary optical lens is not.
* radius of curvature of anterior surface of lens > radius of posterior surface.
* The lens in fiber follows certain things.
  i) Focus distance objects (>3m) lens made flatted.
  ii) Focus nearer object, muscles allow lens to become thicker & most strongly refractive.
* <u>Focal length</u>:-
  Distance b/w center of lens & Retina called focal length
* <u>To calculate Retinal image size</u>:-



→ object distance > 3m.
→ Focal length = 17mm.

$$= \frac{20}{110} = \frac{h}{17\times10^{-3}}$$

size of Retinal $\boxed{h = 3.09mm}$

* <u>Perceiving an object</u>:-
  i) 1st → Retinal image height 'h' reflected in area of fovea.

ii) Perception takes place by the relative excitation of light.

iii) Receptors transform the relative excitation of light receptors.

iii) Receptors transform the radiant energy into electrical impulses.

iv) These electrical impulses are decoded by brain

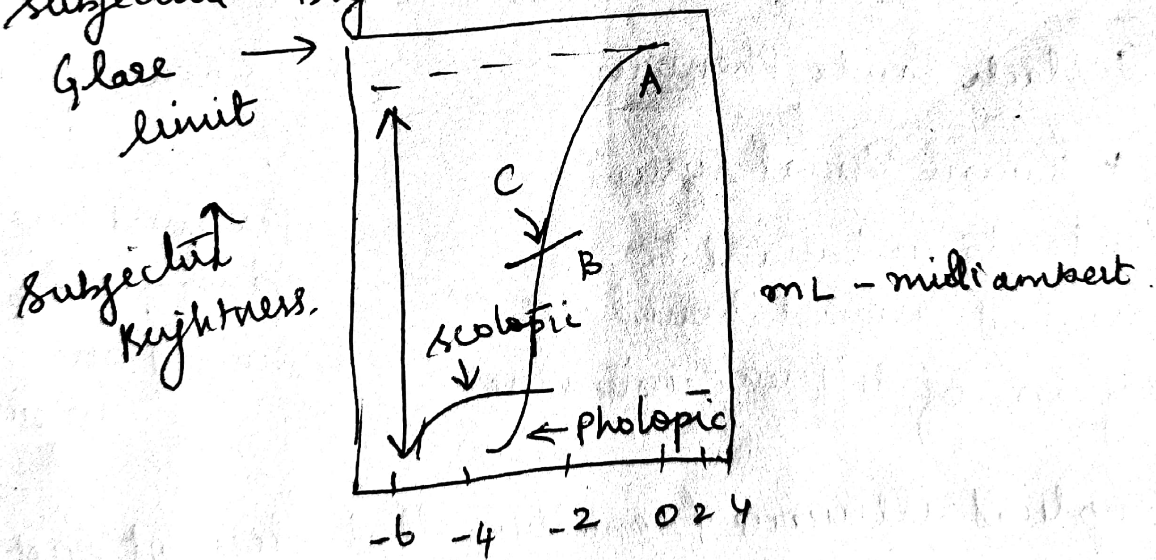
3) Brightness Adaptation :-

Large Variation is accomplished by changing its overall sensitivity. The Phenomenon is known as 'Brightness adaptation'.

i) Brightness Adaptation level:
current sensitivity level of visual system for given set of condition called brightness adaptation level.

ii) Adaptable Range: order 10^{10}.

iii) Subjective Brightness: Human Visual system known a Subjective Brightness.



A → Represents range of intensities

(B-C) → range of subjective

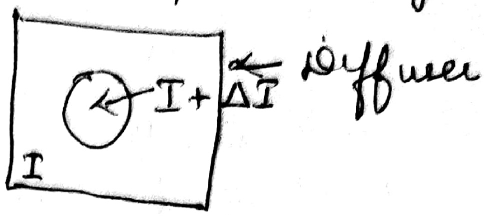
Double Branches → scotopic to phototopic vision.

Scanned with OKEN Scanner

4) Brightness Discrimination

* Total range of intensity levels can be discriminated simultaneously its smaller total adaptation range.

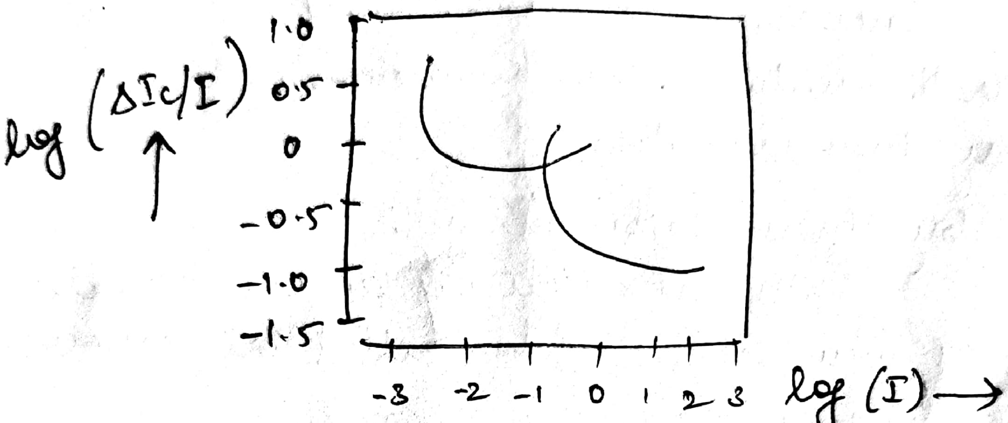
i) Experimental setup :-



Case i) :-

Field = $I + \Delta I$.

weber ratio = $\frac{\Delta I_c}{I}$ – Increment of Illumination 50% of time



Case ii) Observer can see 24 different intensity change.

ii) Perceived Brightness & Intensity :-

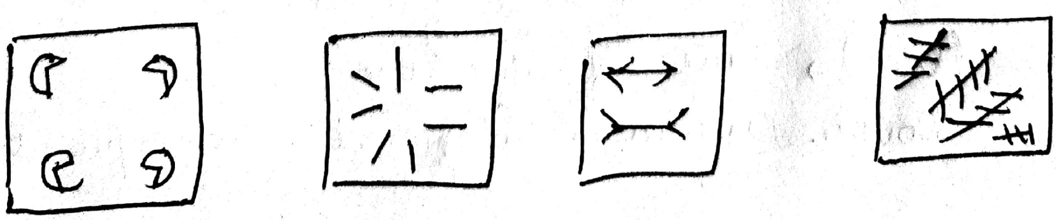
i) Mach Bands Phenoma

* human Visual system tends to understand to overshoot around boundary region of different intensities

ii) Simultaneous Contrast.

Perceived brightness does not depend simply on intensity.

iii) optical illusions (something that does not exist)



IV Image sensing & Acquistion :-

* Process of capturing or generating digital images using imaging sensors.
* Combination of 2 things in images.
  - (i) Illumination source
  - (ii) Reflection or absorption of elements

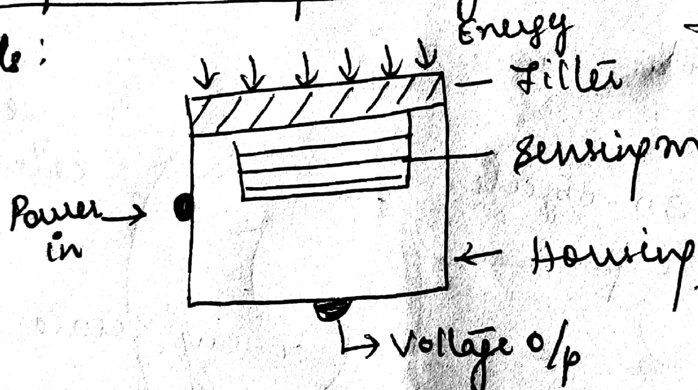
Types of sensors:

(i) single imaging sensor
(ii) Line sensor
(iii) Array sensor

Imaging process.

1) Image Acquistion using a single sensor

Components:

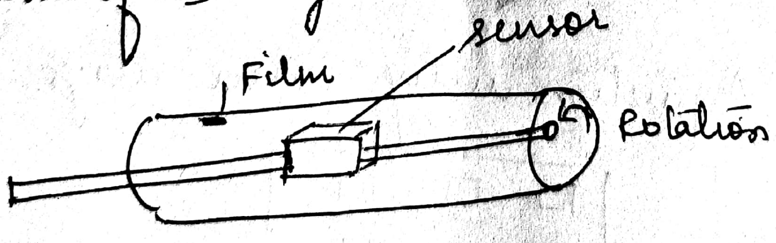


eg:- photodiode

Sensor working Mtd.

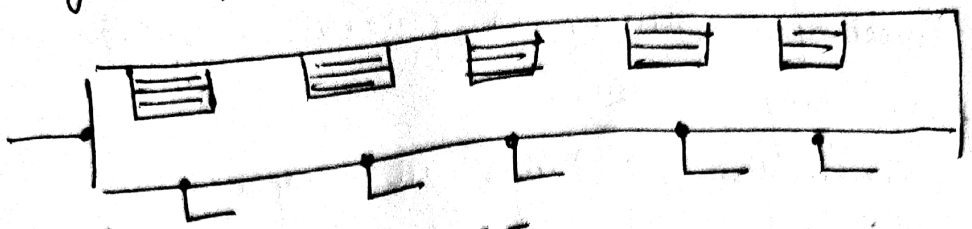
1) Incoming energy is Transformed into a Vge by the combination of i/p electrical power
2) Sensor material is responsive to particular type of energy being detected
3) o/p Vge waveform is response of sensor & digital Qty is obtained from each sensor by digitizing its response.

Generation of 2D images



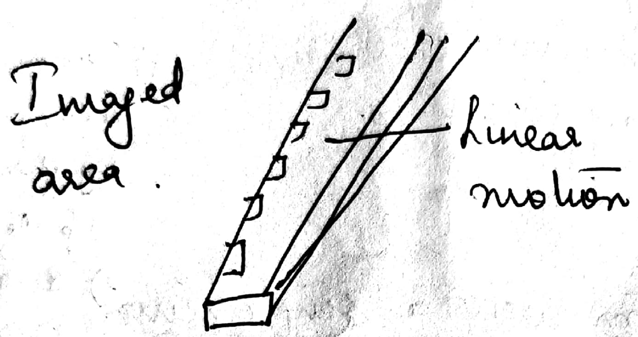
other arrangement:

2) Image Acquisition using sensor strips.

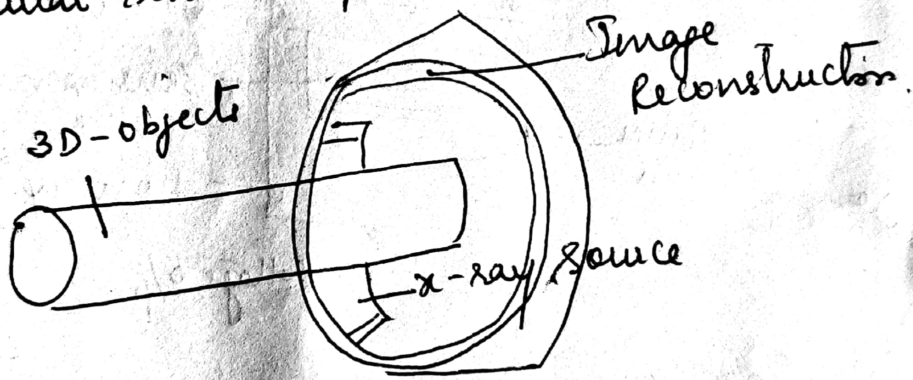


i) Linear sensor strip
ii) Circular sensor strip

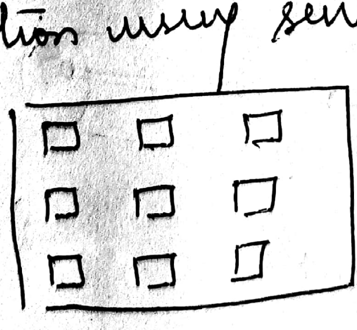
i) Linear sensor strip



ii) Circular sensor strip



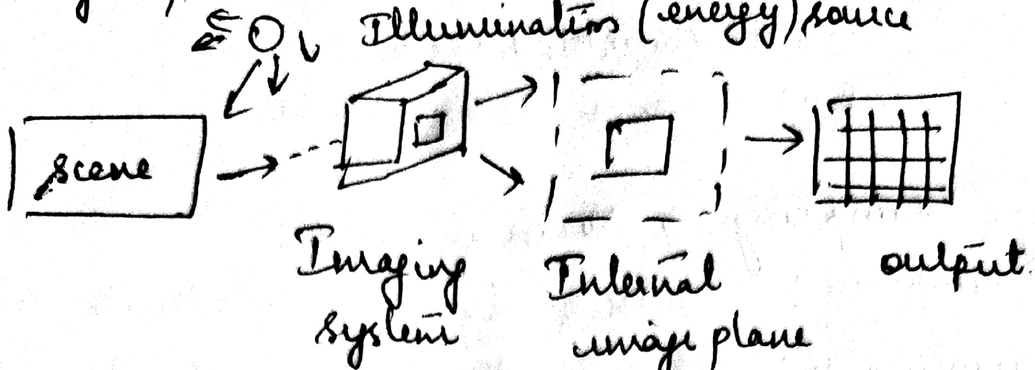
3) Image Acquisition using sensor arrays:



* sensor array used in digital cameras & other light sensing instruments is known as 'CCD array'

Scanned with OKEN Scanner

i) Image Generation



ii) Image formation model:-

* 2 dimensional image is denoted by $f(x, y)$ — fn.

ie; $0 < f(x, y) < \infty$

Components:

i) Illumination $i(x, y)$: amount of source illumination incident on scene.

ii) Reflectance $r(x, y)$: amt of illumination reflected in scene.

$$f(x, y) = i(x, y) \cdot r(x, y)$$

$0 < i(x, y) < \infty$

$0 < r(x, y) < 1$ (Represents reflectance component is bounded by total absorption)

* Intensity or gray level of monochrome image

$$l = f(x_0, y_0)$$

Range l is $L_{min} \leq l \leq L_{max}$

In practice $L_{min} = i_{min} \cdot r_{min}$

$L_{max} = i_{max} \cdot r_{max}$

$[L_{min}, L_{max}]$ called gray scale or intensity scale

Range of $l = [L_{min}, L_{max}] = [0, L-1]$

$l = 0$ – considered black

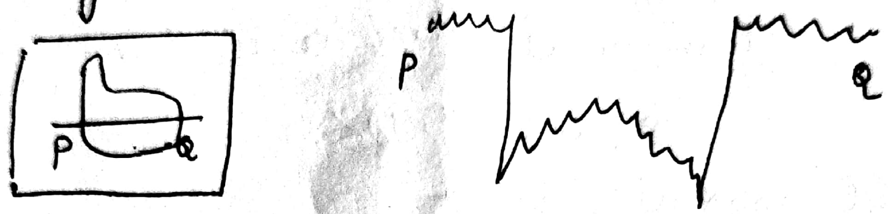
$l = L-1$ – considered white

v) Image sampling & Quantization :-

Sampling : Process of digitizing coordinate values

Quantization : Process of digitizing amplitude values.

i) Generating a Digital Image:



continuous image

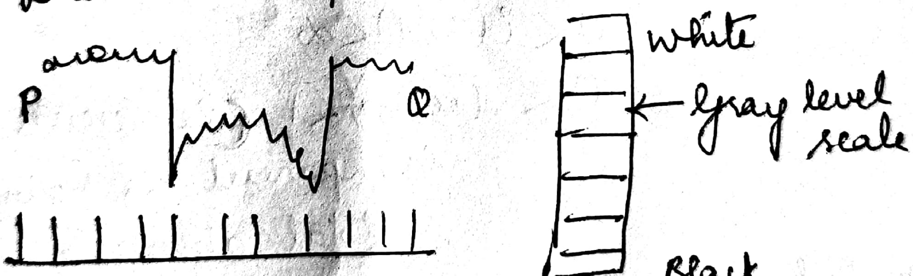
gray level point

sampling & Quantization:

* equally spaced samples taken along line PQ

samples

* Discrete samples are Quantized



sampling

Methods of Image Acquisition :-

* methods differ in sensor arrangement

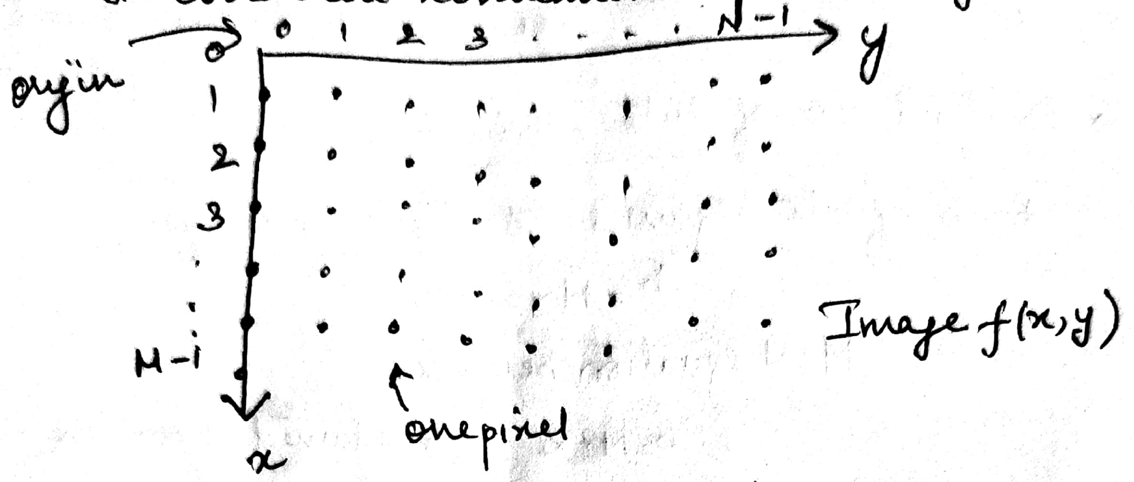
① Image used in single sensing elements combined with mechanical motion

② Sensing strip for image acquisition

③ Sensing array for acquisition : no. of array establishes sampling limitations.

2) Digital Image Representation:

* 2 ways represent digital images
* $f(x,y)$ result digital image M rows, N columns
* coordinate conventions represent digital images



* Value of origin $(x,y) = (0,0)$
* 2nd coordinate $(x,y) = (0,1)$ 1st row
* 2nd value $(x,y) = (1,0)$ 1st column.

Method 1:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & & & \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

* matrix array called image element, picture element, pixel.

Method 2: M×N digital image,

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & & & \\ a_{M-1,0} & a_{M-1,0} & \cdots & a_{M-1,N-1} \end{bmatrix}$$

$$a_{ij} = f(x=i, y=j) = f(i,j)$$

→ To find the no. of gray levels:-

* Integral power of 2

$$L = 2^k$$

* 2^K gray level is K-bit image.

→ **Dynamic Range** :-

* gray scale $[0, L-1]$ called dynamic range of image
* Dynamic range is high → contrast
* " low → dull washed out gray look.

→ **To find no. of bits required** :-

* no. of bits required to digitized image,

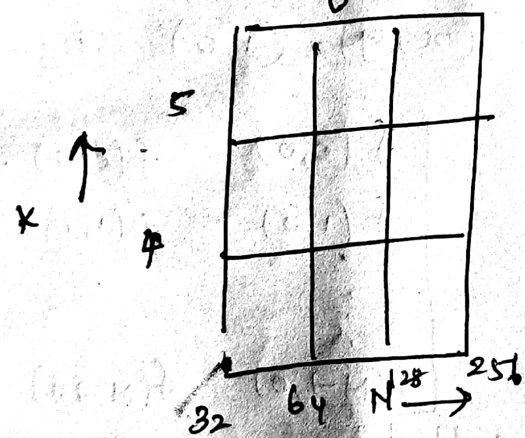
$$b = M \times N \times K$$

M=N equation becomes

$$b = N^2 K \quad , K \to \text{obtained from eqn.}$$

Isopreference curves :-

Values of N & K, a set of curves generated



3) **Spatial Resolution** :-

* A smallest no. of discernible ie; recognizable line pairs per unit distance.
* image of size M×N with L gray levels.
* Line pair per unit distance $= \frac{1}{2W}$

False contouring :

4) Gray-level Resolution:-

* Defined as smallest discernible change in gray level.
* Gray level depends on human perception "subjective Process"

* L-level digital image size $M \times N = L$ levels.

5) Aliasing Effect:

a) Band-limited functions:

* It represents in terms of sine & cosine of various frequencies. The sine & cosine component have highest frequency determines 'highest frequency content of function.

b) Sampling rate:

no. of samples taken per unit distance.

c) Shannon's sampling Theorem:

It states that if fn is sampled at a rate greater or equal to twice its higher frequency, original fn completely recovered from samples.

$$f_s \geq 2f_{max}$$

f_s — sampling frequency

f_{max} — highest i/p frequency

Aliasing:- A fn is undersampled if sample freq is very low, so it cannot satisfy the shannon theorem. In Additional frequency components introduced into the sampled fn which sampled images. This effect is called aliasing.

6) Moire Patterns:

Fn of finite duration can be sampled over finite interval without violate Shannon

Shannon sampling theorem.

7) Zooming of Digital images:

Zooming indicates oversampling.

steps: 2 steps a) Creation of new pixel location

b) assignment of gray level location.

Methods: 1) Nearest neighbor interpolation.

2) Pixel replication

3) Bilinear interpolation

Pixel replication → duplication

Bilinear interpolation → 4 nearest neighbour pt.

