| Page No. | Teacher's Sign / Remarks |
|---|---|
| | |

# Digital Image Processing

Processing the image that are in digital nature by a digital System.

Why?

→ Improvement of pictorial information for human perception

→ Image Processing at automation.

→ Storage & transmission

→ Noise filtering

→ Enhancing

- Contrast
- Blurring / Defocused

→ ~~Aerial Image~~ / Medical Images / Remote sensing / Atmospheric /

→ Here the interest is on procedures for extraction of image information suitable for computer processing

→ Automated Inspection / Boundery Information.

→ ~~Boundery Information~~

→ Video sequence processing (Movement detection)

→ Image Compression (An Image usually contains lot of redundancy that can be exploited to achieve compression)

→ Pixel redundancy

→ Coding redundancy

→ Psychovisual redundancy.

→

- Application
  - Reduce storage
  - Reduction of bandwidth.

Bartlane Systems

- Telegraph printer (1920)
- tape (1921)
- (1929) — 15

Image Representation

- An Image is an 2D light intensity function $f(x,y)$.
- A digital Image $f(x,y)$ is discretized both the spatial coordinates and brightness.
- It can be considered as a matrix whose row, column indices specify a point in the image and the element values identifies gray level value at that point.
- These elements are referred to as pixels (or) pels.

$$f(x,y) = \underbrace{r(x,y)}_{\text{reflectivity}} * \underbrace{i(x,y)}_{\text{intensity incident light}}$$

✓ $0 < f(x,y) < \infty$

→ sample ( 1. spatial discretization by grids
2. Intensity discretization by Quantization )

$$\begin{bmatrix} f(0,0), & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdot & f(1,N-1) \\ & & & \\ f(M-1,0) & f(M-1,1) & & f(M-1,N-1) \end{bmatrix}$$

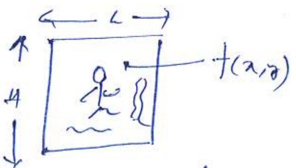
# Steps in Digital Image Processing

1. Image Acquisition: ✓
2. Preprocessing ✓
3. Segmentation ✓
4. Description / feature selection ✓
5. Recognition & Interpretation ✓
6. Knowledge Base ✓

## Digitization (

1. Why image digitization is nessary?
2. What is meant by signal bandwidth
3. Selecting the sampling frequency of a given signal
4. Exlpain image reconstruction from sampled values.

Why?
What?
How?




$0 \le x \le H$

$0 \le y \le L$

$0 \le r(x,y) \le 1$

$0 \le I(x,y) \le \infty$

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

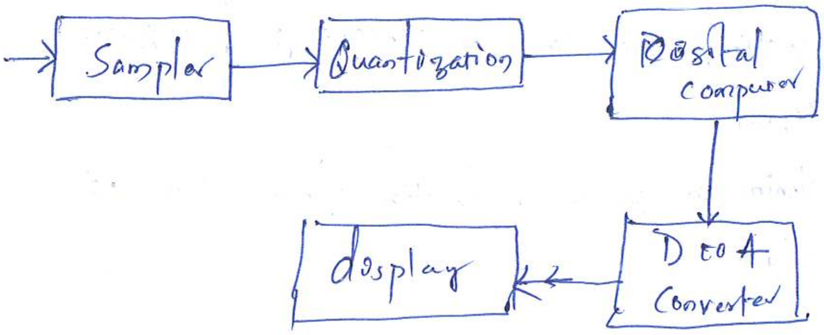
$I_{min} \le f(x,y) \le I_{max}$

* Infinal number of points.

An image to be represented in the form of a finite 2D matrix — Sampling

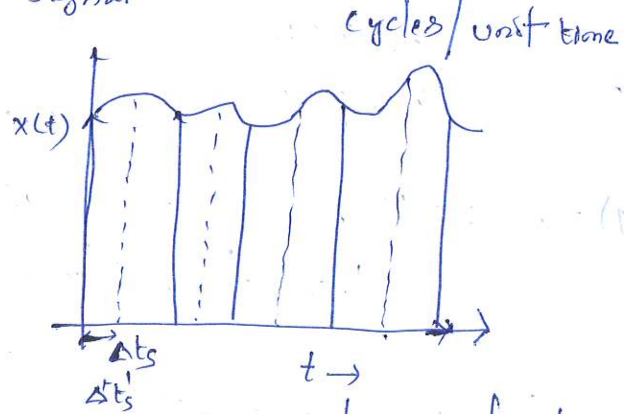
$$\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

Each matrix element represented by of the finite set of discrete values — Quantization




Sampling

1-D signal




Sampling frequency $f_s = \frac{1}{\Delta t_s}$ $\left[\because f = \frac{1}{T}\right.$

$\Delta t'_s = \frac{\Delta t_s}{2}$

finite 2D

ne finite

$\therefore f = \frac{1}{T}$

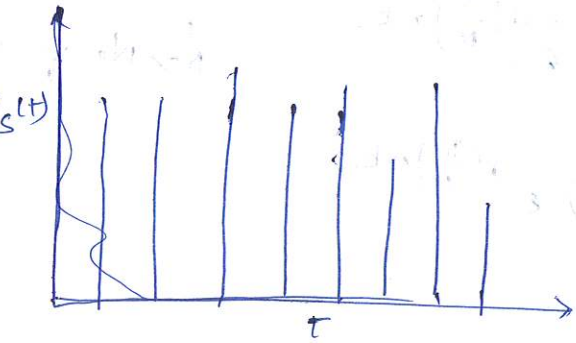
Sampling frequency $f'_s = \frac{1}{\Delta t'_s} = \frac{2}{\Delta t_s} = 2 f_s$

What is proper sampling frequency?

Sampling function : 1D array (delta function)


$$comb(t;\Delta t) = \sum_{m=-\infty}^{\infty} \delta(t - m\Delta t)$$

$t = m\Delta t$, then 1




$$X_s = X(t)\, comb(t, \Delta t)$$
$$= \sum_{m=-\infty}^{\infty} x(m\Delta t)\,\delta(t - m\Delta t)$$

Aperiodic signal

Continuous signal

$x(t)$ in time. $f$

frequency components: $f(x(t)) = \int_{-\infty}^{\infty} X(\omega)$

$$= \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

Periodic Signal

$$v(t) = \sum_{n=-\infty}^{\infty} C(n)\, e^{jn\omega_0 t}$$

$\omega_0 \rightarrow$ fundament frequency.

$$C(n) = \frac{1}{T_0} \int_{T_0} v(t)\, e^{-jn\omega_0 t}\, dt$$

$C(n) \rightarrow$ fourier coefficient

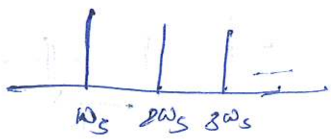
$T_0 \rightarrow$ Period of periodic signal

$T_0 = \Delta T_s$ $\qquad v(t) = 1$ for $t = 0$

$\qquad\qquad = 0$; otherwise.

$c(n) = \frac{1}{\Delta T_s} = \omega_s$

$$\boxed{v(t) = \frac{1}{\Delta T_s} \sum_{n=-\infty}^{\infty} e^{j n \omega_0 t}}$$




Discrete signal $x(n)$

$x(n)$ *

$$X(K) = \sum_{n=0}^{N-1} x(n) e^{-j\left(\frac{\pi}{N}\right) nK}$$

$N \rightarrow$ No. of samples

$$x(n) = \sum_{K=0}^{N-1} X(K) e^{j\left(\frac{\pi}{N}\right) nK}$$

Convolution

$$x_s(t) = x(t) \cdot \mathrm{comb}(t, \Delta t)$$

$x(t) \quad h(t)$

$h(t) * x(t)$

$$F(h(t) * x(t)) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} [h(\tau) x(t-\tau) \, d\tau] \right] e^{-j\omega t} \, dt \qquad \int_{-\infty}^{\infty} h(\tau) x(t-\tau) \, d\tau$$

