Digital Image Processing
[DIP]

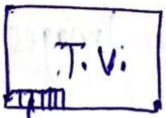
An Introduction to DIP ⇒

Image
- Analog
- Digital

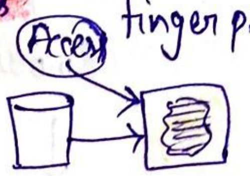
Capture → Process → Display

Process → Sampling

Process → Quantisation



Collection of pixels

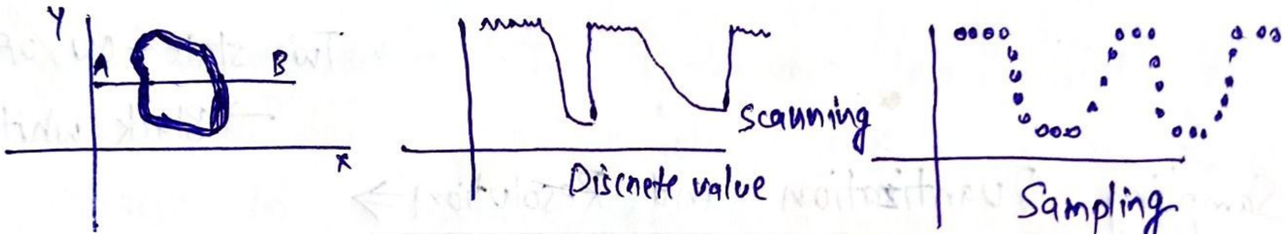


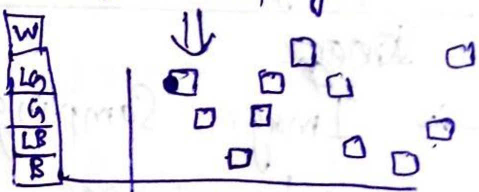
Pixel

RGB

millions

Sampling & Quantisation :-

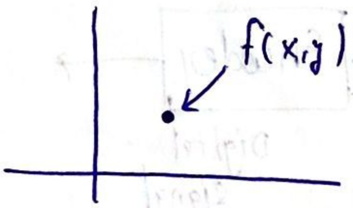




Quantisation

Image Representation :-

The image is 2-D light intensity function $f(x,y)$. It is certified in both spatial coordinates and brightness.



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

$$f(x,y) = r(x,y) * i(x,y)$$

where

$$0 < r(x,y) < 1 \quad \text{and} \quad 0 < i(x,y) < \infty$$

$r(x,y)$ = Reflectivity of surface of corresponding image point.

$i(x,y)$ = Intensity of incident light.

$$f(x,y) = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ \vdots & \vdots & & \vdots \\ f(M,1) & f(M,2) & \cdots & f(M,N) \end{bmatrix}$$

→ The no. of rows in DI is called Vertical Resolution.
→ The no. of columns in DI is called Horizental Resolution.

Type of Digital Images

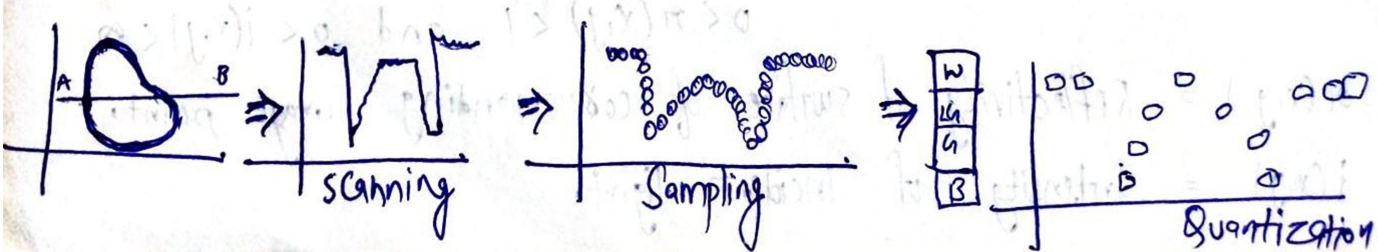
- Gray Scale → (0-255)
- Colour → RGB
- Binary → Only a single bit to represent each pixel. Two state → ON, OFF → Black white

* Sampling, Quantization and Resolution ⇒

→ Image Sampling & Quantization is used to convert analog image into Digital Image.
Digitizing coordinates value (x,y) is called sampling and digitizing the amplitude values (F) is called Quantization.

Analog Signal $X_a(t)$ → Sampler → Quantizer → Encoder →

Sampler: Discrete time signal; Quantizer: Quantized signal; Encoder: Digital signal



Resolution ⇒ Optical Resolution of Lens

Spatial Resolution

↓ no. of pixel → no. of bits

→ no. of bits necessary to encode pixel value is called <u>Bit Depth</u>

Bit depth is power of two.

No. of bit necessary to represent image is =

No. of rows × no. of columns × Bit depth

**
Steps In Image Processing :-

Image Acquisition :- Image is captured by a sensor & digitized if o/p of sensor is not already in digital form.

Image Enhancement :- To bring out details that are hidden or simple to highlight certain features of interest in an image.

Image Restoration :- Improving the appearance of an image.

Color Image processing :- Use the color of image to extract features of interest in an image.

Wavelets :- It is foundation of representing image in various degrees of resolution and used for image data compression.

Compression :- To reduce storage required to save an image or bandwidth required to transmit.

Morphological :- There would be transition from processes that output images, to processes that output image attributes.

Image Segmentation :- Partition an image into its constituent parts or objects.

Representation & Description :- Boundary
Region
(Make Diagram It is compulsory)

Recognition & Interpretation :- The process that assigns level to an object based on the information provided by its description.

Knowledge Base :- knowledge about a problem domain is coded into an image processing system in the form of a knowledge database.

Introduction to Color Image Representation ⟹

Color is a powerful descripton which simplifies object identification & extraction from a scene.

It is divided into 2 major categories:-

(i) Full color :- Eg- TV

(ii) Pseudo-color :- to a particular monochrome intensity or range of intensities. It is grayscale.

* Characteristic of light:-

Radiance (Watts-W) :- Total amount of energy coming out of light source.

Luminance (Lumens-lm) :- measure of amount of energy an observer perceives from light source.

Brightness (no unit) :- It is subjective measure that is practically impossible to measure.
It is corresponds to achromatic attribute of intensity.