(x', y') co-ordinates of points, $v(x', y')$ grey level assigned to that point.

$$v(x', y') = ax' + by' + cx'y' + d.$$

8) Shrinking of Digital images:-

Viewed by under sampling, applied to image of digitizing

Method I:

* Row - column deletion method.
* Pixel replication method in zooming
* Required size of image.

Method II:-

* grid analogy for zooming.

i) Keep an imaginary grid of required size over original image.

ii) Expand the grid to fit over original image

iii) Gray level assignment
iv) Shrink the grid back to its original specified size which gives shrunk image of required size.

vi Relationships between pixels:-

* Image $f(x, y)$ are explained.

1) Neighbors of a pixel.

3 types

1) 4- Neighbors, $N_4(P)$
2) Diagonal Neighbors $N_D(P)$
3) 8 - neighbors $N_8(P)$

① 4- Neighbors $N_4(P)$

$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$

$x-1, y-1$	$x-1, y$	$x-1, y+1$
$x, y-1$	x, y	$x, y+1$
$x+1, y-1$	$x+1, y$	$x+1, y+1$

② Diagonal Neighbors, $N_D(P)$:

$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$

③ 8- Neighbors, $N_8(P)$

* 4 neighbors called 8- neighbors.

2) Adjacency:

Values 0 to 255

① 4 - Adjacency:

P & q values from $\{V\}$, $N_4(P)$

② 8- Adjacency:

* Two pixels p & q with values {V} are 8-adjacent

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \Rightarrow \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Here, $\{V\} = 1$

③ m-Adjacency (or) mixed Adjacency :-

2 pixels

a) q is in $N_4(p)$ or

b) q is in $N_D(p)$ set $\{N_4(p) \cap N_4(q)\}$

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \Rightarrow \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

$\{V\} = 1$

3) Path

* known as digital path or curve.
* pixel p with coordinates (x, y) with pixel q

$(x_0, y_0), (x_1, y_1), \ldots (x_n, y_n)$

where $(x_0, y_0) = (x, y)$

$(x_n, y_n) = (s, t)$.

Path length: no. of pixels present in a path

closed path: $(x_0, y_0) = (x_n, y_n)$

Types: ① 4-path

② 8-path

③ m-path.

4) Connectivity:

condition:- 2 pixels said to connected,

a) They are neighbours

b) their gray levels satisfy a specified similarity criterion.

Def: 'S' represent a subset of pixels in an image
connected set:
For pixel p, the set pixels that are connected p called connected component of S.

5) Region:-
Let R represent a subset of pixels in an image.

6) Boundary:-
* Known as border (or) contour.
* Boundary of a finite region forms a closed path, it is a global concept.

7) Edge:-
* gray level or intensity discontinuities
Edge segments.

8) Distance measures
Distance b/w different pixels.
conditions
a) $D(p,q) \geq 0$, $[D(p,q) = 0$ if $p=q]$
b) $D(p,q) = D(q,p)$ and
c) $D(p,z) \leq D(p,q) + D(q,z)$

Types:
① Euclidean distance $= D_e(p,q) = \sqrt{(x-s)^2+(y-t)^2}$
② city - Block (or) D_4 distance $= D_4(p,q) = |x-s|+|y-t|$
③ chessboard (or) D_8 distance
④ D_m distance.

9) Image operations on a pixel Basis :-

* Image are represented in form of matrices
* Dividing one image by another.
* Arithmetic operations.

Addition = p+q
subtraction = p-q
multiplication = p * q (or) p.q (or) p×q
Division = p ÷ q.

* Principal of logic operations.

VII color Image Fundamentals :-

* 2 broad categories :

i) Pseudo color processing.
   * monochrome intensity or range of intensities

ii) Full color processing :
   * Full color sensors.

1) characterization of light -

i) Achromatic (or) monochromatic light :
   * " is light seen on a black & white television set.
   * characterized by one attribute called intensity or amount or graylevel

ii) chromatic light
   3 basic Quantities ;

a) Radiance - Total amount of energy flow from light source.

b) luminance - measures amount of energy Perceived from a light source by observer

c) Brightness : Key factor of describing color sensation which is an achromatic notion of intensity.

(11)

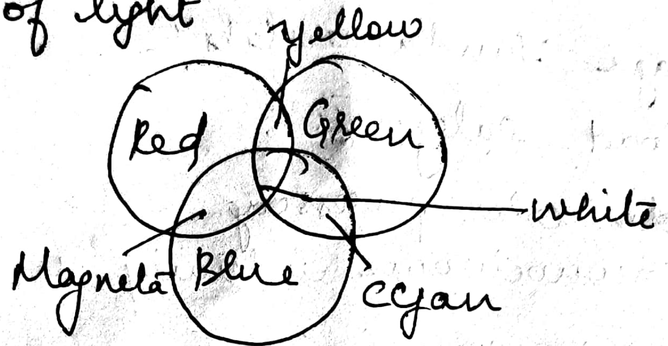
iii) colors:

* Brightness
* Hue
* Saturation

2) Primary & secondary colours

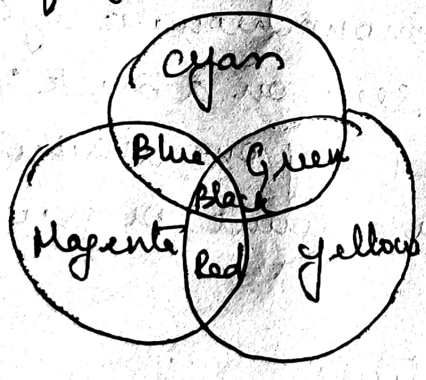
Primary colors — Red (R), Green (G), Blue (B)

i) mixtures of light



additive primaries

ii) Mixtures of pigments



subtractive Primaries.

3) Tri chromatic coefficients:

* Red, green, blue needed any particular color called Tristimulus Values.

$$r = \frac{R}{R+G+B}$$

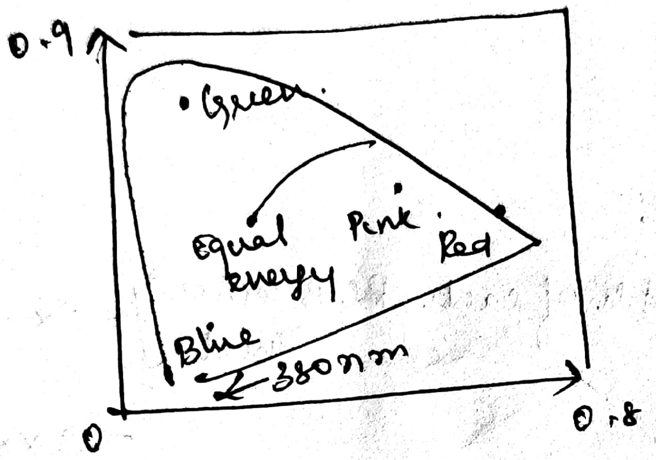
$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

eqn as, $r+g+b = 1$.

4) chromaticity Diagram:-

* Value of b (blue) for any value of r & g.



uses:-

* color mixing as a straight
* center pt of chart to any pt on boundary will give all shades.
* Range of color acquired 3 given colors.

VIII) color models:-

* A specification of a coordinate system & subspace within that system.
* Color model called color space or color system.

classification

i) RGB (Red, Green, Blue)

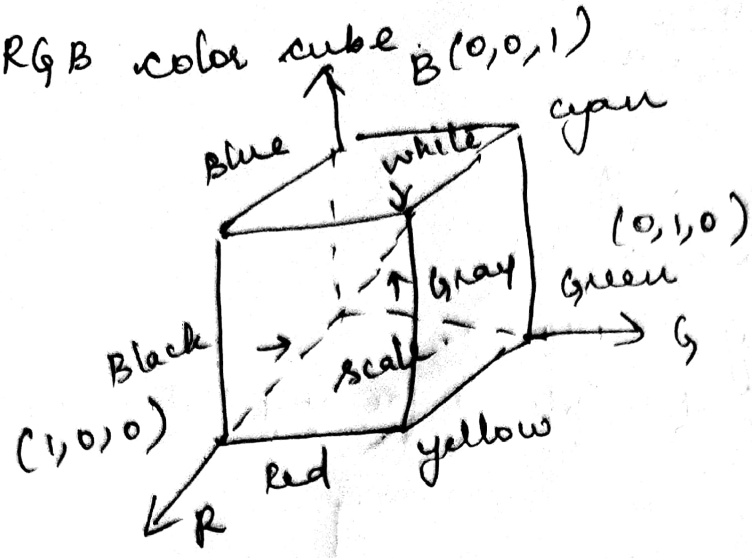
ii) CMY (cyan, magneta, yellow) & CMYB Black

iii) HSI (Hue, saturation, Intensity) model

RGB:-

* Color appears in Primary spectral components

① RGB color cube



Pixel Depth:

Pixel Depth of each RGB color Pixel $\}$ = 3x no. of bits / plane

= 3 X 8 = 24

Full color image:-

Total no. of colors = $(2^8)^3$ = 16,777,216

② safe RGB colors:-

* application of few hundred or fewer colors 256 colors.
* set of all systems safe colors.

→ standard safe colours;

→ component values of safe colors;

∴ Total no. of possible values = 6 × 6 × 6 = 216

→ Hexadecimal Representation :-

Decimal	Hexadecimal
0	00

51	33
102	66
153	99
204	CC
255	FF

(3) Applications.
i) color monitors
ii) color video cameras.

(4) Advantages :-
* Image code generation.
* other models such as CMY straight forward.

(5) Disadvantages:
* Combining 3 primary images.
* not suitable for describing colors in a way which is pratical for human interpretation.

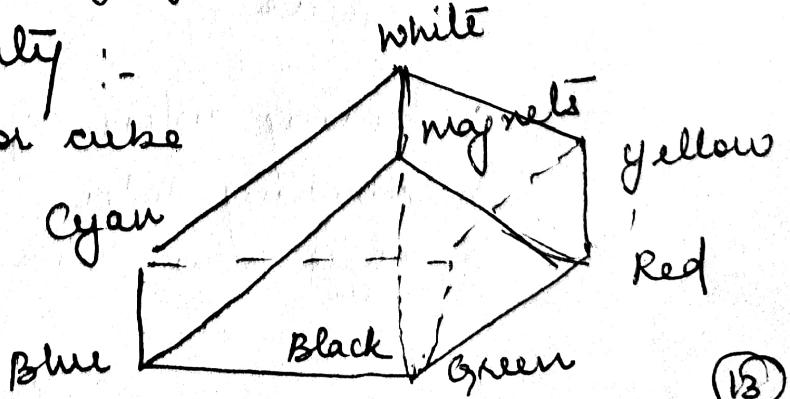
2) The HSI (Hue saturation Intensity) color model:-

Hue: a color that describe a pure color

Saturation: a measure of degree to which a pure color is diluted by white light

Intensity: monochromatic images also called gray level.

(1) To find intensity :-
* RGB color cube



Intensity axis: black Vertex (0,0,0) & white Vertex (1,1,1)

Determining Intensity component:

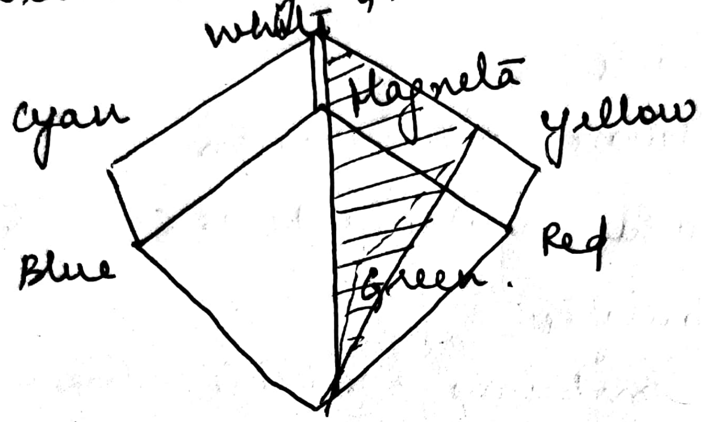
* Intensity Value range [0,1]

② To find saturation:

* intensity axis are gray means saturation.

③ To find Hue:

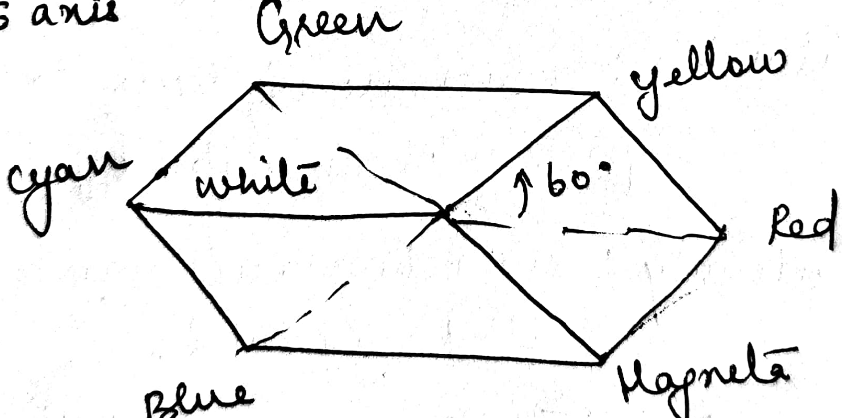
* Determine by RGB color cube.



④ HSI color space:

* A vertical intensity axis &
* the locus of color points that lie on planes

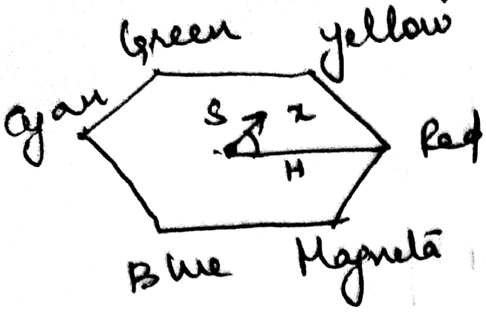
⊥re to axis

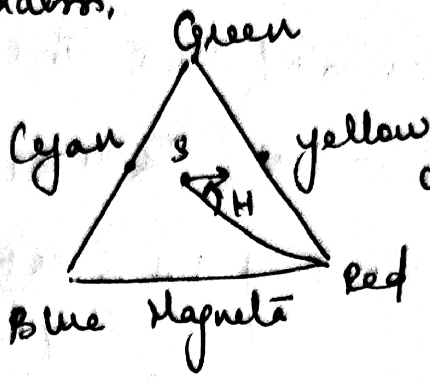


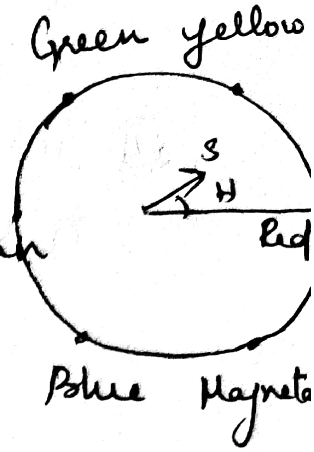
* Primary colors separated by 120°
* secondary color by 120°
* angle b/w secondaries & primaries is 60°

→ Representation of Hue:

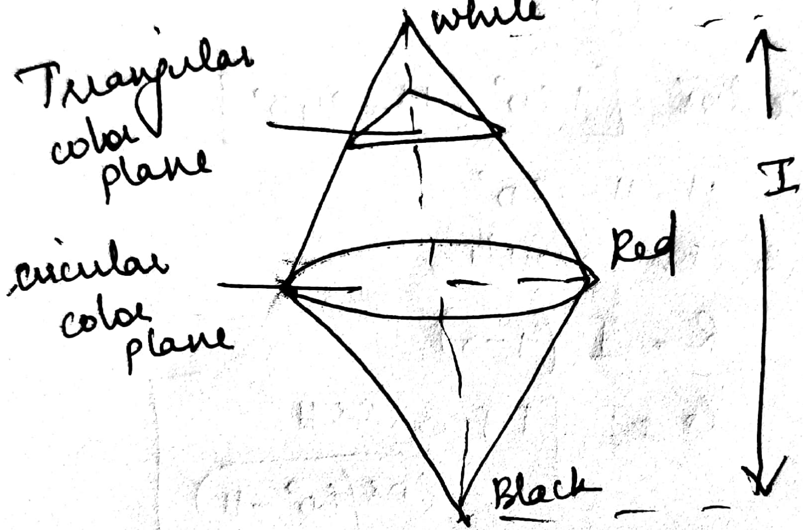
→ Representation of saturation:







Triangular & circular color planes in HSI models.



3) Image format conversion

(I) Conversion of RGB to HSI.

a) Hue component, H of each RGB pixel as,

$$H = \begin{cases} \theta & , B \leq G \\ 360 - \theta & , B > G \end{cases}$$

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{1/2}} \right\}$$

b) Saturation component, S

$$S = 1 - \left\{ \frac{3}{(R+G+B)} \cdot (R, G, B)_{min} \right\}$$

c) Intensity component I,

$$I = \frac{1}{3}(R+G+B)$$

(2) conversion of HSI to RGB:

(a) RG sector $[0° \leq H < 120°]$

$$R = I\left[1 + \frac{S\cos H}{\cos(60° - H)}\right]$$

$$G = 1 - (R+B)$$

$$B = I(1-S)$$

(b) GB sector $[120° \leq H < 240°]$

$$H = H - 120°$$

RGB components,

$$R = I(1-S)$$

$$G = I\left[1 + \frac{S\cos H}{\cos(60° - H)}\right]$$

$$B = I - (R+G)$$

(c) BR sector $[240° \leq H < 360°]$

$$H = H - 240°$$

RGB component

$$R = 1 - (G+B)$$

$$G = I(1-S)$$

$$B = 1\left[1 + \frac{S\cos H}{\cos(60° - H)}\right]$$

(3) Modifying RGB images:-

(a) To change the Hue or individual color:

(b) To change the saturation or Purity of a color.

(c) To change Average intensity.

IX Two Dimensional mathematical Preliminaries:-

1) Array Versus Matrix operations.

array product

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

matrix product of 2 images,

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22}. \end{bmatrix}$$

2) Linear Versus non-linear operations:

* H produces an o/p image $g(x,y)$, i/p image $f(x,y)$

$$H[f(x,y)] = g(x,y)$$

* Property of additivity & homogeneity,

$$H[a_i f_i(x,y) + a_j f_j(x,y)] = a_i H[f_i(x,y)] + a_j H[f_j(x,y)]$$

$$= a_i g_i(x,y) + a_j g_j(x,y)$$

i) Additivity Property,

$$H[f_i(x,y) + f_j(x,y)] = H[f_i(x,y)] + H[f_j(x,y)]$$

Homogeneity property:

$$H[a f(x,y)] = a H[f(x,y)]$$
$$= a g(x,y)$$

3) Arithmetic operations:-

* Basic operations are,

$$S(x,y) = f(x,y) + g(x,y)$$
$$d(x,y) = f(x,y) - g(x,y)$$
$$P(x,y) = f(x,y) \times g(x,y)$$
$$q(x,y) = f(x,y) \div g(x,y)$$

4) Logical operations:-

* Basic are,

$C = (A)$ OR (B) – union

$D = (A)$ AND (B) – Intersection

$E =$ NOT (A) – Complement

* 2D Transforms – DFT, DCT:-

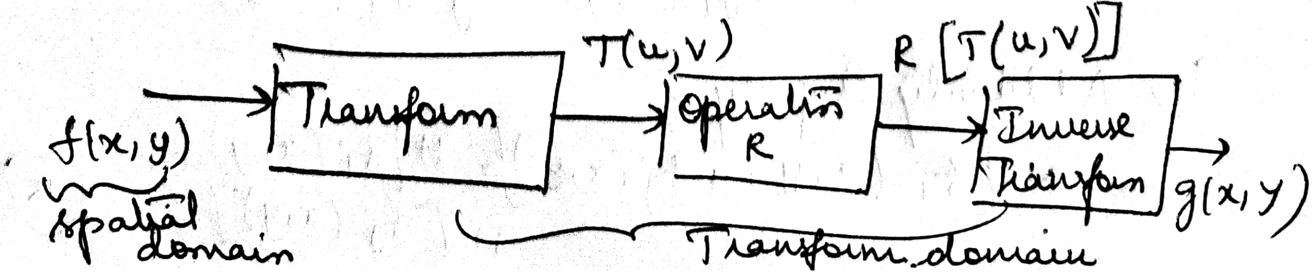
* It works by spatial domain,
* 2-D linear transforms, denoted $T(u,v)$ can be as,

$$T(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\, r(x,y,u,v)$$

I/p image

fwd Transformation Kernel

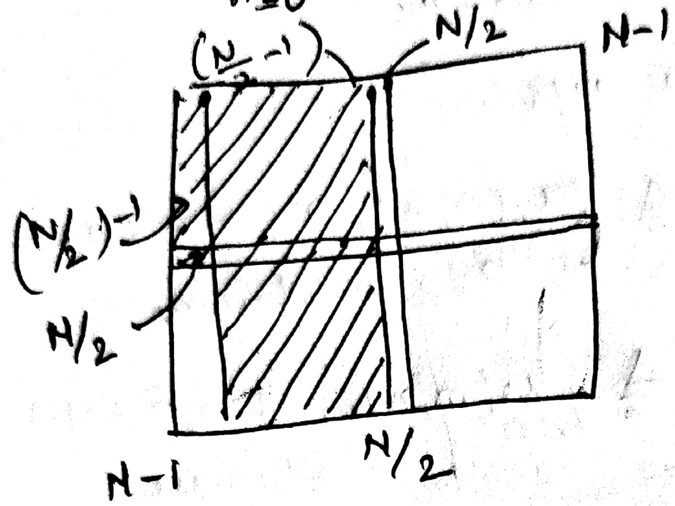
$T(u,v) \rightarrow$ forward transform of $f(x,y)$.