Time inversion

$$= \int_{-\infty}^{\infty} h(\tau) \left[ \int_{-\infty}^{\infty} x(t-\tau) e^{-j\omega(t-\tau)} \, d\tau \right] dt$$

$$= \int_{-\infty}^{\infty} h(\tau) X(\omega) e^{-j\omega \tau} \, d\tau$$

$$= X(\omega) \int_{-\infty}^{\infty} h(\tau) e^{-j\omega \tau} \, d\tau$$

$= X(\omega) \cdot H(\omega)$

$\Rightarrow \boxed{x(t) * h(t) \Leftrightarrow X(\omega) \cdot H(\omega)}$

$\boxed{X(\omega) * H(\omega) = x(t) * h(t)}$

$$= \int_{-\infty}^{\infty} h(\tau) \left[ \int_{-\infty}^{\infty} x(t-\tau) e^{-j\omega(t-\tau)} \, dt \right] e^{-j\omega\tau} \, d\tau$$

$$= \int_{-\infty}^{\infty} h(\tau) X(\omega) e^{-j\omega\tau} \, d\tau$$

$$= X(\omega) \int_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} \, d\tau$$

$= X(\omega) \cdot H(\omega)$

$\therefore \boxed{x(t) * h(t) = X(\omega) \cdot H(\omega) \Rightarrow X(\omega) * H(\omega) = x(t) \cdot h(t)}$

$X_s(t) = X(t) \cdot \text{comb}(t, \Delta t_s)$

$X_s(\omega) = X(\omega) \otimes \mathcal{F}(\text{comb}(t, \Delta t_s))$

$h(n)$ : comb$(t, \Delta t)$

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

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

$x(n)$

| | | | | |
|---|---|---|---|---|
| 2 | 5 | 7 | 9 | 3 |
| -2 | -1 | 0 | 1 | 2 |

$$y(n) = \sum_{m=-\infty} h(m) \cdot x(n-m)$$

$x(-n)$

| | | | | |
|---|---|---|---|---|
| 3 | 9 | 7 | 5 | 2 |

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)\, x(n-m)$$

$x(-m)$

| | | | | |
|---|---|---|---|---|
| 3 | 9 | 7 | 5 | 2 |

$h(m)$

-13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12

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

$$y(n) = \sum_{-\infty}^{\infty} x(n-m) \cdot h(m)$$

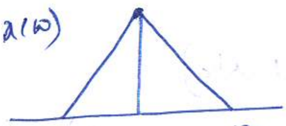
12 11 10 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 10 11 12

2 5 7 9 3 0 0 0 0 2 5 7 9 3 0 0 0 0 2 5 7 9 3 0

Convolution

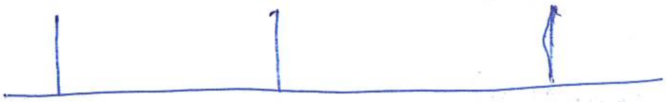
$$X_s(\omega) = X(\omega) \circledast COMB(\omega)$$

$x(\omega)$

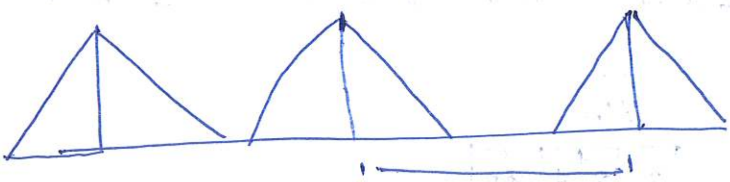


$w_0$

$\circledast$



$\Downarrow$



0 $w_0$ $W_0 = \frac{1}{\Delta t_s}$

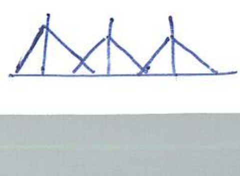
$$\frac{1}{\Delta t_s} - W_0 > W_0$$

$$\Rightarrow \frac{1}{\Delta t_s} > 2W_0$$

$$\Rightarrow \boxed{f_s > 2W_0}$$

Nyquist rate

$\boxed{\text{Aliasing: } f_s < 2W_0}$



8 9 10 11 12

0 0 0 0 0

8 9 10 11 12

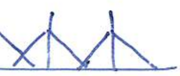
5 7 9 3 0

$$\frac{1}{\Delta t_s} - W_0 > W_0$$

$$\Rightarrow \frac{1}{\Delta t_s} > 2W_0$$

$$\Rightarrow \boxed{f_s > 2W_0}$$

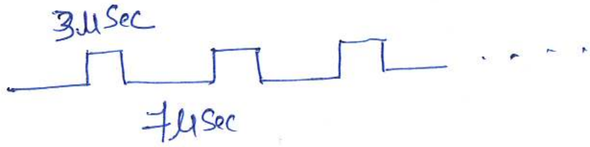
Nyquist rate



1. What are the steps involved in image digitization Process?

2. What is sampling?

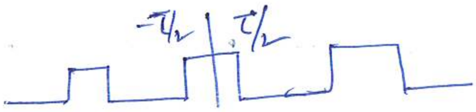
3. Find the frequency spectrum of the following periodic signal :



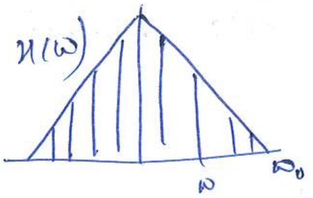

4. Speech signal has a bandwidth of 4 KHz. If every sample is digitized using 8 bits and the digital speech is to be transmitted over a Communication channel, what is the minimum bandwidth requirement of the channel?

Hz → cycles/unit length.

3.



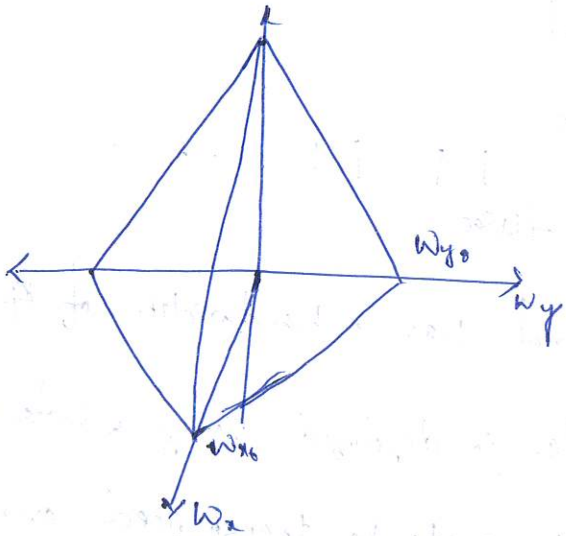
1D signal




Bandlimited $\Rightarrow X(\omega) = 0$

$|\omega| > \omega_0$

2D signal




Bandlimited

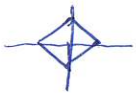
$F(\omega_x, \omega_y) = 0$

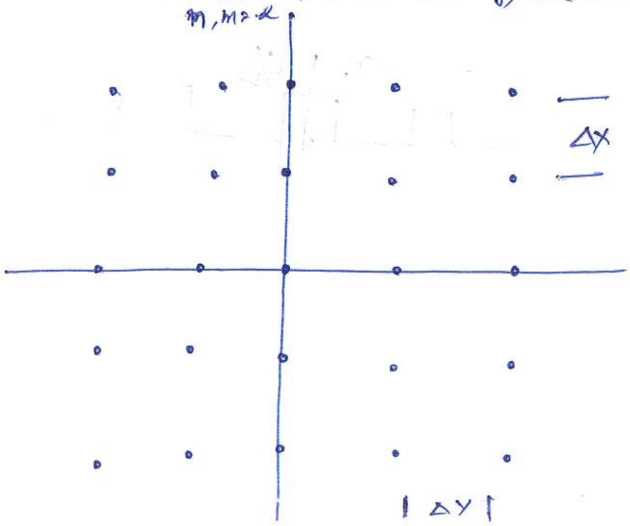
for $|\omega_x| > \omega_{x0}$

$|\omega_y| > \omega_{y0}$

2D Sampling

$$f(x,y) = f(x,y)\,\mathrm{comb}(x, y; \Delta x, \Delta y)$$
$$= \sum\sum_{m,n=-\infty}^{\infty} f(m\Delta x, n\Delta y)\,\delta(x - m\Delta x, y - n\Delta y)$$





Quantization is a mapping of a continuous variable u to a discrete variable u'.