* Primary and Secondary Colors:—

Blue = 435.8nm , Green = 546.1nm , Red = 700nm

Color Image Representation

Magneta = Red + Blue

Cyan = Green + Blue

Yellow = Red + Green

3 attributes of color:—

→ Luminance (brightness)
→ Chrominance – Hue & Saturation
→ Represented by a "color cone"

RGB	vs	CMY
Magneta = Red + Blue		Magneta = White – Green
Cyan = Blue + Green		Cyan = White – Red
Yellow = Green + Red		Yellow = White – Blue

Color Representation Models:—

→ Three primary colors:- RGB, CMY, XYZ
→ Luminance & chrominance:- HSI , YIQ , YC_bC_r

HSI: Hue, Saturation, Intensity

Pseudo color :- To artifically assign colors to a grey scale.

→ In intensity slicing, assigns a shade of color to all grey levels that fall under a specified value and a different shade of color.
→ Majority of techniques perform a grey level to color transformation.
→ Feed this to 3 color inputs (RGB) of a color monitor.

* Conversion of RGB to HSI model ⇒

The H is given by

(Hue) $$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

where

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

S is given by

(Saturation) $$\boxed{S = 1 - \frac{3}{R+G+B}\,[\min(R,G,B)]}$$

(Intensity) I is given by $$\boxed{I = \frac{1}{3}(R+G+B)}$$

All RGB values are normalized to range [0, 1].

* Conversion of HSI to RGB model →

If $0^\circ \leq H < 120^\circ$

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

$$R = I\left(1 + \frac{S\cos H}{\cos(60^\circ - H)}\right)$$

$$G = 3I - (R+B)$$

if $120^\circ \leq H < 240^\circ$

→ $H = H - 120^\circ$

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

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

$$B = 3I - (R+G)$$

if $240^\circ \leq H \leq 360^\circ$: → $H = H - 240^\circ$

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

$$B = I\left(1 + \frac{S\cos H}{\cos(60^\circ - H)}\right)$$

$$R = 3I - (R+G)$$

YIQ Color Coordinate System

Y → Luminance
I, Q → chrominance

Conversion of RGB to YIQ ⇒

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Conversion of YIQ to RGB ⇒

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0.956 & 0.621 \\ 1.0 & -0.272 & -0.649 \\ 1.0 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

YUV / YCbCr Coordinate :-

YUV is color coordinate used in color TV in PAL system.

YCbCr is digital equivalent of YUV used for digital TV, with 8 bit for each component, in range of 0–255

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0.000 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0.000 \end{bmatrix} \begin{bmatrix} Y - 16 \\ C_b - 128 \\ C_r - 128 \end{bmatrix}$$

Image Processing Domains → (Image Enhancement) *

brightness, contrast, appearance

(i) **Spatial Domain** :- It refers to image plane itself & image processing is based on direct manipulation of pixels in an image.

Categories of spatial Domain

(a) **Intensity transformations** :- It operates on single pixels of an image for purpose of contrast manipulation & image thresholding.

(b) **Spatial filtering** :- It deals with performing operations, such as image sharpening, by working in a neighborhood of every pixel in an image.

(ii) **Frequency Domain** :- we transfrom an into transform domain (fourier) doing processing there, and obtaining the inverse transform to bring the results back into spatial domain.

$$\boxed{g(x,y) = T(f(x,y))}$$

$g(x,y)$ → o/p image; $f(x,y)$ → i/p image

Input image $f(x,y)$ ⇒ preprocessing ⇒ Fourier Transform ⇒ Filter Function ⇒ Inverse fourier Transform ⇒ Post Processing ⇒ $g(x,y)$ Enhance Image

* Intensity Transformation :- It operates on single pixels of an image for purpose of contrast manipulation & image thresholding.

$$\boxed{S = T(r)}$$

S → o/p intensity of pixel after trans.

r → input intensity of pixel before trans.

(A) Linear (negative & identity) transformation

(B) Logarithmic (log & inverse log) transformation

(C) Power-Law (nth power & nth root) transformation

Linear Transformation :-

(A) Image Negatives ⇒ Negative of an image with intensity levels in range $[0, L-1]$ is obtained by using negative transformation

$$\boxed{S = L-1-r}$$

If $r=0$, $S = L-1$
If $r = L-1$, $S=0$

(B) Identity transformation ⇒

Intensity level of the image after transmission is identical to that of the original image.
Each value of input image is directly mapped to each other value of o/p image. [for enhancing white on gray detail embedded in dark areas of an image]

Log Transformation :-

The low intensity values are mapped into higher intensity values.

$$\boxed{S = c\log(1+r)}$$

By using this, we can compress or expand gray level.

→ The inverse log transformation is opposite to log transform so will produce low contrast image.

Power Law Transformation :-

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

C & γ are +ve constants.

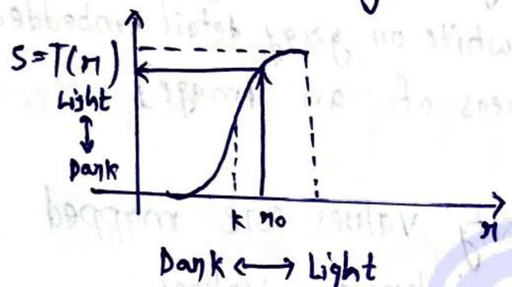
→ We use gamma transformation where we need to expand or compress darker region.

→ If $\gamma = 1$, No change
→ If $\gamma > 1$, It provides compression of brighter region.
→ If $\gamma < 1$, It provides expansion of darker region.

Image Enhancement Technique

Contrast stretching

- Used to Enhance quality of image
- $r<k$, it will be mapped to narrow range of intensity levels towards dark region.
- To increase dynamic range of modified image.
- $r>k$, it will be mapped towards brighter region.



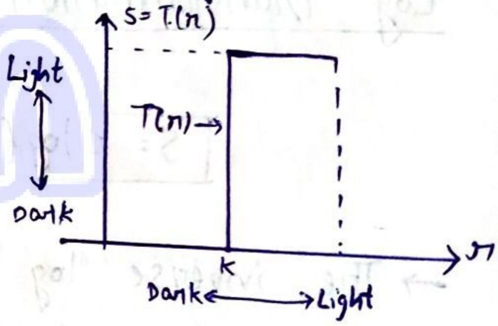
- It expands range of intensity levels in image so that it spans the full intensity range of recording medium or display device.

Image thresholding

It is applied once to an input image it gives Binary o/p image. Based on value of k.

$r<k$, assigned with black intensity level.

$r>k$, assigned with white intensity level.



Histogram Equalization

Histogram is a graph showing no. of pixels in an image at each different value found in that image.