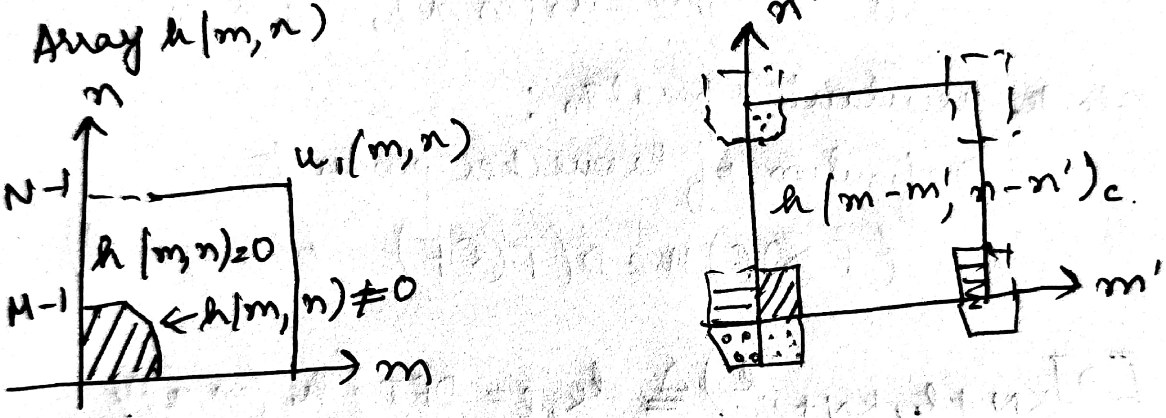
1) 2-dimensional DFT,

* Two dimensional DFT as, $\{u(m,n)\}$

$$V(k,l) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} u(m,n) W_N^{km} W_N^{ln}, \quad 0 \le k, l \le N-1$$



Array $h(m,n)$



* N×N region with $u_1(m,n)$.

$$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h(m-m', n-n')_c W_N^{(mk+nl)}$$

$$= W_N^{(m'k+n'l)} \sum_{i=-m'}^{N-1-m'} \sum_{j=-n'}^{N-1-n'} h(i,j)_c W_N^{(ik+jl)}$$

$$= W_N^{(m'k+n'l)} \text{DFT}\{h(m,n)\}_N$$

DFT of both sides,

$DFT\{u_2(m,n)\}_N = DFT\{h(m,n)\}_N \; DFT\{u_1(m,n)\}_N$

Calculating 2 dimensional convolution,

$$x_3(m,n) = \sum_{m'=0}^{M-1} \sum_{n'=0}^{M-1} x_2(m-m', n-n')\, x_1(m',n')$$

define $N \times N$ arrays

$$\bar{h}(m,n) \triangleq \begin{cases} x_2(m,n), & 0 \le m, n \le M-1 \\ 0, & \text{otherwise} \end{cases}$$

$$u_1(m,n) \triangleq \begin{cases} x_1(m,n), & 0 \le m, n \le M-1 \\ 0 & \text{otherwise} \end{cases}$$

From fig $x_3(m,n) = u_2(m,n)$, $0 \le m, n \le 2M-2$

Block circulant operation:

Definition of Kronecker product

$$(F \otimes F) H = D (F \otimes F) \quad D(\text{diagonal})$$

$$[D]_{kN+l,\, kN+l} \triangleq d_{k,l} = DFT\{h(m,n)\}_N, \quad 0 \le k, l \le N-1$$

Block Toeplitz operations:

* doubly block circulant operation.

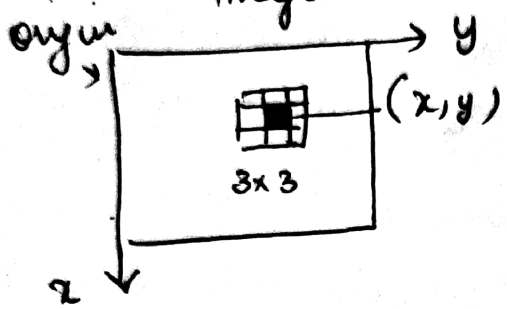
UNIT - II

IMAGE ENHANCEMENT

I SPATIAL DOMAIN METHODS:

* spatial domain process are denoted by expression $g(x,y) = T[f(x,y)]$ i/p image

Processed image — operator on f.



II Gray level transformation :-

1) Basic Gray level Transformation function.

* neighborhood of size 1x1, functions as,

$$s = T(r)$$

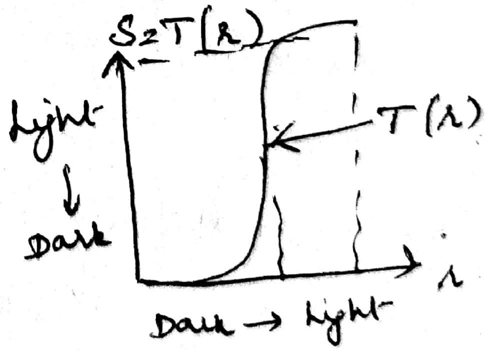
r – denotes gray level $f(x,y)$

s = denotes gray $g(x,y)$ at pt (x,y).

① Contrast stretching

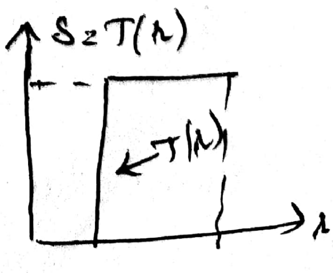
② Thresholding

① Contrast stretching :-



* Darkening of level below m, & brightening the level m.

② Thresholding function :



* T(r) produces a two level binary images

(8) Point processing:-

The contrast stretching & Thresholding process are called point processing.

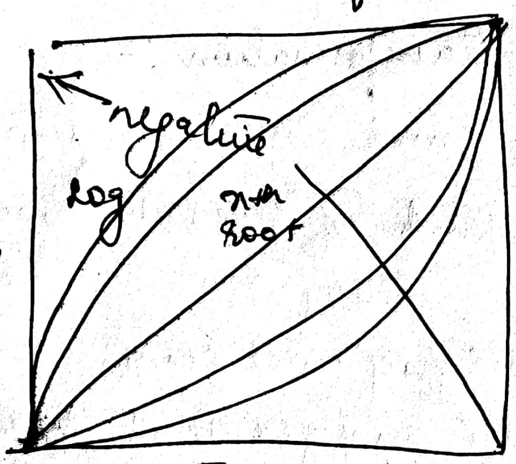
2) Basic gray level Transformation

* simplest image processing techniques,

$$S = T(r)$$

* 3 Basic types of fn,

1. Linear (-ve & identity transformation)
2. Logarithmic (Log an inverse log Transformation)
3. Power law (nth power & nth root)
4. Piece wise linear Transformation functions



O/P intensity level, s

Input intensity level, r.

(1) Image negatives:

range $[0, L-1]$, $\boxed{S = L-1-r}$

(2) Log Transformations:

$$S = c\log(1+r)$$

c – constant

$r \geq 0$.

(3) Power law (Gamma) Transformations:

$$S = Cr^{\gamma}.$$

* & map a narrow range of dark i/p values

Gamma correction:

$\gamma > 1$

$\gamma < 1$

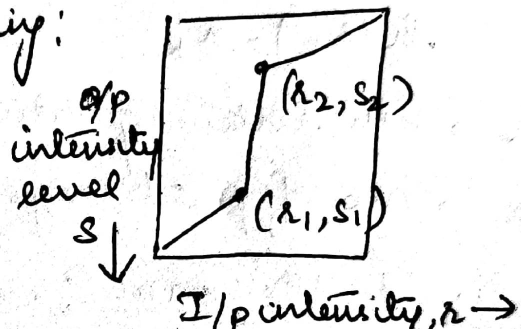
uses of gamma correction:

(4) Piecewise linear Transformation Functions:

* Complementary approach for image negatives, log transformation & power law transformations

<u>Types</u>:

(1) contrast stretching:



* Location of pts (r_1, s_1) & (r_2, s_2)

a) $r_1 = r_2$ & $s_1 = s_2$ → Transformation is linear for

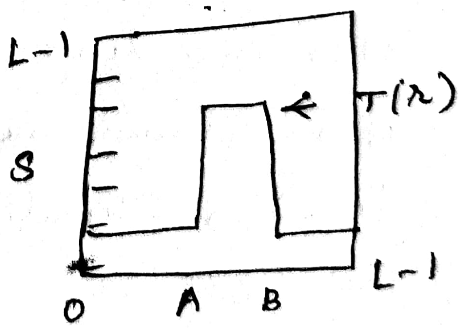
b) $r_1 = s_1$, $s_1 = 0$, & $s_2 = L-1$

c) Intermediate values (r_1, s_1) & (r_2, s_2)

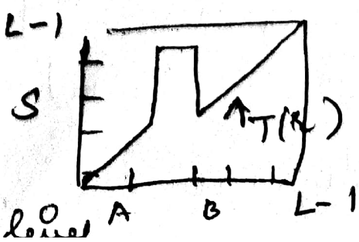
(2) Gray level slicing:-

2 ways

i) 1st Mtd → high value of all gray level in range of interest & low value all gray level

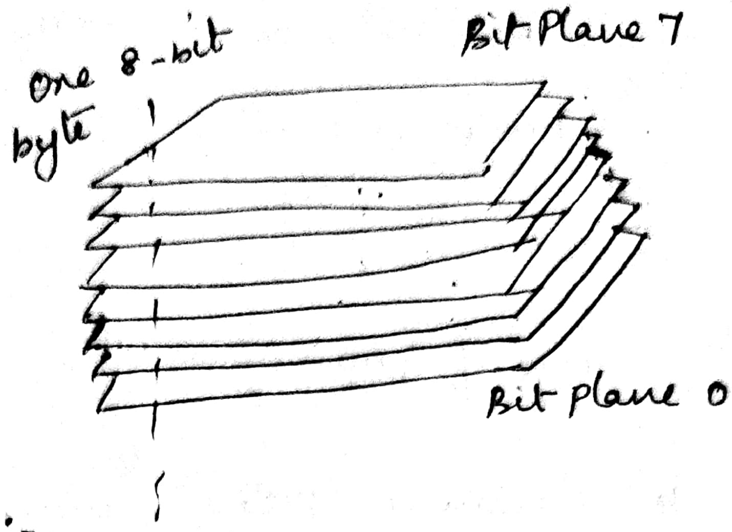


ii) 2nd Mtd: desired range of gray level Preserve the background & gray level



(8) Bit plane slicing

* 8 bit bytes.
* 7 contain all high order bit



III Histogram Processing :-

* Basis for numerous spatial domain processing techniques. It is used for image enhancement.
* gray level [0, L−1] is discrete fn,

$$h(r_k) = n_k$$

$r_k \rightarrow$ kth intensity value.

$n_k \rightarrow$ no. of pixel image with intensity r_k.

Normalized Histogram :-

* Obtained by dividing each components by total no. of pixels in image.

$$P(r_k) = \frac{n_k}{MN} \text{ for } k = 0, 1, 2 \ldots L-1$$

Total no. of pixels in image

M → Row dimension of image

N → column dimension of image.

$P(r_k) \rightarrow$ estimate of probability of occurrence, of gray level r_k.

* Histogram may view graphically by $h(r_k) = n_k$

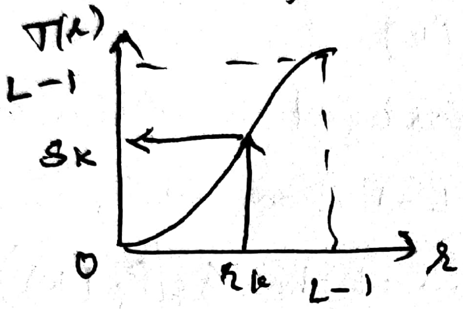
$$P(r_k) = \frac{n_k}{MN} \text{ versus } r_k.$$

(1) Histogram Equalization.

* Enhancing the appearance of images.
* range r is $[0,1]$, $r=0$ & $r=1$.
* Transformation fn,

$$s = T(r) \quad 0 \le r \le L-1$$

* 2 condition:-
  a) $T(r)$ is monotonically increasing fn in interval $0 \le r \le L-1$.
  b) $0 \le T(r) \le L-1$ for $0 \le r \le L-1$.



$$r = T^{-1}(s), \quad 0 \le s \le L-1$$

① Probability Density Function (PDF):-
* Interval $[0,1]$
* $P_r(r)$ & $P_s(s)$ denote probability density fn.
* $P_s(s) = P_r(r) \left| \frac{dr}{ds} \right|$
* Transformation fn of particular importance in image processing is,

$$s = T(r) = (L-1) \int_0^r P_r(w) \cdot dw$$

* cumulative distribution fn (CDF) of Random Variable r.

$$\frac{ds}{dr} = P_r(r)$$

sub expression for P_s,

$$P_s(s) = P_r(r) \frac{1}{P_{(r)}} = 1$$

* Probability occurrence of gray level r_k

$$P_r(r) = \frac{n_k}{HN}, \quad k = 0, 1, 2 \dots L-1.$$

* discrete Transformation,

$$s_k = T(r_k) = \sum_{i=0}^{k} \frac{n_i}{HN}$$

$$= \sum_{i=0}^{k} P_r(r_i)$$

② Histogram Matching (specification)

Algorithm :

S1 : Compute $s_k = P_f(k)$ where $k = 0, 1 \dots L-1$, cumulative normalized histogram of f

S2 : $G(k)$, $k = 0, 1 \dots L-1$ from histogram h_z

S3 : $G_l(s_k)$, $k = 0, 1, \dots L-1$, $G_l(P_f(k))$

S4 : Transform f using $G_l(P_f(k))$

③ Local Histogram Processing (Local Enhancement)

i) Calculate Histogram of the pts in the neighborhood

ii) obtain histogram equalization / specification fn.

iii) Map gray level of pixel centered in neighborhood

iv) Center of neighborhood region is then moved to an adjacent pixel location & procedure is repeated

④ Using Histogram statistics for image enhancement.

* Let r denote discrete random variable intensity values in range $[0, L-1]$. The nth moment of r its mean,

$$\mu_n(r) = \sum_{i=0}^{L-1} (r_i - m)^n P(r_i)$$

where,

m is the mean value of r.

$$m = \sum_{i=0}^{L-1} r_i\, P(r_i)$$

2nd moment,

$$\mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2\, P(r_i).$$

* Intensity Variance is denoted by σ^2,

$$\sigma^2 = \mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 (P_{r_i})$$

* working with mean & Variance,

$$m = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y).$$

$$\text{\&}\ \sigma^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - m]^2.$$

$x = 0, 1, 2 \ldots M-1$

$y = 0, 1, 2 \ldots N-1$

* mean Value of pixels,

$$m_{S_{xy}} = \sum_{i=0}^{L-1} r_i\, P_{S_{xy}}(r_i)$$

* Variance of pixels,

$$\sigma_{S_{xy}}^2 = \sum_{i=0}^{L-1} (r_i - m_{S_{xy}})^2\, P_{S_{xy}}(r_i).$$

$$\therefore g(x,y) = \begin{cases} E \cdot f(x,y), \text{ if } m_{S_{xy}} \le k_0 m_G \text{ \& } k_1 \sigma_G \le \sigma_{S_{xy}} \le k_2 \sigma_2 \\ f(x,y), \text{ otherwise} \end{cases}$$

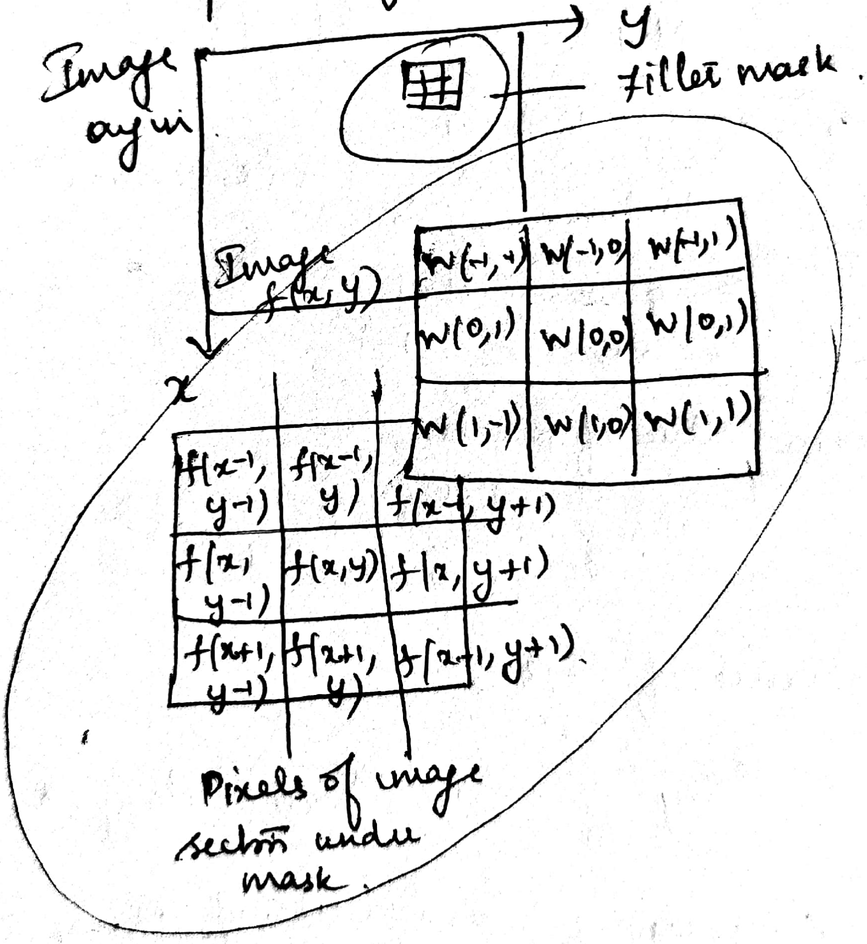
IV Basics of Spatial Filtering :-

* It is used in broad spectrum of applications.
* Filter passes low frequency called low pass filter.

(1) Mechanics of spatial filtering:

* spatial filtering is eg of neighborhood operations; the operation are done at values of image pixels,
* sub image called a filter, mask, kernel, template or window.

i, Linear spatial filter.



* The Result R of linear filtering, with filter mask at pt (x,y)

$$R = w(-1,-1)\, f(x-1, y-1) + w(-1,0)\, f(x-1, y) + \dots + w(0,0)\, f(x,y) + \dots + w(1,0)\, f(x+1, y) + w(1,1)\, f(x+1, y+1)$$

* Filtering of image of size $M*N$ with filter mask of size $m*n$.

$$g(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t)$$

* x & y varied so each pixel in w,

$$a = \frac{(m-1)}{2}$$

$$b = \frac{(n-1)}{2}$$

② Spatial correlation & convolution.

* Correlation is the process of moving a filter mask over the image & computing the sum of product at each location
* Filter mask called as convolution mask.
* Convolution also perform same mechanism of correlation except that in convolution filter is first rotated by 180°.

origin correlation

a) 0 0 0 1 0 0 0 0 (f) 1 2 3 2 8 (w)

0 0 0 1 0 0 0 0

b) 1 2 3 2 8

↑ starting position

c) $\overbrace{0\ 0\ 0\ 0}^{\text{zero padding}}$ 0 0 0 1 0 0 0 0 $\overbrace{0\ 0\ 0\ 0}$ alignment

1 2 3 2 8

d) $\overbrace{0\ 0\ 0\ 0}^{\text{zero padding}}$ 0 0 0 1 0 0 0 0 $\overbrace{0\ 0\ 0\ 0}$

~~1 2 3 2 8~~

1 2 3 2 8

↑ Position after one shift.

origin Convolution

0 0 0 1 0 0 0 0 (f) w rotated 180°: 8 2 3 2 1

0 0 0 1 0 0 0 0

8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0

8 2 3 2 1

0 0 0 0

0 0 0 0 0 0 0 1

8 2 3 2 1 0 0 0 0 0 0 0

⑤

e) 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

1 2 3 2 8 ↑ Position after 4 shifts   8 2 3 2 1

f) 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0   0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

1 2 3 2 8   8 2 3 2 1

g) Full correlation result.

0 0 0 8 2 3 2 1 0 0 0 0 0

Full convolution result.

0 0 0 1 2 3 2 8 0 0 0 0 0

h) Cropped correlation result.

0 8 2 3 2 1 0 0

cropped convolution result.

0 1 2 3 2 8 0 0

* solution for the problem,

1. Pad f with enough 0's on either side to allow each pixel in w to visit every pixel in f.
2. Filter of size m, need $m-1$, 0's on either side of f.