Quantization Rules →

Define a set of decision (or) transition levels

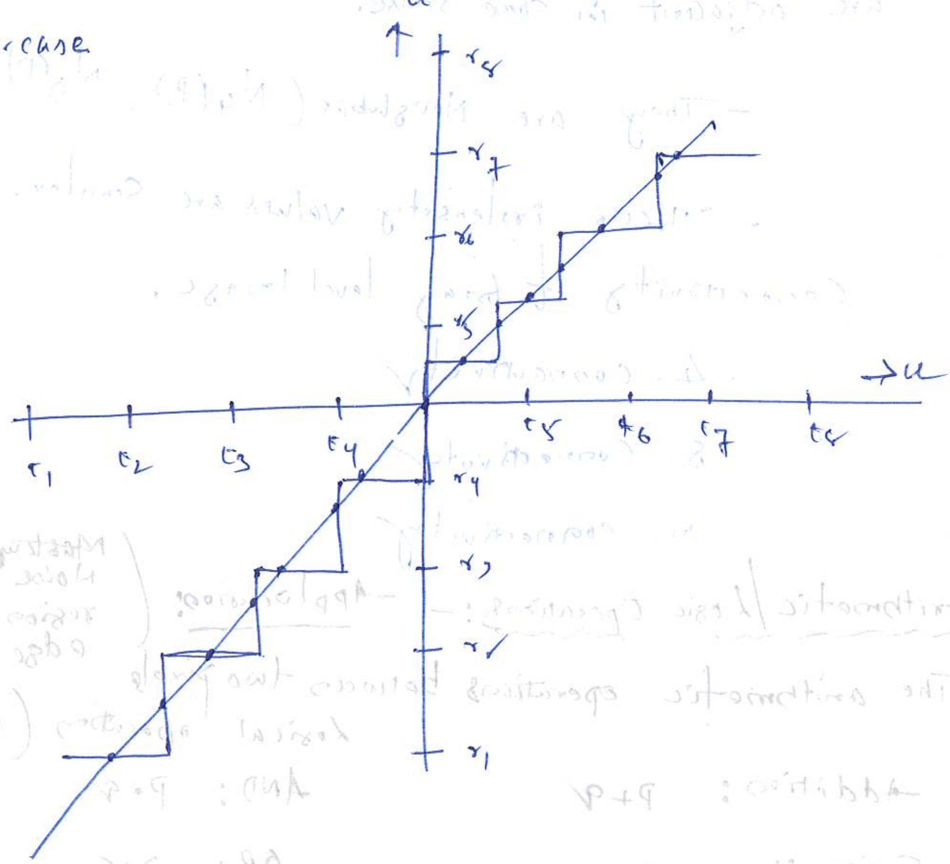
$\{t_k, \; k=1,2,\dots L+1\}$

where $t_1$ is minimum values

$t_{L+1}$ is maximum values

$u' = r_k$ if $t_k \le u < t_{k+1}$

staircase




Quantization Error / Noise

1. What is aliasing?

Ans.

## Pixel Relationship

1. Neighbor

2. Connectivity and adjacency

3. Labeling

→ What is connectivity?

— Two pixels are said to be connected if they are adjacent in some sense.

- They are Neighbor ( $N_4(P)$, $N_D(P)$ (or) $N_8(P)$ )

- Their intensity values are similar.

Connectivity of Gray level Images

- 4 - connectivity

8 - Connectivity

m - connectivity

**Arithmetic/Logic Operations:-** Application: ( ~~Masking Operation~~ Noise reduction region thinning edge detection. )

The arithmetic operations between two pixels

Logical operation (only binary)

Addition : $p+q$ AND : $p \cdot q$

Subtraction : $p-q$ OR : $p+q$

Multiplication : $p*q$ Complement : $\bar{q}$

Division : $p \div q$

$$Z = \sum_{i=1}^{9} w_i z_i$$

$$Z = \frac{1}{9}(z_1 + z_2 + \ldots + z_9) \quad \text{—(1)}$$

$$Z = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 \quad \text{—(2)} \quad w_i = i,\ i = 1, 2, \ldots$$

...onnected if they

..., $N_D(P)$ (or) $N_8(P)$

...similar.

~~Masking Operation~~
Noise reduction
region thinning
edge detection.

...ation (only binary)

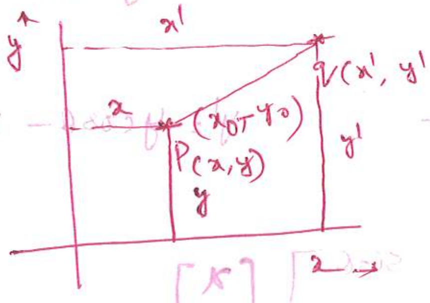
...$z_9$) —①

...$w_9 z_9$ —② $w_i = \frac{1}{9}$, 1, 2 ...

# Basic ~~(O~~ Transformation

Image Geometry

## Pixel ~~Transform~~ Translation :




$$x' = x + x_0$$
$$y' = y + y_0$$

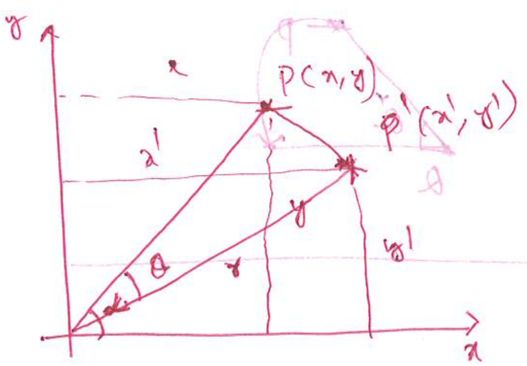
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}$$

$$\Rightarrow \boxed{\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}}$$

unified Expression

## Pixel Rotation with origin




$x' =$

$$x = r\cos\alpha$$
$$y = r\sin\alpha$$
$$x' = r\cos(\alpha - \theta)$$
$$y' = r\sin(\alpha - \theta)$$

$x' = r\cos(\alpha - \theta)$

$= r\cos\alpha \cos\theta + r\sin\alpha \sin\theta$

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

$y' = r\sin(\alpha - \theta)$

$= r\sin\alpha\cos\theta - r\cos\alpha\sin\theta$

$= y\cos\theta - x\sin\theta$

$x' = x\cos\theta + y\sin\theta$  $y' = y\cos\theta - x\sin\theta$

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

Pixel scaling

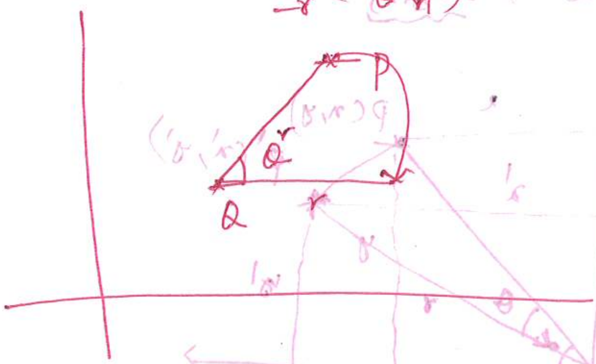
Scaling in x direction $S_x$

Scaling in y direction, $S_y$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

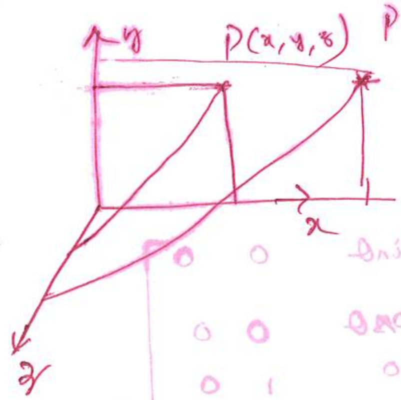
Pixel Rotation with point q.