Eg – 8 bit grayscale image & 256 different possible intensities

Histogram Equalization:— This method is to boost the global contrast of an image it look more visible.

$$h(v) = \text{round}\left(\frac{cdf(v) - cdf_{min}}{(M \times N) - cdf_{min}} \times (L-1)\right)$$

CDF → cumulative distribution function

L → max. intensity value (256)

M → image width, N → image height

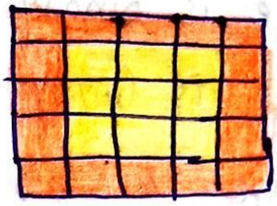
h(v) → equalized value.

Spatial Filtering:-

Spatial Filtering is just moving the filter mask from point to point in an image.

The filter mask may be 3x3 or 5x5 or 77.

eg- 3x3 mask in a 5x5 image



types of Filter

- Linear Spatial Filter
- Nonlinear Spatial Filter

If each pixel in an image can be replaced with constant value then it is linear otherwise nonlinear.

Smoothing Spatial Filter ⇒ It is used for blurring & for noise reduction.

Sharpening Spatial Filter ⇒ To highlight the transactions in intensity.

eg- medical images, military systems etc.

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

Q: RGB values (29, 98, 128) convert into CMY.

Ans- $R = \frac{29}{255} = 0.113$, $G = \frac{98}{255} = 0.384$, $B = \frac{128}{255} = 0.501$

C = Blue+Green | White - Red

M = Red+Blue | White - Green

Y = Green+Red | White - Blue

White = 1

Cyan = 1 - 0.113 = 0.887

Magneta = 1 - 0.384 = 0.616

Yellow = 1 - 0.501 = 0.499

RGB(0.113, 0.384, 0.501)

↓

CMY (0.887, 0.616, 0.499)

Q. Convert RGB values to HSI in range [0-1]

$R=24, \quad G=98, \quad B=118$

Ans-

$R = \frac{24}{255} = 0.09, \quad G = \frac{98}{255} = 0.38, \quad B = \frac{118}{255} = 0.46$

HSI

$I :- \frac{1}{3}(R+G+B) = \frac{1}{3} \times (0.09 + 0.38 + 0.46)$

$$\boxed{I = \frac{0.93}{3} = 0.31}$$

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

$S = 1 - \frac{3}{0.09+0.38+0.46} \min(0.09, 0.38, 0.46)$

$S = 1 - \frac{\cancel{3}}{\cancel{0.93}} \times 0.09$

$S = 1 - \frac{0.09}{0.31} = 1 - 0.29 = 0.71$

$$\boxed{S = 0.71}$$

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

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

$\theta = \cos^{-1}\left(\frac{\frac{1}{2}((0.09-0.38)+(0.09-0.46))}{\sqrt{(0.09-0.38)^2 + (0.09-0.46)(0.38-0.46)}} \right)$

$\theta = \cos^{-1}\left(\frac{\frac{1}{2}((-0.29)-0.37))}{\sqrt{(-0.29)^2 + (-0.37)(-0.08)}} \right) = \cos^{-1}\left(\frac{-0.33}{\sqrt{0.1137}} \right)$

$\theta = \cos^{-1}\left(\frac{-\cancel{0.33}}{\cancel{0.33}} \right) = \cos^{-1}(-1) = 180°$

$H = 360° - \theta = 360° - 180° = \boxed{180° = H}$

Q. An image of 512 × 512 dimension. The no. of bits needs to store the file size of this image.
(a) If image is binary
(b) If image is Gray
(c) If image is RGB

Ans- (a) Binary:- 1 pixel = 1 bit

$$512 \times 512 \times 1 \text{ bit} = \frac{512 \times 512}{8} \text{ byte}$$

$$= 32768 \text{ byte}$$

$$\frac{32768}{1024} = 32\text{KB} \rightarrow \frac{32}{1024} = 0.03125\text{MB}$$

(b) Gray:- 1 pixel = 8 bit

$$512 \times 512 \times 8 \text{ bit} = \frac{512 \times 512 \times 8}{8} = \frac{264144 \text{ byte}}{1024}$$

$$256\text{KB} = 0.25\text{MB}$$

(c) RGB:-

$$1 \text{ pixel} = 8 \times 3 \text{ bit}$$

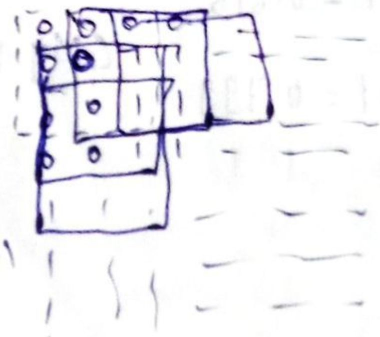
$$\frac{512 \times 512 \times 8 \times 3}{8} \text{ byte} = \frac{786432 \text{ byte}}{1024}$$

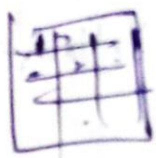
$$\frac{786\text{KB}}{1024} \rightarrow 0.75\text{MB}$$

Q. Convert RGB to HSI values to range [0-1] R=24, G=98, B=118

Spatial Operations :-

3×3





$0 + 0 + \cdots 7 + 1$

$= \frac{8}{9}$

$= 0.8 \rightarrow 1$

Fourier Transformation ⇒ It is an image processing tool which is used to decompose an image into its sine and consine components.

→ The o/p of transformation represents image in fourier or frequency domain, while i/p image is spatial domain.

- Continuous & Discrete Fourier Transform
- Properties of fourier transform

Fourier transformation:-

$$F(f(x)) = F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux} . dx \quad , j = \sqrt{-1}$$

Inverse fourier transformation:-

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

where x is time in seconds, units of u are Hz or cycles/sec.

Fourier transformation Pair

$F(u) \rightarrow$ fourier transform of signal $f(x)$

$F(x) \rightarrow$ Original Signal

Complex function $F(u) = R(u) + j\, I(u)$

Magnitude of FT :- $|F(u)| = \sqrt{R^2(u) + I^2(u)}$

Phase :- $\phi(F(u)) = \tan^{-1} \frac{I(u)}{R(u)}$

Magnitude - Phase:- $F(u) = |F(u)|\, e^{j\phi(u)}$

Power of $f(x)$:- $|F(u)|^2 = R^2(u) + I^2(u)$

Fourier Transformation (2-D continuous signal) :-

FT:-

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

Inverse FT:-

$$f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u,v)\, e^{j2\pi(ux+vy)}\, du\, dv$$