* correlation of filter $w(x, y)$ of size $m \times n$ of image $f(x,y)$,

$$w(x, y) \star f(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t)$$

||rly mannerx,

$$w(x,y) \star f(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x-s, y-t)$$

(3) Vector representation of linear filtering:-

* sum of product,

$$R = w_1 z_1 + w_2 z_2 + \cdots + w_{mn} z_{mn}$$

$$= \sum_{k=1}^{mn} w_k z_k$$

$$\boxed{R = W^T Z}$$

∴ 3x3 filter mask.

$$\begin{bmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \\ w_7 & w_8 & w_9 \end{bmatrix}$$

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9$$

$$= \sum_{k=1}^{9} w_k z_k$$

$$\boxed{R = W^T Z}$$

(4) Generalizing spatial filter masks:-

* $m \times n$ linear spatial filter requires $m.n$ masks coefficients.
* Let z_i, $i = 1, 2 \dots 9$, average is

$$R = \frac{1}{9} \sum_{i=1}^{9} z_i . \quad [w_i = \frac{1}{9}.$$

Gaussian function:

2 Variables:

$$h(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Non linear filter:-

1. size of a neighborhood
2. operation(s) to be performed on image pixels.

c) Smoothing Spatial filtering :-

* It is used for blurring & for noise reduction.

Blurring :- Removal of small details from a image small gaps in lines or curves

Noise Reduction : linear filter & also by non-linear filtering.

① smoothing By linear Filters:

* avg of pixel contained in neighborhood of filter mask.

Operation :-

Replacing the value of every pixel in image by average of gray levels.

Types :

1. Box Filter
2. weighted average Filter.

i) Box Filter :

Spatial averaging filter in which all co-efficients are equal called box filter.

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$1 + 1 + \dots + 1$

Sum of all co-efficients = 9.

$$R = \frac{1}{9} \sum_{i=1}^{9} z_i$$

ii) weighted Average filter.

: A w.avg. filter is one in which pixels are multiplied by different coefficients.

$$= \frac{1}{16} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Sum = 16

$$g(x,y) = \frac{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)\, f(x+s, y+t)}{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)}$$

(2) Smoothing by non-linear filters:-

* Types of non-linear spatial filters are known as order statistic filters.

1) Median filter:

The value of pixel by the median filter

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\text{median}} \{g(s,t)\}.$$

2) Max & Min filter:-

* median filter for the order statistics filter. 100th Percentile results.

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\max} \{g(s,t)\}.$$

(VI) Sharpening Spatial filters:

* To highlight fine details in an image or to enhance details.
* Common practice to approximate the magnitude of gradient by using absolute values instead of square & square roots.

(1) 1st order Derivative:-

1. Must be zero in areas of constant intensity
2. Non-zero at the onset of intensity step & ramp
3. Must be non-zero along ramps

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

(2) 2nd order Derivative:

1. Must be zero in constant area
2. Must be non-zero at onset & end of intensity
3. Must be zero along ramps of constant slope

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) - f(x-1) - 2f(x)$$

(3) using the 2nd Derivative for image sharpening - the laplacian.

* The 2 dimensional fn $f(x,y)$ is defined as,

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

x direction

$$\frac{\partial^2 f}{\partial^2 x^2} = f(x+1, y) + f(x-1, y) - 2f(x,y)$$

|||ly y direction

$$\frac{\partial^2 f}{\partial^2 y^2} = f(x, y+1) + f(x, y-1) - 2f(x,y)$$

* Two dimensional laplacian obtained by,

$$\nabla^2 f = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x,y)]$$

Eqn as,

0	1	0
1	-4	1
0	1	0

1	1	1
1	-8	1
1	1	1

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

* Laplacian for image sharpening,

$$g(x,y) = f(x,y) + c\left[\nabla^2 f(x,y)\right]$$

(4) Unsharp Masking and High Boost filtering.

(i) Blur the original image

ii) Subtract the blurred image from original

iii) Add mask to the original.

* $\bar{f}(x,y)$ denote the blurred image,

$$g_{mask}(x,y) = f(x,y) - \bar{f}(x,y)$$

* weighted portion of mask back to original image

$$g(x,y) = f(x,y) + k * g_{mask}(x,y).$$

(5) using 1st order Derivatives for (non-linear) image sharpening - the gradient.

* two dimensional column vector,

$$\nabla f = grad(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

* magnitude of vector ∇f, denoted as $M(x,y)$

$$M(x,y) \approx |g_x| + |g_y|.$$

ie; $M(x,y) = mag(\nabla f) = \sqrt{g_x^2 + g_y^2}$

* 3x3 region

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

i) Robert cross gradient operator :-

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

* Robert cross gradient as,

$$g_x = (z_9 - z_5)$$
$$g_y = (z_8 - z_6)$$
$$M(x,y) = \sqrt{|z_9 - z_5|^2 + |z_8 - z_6|^2}$$

ii) Sobel operators:
g_x & g_y using 3×3 neighborhood centered on z_5 are

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

VII) Frequency domain: Introduction to fourier transform:

1) Fourier transform & frequency domain.

* sum of sines & cosines of diff. frequencies multiplied by diff. co-efficients.

(i) 1D Fourier Transformation & its inverse

* Fourier Transformation $F(U)$

$$F\{f(x)\} = F(U) = \int_{-\infty}^{\infty} f(x)\exp(-j2\pi ux)\,dx \qquad j = \sqrt{-1} \quad (1)$$

and the reverse process to recover $f(x)$ from $F(U)$ is,

$$F^{-1}\{F(U)\} = f(x) = \int_{-\infty}^{\infty} F(U)\exp(j2\pi ux)\,du.$$

From above 2 eqn, (1) & (2) ⌞(2)

$f(x), x = 0, 1, 2 \cdots m-1$

$$F(U) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp[-j2\pi ux/N] \text{ for}$$

$u = 0, 1 \cdots N-1$ ⌞(3)

obtain $f(x)$ from $F(U)$

$$f(x) = \sum_{u=0}^{N-1} F(U)\exp[j2\pi ux/N] \text{ for} \quad (4)$$

$x = 0, 1, 2 \cdots N-1$.

* Euler's formula,

$$e^{jx} = \cos x + j\sin x. \quad (5)$$

sub (5) in (3).

$$F(U) = \sum f(x)\left[\cos\frac{2\pi ux}{N} + j\sin\frac{2\pi ux}{N}\right] \text{ for}$$

$u = 0, 1 \cdots N-1$.

⌞(6)

* $F(U)$ in polar co-ordinates,

$$F(u) = R(u) + j(I(u)).$$

(or)

$$F(u) = |F(u)|\, e^{j\phi(u)}$$

$$\left[\phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right]\right.$$

⌞(7)

(I) 2D-Fourier transformation & its inverse

* Size $M * N$ is given by,

$$F\{f(x,y)\} = F(u,v) = \int_{-\infty}^{\infty}\int f(x,y)\exp[j2\pi(ux + vy)]\,dx\,dy$$

⌞(8)

(9)

viii Smoothing By frequency domain filters:-

* Edges & other sharp transitions of gray levels of an image contribute significantly to high frequency contents of its Fourier transformation.

* frequency domain,

$$G(u,v) = H(u,v)F(u,v)$$

* 3 types of low pass filters

(1) Ideal
(2) Butterworth
(3) Gaussian.

(1) Ideal low pass filter:-

* It cuts all high freq. component of fourier transform are at distance } than specified distance doing from origin,

* ILPF, $H(u,v) = \begin{cases} 1, & \text{if } D(u,v) \le D_0 \\ 0, & \text{if } D(u,v) > D_0 \end{cases}$

(1) <u>Ideal low pass filter.</u>

* $H(u,v) = \begin{cases} 1, & \text{if } D(u,v) \le D_0 \\ 0, & D(u,v) > D_0. \end{cases}$

* $D(u,v)$ distance from point (u,v)

$$D(u,v) = \left[(u-P/2)^2 + \left(v - \frac{Q}{2}\right)^2\right]^{1/2}$$

* Frequency Rectangle $(u,v) = (M/2, N/2)$,

$$D(u,v) = (u^2+v^2)^{1/2}$$

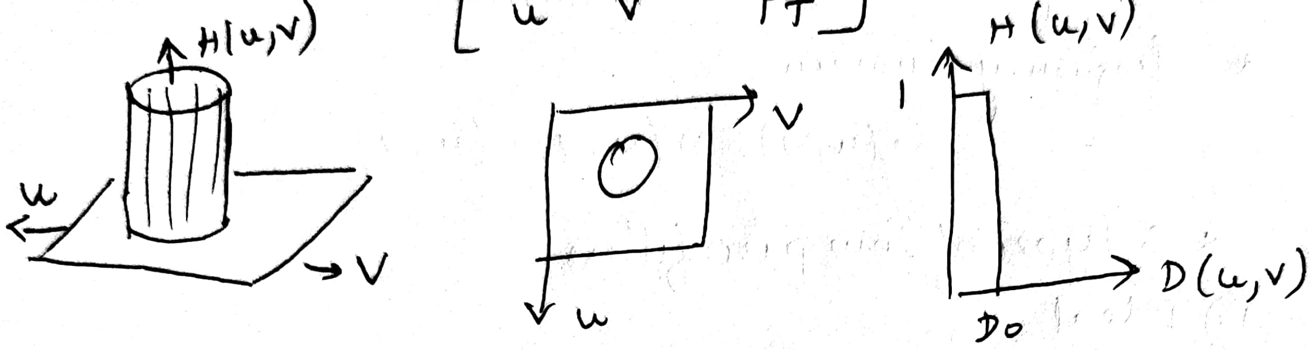
* Point of Transition b/w $H(u,v)=1$ & $H(u,v)=0$, called cut-of frequency.

* Power spectrum,

$$P_T = \sum_{u=0}^{P-1}\sum_{v=0}^{Q-1} P(u,v)$$

* A circle of radius D_0, at centre of frequency rectangle encloses,

$$\alpha = 100\left[\sum_u \sum_v \frac{P(u,v)}{P_T}\right]$$

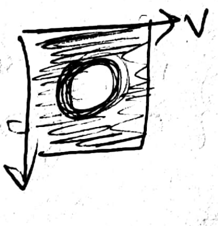


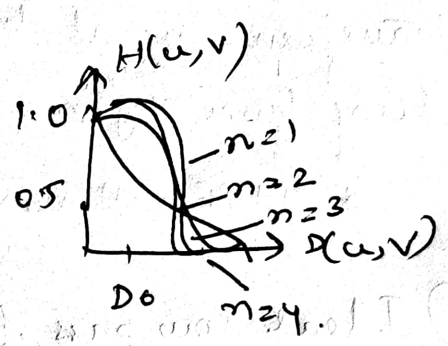
② Butterworth low pass filter.

* distance D_0 from origin is defined as

$$H(u,v) = \frac{1}{1+\left[D(u,v)/D_0\right]^{2n}}$$

Value of n is 2.



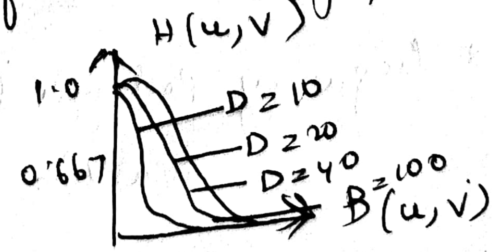


③ Gaussian low pass filter.

* $H(u,v) = e^{-D^2(u,v)/2\sigma^2}$

* where $D(u,v)$ distance of pt (u,v) from center of Transform,

$\sigma = D_0$ - specified cut-off frequency.



14. Sharpening frequency domain filters - Ideal, Butterworth & Gaussian filters:-

* Image sharpening by high pass filtering process.
* $H_{HP}(u,v) = 1 - H_{LP}(u,v)$
* $H_{LP}(u,v)$ — is Transfer fn of low pass filter.

(1) Ideal High pass filters

$$H(u,v) = \begin{cases} 0, & \text{if } D(u,v) \leq D_0 \\ 1, & \text{if } D(u,v) > D_0 \end{cases}$$

$D_0 \rightarrow$ cut-off frequency

(2) Butterworth High pass filter (BHPF)

$$H(u,v) = \frac{1}{1 + [D_0 / D(u,v)]^{2n}}$$

(3) Gaussian high pass filters,

$$H(u,v) = 1 - e^{-D^2(u,v)/2D_0^2}$$

(or)

$$H(u,v) = 1 - e^{-D^2(u,v)/2\sigma^2}$$

Homomorphic filtering :-

* As per model, image $f(x,y)$ 2 components,
  1) amt of source light incident on the scene being viewed
  2) Amt of light reflected by objects in the scene

* fn $i(x,y)$ & $r(x,y)$.

$$f(x,y) = i(x,y) \cdot r(x,y).$$

* Illuminations & reflections b/c cause fourier Transform $f(x,y)$,

$$F[f(x,y)] \neq F[i(x,y)] \cdot F[r(x,y)]$$

* 2 components taking logarithm 2 sides,

$$\ln f(x,y) = \ln i(x,y) + \ln r(x,y)$$

Taking fourier Transform both sides,

$$F[\ln f(x,y)] = F[\ln i(x,y)] + F[\ln r(x,y)]$$

$$F(x,y) = I(x,y) + R(x,y).$$

* sum of 2 images,
  1. Low frequency illumination image
  2. High frequency reflectance image.

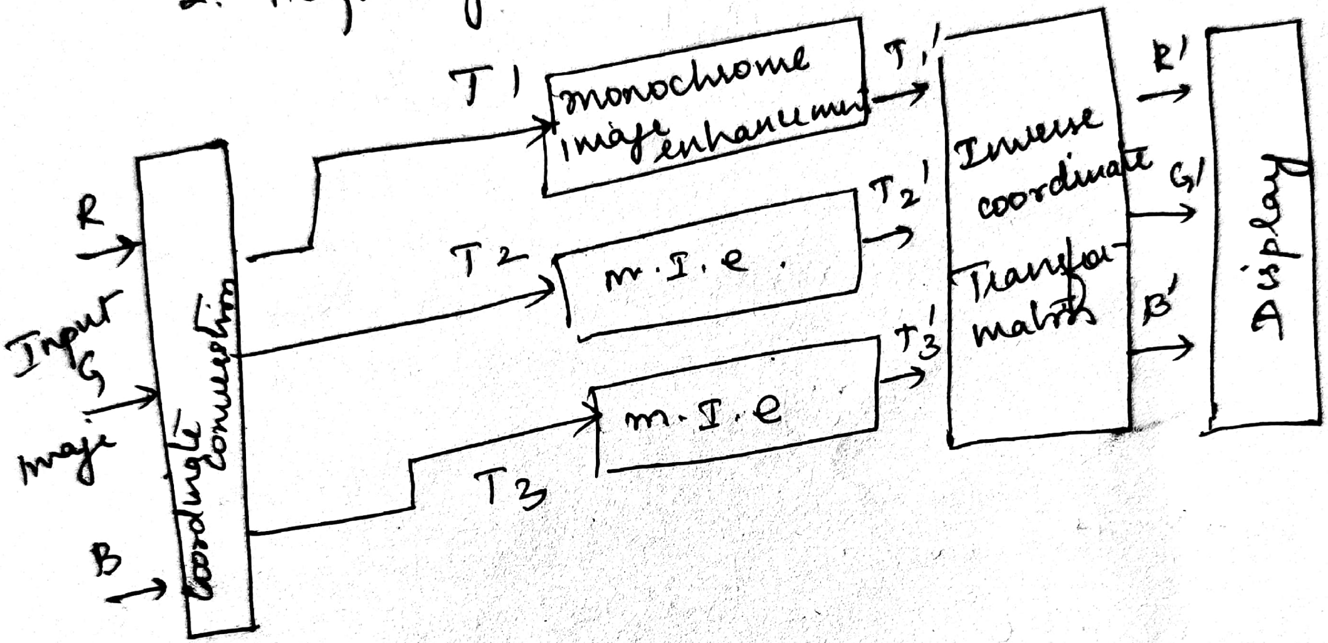
$$\underset{\text{input image}}{f(x,y)} \Rightarrow \boxed{\ln} \Rightarrow \boxed{DFT} \Rightarrow \boxed{H(u,v)} \Rightarrow \boxed{(DFT)^{-1}} \Rightarrow \boxed{exp} \Rightarrow \underset{\text{enhanced Image}}{g(x,y)}$$

XI Color image enhancement:-

* Enhancement techniques can be process an image so that the final result is more suitable than original image.

* 2 Broad categories,

1. Spatial domain Techniques
2. Frequency domain Techniques.



UNIT-III IMAGE RESTORATION:

I Image Restoration:-

* IR refers to a class of methods that aim to remove or reduce the degradation that have occured while digital image.
* some sort of degradation.
  a) During Display mode
  b) Acquisition mode
  c) Processing mode.
* Degradation due to,
  a) sensor noise
  b) Blur due to camera mis-focus
  c) Relative object camera motion
  d) Random atmospheric turbulence.

Types: 1) spatial domain techniques
2) Frequency domain techniques.

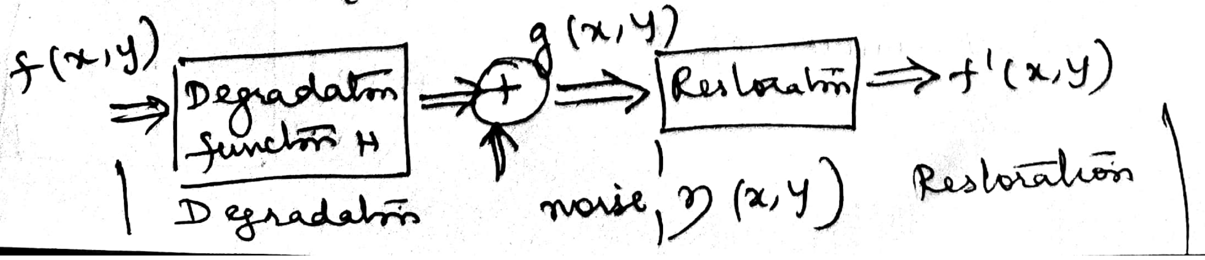
II Degradation model.

* Degradation fn that operates on an i/p with additive noise term.
* Input image notation $f(x,y)$ noise, can be $\eta(x,y)$. two terms combined as $g(x,y)$.
* If linear position invariant process as

$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y) \quad \text{(1)}$$

* frequency domain of equation as,

$$G(u,v) = H(u,v)F(u,v) + N(u,v) \quad \text{(2)}$$



* Image Restoration process as,

$$\hat{G}(u,v) = \frac{F(u,v) - N(u,v)}{H(u,v)} = \frac{F(u,v)}{\hat{H}(u,v)}$$

III Properties, noise models:-

* source of noise in digital images arises during image acquisition and/or transmission.
* Assumptions are made as,
  - → noise model is spatial invariant
  - → noise model is uncorrelated with object function.

1) Spatial & Frequency properties of noise:-

* Frequency property refers to frequency content of noise in fourier sense
* Fourier spectrum of noise is constant called white noise.

2) Probability density functions:-

(i) Gaussian noise:

* Tractability in both spatial & frequency domain called normal noise model.

PDF, z is given by,

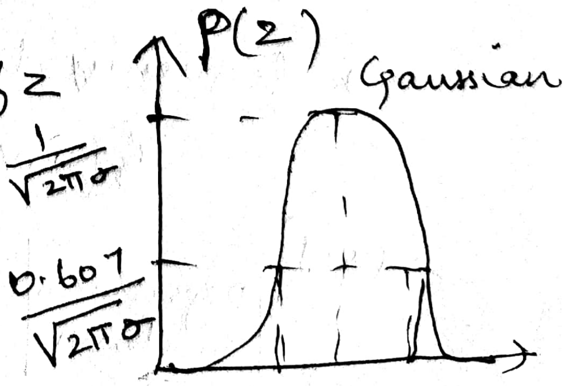
$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(z-\bar{z})^2}{2\sigma^2}} \quad \text{--- (3)}$$

z - gray level

$\bar{z}$ → mean of avg. value of z

σ - std deviation

σ^2 - variance of z



② Rayleigh noise:-

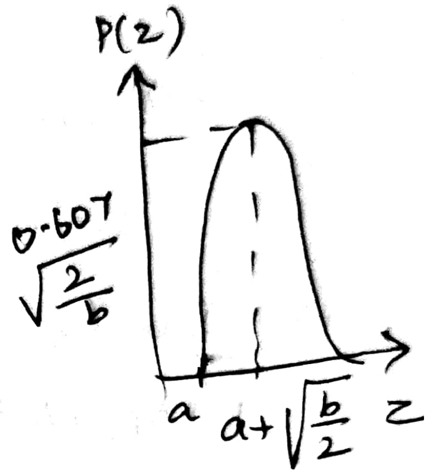
* It is not symmetric,

$$P(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & \text{for } z \geq a \\ 0 & \text{for } z < a \end{cases} \quad — (4)$$

The mean & Variance of density,

$$\bar{z} = a + \sqrt{\frac{\pi b}{4}} \quad — (5)$$

$$\sigma^2 = \frac{b(4-\pi)}{4} \quad — (6)$$



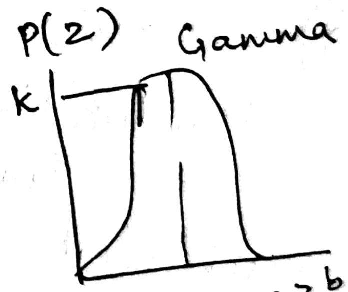
③ Erlang (gamma) Noise:

* PDF of erlang noise,

$$P(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az}, & \text{for } z \geq 0 \\ 0, & \text{for } z < 0 \end{cases} \quad — (7)$$

mean & variance (8) (9)

$$\bar{z} = \frac{b}{a}, \quad \sigma^2 = \frac{b}{a^2}.$$



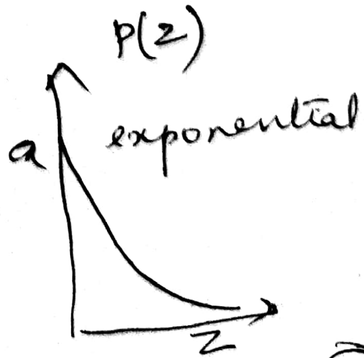
$$K = \frac{a(b-1)^{b-1}}{(b-1)!} e^{-b-1}$$

④ Exponential Noise:-

$$P(z) = \begin{cases} ae^{-az}, & \text{for } z \geq 0 \\ 0, & z < 0. \end{cases} \quad — (10)$$

* $a > 0$, $\bar{z} = \frac{1}{a}$ — (11)

$$\bar{z} = \frac{1}{a}, \quad \sigma^2 = \frac{1}{a^2} \quad — (12)$$

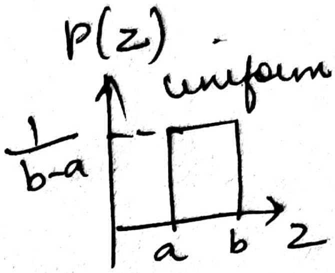


5f) uniform noise

* $p(z) = \begin{cases} \frac{1}{b-a}, & \text{if } a<b<c \\ 0, & \text{otherwise} \end{cases}$ — (13)

* mean of density,

$$\bar{z} = \frac{a+b}{2} \quad \text{— (14)}$$

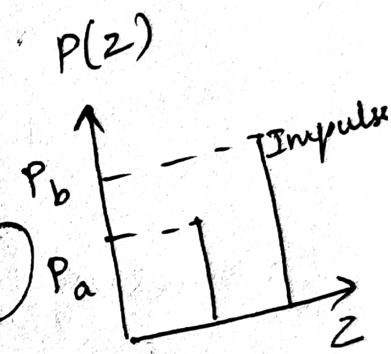


* Variance, $\sigma^2 = \frac{(b-a)^2}{12}$ — (15)

6) Impulse noise:-

* PDF of bipolar noise given by,

$$P(z) = \begin{cases} P_a, & z=a \\ P_b, & z=b \\ 0, & \text{otherwise} \end{cases} \quad \text{— (16)}$$



3) Comparison B/W Various noise:-

Periodic Noise:-

Image arises from electrical or electro-mechanical interference during image acquisition

Restoration in the Presence of noise only-

spatial filtering:-

* Degradation present in image can have,

$g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$ into

$$g(x,y) = f(x,y) + \eta(x,y) \quad \text{--- (17)}$$

$G(u,v) = H(u,v) * F(u,v) + N(u,v)$ into

$$G(u,v) = F(u,v) + N(u,v) \quad \text{--- (18)}$$

IV Mean Filters:-

1. Arithmetic mean filter
2. Geometric mean filter
3. Harmonic Mean filter
4. Contra Harmonic mean filter.

(1) Arithmetic mean filter:-

in sub image $m \times n$,

* S_{xy} represents set of coordinates defined as $S_{x,y}$,

* Restored image $\hat{f}$ at point (x,y),

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s,t) \quad \text{--- (19)}$$

coefficient $= \frac{1}{mn}$.

weight $w_k = \frac{1}{mn}$.