$T_r(R_\theta T_p)$



## 3D

### Translation




$x' = x + x_0$

$y' = y + y_0$

$z' = z + z_0$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}$$

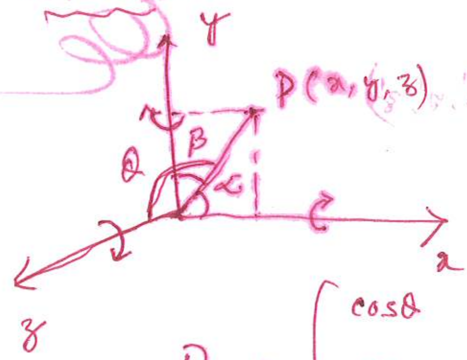
$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

### Scaling

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Rotation




$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_\beta = \begin{bmatrix} \cos\beta & \overset{0}{\cancel{\sin\beta}} & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

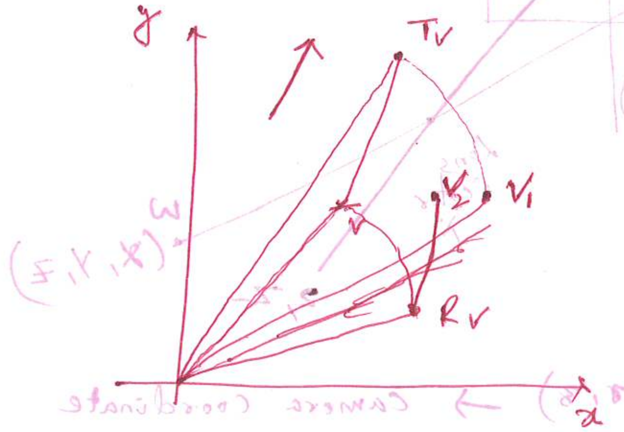
$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Concatenation ~~Non~~

$$V^* = R_\theta(S(Tv))$$

$$A = R_\theta ST$$

## Non commutative

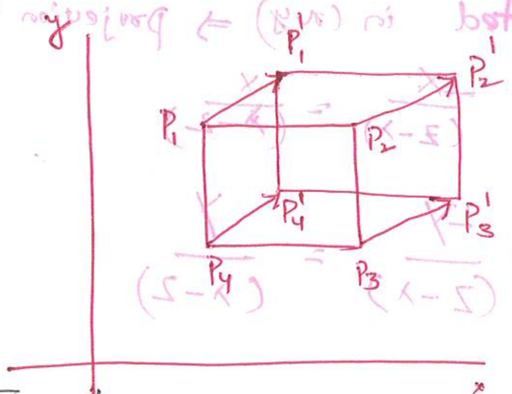



$V_2 = T(Rv)$

$V_1 = R(Tv)$

$V_1 \neq V_2$

## Transformation set of point



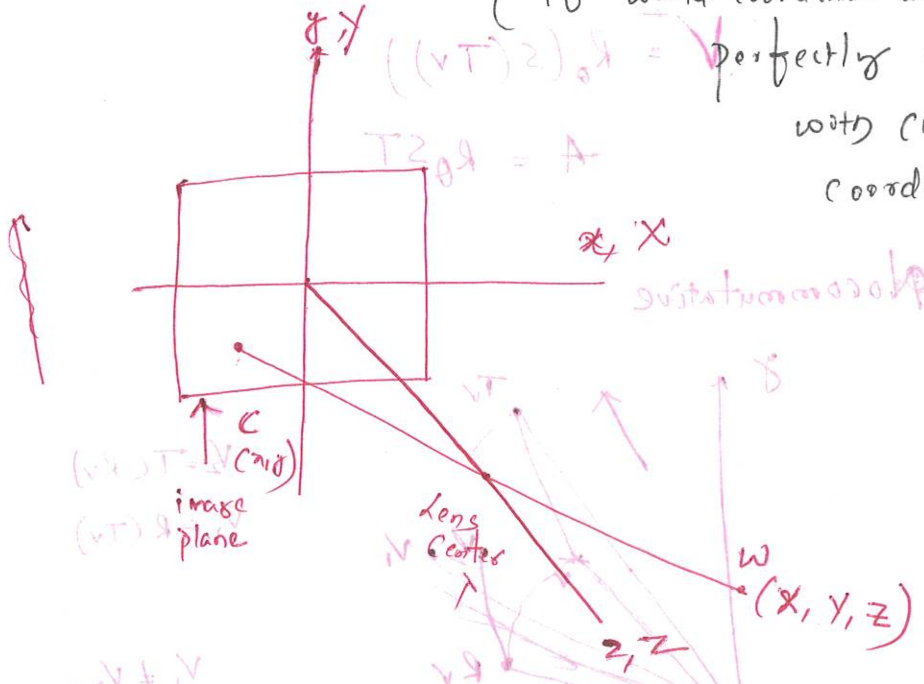

## Inverse Transform

$$T^{-1} = \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad S^{-1} = \begin{bmatrix} -S_x & 0 & 0 & 0 \\ 0 & -S_y & 0 & 0 \\ 0 & 0 & -S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_\theta^{-1}(\text{with } z) = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) & 0 & 0 \\ -\sin(-\theta) & \cos(-\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Perspective Transformation (or) Image Transformation

(If world coordinate are perfectly alligned with camera coordinates)




$(x, y, z) \rightarrow$ camera coordinate

$(X, Y, Z) \rightarrow$ world coordinate

$(0, 0, \lambda) \rightarrow$ center of lens

We are interested in $(x, y) \Rightarrow$ projection of $(X, Y, Z)$.

$$\frac{x}{\lambda} = \frac{-X}{(Z-\lambda)} = \frac{X}{(\lambda - Z)}$$

$$\frac{y}{\lambda} = \frac{-Y}{(Z-\lambda)} = \frac{Y}{(\lambda - Z)}$$

$$\left.\begin{aligned} x &= \frac{\lambda X}{(\lambda - Z)} \\ y &= \frac{\lambda Y}{(\lambda - Z)} \end{aligned}\right\} \longrightarrow$$ can be expressed as matrix representation of Homogeneous coordinates

formation

rdinate are
jectly alligned
with camera
coordinates)

, Y, Z)

nate

te

f (X, Y, Z).

xpressed as
preservation of
coordinates

Cartesian Coordinate - Homogeneous coordinate

$$(x, y, z) \Rightarrow (kx, ky, kz, k)$$

k is arbitrary constant.

Define perspective transform matrix $P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix}$ $\quad P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/\lambda & 1 \end{bmatrix}$

$$C_h = P w_h = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix}.$$

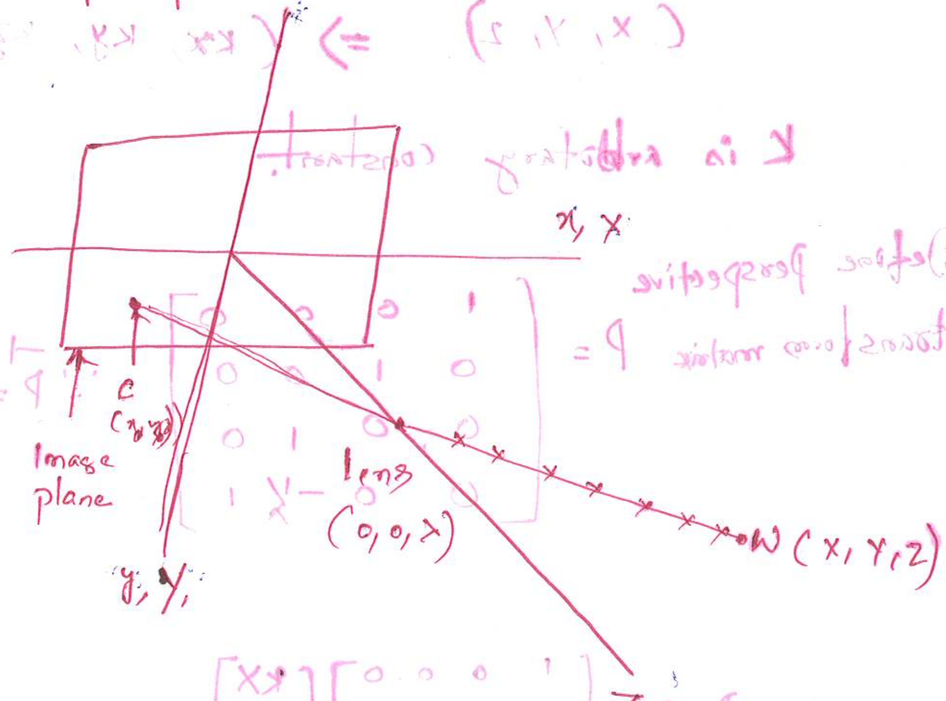
$$= \begin{bmatrix} kX \\ kY \\ kZ \\ -\frac{kZ}{\lambda} + k \end{bmatrix} = \begin{bmatrix} kX \\ kY \\ kZ \\ k\left(\frac{\lambda - Z}{\lambda}\right) \end{bmatrix}$$

$$C = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \lambda X/(\lambda - Z) \\ \lambda Y/(\lambda - Z) \\ \lambda Z/(\lambda - Z) \end{bmatrix}$$

$$x = \lambda X/(\lambda - Z)$$

$$y = \lambda X/(\lambda - Z)$$

# Inverse perspective transform:




For an image point $C(x_0, y_0, 0)$

$$C_h = \begin{bmatrix} kx_0 \\ ky_0 \\ 0 \\ k \end{bmatrix} \Rightarrow w_h \Rightarrow \begin{bmatrix} kx_0 \\ ky_0 \\ 0 \\ k \end{bmatrix}$$