Spectrum:- $|F(u,v)| = \left[R^2(u,v) + I^2(u,v)\right]^{2}$

phase angle:- $\phi(u,v) = \tan^{-1}\left(\frac{I(u,v)}{R(u,v)}\right)$

Power Spectrum:-

$$P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v)$$

Discrete Fourier Transform (DFT) ⇒

Forward DFT:-

$$F(k) = \sum_{k=0}^{N-1} f(x)\, e^{-j2\pi kx/N}, \quad 0 \le k \le N-1$$

IDFT :-

$$F^{-1}(F(k)) = F(n) = \frac{1}{N} \sum_{k=0}^{N-1} F(k) \cdot e^{j\frac{2\pi}{N}kn}, \quad 0 \le n \le N-1$$

2D DFT :-

$f(x,y) \xrightarrow{DFT} F(u,v)$

DFT:- $$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\, e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}$$

IDFT:- $F(u,v) \xleftrightarrow{IDFT} f(x,y)$

$$f(x,y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v)\, e^{j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}$$

u & v → transform grey variables

x & y → spatial Image variables

Properties of Fourier Transformation:-

(i) Seperability
(ii) Translation
(iii) Periodicity
(iv) Conjugate → To visualize fourier Spectrum
(v) Rotation → To ROtate
(vi) Distributive → addition, not on multiplication
(vii) Scaling
(viii) Convolution → $f(x) \cdot g(x) \Longleftrightarrow f(u) \cdot g(u)$
(ix) Corelation → $f(x,y) \odot g(x,y) \Longleftrightarrow f^*(u,v) \circledast \cdot G(u,v)$

* **1-D Wavelet Transform :-** Wavelets are often used to denoise 2 dimensional signals. It is a time-frequency analysis method which selects the appropriate frequency band adaptively based on characteristics of the signal.

$$f(x) = \sum_K C_{j_0}(K)\, \phi_{j_0,K}(x) + \sum_{j=j_0}^{\infty} \sum_K d_j(K)\, \psi_{j,K}(x)$$

Scaling coefficients:

$$C_{j_0}(K) = \int f(x)\, \phi_{j_0,K}(x)\, dx$$

wavelet coefficients: $d_j(K) = \int f(x)\, \psi_{j,K}(x)\, dx$

DWT :- $$f(n) = \frac{1}{\sqrt{M}} \sum_K W_\phi(j_0,K)\, \phi_{j_0,K}(n) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_K W_\psi(j,K)\, \psi_{j,K}(n)$$

Scaling coefficients $W_\phi(j_0,K) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n)\, \phi_{j_0,K}(n)$

Details coefficients $W_\psi(j,K) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n)\, \psi_{j_0,K}(n)$ for $j \geq J_0$

$M = 2^J$, $J_0 = 0$

2-D Wavelet Transform:- We need 2D scaling function & 3 2-D wavelet functions are required.

$$\phi(x,y) = \phi(x)\,\phi(y)$$

variations along columns $\psi^H(x,y) = \psi(x)\,\phi(y)$

variations along rows $\psi^V(x,y) = \phi(x)\,\psi(y)$

variations along diagonals $\psi^D(x,y) = \psi(x)\,\psi(y)$

$$W_\phi(j_0,m,n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)\,\phi_{j_0,m,n}(m,n)$$

$$W_\phi^i(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)\,\psi^i_{j_0,m,n}(m,n), \quad i = \{H,V,D\}$$

Inverse:-

$$f(m,n) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\phi(j_0,m,n)\,\phi_{j_0,m,n}(m,n) + \frac{1}{\sqrt{MN}} \sum_{i=H,D,V} \sum_{j=j_0}^{\infty} \sum_m \sum_n W_\psi^i(j,m,n,k)\,\psi^i_{j,m,n}(m,n)$$

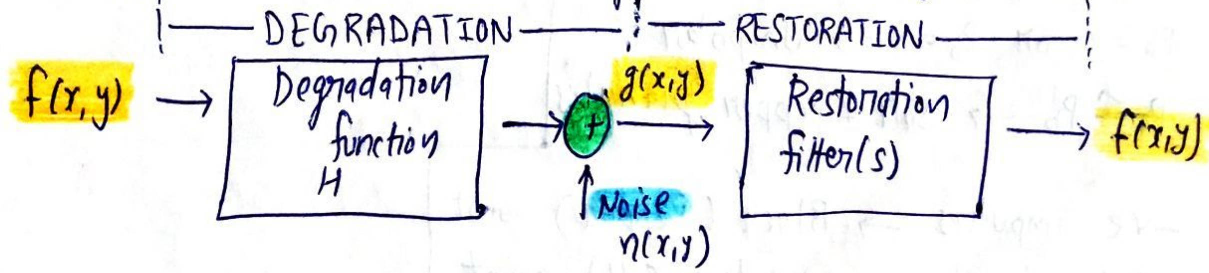
$j_0 = 0, \quad N = M = 2^J$

$m = n = 0,1,2 \ldots M-1, \quad j = 0,1,2 \ldots J-1, \quad k = 0,1,2 \ldots 2^J-1$

Image Degradation & Restoration:-

→ Degradation comes in many forms such as motion blur, noise and camera misfocus.

→ Restoration tech. is to improve an image in some predefined sense. The purpose of image restoration is to "compensate for" or "undo" defects which degrade an image.



$$g(x,y) = h(x,y) * f(x,y) + n(x,y)$$

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

Noise :- It is random variation of brightness or color info in image captured.

(A) Sources of image Noise :- Sending, Sensor heat, ISO factor

Types of Noise Models $\Rightarrow$ Probability density function (PDF)

(i) Gaussian Noise: [Normal noise model]

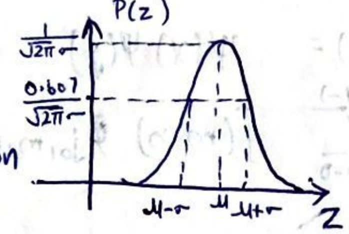
$$P(z) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(z-\mu)^2/2\sigma^2}$$

$z \rightarrow$ gray level

$\mu \rightarrow$ mean of z

$\sigma \rightarrow$ standard deviation

$\sigma^2 \rightarrow$ variance



70% of values $\rightarrow [(\mu-\sigma), (\mu+\sigma)]$

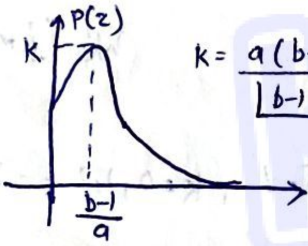
$m = \frac{a+b}{2}$, $\sigma^2 = \frac{(b-a)^2}{12}$