(2) Geometric Mean filter:-

$$\hat{f}(x,y) = \left[\prod_{(s,t) \in S_{xy}} g(s,t) \right]^{\frac{1}{mn}} \quad \text{--- (20)}$$

(3) Harmonic Mean filter:-

$$\hat{f}(x,y) = \frac{mn}{\sum_{(s,t)\in S_{xy}} \frac{1}{g(s,t)}} \quad \text{—— (21)}$$

(4) Contra Harmonic mean filter:-

$$\hat{f}(x,y) = \frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}} \quad \text{—— (22)}$$

Q → called order of filter.

Q +ve → pepper noise eliminated
Q -ve → salt noise eliminated
Q = 0 → arithmetic mean filter
Q = -1 → harmonic mean filter.

Order statistics filters:-

(1) median filter:-
* best order statistic filter, replaces value of a pixel by median of gray level.

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\text{median}} \{g(s,t)\} \quad \text{—— (23)}$$

(2) max & min filters:-
* 100th percentile of ranked set of no's called max filter

$$\hat{f}(x,y) = \max_{(s,t)\in S_{xy}} \{g(s,t)\} \quad \text{—— (24)}$$

* 0th percentile filter of min filter

$$\hat{f}(x,y) = \min_{(s,t)\in S_{xy}} \{g(s,t)\} \quad \text{—— (25)}$$

(3) Midpoint filter :-

* max & min values,

$$\hat{f}(x,y) = \frac{1}{2}\left[\max_{(s,t)\in S_{xy}}\{g(s,t)\} + \min_{(s,t)\in S_{xy}}\{g(s,t)\}\right]$$

(2b)

(4) Alpha trimmed mean filter :-

* delete the $\frac{d}{2}$ lowest & $\frac{d}{2}$ highest intensity values of $g(s,t)$ in neighbourhood S_{xy}.

$$\hat{f}(x,y) = \frac{1}{mn-d}\sum_{(s,t)\in S_{xy}} g_r(s,t).$$

* If $d = mn-1$, filter becomes a median filter.

VI Adaptive filters :-

* Capable of performance superior & behaviour changes based on statistical characteristics

(1) Adaptive, local noise reduction filter.

* Filter operate on local region S_{xy},

It has 4 quantities,

a) $g(x,y)$ value of noisy image at (x,y)

b) σ_η^2 variance of noise corrupting $f(x,y)$ to form $g(x,y)$

c) m_L, local mean of pixels in S_{xy}.

d) σ_L^2 local variance pixels in S_{xy}.

* Behaviour of filter :-

1. σ_η^2 is zero should return value $g(x,y)$, trivial zero noise case $g(x,y)$ equal to $f(x,y)$

2. Local variance high relative to σ_η^2.

(4)

3.2 Variance are equal, we want filter to return arithmetic mean value of pixel S_{xy}.

$$\hat{f}(x,y) = g(x,y) - \frac{\sigma_\eta^2}{\sigma_L^2}\left[g(x,y) - m_L\right] \quad (28)$$

(2) Adaptive median filter :-

* It works in rectangular window area S_{xy} & size of S_{xy} during filter operation,

Z_{min} = min intensity value in S_{xy}

Z_{max} = max " " S_{xy}

Z_{med} = median " S_{xy}

Z_{xy} = intensity value at coordinates (x,y)

S_{max} = max allowed size of S_{xy}.

Algorithm :-

* works in 2 stages,

Stage A, $A_1 = Z_{med} - Z_{min}$

$A_2 = Z_{med} - Z_{max}$

$A_1 > 0$ & $A_2 < 0$.

Stage B = $B_1 = Z_{xy} - Z_{min}$

$B_2 = Z_{xy} - Z_{max}$

If $B_1 > 0$ and $B_2 < 0$, output Z_{xy}.

Output Z_{med}

3 main purposes,

1. To remove salt & pepper noise
2. To provide smoothing to other noise that may not impulsive
3. To reduce distortion

Periodic noise Reduction by frequency domain filtering:-

VII · Band reject filter :-

* It removes band of frequencies about the origin of fourier transformer.

* 3 types — Ideal band reject filter
— Butterworth band reject filter
— Gaussian band reject filter.

① Ideal band reject filter:

$$H(u,v) = \begin{cases} 1 & , D(u,v) < D_0 - \frac{W}{2} \\ 0 & , D_0 - \frac{W}{2} \leq D(u,v) \leq D_0 + \frac{W}{2} \\ 1 & , D(u,v) > D_0 + \frac{W}{2} \end{cases}$$

② Butterworth Bandreject filter:

$$H(u,v) = \frac{1}{\left[1 + \left(\frac{D(u,v)W}{D^2(u,v) - D_0^2}\right)^{2n}\right]}$$

③ Gaussain Bandreject filter.

$$H(u,v) = 1 - \exp\left[\frac{-1}{2}\left(\frac{D^2(u,v) - D_0^2}{D(u,v)W}\right)^2\right]$$

VIII Bandpass filter :-

* Opp to Bandreject filter.

* Transfer fn, $H_{BP}(u,v) = 1 - H_{BR}(u,v)$.

Notch filters :-

* A notch filter rejects frequencies in pre-defined neighborhoods about a center frequency.

$$H(u,v) = \begin{cases} 0, & D_1(u,v) \leq D_0 \text{ or } D_2(u,v) \leq D_0 \\ 1, & \end{cases}$$

where,

$$D_1(u,v) = \sqrt{\left(u - \frac{M}{2} - u_0\right)^2 + \left(v - \frac{N}{2} - v_0\right)^2}$$

$$D_2(u,v) = \sqrt{\left(u - \frac{M}{2} + u_0\right)^2 + \left(v - \frac{N}{2} + v_0\right)^2}$$

Butterworth notch reject filter of order n,

$$H(u,v) = 1 - \exp\left[-\frac{1}{2}\left(\frac{D_1(u,v)\, D_2(u,v)}{D_0^2}\right)\right]$$

* Gaussian notch reject filter formula,

$$H(u,v) = \frac{1}{\left[1 + \left(\frac{D_0^2}{D_1(u,v)\, D_2(u,v)}\right)^n\right]}$$

* Transfer fn given as,

$$\underbrace{H_{NP}(u,v)}_{\text{Tf of pass filter}} = 1 - \underbrace{H_{NR}(u,v)}_{\text{Tf of notch reject filter}}$$

Optimum notch filtering :-

* minimize local variances of restored estimate $\hat{f}(x,y)$

i) 1st step is to extract principal freq. components of interference pattern.

ii) Filter constructed to pass only components associated with interference pattern.

$$N(u,v) = H_{Np}(u,v)\,G(u,v)$$
└ F/T corrupted image.

* Particular filter selected as,

$$\eta(x,y) = \ln^{-1}\{H_{Np}(u,v), G(u,v)\}$$

$$\therefore \hat{f}(x,y) = g(x,y) - w(x,y)\,\eta(x,y)$$

where $w(x,y)$ called a weighting or modulation function.

$$w(x,y) = \frac{\overline{g(x,y)\,\eta(x,y)} - \bar{g}(x,y)\,\bar{\eta}(x,y)}{\overline{\eta^2}(x,y) - \bar{\eta}^2(x,y)}$$

xi Inverse filtering:-

* Process of restoring an image degraded by a degradation fn H.
* Inverse filtering provides estimate $F(u,v)$, degraded image $G(u,v)$ of degradation fn.

$$G(u,v) = H(u,v)\,F(u,v) + N(u,v)$$

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)}$$

$$= \frac{H(u,v).F(u,v) + N(u,v)}{H(u,v)}$$

$$\hat{F}(u,v) = F(u,v) + \frac{N(u,v)}{H(u,v)}$$

* If degradation value has zero or very small values

ratio, $\frac{N(u,v)}{H(u,v)}$.

xii) Wiener Filtering (or) minimum mean square error :-

* Filter incorporates both degradation & statistical behavior of noise into restoration process.

$$\hat{f}(x) = \sum_{s=-\infty}^{\infty} h_W(x-s)\, g(s)$$

error measure,

$$e^2 = E\{(f-\hat{f}^2)\}$$ [∵ $E\{\cdot\}$ expected value of argument).

* min. error fn as,

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)\, S_f(u,v)}{S_f(u,v)\,|H(u,v)|^2 + S_\eta(u,v)}\right] G(u,v)$$

$$= \left[\frac{H^*(u,v)}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v)$$

$$= \left[\frac{1}{H(u,v)} \cdot \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v)$$

* Bracket is min. mean square error filter.

$H(u,v)$ = FT of degradation fn

$H^*(u,v)$ = Complex conjugate of $H(u,v)$

$|H(u,v)|^2 = H^*(u,v)\,H(u,v)$

$S_\eta(u,v) = |N(u,v)|^2$ = Power spectrum of noise

$S_f(u,v) = |F(u,v)|^2$ = power spectrum of undegraded image

$G(u,v)$ = F.T of degraded image.

* Mean square error,

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \cdot \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] G(u,v)$$

* k is specified constant of $|H(u,v)|^2$.

* Signal to noise ratio

$$SNR = \frac{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F(u,v)|^2}{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |N(u,v)|^2}$$

* mean square error,

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[f(x,y) - \hat{f}(x,y)\right]^2$$

* SNR in spatial domain,

$$SNR = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[f(x,y) - \hat{f}(x,y)\right]^2}$$

UNIT-4

IMAGE SEGMENTATION:-

* Image is preprocessed to remove noise & artifacts Segmentation is often key step in interpreting the image.
* Region based techniques on common patterns in intensity values. Cluster is referred as region.
* Edge base techniques discontinuities in image values b/w distinct region.

Detection of Discontinuities:

① Point detection : pt in which gray level or intensity is entirely diff from background.

② Line detection : boundaries b/w region distinct gray levels.
The Hough Transform is used to detect lines.

③ Edge detection : identifying edges in images.
1. step discontinuities
2. line discontinuities.

I Edge detection :-

* Process of identifying edges in an image to be used as fundamental asset in image analysis.

Step discontinuities :-

Image intensity abruptly changes from one value on one side of discontinuity to a different value on opposite side.

Line discontinuities :-

* changes value but then returns to starting value with short distance.

→ Step edge :-
Involves a transition b/w two intensity

levels occurring ideally over the distance of 1 pixel.

→ Ideal edges:- occurs over the distance of 1 pixel, no additional processing to make "Real".

→ Roof edges: nothing more than a 1 pixel thick line running ~~...~~ region.

* 2nd derivatives is +ve at the beginning of ramp
-ve at end of ramp.

1) Image gradient and its properties,

* Image gradient ~~tool~~ used for edge strength & direction at location (x, y) of image f,

$$\nabla f \equiv grad(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

* magnitude length of vector ∇f, denoted $M(x,y)$

$$M(x,y) = mag(\nabla f) = \sqrt{g_x^2 + g_y^2}$$

* direction of gradient vector,

$$\alpha(x,y) = \tan^{-1}\left[\frac{g_y}{g_x}\right]$$

<u>Roberts operator</u>:-

* diagonal edge direction of interest, a 2D mask.

$$g_x = \frac{\partial f}{\partial x} = (z_9 - z_5)$$

$$g_y = \frac{\partial f}{\partial y} = (z_8 - z_6)$$

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

3x3 image

-1	0
0	1

Mask

0	-1
1	0

Mask

Prewitt operators :-

* To find g_x and g_y,

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

$$g_x = \frac{\partial f}{\partial x} = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

Edge linking and Boundary detection :- Via Hough Transform :-

* Pixels lying on boundaries b/w regions.
i) noise   ii) Breaks in boundary due to unifom illumination
iii) other effects that introduce spurious discontinuities in intensity value.

* Edge linking & Boundary detection.

① Local processing :-
* Analyze the characteristics of pixels in a small neighborhood (3x3, 5x5 etc).

Principal properties :
1. The strength of response of gradient operator used to produce edge pixels
2. direction of gradient

$$\alpha(x,y) = \tan^{-1}\left[\frac{g_y}{g_x}\right]$$

coordinates (s,t) in S_{xy},
* An edge pixel

$$|M(s,t) - M(x,y)| \leq E$$

* An edge pixel with co-ordinates (s,t) in S_{xy} to pixel (x,y),

$|\alpha(s,t) - \alpha(x,y)| \leq A.$

1. Compute the gradient magnitude & angle arrays $M(x,y)$ & $\alpha(x,y)$ of i/p images $f(x,y)$.

2. Form a binary image g,

$$g(x,y) = \begin{cases} 1, & \text{if } M(x,y) > T_M \\ 0, & \text{otherwise.} \end{cases}$$

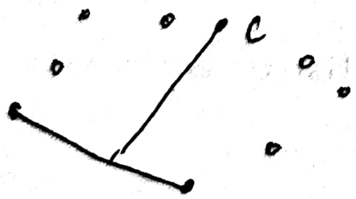
T_M is Threshold.

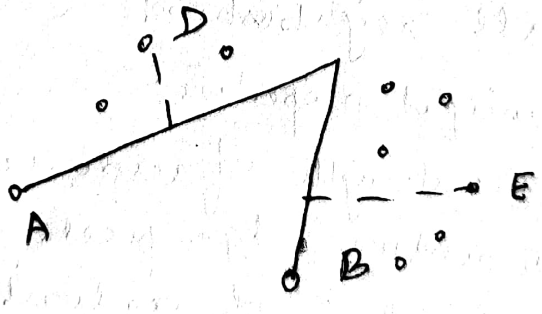
3. Scan the rows of g & fill (set to 1) all gaps (sets of 0s) in each row, do not exceed a specified length k.

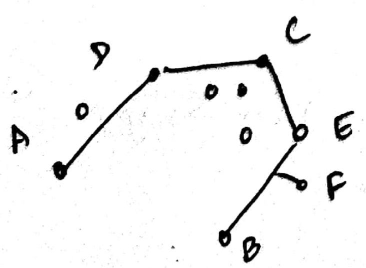
4. To detect gaps in other directions, θ rotate g by this angle & apply horizontal scanning.

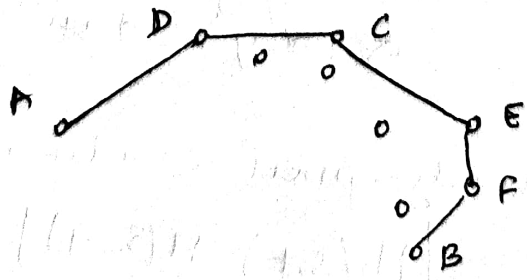
(2) Regional processing:-

* 2 important requirements for polygonal,
  1. starting pt must be specified
  2. All points must be ordered









Algorithm :-

S1: Let P be sequence of ordered, distinct, value of binary image

S2: Specify a Threshold T & 2 empty Stacks

S3: If pt P correspond to closed curve.

S4: Compute the parameters of line passing from last vertex in CLOSED to last vertex OPEN

S5: Compute the distance from line S4 to all pts in P.

S6: $D_{max} > T$, place V_{max} at end of OPEN stack. Go to S4

S7: Remove the last Vertex from OPEN & insert it as last Vertex of CLOSED

S8: OPEN is not empty, go to S4.

S9: Else exit. The vertices of polygonal fit into pts in P.

③ Global Processing using the HOUGH Transform:-

* Finding straight-line points:-

* Given n pt in image, we want to find subsets of these pts that lie on straight line.

* 2 possible soln,

Md 1: * First, find all lines determined by every pair of pts

* Next find all subsets of pts that are close to particular line.

Drawback :-

Finding $n(n-1)/2 \sim n^2$ lines & performing $(n)(n(n-1))/2 \sim n^3$ comparison at every point to all lines.

M2 : Hough Transform is an alternative approach.

Hough Transform:-

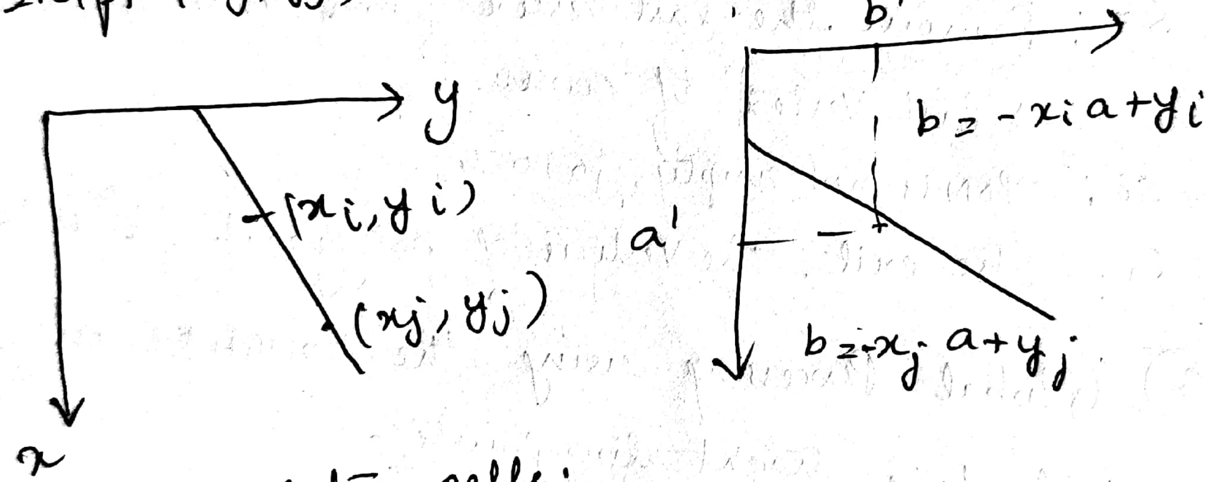
* consider pt (x_i, y_i) in xy-Plane & general eqn of st line in slope-intercept form $y_i = ax_i + b$. Infinitely many lines pass through (x_i, y_i) eqn as,

$y_i = ax_i + b$

* Single line for fixed pair (x_i, y_i)

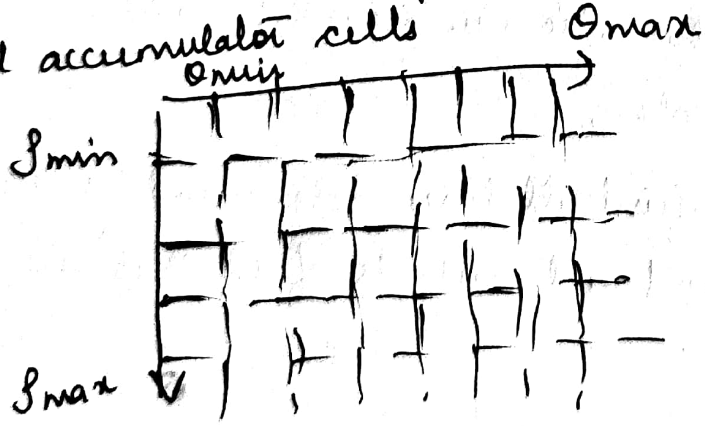
$$b = -x_i a + y_i$$

* 2nd pt (x_j, y_j) also line parameter space associated



→ Accumulator cells:

* Parameter space can sub-divided into cells called accumulator cells



* where (Smin, Smax) & (θmin, θmax) are expected range of Parameters Values

* Solve corresponding S eqn, $S = x_k \cos\theta + y_k \sin\theta$

* Transform is applicable at any fn $g(v, c) = 0$,

v is vector of coordinates
c is vector of co-efficients.