Inverse ~~transform~~ perspective transformation is formulated by using Z component of $C_h$ as a free variable.

$$C_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ k \end{bmatrix}$$

$$w_h = P^{-1} C_h$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/\lambda & 1 \end{bmatrix} \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ k \end{bmatrix}$$

Homogeneous Coordinates $w_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ kz/\lambda + k \end{bmatrix}$

$$W_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ \frac{k(z+\lambda)}{\lambda} \end{bmatrix}$$

Cartesian coordinate

$$W = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x_0\lambda/(\lambda+z) \\ y_0\lambda/(\lambda+z) \\ z\lambda/(\lambda+z) \end{bmatrix}$$

$$X = \frac{x_0\lambda}{(\lambda+z)} \qquad Y = \frac{y_0\lambda}{(\lambda+z)} \qquad Z = \frac{z\lambda}{(\lambda+z)}$$

$$\Rightarrow Z(\lambda+z) = z\lambda$$

$$\Rightarrow z = \frac{Z\lambda}{(\lambda-Z)}$$

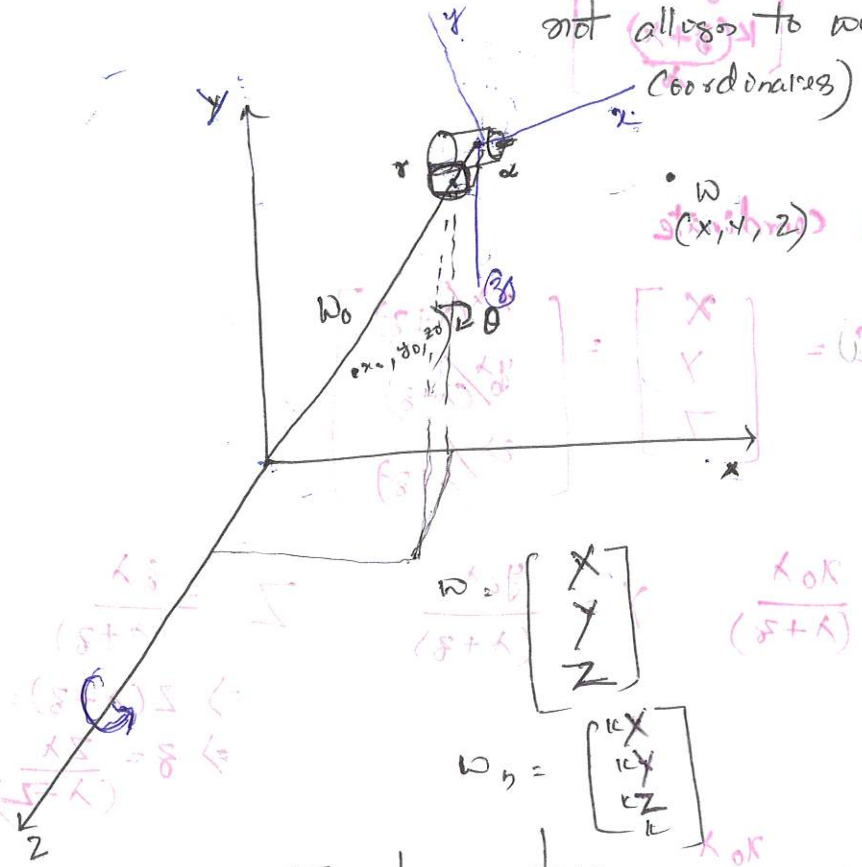
$$X = \frac{x_0\lambda}{\lambda + \frac{Z\lambda}{(\lambda-Z)}}$$

$$X = \frac{x_0\lambda}{\frac{\lambda(\lambda-Z)+Z\lambda}{(\lambda-Z)}} = \frac{x_0\cancel{\lambda}(\lambda-Z)}{\cancel{\lambda}(\lambda-\cancel{Z}+\cancel{Z})}$$

$$X = \frac{x_0}{\lambda}(\lambda-Z)$$

$$Y = \frac{y_0}{\lambda}(\lambda-Z)$$

Perspective transformation (If world coordinares are not allign to world coordinares)

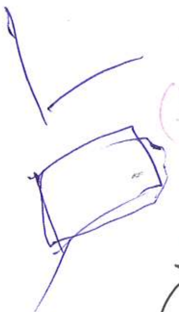



$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$w_h = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix}$$

— Transform steps

- Displasement of gimbal center form original by $w_0$



- Pan of ~~x axis~~ Z axis by θ
- Tilt of ~~x axis~~ x axis by α
- Displacement of the image plane with respective to gimbal center by α.

$C_h = \cancel{P R_{\alpha} R_{\theta} G W_h}$ $\quad PTRGw$

$$G = \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}$$

$4+4$

$$R_{\alpha} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 4+12 & 2+3 \\ 8+8 & 4+12 \end{bmatrix}$$

$$\begin{bmatrix} 16 & 5 \\ 16 & 6 \end{bmatrix}$$

$$R_{\theta} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R = R_{\alpha} * R_{\theta}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\cos\alpha\sin\theta & \cos\alpha\cos\theta & \sin\alpha & 0 \\ \sin\alpha\sin\theta & -\sin\alpha\cos\theta & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 & -r_1 \\ 0 & 1 & 0 & -r_2 \\ 0 & 0 & 1 & -r_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C_h = P\,T\,R\,G\,w_h$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & -r_1 \\ 0 & 1 & 0 & -r_2 \\ 0 & 0 & 1 & -r_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} *$$

$$\begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\cos\alpha\sin\theta & \cos\alpha\cos\theta & \sin\alpha & 0 \\ \sin\alpha\sin\theta & -\sin\alpha\cos\theta & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} *$$

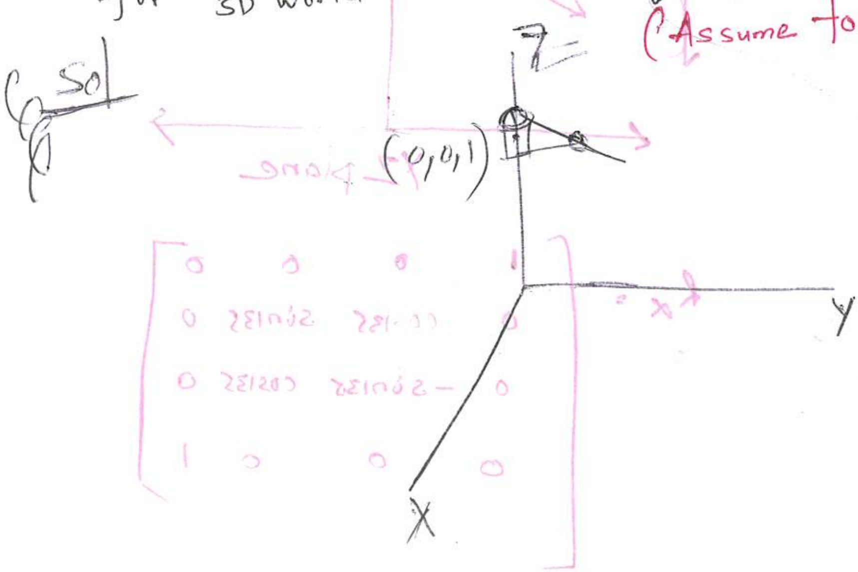
$$\begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix}$$

$$x = \lambda * \frac{(X-X_0)\cos\theta + (Y-Y_0)\sin\theta - r_1}{-(X-X_0)\sin\theta\sin\alpha + (Y-Y_0)\cos\theta\sin\alpha - (Z-Z_0)\cos\alpha + r_3 + \lambda}$$