(ii) Gamma Noise:-

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

mean: $\mu = b/a$

variance: $\sigma^2 = b/a^2$



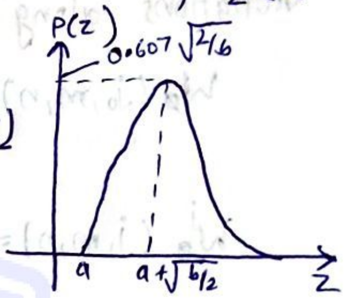
$$K = \frac{a(b-1)^{b-1}}{(b-1)!} e^{-(b-1)}$$

(iii) Rayleigh Noise:-

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

mean: $\mu = a + \sqrt{\pi b/4}$

Variance: $\sigma^2 = \frac{b(4-\pi)}{4}$

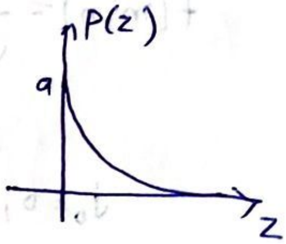


(iv) Exponential Noise:-

$$P(z) = \begin{cases} a \cdot e^{-az}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

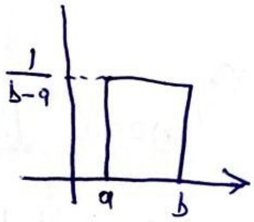
$\mu = 1/a$

$\sigma^2 = \frac{1}{a^2}$



(v) Uniform Noise:-

$$P(z) = \begin{cases} \frac{1}{b-a}; & a \leq z \leq b \\ 0; & \text{otherwise} \end{cases}$$

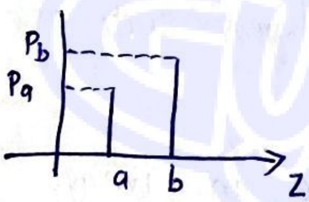


$\mu = (a+b)/2$

$\sigma^2 = \frac{(b-a)^2}{12}$

(vi) Salt & Pepper Noise [Impulse Noise]:-

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



if $b > a \Rightarrow$ graylevel $b \rightarrow$ light dot

gray level $a \rightarrow$ dark dot

$P_a = 0$ or $P_b = 0 \rightarrow$ unipolar

$P_a \simeq P_b \rightarrow$ Salt & pepper granules

$-$ve impulses $\rightarrow$ Black (Pepper) point

$+$ve impulses $\rightarrow$ White (Salt) point

8 bit = $a = 0$ (Black)

$b = 255$ (white)

Noise Filtering ⇒

we know

$$g(x,y) = h(x,y) \cdot f(x,y) + n(x,y)$$

degradation → only due to noise

$$g(x,y) = f(x,y) + n(x,y)$$

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

Ay – 2

(a)

4

Image Noise Filters

- **Time Domain**
  - Mean
  - Median
  - Gaussian
  - Bilateral
  - Weiner
- **Frequency Domain**
  - Fourier Transform
  - Wavelet transform

(i) Mean Filter ⇒ Replace the center value with average of all pixel values in the window

5	3	6
2	1	9
8	4	7

⟶

	5	

(ii) Bilateral Filter ⇒ It uses Gaussian filter but it has one more multiplicative component which is a function of pixel intensity difference.

(iii) Gaussian Filter ⇒ It is widely used effect in graphics s/w typically to reduce image noise & reduce detail.

(iv) Median :- We replaces pixel value with the median value.

6	2	0
3	97	4
19	3	10

⟹

0	2	3
3	4	

(v) Weiner :- To reduce mean square error (MSE) as much as possible

Inverse Filtering :- It is to recover the original image from the blurred image.

We explore 2 methods of inverse filtering -

→ a thresholding method

→ An Iterative Method

Method 1: Thresholding :-

$$g(n_1, n_2) = f(n_1, n_2) ** b(n_1, n_2)$$

$$f(n_1, n_2) = g(n_1, n_2) ** h(n_1, n_2)$$

Method 2: Iterative Procedure :-

$$f_0(n_1, n_2) = \lambda g(n_1, n_2)$$

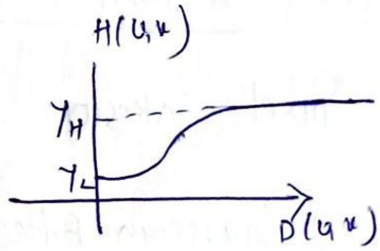
$$\hat{f}_{K+1}(n_1, n_2) = \hat{f}_K(n_1, n_2) + \lambda(g(n_1, n_2) - \hat{f}_K(n_1, n_2) ** b(n_1, n_2))$$

$$\hat{F}_K(\omega_1, \omega_2) = \lambda G(\omega_1, \omega_2)[1 + (1 - \lambda B(\omega_1, \omega_2)) + \dots + (1 - \lambda B(\omega_1, \omega_2))^K]$$

$$= \frac{G(\omega_1, \omega_2)}{B(\omega_1, \omega_2)}\left[1 - (1 - \lambda B(\omega_1, \omega_2))^{K+1}\right]$$

* Homomorphism filtering :-

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



$0 < i(x, y) < \infty$ ↑ illumination

$0 < r(x, y) < 1$ ↑ Reflectance

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

DFT :- $Z(u, v) = F_i(u, v) + F_r(u, v)$

H(u,v) :- $S(u, v) = H(u, v) Z(u, v)$

$(DFT)^{-1}$:- $s(x, y) = i(x, y) + r(x, y)$

exp :- $g(x, y) = e^{s(x, y)} = i_0(x, y)\, r_0(x, y)$

$f(x, y) \Rightarrow \boxed{\ln} \Rightarrow \boxed{DFT} \Rightarrow \boxed{H(u,v)} \Rightarrow \boxed{(DFT)^{-1}} \Rightarrow \boxed{\exp} \Rightarrow g(x, y)$

Image Compression :— *

→ Image compression refers to the process of redundancy amount of data required to represent the given quantity of information for digital image. The basis of reduction process is removal of redundant data. JPEG → pixel

Compression Methods

(i) Run Length Encoding (RLE)
(ii) Arithmetic coding
(iii) Huffman coding
(iv) Transform coding

Original data N1 bits → Compressed Data N2 bits

$f(x,y)$ → Source Encoder → Channel Encoder → Channel → Channel Decoder → Source Decoder → $\hat{f}(x,y)$

Source Encoder ⇓ Removes redundancy

Channel Encoder ⇓ increases noise immunity

fig:- Image Compression Model

Stages of Encoding:-

$f(x,y)$ → Mapper → Quantizer → Symbol Encoder → $\hat{f}(x,y)$

Mapper ⇒ Reduces interpixel redundancy, reversible operation

Quantizer ⇒ Reduces Phycovisual redundancy, non-reversible operation

Symbol Encoder ⇒ Create a fixed or variable length code, reduces coding redundancy
Reversible operation

Source Decoder :—

channel → Source Decoder → Inverse Mapper → $\hat{f}(x,y)$

Redundancy ⇒ Redundancy means repetitive data.