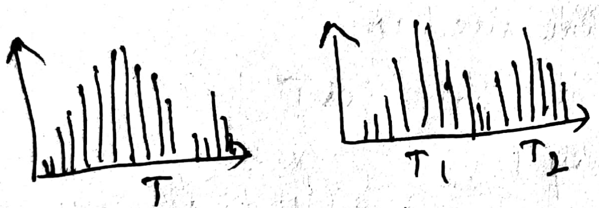
Approach based on Hough Transform as follows.

i) obtain a binary edge image
ii) specify subdivisions in $\rho\theta$-plane
iii) Examine the counts of the accumulator cells for high pixel concentrations.
iv) Examine the relationship.

II Thresholding :-

* simple form of image segmentation.
* thresholding is a technique for partitioning image directly into region based on intensity value & property of value

→ Because of intuitive property
→ simplicity of implementation
→ Computational speed.



1) Thresholding foundation:

* Grey level corresponds to image $f(x,y)$
* Threshold 'T' that separates these modes, pt (x,y) for which $f(x,y) > T$ called object points., The pt called background modes.

* $$g(x,y) = \begin{cases} 0 & f(x,y) < T \\ 1 & f(x,y) => T \end{cases}$$

* 2 modes
1) histogram has to be partitioned by multiple thresholds.
2) Multilevel thresholding classifies pt (x,y) if $T_1 < (x,y) <= T_2$.

4

* other object if $f(x,y) > T_2$
* background,

$$f(x,y) <= T_1$$

* segmented image,

$$g(x,y) = \begin{cases} a, & \text{if } (x,y) \; \epsilon > T_2 \\ b, & \text{if } T_1 < f(x,y) \le T_2 \\ c, & \text{if } f(x,y) \le T_1 \end{cases}$$

2) Basic Global & local thresholding :-

* $T = T[x, y, p(x,y), f(x,y)]$

$f(x,y) \rightarrow$ gray level & $p(x,y) \rightarrow$ local property.

* Thresholding schemes compare each pixel gray level with single global threshold it is as global Thresholding
* Interactive algorithm used as,

1. Select an initial estimate for T
2. Segment image using T
3. Compute the average grey level values μ_1 & μ_2 in region G_1 & G_2
4. Compute new threshold value $T = \frac{1}{2}[\mu_1 + \mu_2]$

3) optimum global thresholding using Otsu's Method :-

* To minimize the avg error incurred in assigning pixels to two or more group called classes.
* The total number MN of pixel in image

$$MN = n_0 + n_1 + n_2 + \cdots + n_{L-1}$$

* Normalised Histogram has Component $P_i = n_i / MN$

$$\sum_{i=0}^{L-1} P_i = 1, \quad P_i \geq 0 \qquad (1)$$

$$P_1(K) = \sum_{i=0}^{k} P_i \qquad (2)$$

Probability of C_1 occuring, C_2 occuring,

$$P_2(K) = \sum_{i=k+1}^{L-1} P_i = 1 - P_1(K) \qquad (3)$$

* mean intensity value of pixels to class C_1 is,

$$m_1(K) = \sum_{i=0}^{k} i\,P(i/C_1)$$

$$= \sum_{i=0}^{k} i\,P(C_1/i)\,P(i)/P(C_2)$$

$$= \frac{1}{P_1(K)} \sum_{i=0}^{k} i\,P_i. \qquad (4)$$

* 2nd line of equation as, $P(A/B) = P(B/A)\,P(A)/P(B)$.
* mean intensity value of pixels to class C_2,

$$m_2(K) = \sum_{i=k+1}^{L-1} i\,P(i/C_2)$$

$$= \frac{1}{P_2(K)} \sum_{i=k+1}^{L-1} i\,P_i \qquad (5)$$

* cumulative mean upto level k, $m(K) = \sum_{i=0}^{k} i\,P_i$ (6)
* avg intensity of entire image, $m_G = \sum_{i=0}^{L-1} i\,P_i$ (7)
* Validity of 2 eqn as, $P_1 m_1 + P_2 m_2 = m_G$ (8)

$$P_1 + P_2 = 1 \qquad (9)$$

* Dimensionless metric $\eta = \frac{\sigma_B^2}{\sigma_G^2}$ (10)
* σ_G^2 is global variance, $\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2\, P_i$ (11)
* σ_B^2 is b/w class variance, $\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2$ (12)
* Expression written as, $\sigma_B^2 = P_1 P_2 (m_1 - m_2)^2$.

5

$$= \frac{(m_G P_1 - m)^2}{P_1(1-P_1)} \quad \text{——— (13)}$$

$$\therefore \eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2}$$

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1-P_1(k)]}$$

4) Multiple Thresholds :-

* It can be extended by arbitary no of Threshold because measure on which based also extends to an arbitrary no. of classes.

$$\sigma_B^2 = \sum_{k=1}^{K} P_K (m_K - m_G)^2 \quad \text{———.}$$

where $P_K = \sum_{i \in C_K} i P_i$

$$m_K = \frac{1}{P_K} \sum_{i \in C_K} i P_i$$

* m_G is global mean,

$$\sigma_B^2(k_1^*, k_2^* \ldots k_{K-1}^*) = \sum_{K=1}^{K} \underset{(k_1, k_2 \ldots k_{K-1})}{\max\ 0 < k_1 < k_2 < \ldots < k_{n-1} < L-1}$$

* 3 classes consisting of 3 intensity interval,

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 + P_3(m_3 - m_G)^2.$$

where, $P_1 = \sum_{i=0}^{k_1} p_i$, $P_2 = \sum_{i=k_1+1}^{k_2} p_i$, $P_3 = \sum_{i=k_2+1}^{L-1} p_i$

& $m_1 = \frac{1}{P_1} \sum_{i=0}^{k_1} i p_i$, $m_2 = \frac{1}{P_2} \sum_{i=k_1+1}^{k_2} i p_2$, $m_3 = \frac{1}{P_3} \sum_{i=k_3+1}^{L-1} i p_i$

* The following, $P_1 m_1 + P_2 m_2 + P_3 m_3 = m_G$
  & $P_1 + P_2 + P_3 = 1$.
* Threshold image, $g(x,y) = \begin{cases} a, & \text{if } f(x,y) \le k_1^* \\ b, & \text{if } k_1^* < f(x,y) \le k_2^* \\ c, & \text{if } f(x,y) > k_2^* \end{cases}$

$$\therefore \eta(k_1^*, k_2^*) = \frac{\sigma_B^2(k_1^*, k_2^*)}{\sigma_G^2}.$$

5) Variable Thresholding :-
* Noise & non-uniform play major role performance, of a thresholding algorithm.
* Local Thresholds, $T_{xy} = a\sigma_{xy} + b m_{xy}$
* a and b non-negative constant,
  $T_{xy} = a\sigma_{xy} + b m_G$
* m_G global image mean, $g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T_{xy} \\ 0 & \text{if } f(x,y) \le T_{xy} \end{cases}$.

* Parameters computed in neighbourhoods of (x,y),
  $g(x,y) = \begin{cases} 1, & Q \text{ is True} \\ 0, & Q \text{ is false} \end{cases}$

* Standard deviation,
  $Q(\sigma_{xy}, m_{xy}) = \begin{cases} \text{True}, & \text{if } f(x,y) > a\sigma_{xy} \text{ and } f(x,y) > b m_{xy} \\ \text{False}, & \text{otherwise} \end{cases}$

<u>local Thresholding using moving averages</u> :-
* moving avg (mean) at this pt,

$$m(k+1) = \frac{1}{n} \sum_{i=k+2-n}^{k+1} z_i$$

$$= \frac{m(k)+1}{n(z_{k+1} - z_{k-n})}$$

* Thresholding based on moving average

$$g(x,y) = \begin{cases} 1, & \text{if } f(x,y) > T_{xy} \\ 0, & f(x,y) \le T_{xy} \end{cases}$$

iii Region based segmentation :-

* Segmentation is Process of partitioning an image into multiple regions group of connected pixels with similar properties.

* Properties are
  i) connectivity & compactness
  ii) Regularity of boundaries
  iii) Homogeneity interms of color or textures
  iv) Differentiation from neighbor regions.

* we want partition R into n,
  a) Summation $R_i = R$ - all pixel belong to a region
  b) R_i connected region $i = 1, 2 \ldots n$ - pixels
  c) $R_i \cap R_j \neq 0$, for $i \neq j$
  d) $P(R_i)$
  e) $P(R_i \cup R_j) =$ False.

2 Types — Region Growing
— Region splitting & merging.

iv Region growing :-

* Procedure that group pixels or sub region into larger region based on predefined criteria for growth.

Region Growing algorithm :-

$f(x,y)$ denote i/p array $s(x,y)$ denote seed array

S1: Find all connected component in $s(x,y)$

S2: image f_Q pair of coordinates (x,y), $f_Q(x,y) = 1$

S3: Let g be an image formed from appending to each seed point in s all the 1-valued pt in f_Q

Scanned with OKEN Scanner

S4: Label each connected component in g with different region label.

Simultaneous Region Growing:

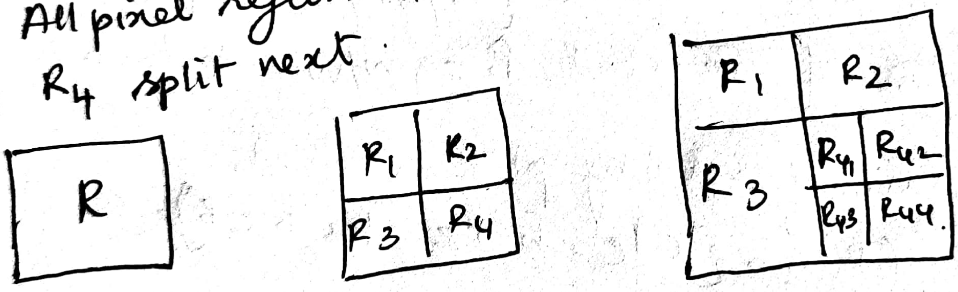
* neighbouring region are taken into account in growing process.
* no single region is allowed
* no. of region is allow to grow @ same time.
* IIIrl region will gradually coalesce into expanding region.

ii) Region splitting & merging :-

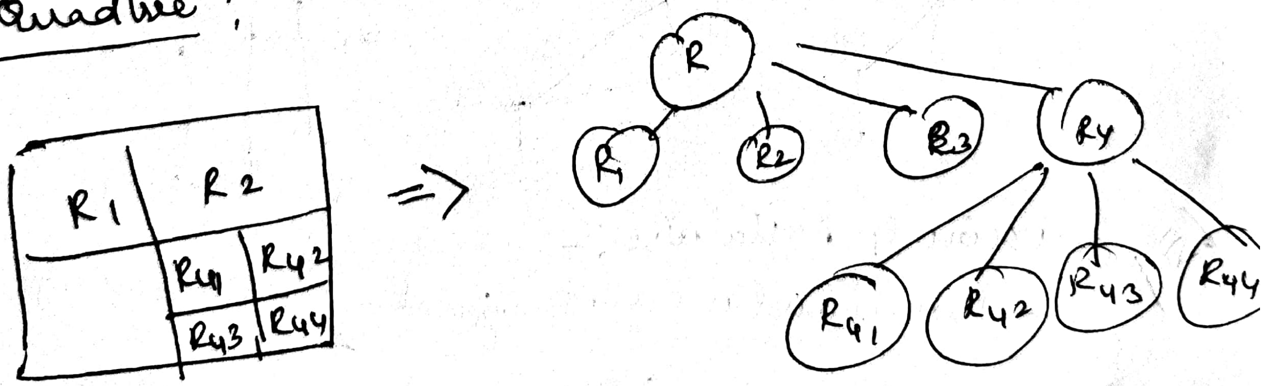
* It is segmentation process in which an image initially subdivided into set of arbitrary.

Note:

1) R Represents whole image
2) Not all the pixels are similar. so region split is $Q(R) = FALSE$
3) All pixel region R_1, R_2 & R_3 are IIIrl, R_4 not.
4) R_4 split next.



Quadtree:



$Q(R_j \cup R_k) = TRUE$.

Split & Merge algorithm:-

S1: $Q(R_i) = FALSE$

S2: $Q(R_j \cup R_k) = TRUE$

S3: Stop when no further merging

9

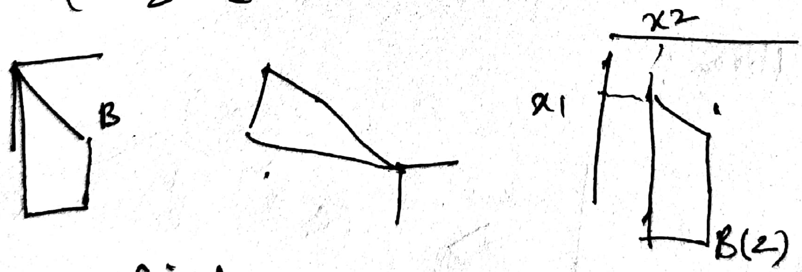
VI Morphological processing:-

* Refers to scientific branch that deals the form & structure of animals & plants.
* Tool for extracting image component.

1) <u>Basics of set theory</u>:-

* Let A be set in Z^2, & $a = (a_1, a_2)$, a is element A : $a \in A$.
* a is not element of A, $a \notin A$
* If every element of set A also element of set B, $A \subseteq B$
* Union A & B are collection of all elements, $C = A \cup B$
* Intersection of A & B, ~~C=A∪B~~ $D = A \cap B$
* no common element in A and B, $A \cap B = \phi$.
* Complement of set A, $\bar{A} = \{w / w \in A\}$.
* diff of two set $A - B = \{w / w \notin A, w \notin B\}$
* Reflection of two set B, $\hat{B} = \{w / w = -b \text{ for } b \in B\}$
* Translation of set B by pt $z = (z_1, z_2)$,

$$(B)_z = \{c / c = b + z \text{ for } b \in B\}$$



VII Erosion & Dilation:-

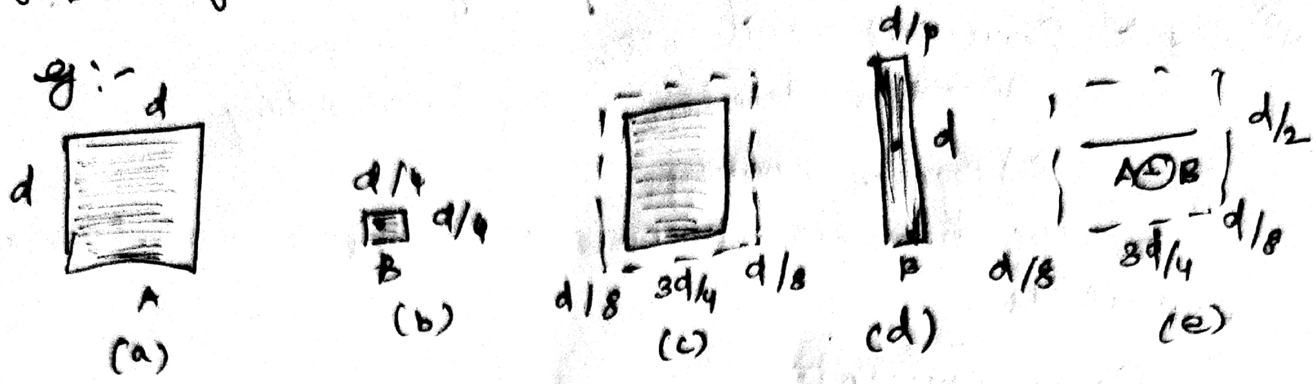
* Two fundamental operations used in image Processing.

<u>Erosion</u>:

~~If~~ erosion shrinks an image object, ~~the~~ erosion of sets A by B denoted $(A \ominus B)$

* B is defined as, $A \ominus B = \{z \mid (B)_z \subseteq A\}$

eg:-



Structuring elements:

* Also called kernel. It consist of pattern specified as co-ordinates of a no of discrete points.
* 2 main character, i) shape ii) size.

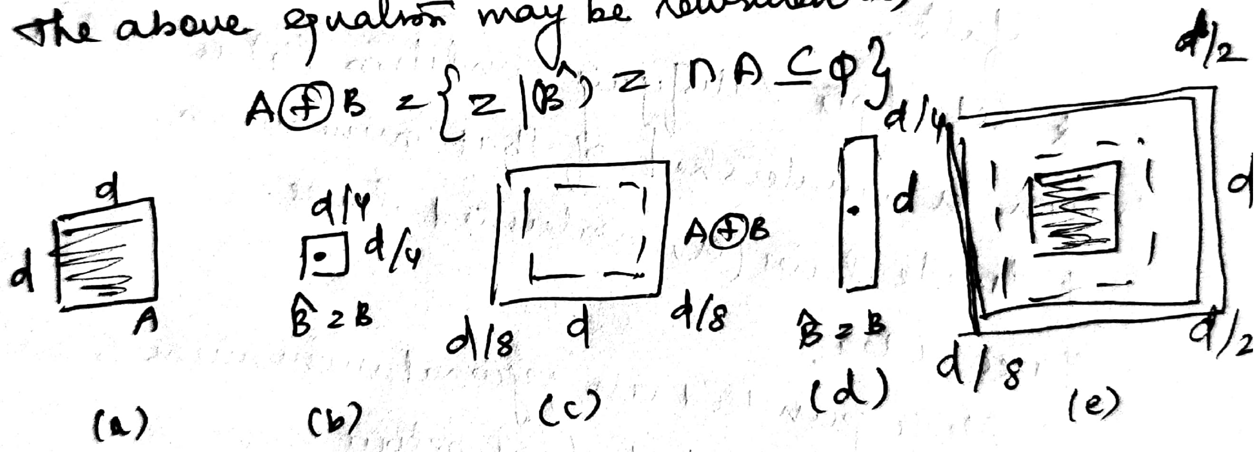
Shape:- * element may be ball or line, convex or ring

size → 3×3 or a 21×21 square.

Dilation → $A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \phi\}$.

The above equation may be rewritten as,

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \subseteq \phi\}$$



3) Duality:-

* Erosion & dilation are duals of each other to set complementation & reflection,

$$(A \ominus B)^c = A^c \oplus \hat{B}$$

$$(A \oplus B)^c = A^c \oplus \hat{B}$$

VIII

Segmentation by morphological watersheds:-

* 3 Principal concepts;
  a) edge detection b) Thresholding
  c) Region growing.
* approach based concept called morphological watersheds.

Basic concepts:-

* Visualizing an image in 3 dimensions, two spatial Coordinates Versus intensity. In such "Topographic" interpretation, we have 3 types.
  a) Pt belonging to regional minimum.
  b) Pt at which a drop of water, ~~fall~~ to placed at location at any points.
  c) pt at which water would equally likely to fall.
* Set of pts satisfying condition (b) is called catchment basin a watershed of that minimum.
* divide lines (a) watershed lines.

Basic idea:

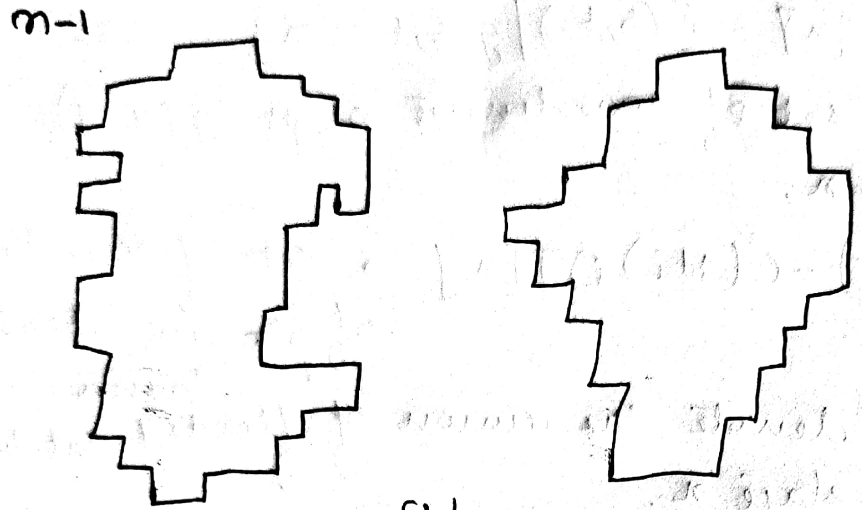
1. A hole is punched in regional minimum & entire topography is flooded from below
2. when the rising water in distinct catchment basins is about to merge.
3. Flooding will eventually reach a stage.
4. Dam line corresponds to ~~zero~~ divide line of watersheds.