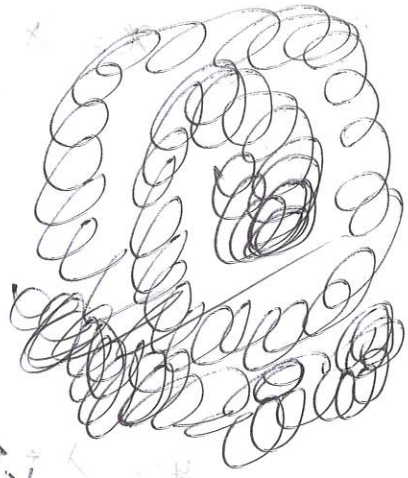
$$y = \lambda * \frac{-(X-X_0)\sin\theta\cos\alpha + (Y-Y_0)\cos\theta\cos\alpha + (Z-Z_0)\sin\alpha - r_2}{-(X-X_0)\sin\theta\sin\alpha + (Y-Y_0)\cos\theta\sin\alpha - (Z-Z_0)\cos\alpha + r_3 + \lambda}$$

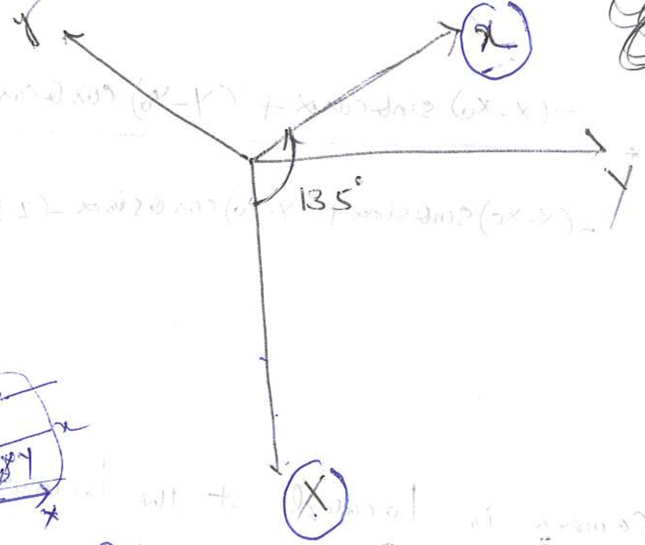
Example:

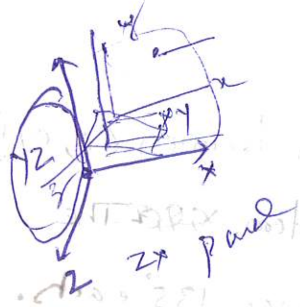
The camera is located at the location (0, 0, 1) with respect to 3D world coordinates system. ~~XQQ~~ The arrangement of pan and the tilt was 135° each. And, we have an object placed in X, Y plane where one of the corners object is located at location P'(1, 1, 0.2). Find out image coordinates for 3D world coordinates of point P? (Assume focal length $\lambda = 0.035$)



$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\phi \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

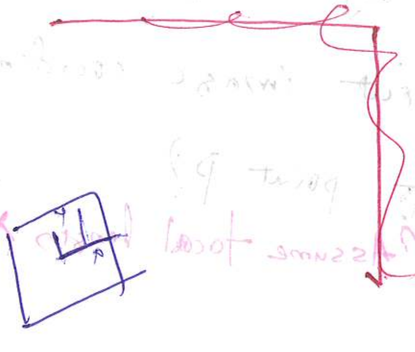


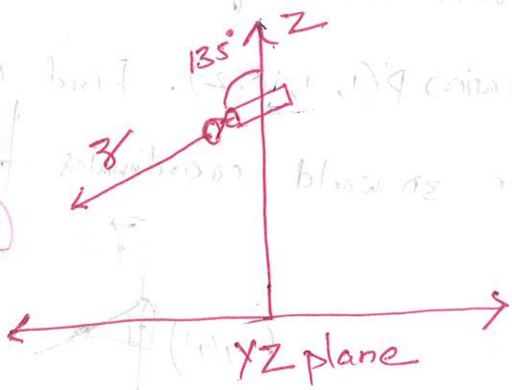



$$R_\theta = \begin{bmatrix} \cos 135 & \sin\theta & 0 & 0 \\ -\sin 135 & \cos 135 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

z' 135°






$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 135 & \sin 135 & 0 \\ 0 & -\sin 135 & \cos 135 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3) World coordinate point $(1, 1, 0.2)$

$$\begin{bmatrix} \cancel{x'} \\ \cancel{y'} \\ \cancel{z'} \\ 1 \end{bmatrix} = \underline{R_\alpha R_\theta T} \begin{bmatrix} 1 \\ 1 \\ 0.2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.707 & 0.707 & 0 & 0 \\ 0.5 & 0.5 & 0.707 & -0.707 \\ 0.5 & 0.5 & -0.707 & 0.707 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0.2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.43 \\ 1.55 \\ 1 \end{bmatrix}$$

$\lambda = 0.035$

$$C_h \Rightarrow \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix}$$

$$x' = \frac{\lambda \hat{x}}{\lambda - \hat{z}} = 0$$

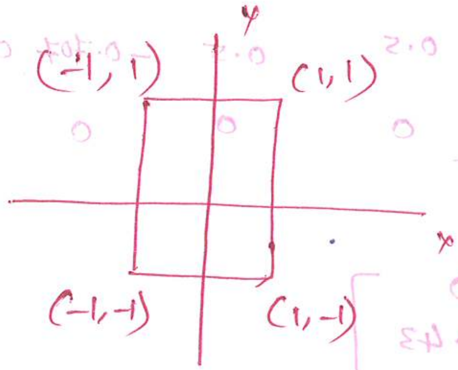
$$y' = \frac{\lambda \hat{y}}{\lambda - \hat{z}} = -0.0099$$

| | $\cos\theta$ | $\sin\theta$ |
|---|---|---|
| $0°$ | $1$ | $0$ |
| $45°$ | $\sqrt{2}/2$ | $\sqrt{2}/2$ |
| $90°$ | $0$ | $1$ |
| $135°$ | $-\sqrt{2}/2$ | $\sqrt{2}/2$ |
| $180$ | $-1$ | $0$ |
| $30°$ | $\sqrt{3}/2$ | $\sqrt{1}/2$ |
| $60°$ | $\sqrt{1}/2$ | $\sqrt{3}/2$ |

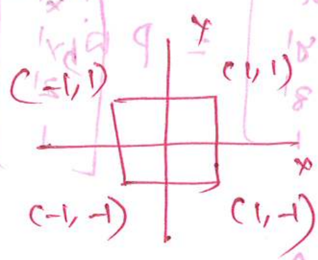
Quiz

1. What is the concatenated transformation matrix for translation by vector [1 1] followed by rotation by angle 45° degrees in 2D.

2. The following figure is first scaled by a factor 2 and then translated by vector [2, 2]. What is the transformed figure?




3. Determine the figure if translation is applied first followed by scaling.




4. For a camera with focal length of 0.05, find out the locus of the points which will be imaged at location (0.2, -0.3) on the image plane. Assume the camera coordinate system and world coordinate system to be perfectly aligned.

5. A camera lens has a focal length of 5. Find out the image point corresponding to a world point at location (50, 70, 100). Assume that the image coordinates system and the world coordinate system to be perfectly aligned.

---

**Camera Calibration:**

The computational procedure used to obtain the camera parameters using these known points often is referred to as camera calibration.

$$C_h = PTRG\, w_h$$

$$\therefore PTRG = A$$

$$\boxed{C_h = A w_h}$$

Estimate 'A' element values.