- Redundancy
  - Statistical
    - inter-pixel
      - Spatial → Correlation b/w neighboring pixel values
      - Temporal → Correlation b/w adjacent frames in a sequence of image in video
    - Coding
  - Psychovisual

(i) **Coding Redundancy** :- • It happens due to poor selection of coding technique.

- Wrong choice of coding technique create unnecessary additional bits. These extra bits are called redundancy.

$$\boxed{\text{Coding Redundancy} = \text{Average bits used to code} - \text{Entropy}}$$

$$CR = \sum_{K=0}^{n} \underbrace{l(r_K)}_{\text{length of code}} \cdot \underbrace{p(r_K)}_{\text{probability}} - \underbrace{H}_{-\sum_{i=1}^{n} P_i \log_2(P_i)}$$

r_K → greylevel

$$\boxed{CR = \sum_{K=0}^{n} l(r_K) \cdot P(r_K) - \sum P_i \log_2\left(\frac{1}{P_i}\right)}$$

Inter-pixel Redundancy :-

The value of any given pixel can be predicted from the value of its neighbors that is they are highly correlated.

→ It is solved by algorithm like : Run length coding, Bit Plane algorithm etc.

Psycho-visual Redundancy :-

- It exists because human perception does not involve quantitative analysis of every pixel in the image.
- It is distinctly vision related and its elimination does result in loss of information.

Compression Algorithm

→ It is to reduce the source data to a compressed form and decompress it to get the original data.

Types of CA :-

(i) Lossless Compression :- Reconstructed data is identical to the original data.

- It means that reducing the size of an image without any quality loss. This is achieved by removing unnecessary metadata from JPEG and PNG files.

Ex techniques :-
- Huffman coding
- Arithmetic coding
- Shanon-fano coding

(ii) <u>Lossy compression</u> :— Reconstructed data is an approximation to the original data.

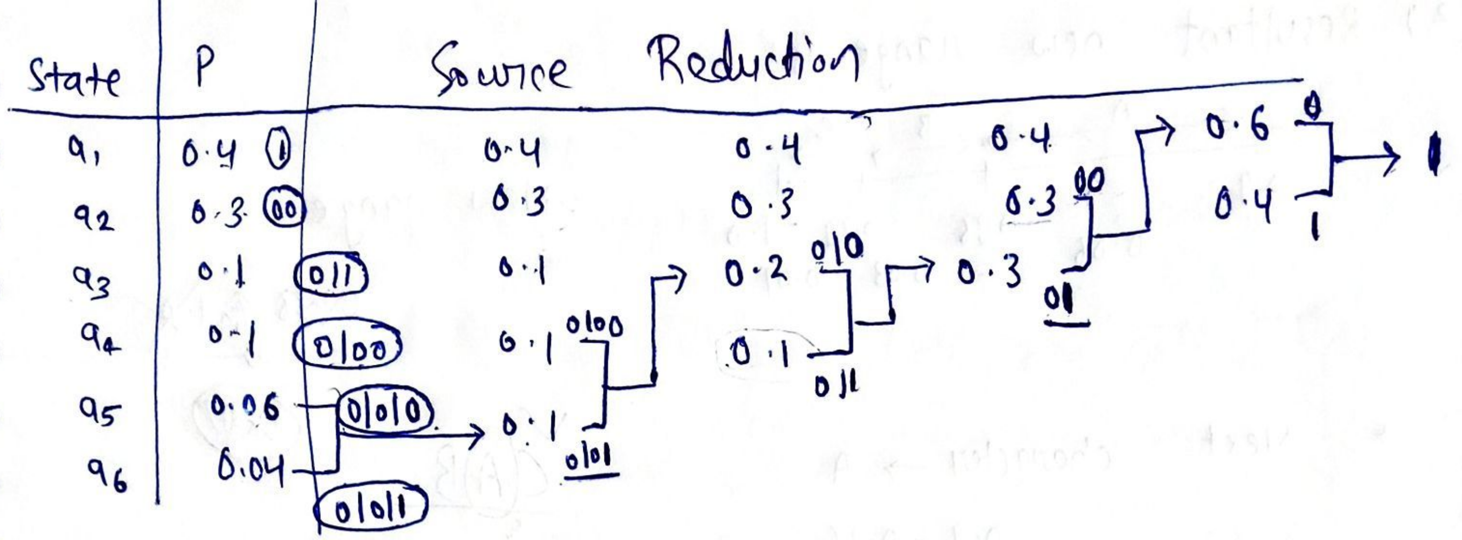
Ex techniques :- Linear Prediction
Transform coding

Lossless Compression vs Lossy Compression

Lossless Compression	Lossy Compression
→ It does not eliminate the data which is not noticeable.	→ It eliminates data which is not noticeable.
→ file can be restored in its original form.	→ File does not restore in its original form
→ Quality doesn't compromised.	→ Data quality is compromised.
→ It does not reduce the size.	→ It reduces the size of data.
→ Eg- Run length coding, Lempel-Ziv coding, Haffman coding	→ Eg- Transform coding, discrete cosine, wavelet.
→ It is used in text, images, sound.	→ It is used in images, audio, video.

Huffman coding :— It reduces an average no. of bits/pixel

- It assigns variable length bits to different symbols
- Achieves compression in 2 steps.
  - Source Reduction
  - Code Assignment



0.1 Average $L = \sum P_K l_K = 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 3 + 0.1 \times 4 + 0.06 \times 5 + 0.04 \times 5 =$ ______ bit/symbol

Entropy $H(x) = \sum_{k=1}^{?} P_K \log_2 \frac{1}{P_K} = 0.4 \times \log_2 \frac{1}{0.4} + 0.3 \times \log_2 \frac{1}{0.3} +$ ______

Coefficient $\eta = \frac{H}{L} \times 100 =$ ______ %

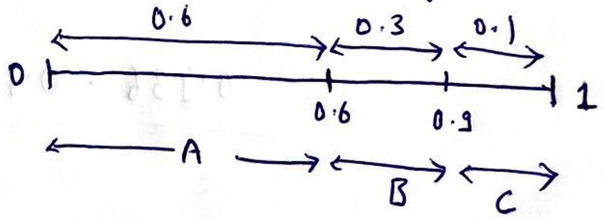
—— X ——

Arithmetic coding ⇒ It is an alternative to Huffman compression. it enables characters to be represented as fractional bit lengths.