IX Dam construction :-

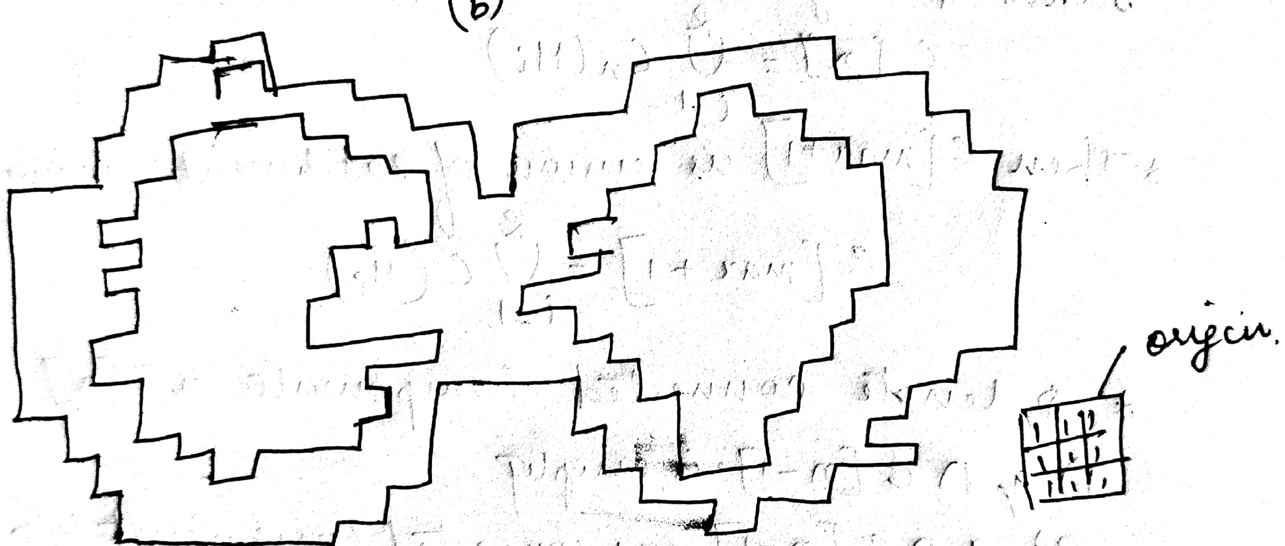
* Based on binary images, which are members of 2D integer space Z^2.

a) Two catchment basins at flooding step

$n-1$

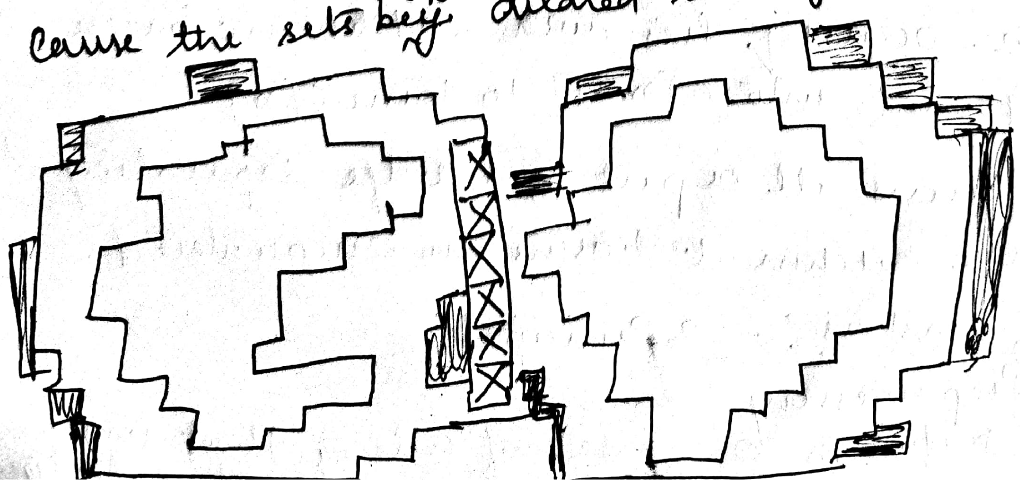


(b)



2 condition

1. dilation has to be constrained to q
2. dilation cannot be performed on pts that would cause the sets being dilated to merge.



7 Watershed Segmentation Algorithm:-

* $M_1, M_2 \ldots M_R$ are coordinates of points in regional image $g(x,y)$. Let $C(M_i)$ set denoting the co-ordinate
* Let $T[n]$ represents set of coordinates (s,t), $g(s,t) < n$,
$$T[n] = \{(s,t) \mid g(s,t) < n\}$$
* $T[n]$ is set of coordinates of pt $g(x,y)$, $g(x,y) = n$,
* $C_n(M_i) = C(M_i) \cap T[n] \rightarrow$ set of coordinate of pts in catchment basin.
* Let $C[n]$ denote the union of flooded catchment basins at stage n.
$$C[n] = \bigcup_{i=1}^{R} C_n(M_i)$$
* Then $C[max+1]$ is union of catchment basins,
$$C[max+1] = \bigcup_{i=1}^{R} C(M_i)$$
* Q denote connected components in $T[n]$,
1) $q \cap C[n-1]$ is empty
2) $q \cap C[n-1]$ one connected component $C[n-1]$
3) $q \cap C[n-1]$ contains " " ' $C[n-1]$
* Condition 1 occur when a new minimum is encountered into $C[n-1]$ to $C[n]$
* Co 2. occurs q lies with catchment basin incorporated into $C[n-1]$ to form $C[n]$
* C3 occurs all or part of a ridge separating two or more catchment basins in encountered.

Use of markers :- 2 Principal
1. Preprocessing
2. Definition of a set of criteria that markers must satisfy

UNIT-V IMAGE COMPRESSION & RECOGNITION:-

* Two types of digital image compression are
  1) Lossless (or error free)
  2) Lossy compression

I Need for data compression:-

* Data compression can dramatically decrease the amount of storage a file takes up.
* Compression optimizes backup storage performance & has recently shown up in primary storage data reductions.

II Huffman coding:-

* Variable length coding is used to reduce only Coding redundancy.

Huffman coding:

* It provides smallest possible no. of code symbols per source symbol, when coding of an information source.

Steps:-

1) create series of source reductions by ordering the Probability of symbols.

* Source symbols,

$\{a_1, a_2, a_3, a_4, a_5, a_6\} = \{0.1, 0.4, 0.06, 0.1, 0.04, 0.3\}$

1st source reduction formed by combining 0.06 & 0.04

$0.06 + 0.04 = 0.1$

Original source		Source Reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_5	0.06	0.1			
a_6	0.04				

S2: minimal length binary codes used 0 & 1

O.S			S.R			
Symbol	Probability	code	1	2	3	4
a_2	0.4	1	0.4 1	0.4 1	0.4 1	0.6
a_6	0.3	00	0.3 00	0.3 00	0.3 00	0.4
a_1	0.1	011	0.1 011	0.2 010	0.3 01	
a_4	0.1	0100	0.1 0100	0.1 011		
a_3	0.06	01010	0.1 0101			
a_5	0.04	01011				

& average length of this code,

$$L_{avg} = (0.4)(1) + (0.3)(2) + (0.1)(3) + (0.1)(4) + (0.06)(5) + (0.04)(5)$$

$= 2.2$ bits/pixel

iii) Run length Encoding :-

* RLE is technique used to reduce the size of repeating string of characters. This repeating string is called a Run.
* Typically RLE encodes a run of symbols into two bytes,
  - → * count
  - → symbol
* RLE can compress any type of data regardless of information content.
* Black white document images, cartoon images etc are quite suitable for RLE.

Principle: Run character is replaced with no. of same characters & a single character.

eg :- R T A A A A S D E E E E E

RLE representation : RT*4A SD*5E

* 3 symbol encoding strategy to represent a repetitive cluster.
  - → CTRL → COUNT → CHAR

iv) Shift codes / Binary shift code:-

* popular technique for removing coding redundancy is huffman coding, when coding symbols of an information source individually.
* Binary shift algorithm :-

1) The source symbol are arranged so that their probabilities are monotonically .
2) The total no. of source symbols is divided

2

into symbol blocks of equal size.

3. The individual source symbols within all blocks are coded identically with natural binary code. Let's designate the huffman code of hypothetical symbol as C.

4. Unique prefix sequence C is chosen and C_{k-1} is concatenated with symbols of block k to identify the symbols within this block.

source symbol	P_i	Binary code
		00
A_0	0.3	01
A_1	0.2	100
A_2	0.15	110
A_3	0.1	1010
A_4	0.08	1011
A_5	0.06	1110
A_6	0.05	11110
A_7	0.04	11111
A_8	0.02	
		2.81
L_{avg}	$H = 2.778$	

average length of Huffman code is,

$$L_{avg} = \sum_{i=0}^{8} P_i l_i$$

$$= 0.5 * 2 + 0.25 * 3 + 0.19 * 4 + 0.06 * 5$$

$$= 2.81 \text{ bit/symbol}$$

∴ Efficiency of shannon - Fano code,

$$\eta = \frac{H}{L_{avg}}$$

$$= \frac{2.778}{2.81}$$

$$= 98.86\%$$

V Arithmetic coding :-

* Arithmetic coding by elias in 1963. Variable length coding M(d) which is used to reduce the coding redundancies present in an image.

* 2 changes happen.

1) The interval used to represent the message becomes smaller according to probability of each symbol

2) no. of information units required to represent the interval becomes larger.

Procedure :-

4 symbols $a_1, a_2, a_3, \& a_4$.

S1: message is assumed to occupy entire half-open interval [0,1]

S2: interval divided into 4 regions

Source symbol	Probability	Initial sub interval
a_1	0.2	[0.0, 0.2]
a_2	0.2	[0.2, 0.4]
a_3	0.4	[0.4, 0.8]
a_4	0.2	[0.8, 1.0]

S3: 1st symbol a_1 of the msg is narrowed to initial subinterval [0, 0.2]

S4: Interval [0, 0.2] subdivided to next symbol a_2

ie; min value + [Diff × subinterval] = new interval

$= 0 + [(0.2) * 0.2] = 0.04$

$0 + [(0.2 - 0) \times 0.4] = 0.08$. a_2 [0.04, 0.08]

S5: a_2 subdivided according probability a_3,

$0.04 + [(0.08 - 0.04) \times 0.8] = 0.072$

S6: Last msg symbol $a_4 = 0.04 + [(0.072 - 0.056) \times 0.4]$

$= 0.0624$

S7: Finally

$0.0624 + [(0.0688 - 0.0624) \times 1] = 0.0688$

$0.0624 + [(0.0688 - 0.0624) \times 0.8] = 0.06752$

Limitations:

1. Addition of end of msg indicator that is needed to separate one msg from another

2. use of finite Precision arithmetic.

Scaling strategy: renormalizes each subinterval to [0,1] range before subdividing

Rounding strategy: Represent the coding subintervals by preventing the transaction effects of finite Precision arithmetic

Image compression standard:

- Still image
  - Binary image compression
    1. CCITT Group 3
    2. CCITT " 4
    3. JBIG or JBIG 1
    4. JBIG 2
  - Continuous Tone still Image compression
    1. JPEG
    2. JPEG-LS
    3. JPEG-2000
- Video
  1. DV
  2. H 261
  3. H 262
  4. H 263
  5. H 264
  6. MPEG-1
  7. MPEG-2
  8. MPEG-4
  9. MPEG-4-AVC

JPEG standard :-

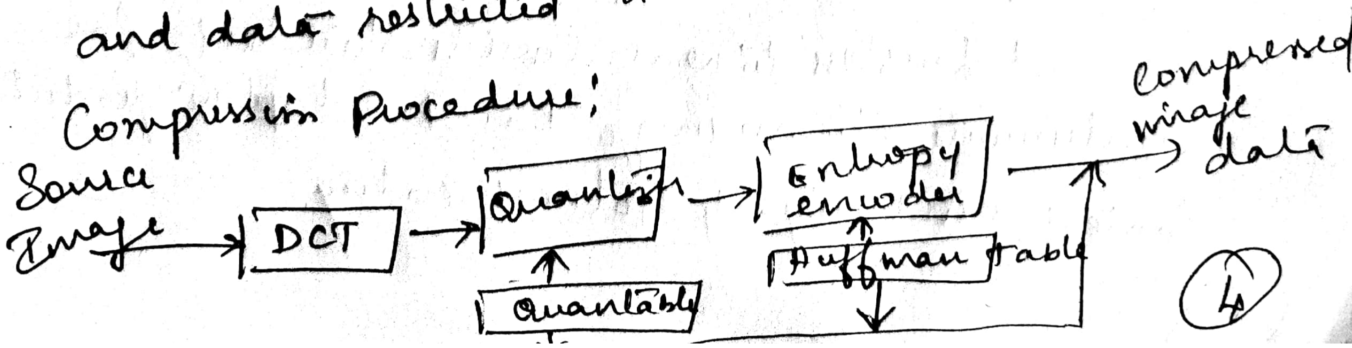
* It has 3 different coding system:

1. A lossy baseline coding system. — Based on DCT and is used in almost compression application
2. An extended coding system. — It is used in greater compression, higher Precision or progressive reconstruction applications
3. A lossless independent coding system. — It is used for reversible Compression

Baseline coding system :-

* Sequential baseline system based on Discrete cosine Transform (DCT), data precision 8 bits and data restricted 11 bits.

Compression Procedure:



3 sequential steps — DCT Computation
— Quantization
— Variable length code assignment

DCT computation :-

1) 1st image is subdivided into pixel blocks size 8x8
2) sub image proceed left to right, top to bottom
3) sub image as 8x8 = 64 pixels, 2^{k-1}, where 2^k is max no. intensity level
4) 2D discrete cosine Transform.

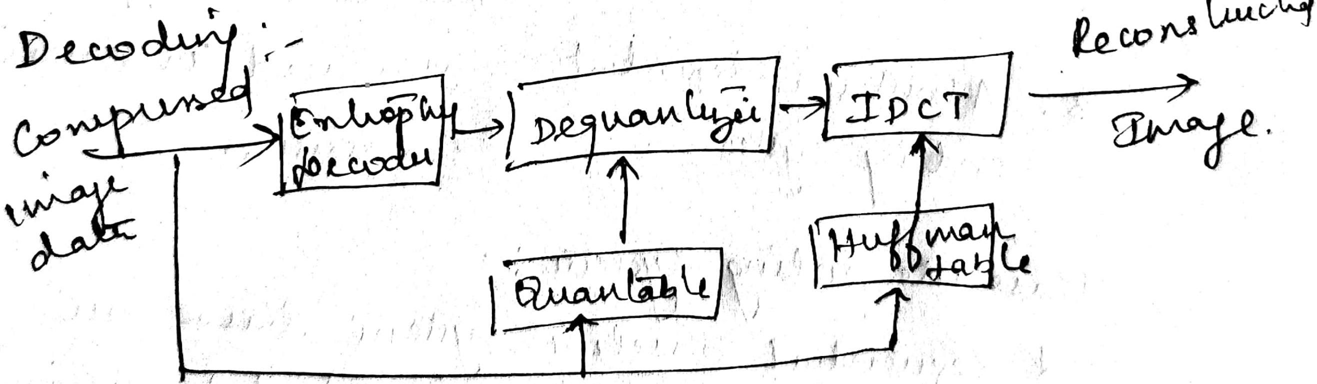
Quantization :-

$$\hat{T}(u,v) = \text{round}\left[\frac{T(u,v)}{Z(u,v)}\right]$$ (Zig Zag Pattern)

Variable length code assignment :-

non zero AC coefficients are coded using variable length code.
DC coefficient is difference coded relative to DC coefficient

Decoding :-



(f) JPEG-LS standard ;

* Lossless to near lossless standard for continuous to no image based adaptive prediction, Context modeling of Golomb coding.

② JPEG – 2000 standard:

* ↑sed flexibility
  1) Compression of continuous tone still images
  2) Access to compressed data.

<u>JPEG encoding</u>:-

S1: Level shifting

S2: optional Decorrelation

S3: Tiling

S4: Transformation

* Six sequential 'lifting' & 'scaling' operations.

$$Y(2n+1) = X(2n+1) + \alpha\left[X(2n) + X(2n+2)\right]$$

$$Y(2n) = X(2n) + \beta\left[Y(2n-1) + Y(2n+1)\right]$$

$$Y(2n+1) = Y(2n+1) + \gamma\left[Y(2n) + Y(2n+2)\right]$$

$$Y(2n) = X(2n) + \delta\left[Y(2n-1) + Y(2n+1)\right]$$

$$Y(2n+1) = -K \cdot Y(2n+1)$$

$$Y(2n) = Y(2n)/K.$$

S5: Quantization:

$$q_b(u,v) = sign\left[a_b(u,v)\right] - floor\left[\frac{|a_b(u,v)|}{\Delta_b}\right]$$

Quantization step size Δ_b,

$$\Delta_b = 2^{R_b - \varepsilon_b}\left(1 + \frac{\mu_b}{2^{11}}\right)$$

S6: co-efficient bit modeling

S7: Arithmetic coding

S8: Bit stream layering

S9: Packetizing

JPEG-2000 decoding:-

S1: Reconstruction of tile components.

S2: Dequantization

* M_b bit planes, N_b bit planes.

$$Rq_{v_b}(u,v) = \begin{cases} \left(\bar{q}_{v_b}(u,v) + r \cdot 2^{M_b - N_b(u,v)}\right) \cdot \Delta b \\ \left(\bar{q}_{v_b}(u,v) - r \cdot 2^{M_b - N_b(u,v)}\right) \cdot \Delta b \\ 0 \end{cases}$$

S3: Inverse Transformation

S4: Reconstruction for Tiles

S5: Inverse component Transformation

S6: DC level shifting.

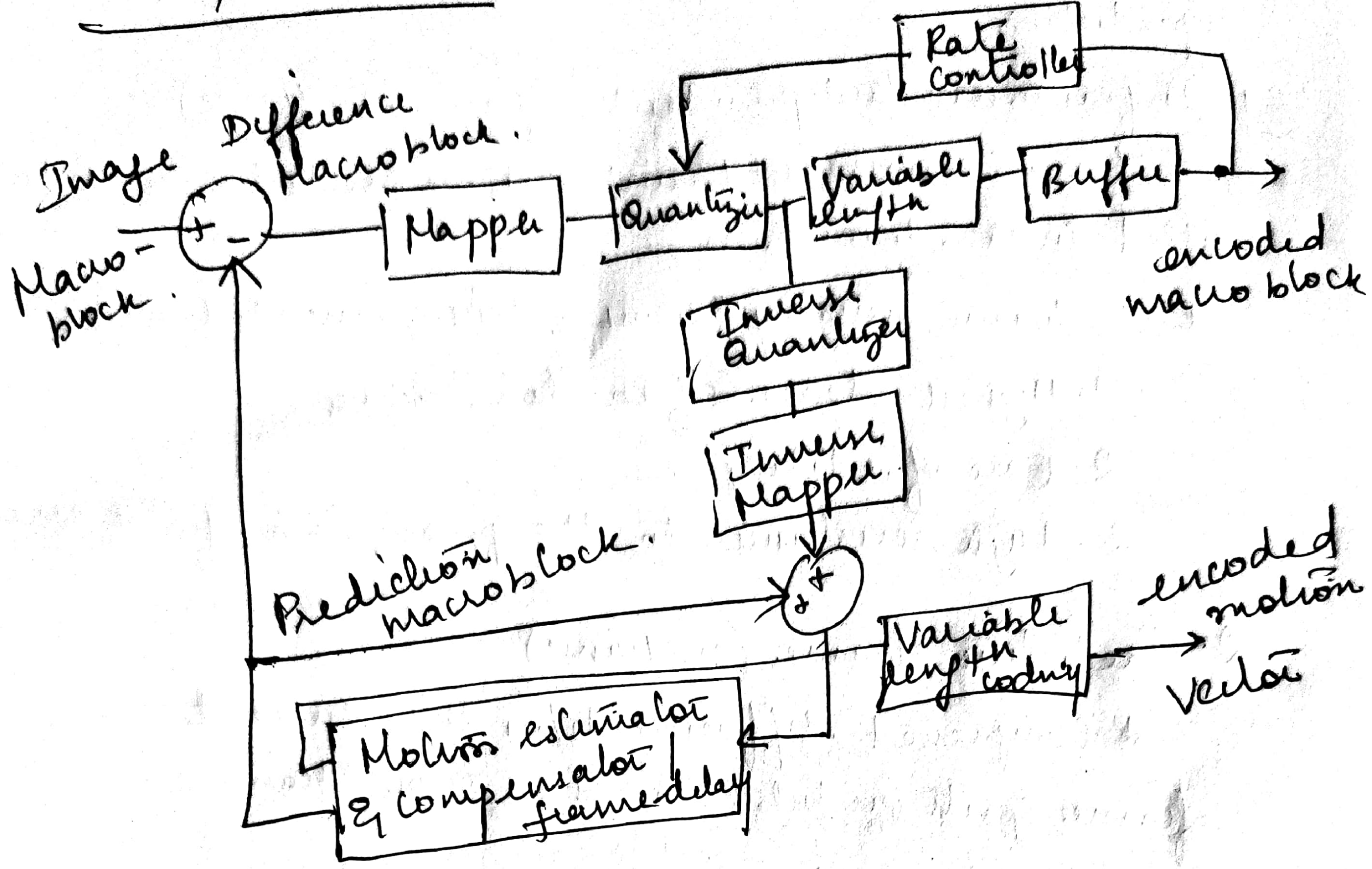
MPEG:- motion picture expert group

1) MPEG-1 → standard CD-ROM applications with video 1.5 Mb/s

2) MPEG-2 → MPEG1 → DVD with transfer rate 15Mb/ HDTV

3) MPEG-4 → * It is extension of MPEG-2.
* bit rate 5 to 64 kbits/sec.