Lets Assume

$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \Rightarrow w_h = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} \qquad C_h = \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix}$$

Lets takes $k = 1$

$$\begin{bmatrix} C_{h1} \\ C_{h2} \\ C_{h3} \\ C_{h4} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x = C_{h1}/C_{h4} \qquad y = C_{h2}/C_{h4}$$

$$\begin{bmatrix} xC_{h4} \\ yC_{h4} \\ C_{h3} \\ C_{h4} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$xC_{h4} = a_{11}X + a_{12}Y + a_{13}Z + a_{14}$$

$$yC_{h4} = a_{21}X + a_{22}Y + a_{23}Z + a_{24}$$

$$C_{h4} = a_{41}X + a_{42}Y + a_{43}Z + a_{44}$$

$$x = \frac{a_{11}X + a_{12}Y + a_{13}Z + a_{14}}{a_{41}X + a_{42}Y + a_{43}Z + a_{44}}$$

$$y = \frac{a_{21}X + a_{22}Y + a_{23}Z + a_{24}}{a_{41}X + a_{42}Y + a_{43}Z + a_{44}}$$

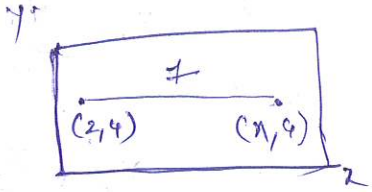
$$\begin{cases} a_{11}X + a_{12}Y + a_{13}Z - a_{41}xX + a_{42}xY - a_{43}xZ - a_{44}x + a_{14} = 0 \\ a_{21}X + a_{22}Y + a_{23}Z - a_{41}yX + a_{42}yY - a_{43}yZ - a_{44}y + a_{24} = 0 \end{cases}$$

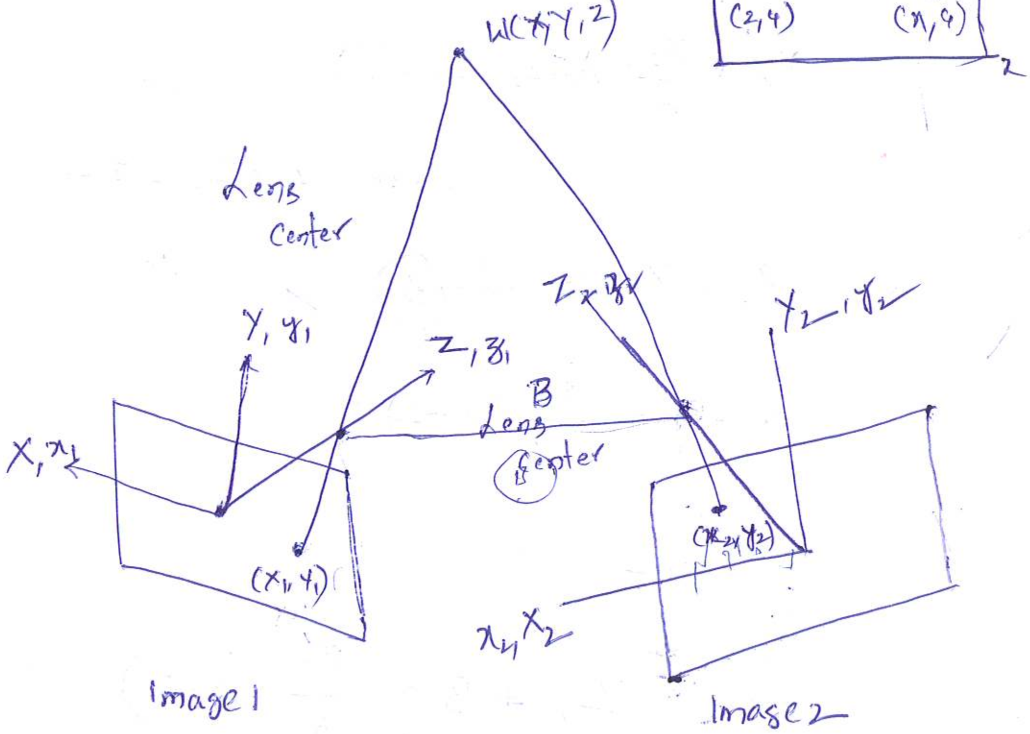
The calibration procedure then consists (1) obtaining $m \geq 6$ world points with known coordinates $(X_i, Y_i, Z_i)$, $i = 1, 2, \ldots m$;

(2) Imaging these points with camera in given position to obtain the corresponding image points $(x_i, y_i)$ $i = 1, 2, \ldots m$

(3) using these results and point solve unknown points.

# Stereo Image Model




$$X_i = \frac{x_1}{\lambda}(\lambda - Z) \qquad Y_i = \frac{y_i}{\lambda}(\lambda - Z)$$

Process of obtaining missing depth information called as stereoscopic imageing technique (or) Stereo imaging.

Stereo imaging involves obtaining two seperate image views of an object. The distance between the centers of two lenses is called the baseline, and the objective is to find the coordinates $(X, Y, Z)$ of the point W having image points $(x_1, y_1)$ and $(x_2, y_2)$.

$$X_1 = \frac{x_1}{\lambda}(\lambda - Z_1) \quad \text{—— (1)}$$

$$X_2 = \frac{x_2}{\lambda}(\lambda - Z_2) \quad \text{—— (2)}$$

where $Z_1 = Z_2 = Z$

$$X_2 = X_1 + B \quad \text{—— (3)} \quad \checkmark$$

$$X_1 + B = \frac{x_2}{\lambda}(\lambda - Z)$$

~~zo=~~

$$\frac{x_1}{\lambda}(\lambda - Z) + B = \frac{x_2}{\lambda}(\lambda - Z)$$

$$(\cancel{\lambda} - \cancel{Z}) = \frac{\lambda B}{(x_2 - x_1)}$$

$$\boxed{Z = \lambda - \frac{\lambda B}{(x_2 - x_1)}}$$

4. Ans Quiz

$\lambda = 0.05 \qquad x_0 = 0.2 \qquad y_0 = -0.3$

$$\vec{P}^{1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 20 & 1 \end{bmatrix}$$

$$W_n = \vec{P} \begin{bmatrix} KX_0 \\ KY_0 \\ KZ_0 \\ K \end{bmatrix}$$

$$\frac{\frac{2}{10}}{\frac{5}{100}}$$

$$X_0 = \frac{x_0}{\lambda}(\lambda - Z) = \frac{0.2}{0.05}(0.05 - Z) = 0.2 - 4Z$$

$$Y_0 = \frac{y_0}{\lambda}(\lambda - Z) = -\frac{0.3}{0.05}(0.05 - Z) = -\frac{\frac{3}{10}}{\frac{5}{100}}(0.05 - Z)$$

$$= -0.3 + 6Z$$

$$X_0 + 4Z - 0.2 = 0$$

$$Y_0 + 6Z + 0.3 = 0$$

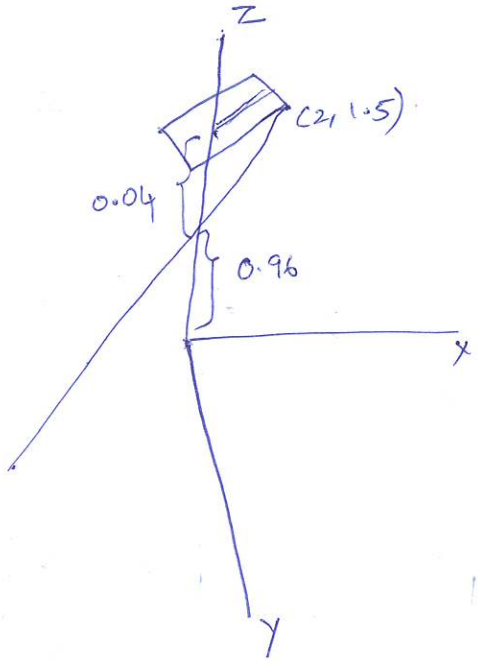
$\frac{\frac{2}{10}}{\frac{5}{100}}$

$) = 0.2 - 4Z$

$= \frac{\frac{3}{10}}{\frac{5}{100}}(0.05 - 2)$

$= -0.3 + 6Z$

1. A camera with focal lengts 0.04m is placed at a heiget of 1.0m and is lookeng vertically downwards to take images of the xy plane. If the size of the image sensor plate is 4mm x 3mm, find the area on the xy plane that can be imaged.




2. What is meant by a set of orthogonal functions?

3. What is different between orthogonality and orthonormality?

4. Determine of the following set of vectors is orthogonal?

$\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$

1. What is the advantage of separable transform) (computational complexity can be reduced)

2. Under what condition a transform is said to be separable? (If

3. Find the kronecker product $A \otimes B$ of the matrices A and B as given below

$$A = \begin{bmatrix} 3 & 7 \\ 8 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 5 \\ 9 & 4 \end{bmatrix} \quad A = \begin{bmatrix} 3B & 7B \\ 8B & 2B \end{bmatrix}$$

4. For the 2x2 transform A and the image U as given below, calculate the transformed image V and basic images

$$A = \frac{1}{2}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix} \quad U = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

Kronecker's product:

If A is an $m \times n$ matrix B is a $p \times q$ matrix, the Kronecker's product $A \otimes B$ is the $mp \times nq$ block matrix.