Eg- code "CAB" using arithmetic coding.

Character	A	B	C
P	0.6	0.3	0.1

1. Divide the range into 0 - 1

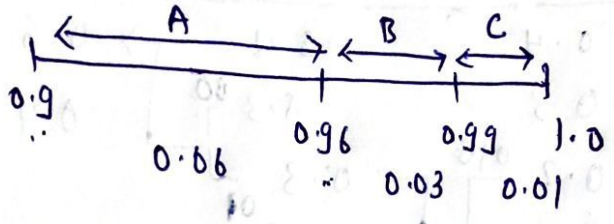


* First character → C

Range → 0.9 − 1 ⇒ $\boxed{0.1}$

Character	Cumulative Probability
A	$0.9 + 0.6 \times 0.1 = 0.96$
B	$0.96 + 0.3 \times 0.1 = 0.99$
C	$0.99 + 0.1 \times 0.1 = 1.0$

(2) Resultant new range is:—



New range

CAB

* Next character → A

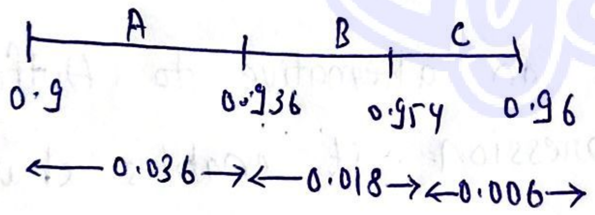
0.9 − 0.96

⇓

0.06

Character	Cumulative Probability
A	0.9 + 0.6 × 0.06 = 0.936
B	0.936 + 0.3 × 0.06 = 0.954
C	0.954 + 0.1 × 0.06 = 0.96

(3) Resultant new range is:—



* Next character → B → end

Range: 0.936 − 0.954

* Final code for string CAB

0.936 − 0.954 Ans

JPEG Compression ⇒ JPEG is a commonly used method of lossy compression for digital images.

- JPEG uses discrete cosine transformation coding (DCT).
- It allows a tradeoff b/w storage size and the degree of compression can be adjusted.

Step-① Input image splits into 8×8 pixel blocks.

Step-② JPEG uses [Y, Cb, Cr] model. so convert RGB to YCbCr.

Step-③ It is forwarded to DCT, DCT uses cosine function to convert into frequency domain.

$$F(w) = \frac{a(u)}{2} \sum_{n=0}^{N-1} f(n) \cos \frac{(2n+1)w\pi}{16}$$

Step-④ Quantization is used to reduce the no. of bits per sample.

$$\hat{F}(u,v) = \text{round}\left(\frac{f(u,v)}{Q(u,v)}\right)$$

Step-⑤ The Zigzag scan is used to scan the 8×8 matrix to 1×64 vector. It is used to group low-frequency coefficients to top level of vector & high-frequency to the bottom.

Step-⑥ Vectoring is used for differential Pulse code modulation (DPCM) is applied to DC component.

Step-⑦ Run length Encoding (RLE) is applied to AC components for removing lot of zeros.

Step-⑧ DC components are coded into Huffman

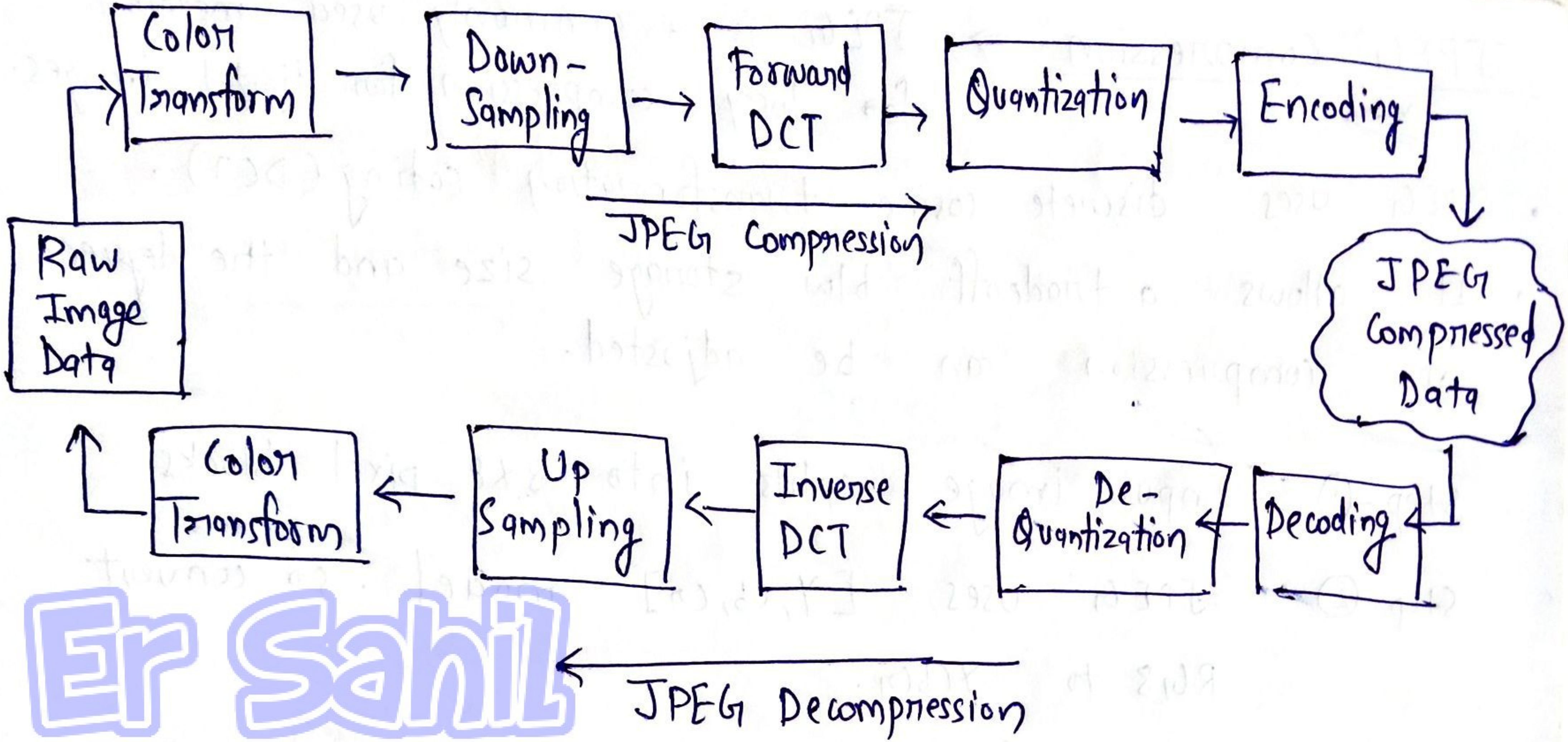

Color Transform
Down-Sampling
Forward DCT
Quantization
Encoding
JPEG Compression
Raw Image Data
JPEG Compressed Data
Decoding
De-Quantization
Inverse DCT
Up Sampling
Color Transform
JPEG Decompression

Image Segmentation

- Segmentation subdivides an image into its constituent regions or objects.