MPEG encoder :-



* Based on a DPCM / DCT coding scheme.
  1. Redundancies within & b/w adjacent video frames
  2. Motion uniformity between frames
  3. The Psycho visual properties of the human visual system.
* 2 Kind of i/p can be given to encoder such as,
  1. conventional macro block of image data
  2. Difference b/w conventional image block & its Prediction based on previous / future video frames.
* 3 types of encoded O/p frame,
  1. I-frame
  2. P-frame
  3. B-Frame.

Decoding :-

Decoder accesses the areas of the reference frames that were used in encoder to form prediction.

1. Inframe (or) Independent frame (I-Frame)
   * Frame compressed without a prediction residual called Intra frames.
   * I-frame usually resemble JPEG encoded images.
     1. Highest degree of Random access
     2. Ease of editing
     3. High resistance to the propogation of Tx error.

2. Predictive Frame (P-Frame)
   * Compressed difference b/w the current frame & its prediction based on I or P frame.

3. Bi-directional frame:
   * Encoded frame based on subsequent frame.

Boundary Representation:-

New

* Representation deals with computation of segmented data into representations that facilitate the computation of descriptors.

(1) Boundary (Border) following:-
   * Algorithm require the pts of region in clockwise direction

Algm:- 1. We work with binary images with object & background 1 & 0.
2. Images are padded with borders of 0's to eliminate the possibility of object merging with image border.

Steps:

1. Starting pt b_0 be uppermost, leftmost point in the image labelled 1.

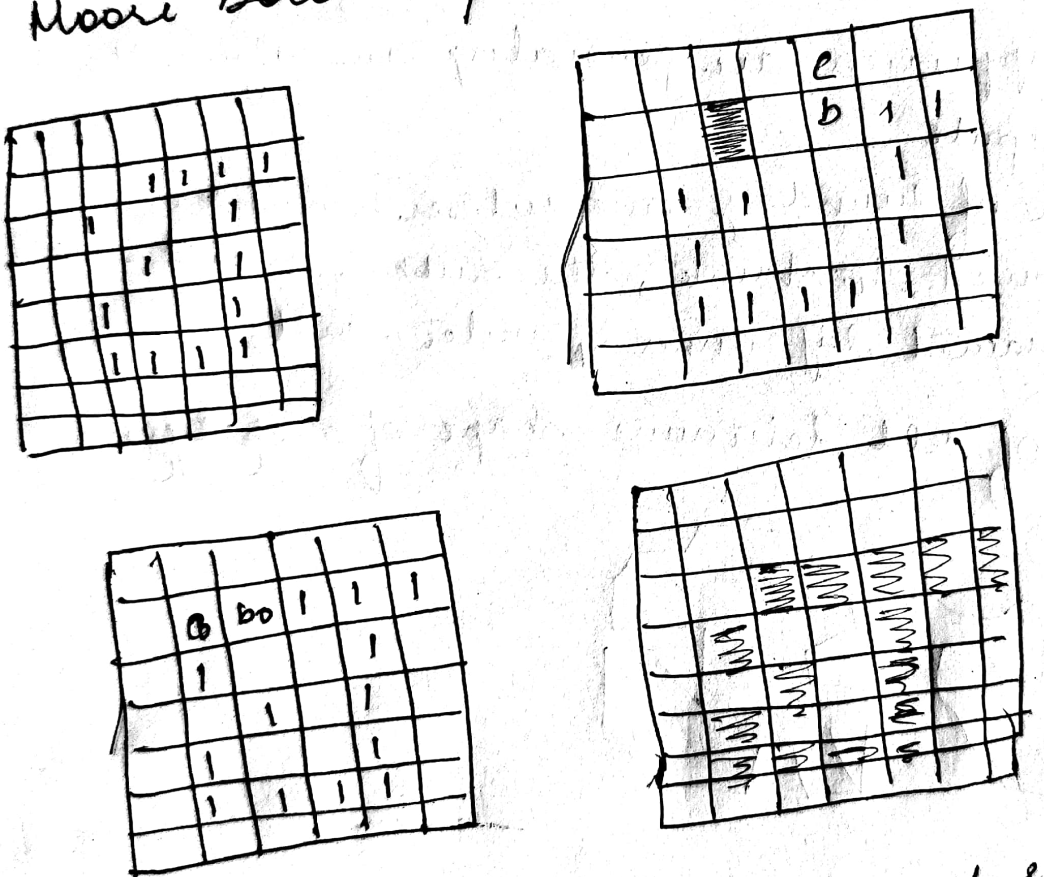
2. let $b = b_1$ & $c = c_1$

3. Let 8 neighbours of b, starting at c & preceding in clockwise direction, denoted $n_1, n_2 \ldots n_8$

4. $b = n_k$, $c = n_{k-1}$

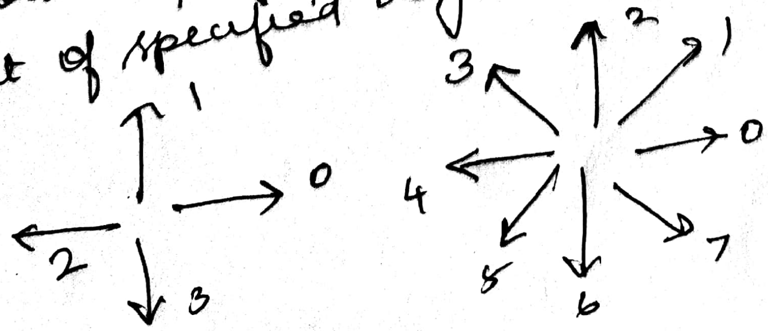
5. Repeat S3 & 4 until $b = b_0$

Moore boundary boundary tracking algorithm:-



② Chain codes.

* Boundary by connected sequence of straight line segment of specified length & direction.



2 Principal reasons

1. Resulting chain tends to be quite long

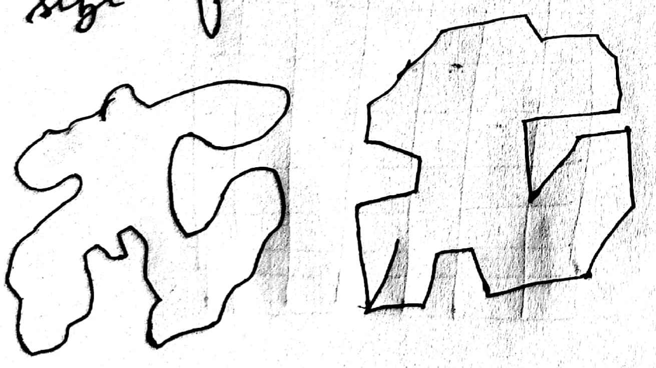
2. Any small disturbance along the boundary due to noise a imperfect segmentation can cause change in code.

⑤ Polygonal approximation using minimum Perimeter Polygons.

$$MPP = \sqrt{2d}.$$

MPP:-

* An approach are generating an algorithm to compute
* Think of boundary as a rubber band, will allowed to shrink, the rubber band
* allowed to shrink, the rubber band will be constrained by inner & outer wall.
* size of cell determine shape of a polygon.



IX Boundary description :-
* The boundary of a region using the features of the boundary.

* 2 types — simple descriptors
  — Fourier descriptors.

* 2 Quantities, 1. shape numbers
  2. statistical moments.

① simple Descriptors :-
* length of a boundary is simple descriptors.
The no. of pixels along a boundary gives a rough approximation of its length.

$$\text{length} = \left\{ \begin{array}{l} \text{no. of vertical} \\ \& \\ \text{Horizontal components} \end{array} \right\} + \sqrt{2} \{ \text{no. of diagonal components} \}$$

* diameter of boundary B,

$$\text{Diam}(B) = \max_{i,j} [D(p_i, p_j)]$$

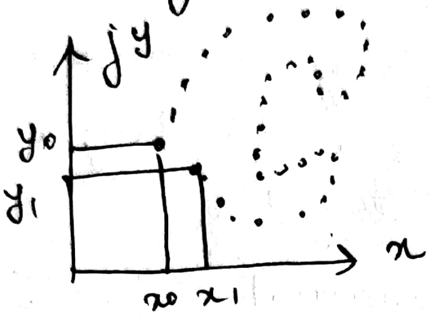
* major axis defined as line segment connecting the two extreme pts
* minor axis defined as line ⊥re to major axis
* Ratio of major axis as called efficiency of a boundary.

$$\text{Efficiency} = \frac{\text{Major axis}}{\text{Minor axis}}.$$

* curvature is defined as rate of change of slope.
Boundary in clockwise can be determined,

1. If change in slope at P is non-negative, Vertex pt P is said convex segment
2. change in slope P is negative → concave segment
3. If change is less than 10°, Vertex pt P.
4. If change is greater than 90°.

Fourier descriptors :-
* digital boundary it as a complex sequence.



* Arbitrary points (x_0, y_0), coordinate points,
  $(x_0, y_0)(x_1, y_1)(x_2, y_2) \ldots (x_{K-1}, y_{K-1})$.

* co-ordinates can be expressed.

$$x(k) = x_k$$
$$y(k) = y_k.$$

* sequence of co-ordinates
  $s(k) = [x(k), y(k)]$ for $k = 0, 1, 2 \ldots K-1$.

* co-ordinate pair can be treated as complex,
  $s(k) = x(k) + j\,y(k)$ for $k = 0, 1, 2 \ldots K-1$

* $s(k) = x(k) + j\,y(k)$ for $0, 1, 2 \ldots K-1$.
  x-axis treated as the real axis.

* DFT of $s(k)$,
$$a(u) = \sum_{k=0}^{K-1} s(k)\, e^{-j2\pi uk/K}$$

* Complex coefficients $a(u)$ called fourier descriptors.
* Inverse fourier Transforms (K).

$$s(k) = \frac{1}{K} \sum_{u=0}^{K-1} a(u) \cdot e^{j 2\pi uk/K} \quad \text{for}$$

$$k = 0, 1, 2 \ldots K-1.$$

* equivalent setting, $a(u) = 0$ for $u > P-1$

Basic properties, 1. Translation
2. Rotation
3. Scale changes
4. Starting point

1. Rotation :-

$$a_r(u) = \sum_{k=0}^{K-1} s(k) \cdot e^{j\theta} e^{-j 2\pi uk/K}.$$

$$= a(u) \cdot e^{j\theta}.$$

2. Translation :-

$$s_t(k) = s(k) + \Delta_{xy}$$

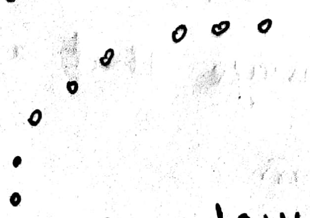
$$\Delta_{xy} = \Delta_x + j \Delta_y.$$

3. Starting pt :-

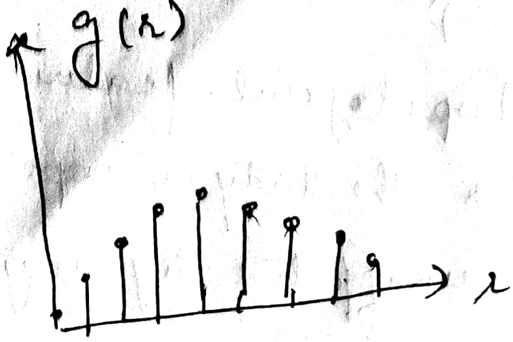
$$s_p(k) = s(k - k_0)$$

$$= x(k - k_0) + j\, y(k - k_0)$$

(4) Statistical moments



Boundary segment



* random variable,

$$\mu_n(r) = \sum_{i=0}^{k-1} (r_i - m)^n g(r_i)$$

$$m = \sum_{i=0}^{k-1} r_i \, g(r_i)$$

Regional descriptors :-

* Describe image region,

1. simple descriptor
2. Topological descriptors
3. Texture
4. Moment invariants

1. simple descriptors;

* no. of pixels in the region.

* 2 descriptors used to measure compactness of a region,

$$\text{Compactness} = \frac{(\text{Perimeter})^2}{\text{area}}.$$

* circularity ratio,

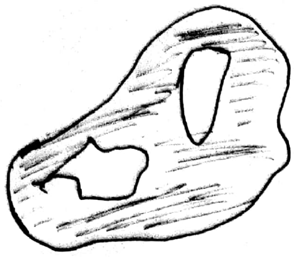
$$R_c = \frac{4\pi A}{P^2}$$ — area of region — length of perimeter.

* Value of measure 1 for circular region & $\frac{\pi}{4}$ for a square.

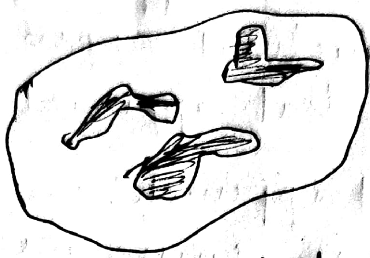
Topological features / Descriptors :-

* The study or properties of an image that are unaffected by any deformation

* Two Topological Properties,
1. no. of holes in region
2. no. of connected component of region



* A region of 3 connected components

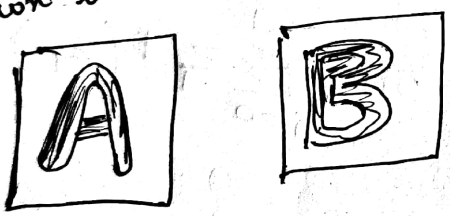


* no. of holes H and connected components C,

$\boxed{E = C - H}$ — Holes

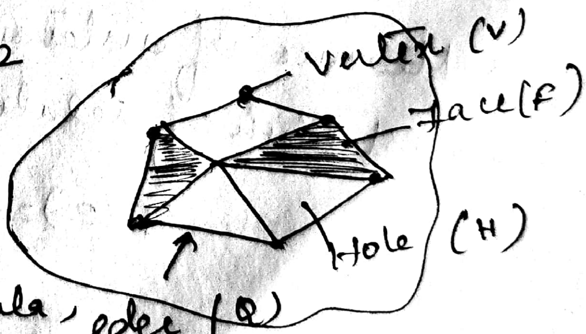
connected component

Region in euler no's equal to 0 & -1



For A, $E = C - H$, $C = 1, H = 1$

For B: $E = C - H$, $C = 1, H = 2$



* no. of edges by Q & no of faces by F called euler formula,

$V - Q + F = C - H$

$E = C - H$

$V - Q + F = E$

XII

Texture :-

Properties,

1. smoothness
2. Coarseness
3. Regularity.

3 principal approaches,

1. statistical approach
2. Structural approach
3. Spectral approach.

① Statistical approach :-

smooth, coarse, grainy.

1) Statistical moments :-

* z random variable,

$P(z_i)$ histogram for $i = 0, 1, 2 \ldots L-1$.

* nth moment of z, $\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n P(z_i)$

* m mean value of z, $m = \sum_{i=0}^{L-1} z_i P(z_i)$

2) second moment :-

* Measure of intensity contrast can be used to establish of relative smoothness

$$R(z) = 1 - \frac{1}{1+\sigma^2(z)}$$

$$\sigma^2(z) = \mu_2(z)$$

3rd moment :-

$$\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 P(z_i)$$

4th moment:-

$$\mu_4(z) = \sum_{i=0}^{L-1} (z_i - m)^4 P(z_i)$$

5th & Higher moments.

uniformity: $U(z) = \sum_{i=0}^{L-1} p^2(z_i)$

Entropy: $e(z) = -\sum_{i=0}^{L-1} P(z_i) \log_2 P(z_i)$

(1) Gray level co-occurrence matrix G

* Total no (n) of pixel Q is equal to sum of elements of G,

$$P_{ij} = \frac{g_{ij}}{n}.$$

Preceding eqn as,

$$m_r = \sum_{i=1}^{k} i P(i)$$

$$m_c = \sum_{j=1}^{k} j P(j)$$

$$\sigma_r^2 = \sum_{i=1}^{k} (i - m_r)^2 P(i)$$

$$\sigma_c^2 = \sum_{j=1}^{k} (j - m_c)^2 P(j)$$

(2) Structural approaches:-

* Arrangement of image primitives, such texture based on regularly spaced ||al lines.

$$A \rightarrow bS$$
$$S \rightarrow a$$

(3) Special approaches:-

* Fourier spectrum ideally suited for a Periodic or almost periodic 2D-patterns in an image.

* Texture description

1. Prominent peaks in the spectrum give the principal direction of texture patterns.
2. Location of peak in the frequency plane gives fundamental spatial period of patterns
3. Eliminating any periodic components via filtering leaves are simplified by non-periodic image elements.

freq. variable

$$S(r, \theta)$$

direction Variable

Spectrum fn

$$S(r) = \sum_{\theta = 0}^{\pi} S_\theta(r)$$

$$= \sum_{r=1}^{R_0} S_r(\theta)$$

(4) Moment Invariants :-

* 2D-moment of order $(p+q)$, digital image $f(x, y)$ of size $M \times N$,

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x,y)$$

* central moment of order $(p+q)$ is defined as,

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x,y)$$

XIII Pattern & Pattern classes :-

* Pattern classes are denoted $w_1, w_2 \ldots w_W$. where W is no. of classes.

* Pattern recognition by machine involves techniques for assigning patterns to their respective classes automatically

1. Vectors
2. Strings
3. Trees.

1) Pattern vectors :-

* Bold letters X, Y & Z, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

x_i = ith descriptor.

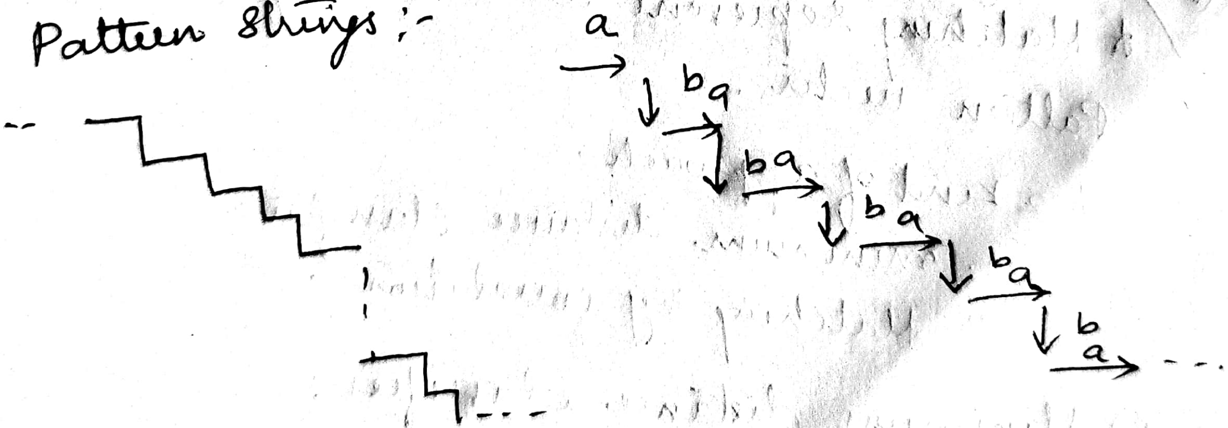
$n \rightarrow$ Total no. of descriptors.

$X = (x_1, x_2 \ldots x_n)^T$.

* 2-D pattern vector form $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

* x_1 & x_2 correspond to petal length & width.

2) Pattern strings :-



* Strings are 1-D structure, their application to image description require.

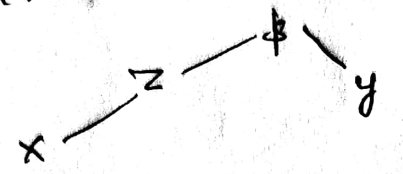
3) Pattern Tree :

* A Tree T is finite set of one or more nodes,

a) unique nodes $ designated the root

b) Remaining nodes in disjointed $T_1 \cdots T_m$, each

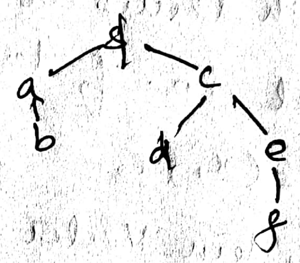
then is a tree called tree of T.



* 2 types of information.

1. Information about a node stored as a set of words describing the node.

2. Information relating a node to its neighbour, stored as a set of pointers to neighbour.



VI Recognition based on matching:-

* Matching represent each class by prototype Pattern vector.

* 2 kind of approach,

1. minimum distance classifier.

2. Matching by correlation.

1) Minimum distance classifier:

* Mean Vector of pattern, $m_j = \frac{1}{N_j} \sum_{x \in w_j} x_j \quad j = 1, 2 \dots W$

$N_j \rightarrow$ no. of pattern vectors

$W \rightarrow$ no. of pattern classes.

* Euclidean distance to distance to determine the closeness reduces the problem to computing the distance measures.

$$D_j(x) = ||x - m_j|| \text{ for } j = 1, 2 \ldots W$$

* smallest distance is equivalent to evaluating the fn's

$$d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \text{ for } j = 1, 2 \ldots W$$

* Decision boundary b/w classes W_i & W_j, min distance,

$$d_{ij}(x) = d_i(x) - d_j(x)$$
$$= x^T(m_i - m_j) - \frac{1}{2}(m_i - m_j)^T(m_i + m_j) = 0$$

② Matching By correlation:-

* Fourier Transform of fn's via convolution theorem, spatial correlation is related to Transform,

$$f(x,y) \star w(x,y) \Leftrightarrow F^*(u,v)\, W(u,v)$$

Spatial convolution, F^* complex conjugate of F.