- Segmentation accuracy determines the eventual success or failure of computerized analysis procedures.

Eg - use of infrared imaging by military to detect objects with strong heat signatures.

It has two approach :- (i) Region Approach
(Group pixels based on common property)

(ii) Boundary Approach (extract regions that differ in properties like intensity, color etc.)

Detection of Discontinuities ⇒

(i) Point (isolated point)
(ii) Line
(iii) Edge

(i) Point Detection :-

An isolated point is point whose grey level is significantly different from its back ground in a Homogeneous area.

3×3

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Mask

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Image

$$\boxed{R = \sum_{i=1}^{9} w_i z_i}$$ Response of mask

if $|R| \geq T$ a point is detected where T is a non negative integer

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

Sample Mask for point Detection.

Line Detection :- four types of masks are used to get responses that is vertical, horizontal, +45°, -45° for R_1 R_2 R_3 R_4.

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Horizontal   Vertical   45°   -45°

$$\boxed{R_K = \sum_{K=1}^{4} W_K Z_K}$$

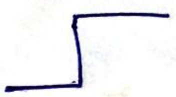
At certain point in image $|R_i| \geq R_j$ for all $j \neq i$ that point is said to be more likely associated with line in direction of mask i.

Edge Detection :- An edge is a set of connected pixels that lies on the boundary b/w two regions which differ in grey value.

→ Edges provide an outline of the object.

→ An edge can be extracted by
- Magnitude of derivate
- Direction of the derivative vector

Step Edge

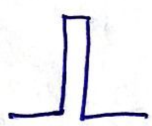


An abrupt change in intensity

Ramp Edge

A slow & gradual change in intensity

Spike Edge



Quick change to original intensity level.

Roof Edge

It is not instantaneous over short distance.

Stages in Edge Detection:-

Start → Input Image → filtering → Differentiation → Localization → Display → end

- filtering ⇓ involves smoothing explicitly or implicitly
- Differentiation ⇓ distinguishes edge pixels from other pixels
- Localization ⇓ exact location of edge

Edge Detection Algorithm:-

Derivative Types → Uses differentiation technique for edge detection.

Template Matching → Uses templates that resembles target shapes

Gaussian Derivatives → Very effective for real time image

Pattern fit Approach → Surface is considered as topographic surface with pixel value representing altitude.

Thresholding :- It is an important technique for image segmentation.

- It produces uniform regions based on threshold criteria.
- From a grayscale image, thresholding can be used to create binary image.

Global Thresholding → Operation depends upon only grey scale value.

Local thresholding → neighboring properties is also taken into account

Dynamic Thresholding → T depends on pixel coordinates

→ Histogram is used to find threshold.

Single level Thresholding :-

The objects can be extracted by comparing pixel values with threshold T.

Multi level Thresholding :-

The objects can be extracted by comparing pixel values with multiple threshold $T_1\ T_2 \dots T_n$.

Edge Linking and Boundary Detection :-

→ Set of pixels from edge detecting algorithms, seldom define a boundary completely because of noise, breaks in boundary etc.

Therefore Edge detecting algorithms are typically followed by Linking.

2 types – Local & Global

Local Processing :- All points that are similar are linked, forming a boundary of pixels that share some common properties.

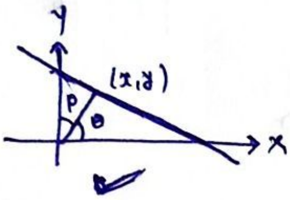
** Hough Transform :- It is a technique which can be used to isolate features of a particular shape within an image.

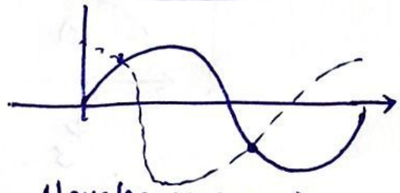
- It is most commonly used for detection of regular curves such as lines, circles, ellipses etc.

- Adv :- It is tolerant of gaps in feature boundary descriptions and is relatively unaffected by image noise.

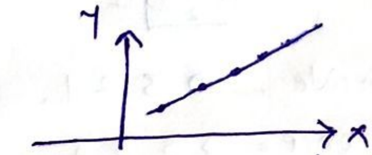
Line's equation $Y = mx + c$ into single point in (p, θ) plane

$$p = x\cos\theta + y\sin\theta$$





Houghspace (p, θ)



(used to connect disjoint edge points)

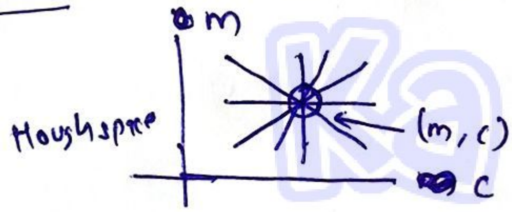
→ Hough Transform approach is to find the points of intersection in the curves, each of which corresponds to a line in the Cartesian xy plane.

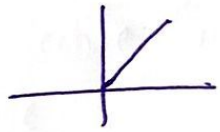
$$p = x\cos\theta + y\sin\theta$$

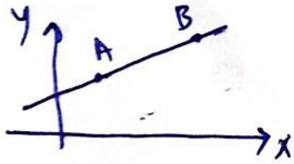
in circle $(x-a)^2 + (y-b)^2 = r^2$

$\Rightarrow$ $y_i = mx_i + c$ $\Rightarrow$ $c = -x_i m + y_i$

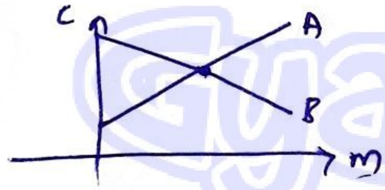
Steps Hough Transform







$\Rightarrow$



in Hough Space

$\Rightarrow$ find local maxima in parameter space

Region Based Segmentation :- Region growing is a procedure that groups pixels into large regions based on predefined criteria for growth.

→ $f(x, y)$ denotes an input image array.

→ $S(x, y)$ denotes a seed array containing 1s at locations of seed points.

→ Q denote a predicate to be applied at each location.

→ Array f & S are same size