# Image Transform

## Fourier Transform

- Continuous ✓
- discrete ✓
- Properties
  - 1) Separability
  - 2) ~~Transform~~ Translation
  - 3) periodicity and Conjugate
  - Rotation
  - 4) Distributivity and scaling
  - 5) Convolution and Correlation
- FFT
- DFT

$f(x) \to$ continuous function $x$

$$\mathcal{F}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux}\, dx$$

Requirement

$f(x) \to$ continuous, integrable

$F(u) \to$ integrable

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

$f(x)$ & $F(u)$ are Fourier transform pairs.

$F(u) \to$ complex

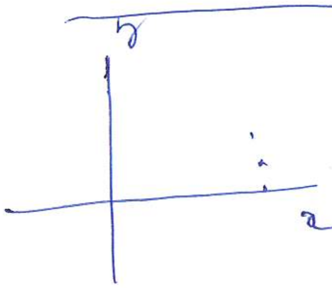
$$F(u) = R(u) + j\, I(u)$$

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

$$|F(u)| = \sqrt{\left( R(u)^2 + I(u)^2 \right)}^{1/2}$$

**Note: Fourier Spectrum:**

A plot of the ~~F(u)~~ magnitude and phase of the fourier Transform of a function. $\Rightarrow$ Fourier spectrum of $f(x)$



$$\phi = \tan^{-1} \frac{I(u)}{R(u)} \Rightarrow \text{phase angle}$$

$$P(u) = |F(u)|^2 \rightarrow \text{power spectrum}$$

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

2D signal Fourier transform

$f(x, y) \rightarrow$ continuous function of ~~x, y~~

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

Inverse fourier transform

$$\mathcal{F}^{-1}(F(u,v)) = f(x,y) = \iint_{-\infty}^{\infty} F(u,v)\, e^{j2\pi(ux+vy)}\, du\, dv$$

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

$$|F(u,v)| = \left[R^2(u,v) + I^2(u,v)\right]^{1/2}$$

↳ Fourier spectrum of $f(x,y)$

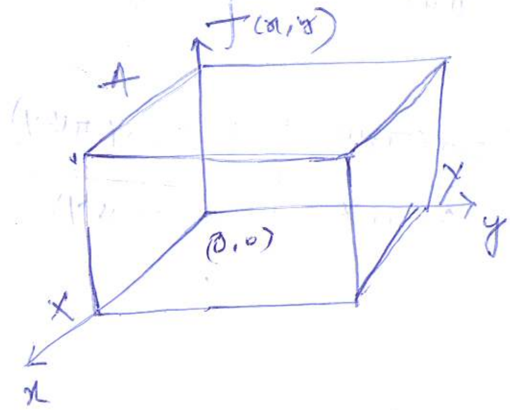
$$\phi = \tan^{-1} \frac{I(u,v)}{R(u,v)} \longrightarrow \text{Phase Angle}$$

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

↳ power spectrum

Example

Find out fourier transform of the following function



$$F(u,v) = \iint_{-\infty}^{\infty} f(x,y)\, e^{-j\frac{2\pi}{\cancel{N}}(ux+vy)}\, dx\, dy$$

$$= A \int_0^{\cancel{X}} e^{-j2\pi ux}\, dx \int_0^{\cancel{Y}} e^{-j2\pi vy}\, dy$$

$$= A \left[\frac{e^{-j2\pi ux}}{-j2\pi u}\right]_0^X \cdot \left[\frac{e^{-j2\pi vy}}{-j2\pi v}\right]_0^Y$$

$$= \frac{A}{-j2\pi u}\left[e^{-j2\pi uX} - 1\right] \cdot \left[e^{-j2\pi vY} - 1\right]$$

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

$$= AXY \left[\frac{\sin(\pi uX) \cdot e^{-j\pi uX}}{\pi uX}\right]\left[\frac{\sin(\pi vY) \cdot e^{-j\pi vY}}{(\pi vY)}\right]$$

$$|F(u,v)| = AXY \left|\frac{\sin(\pi uX)}{(\pi uX)}\right| \left|\frac{\sin(\pi vY)}{(\pi vY)}\right|$$

# 2D Discrete Fourier Transform.

$M \times N$ image $f(x,y)$

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

$u = 0, 1, \cdots M-1$

$v = 0, 1, 2 \cdots N-1$

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

$x = 0, 1, \cdots M-1$

$y = 0, 1, \cdots N-1$

If image is asquared image

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

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

Fourier Specturm

$$|F(u,v)| = \left[R^2(u,v) + I^2(u,v)\right]^{1/2}$$

phase Angle

$$\phi(u,v) = \tan^{-1} \frac{I(u,v)}{R(u,v)}$$

Power spectrum

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

Propertice of Fourier Transform.

1. Separability — 2D DFT can be obtained in two steps by successive applying 1D DFT. F(U,V) can be obtained by applying 1D-DFT along the rows and then along columns.

$$F(u,v) = \frac{1}{N^2} \sum \sum_{x,y=0}^{N-1} f(x,y)\, e^{-j\frac{2\pi}{N}(ux+vy)}$$

$$= \frac{1}{N^2} \sum_{x=0}^{N-1} e^{-j\frac{2\pi}{N}ux} \cdot \frac{1}{N} \sum_{y=0}^{N-1} f(x,y)\, e^{-j\frac{2\pi}{N}vy}$$

x is fixed

$$= \frac{1}{N} \sum_{x=0}^{N-1} e^{-j\frac{2\pi}{N}ux} \cdot F(x,v)$$

$$= \frac{1}{N} \sum_{x=0}^{N-1} F(x,v)\, e^{-j\frac{2\pi}{N}ux}$$

$$= F(u,v)$$

✓

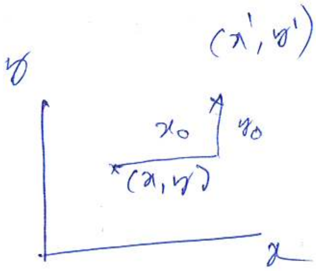
Inverse Fourier transform:

$$f(x,y) = \sum \sum_{u,v=0}^{N-1} F(u,v)\, e^{j\frac{2\pi}{N}(ux+vy)}$$

$$= \sum_{u=0}^{N-1} e^{j\frac{2\pi}{N}ux} \cdot \sum_{v=0}^{N-1} F(u,v)\, e^{j\frac{2\pi}{N}vy}$$

$$= \sum_{u=0}^{N-1} f(u,y)\, e^{j\frac{2\pi}{N}ux}$$

$$= f(x,y)$$




## 2. Translation

$$f(x,y) \xrightarrow{(x_0,\, y_0)} f(x-x_0,\, y-y_0)$$

$$F_t(u,v) = \frac{1}{N^2} \sum_{x,y=0}^{N-1}\sum f(x-x_0,\, y-y_0)\, e^{-j\frac{2\pi}{N}(u(x-x_0)+v(y-y_0))}$$

$$= \frac{1}{N} \sum_{x,y=0}^{N-1}\sum f(x-x_0,\, y-y_0)\, e^{-j\frac{2\pi}{N}(ux+vy)} \cdot e^{j\frac{2\pi}{N}(ux_0+vy_0)}$$

$$= F(u,v) \cdot \boxed{e^{j\frac{2\pi}{N}(ux_0+vy_0)}}$$

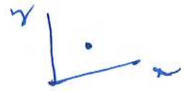
Inverse

$$f_t(x,y) = \frac{1}{N} \sum_{u,v=0}^{N-1}\sum F(u-u_0,\, v-v_0)\, e^{j\frac{2\pi}{N}(x(u-u_0)+y(v-v_0))}$$

$$= \frac{1}{N} \sum_{u,v=0}^{N-1}\sum F(u-u_0,\, v-v_0)\, e^{j\frac{2\pi}{N}(ux+vy)} \cdot e^{-j\frac{2\pi}{N}(u_0x+v_0y)}$$

$$= f(x,y) \cdot \boxed{e^{-j\frac{2\pi}{N}(u_0x+v_0y)}}$$

3. Periodicity and Conjugate:—



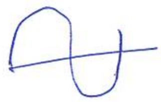
Periodicity

$$\boxed{F(u,v) = F(u+N, v) = F(u, v+N) = F(u+N, v+N)}$$

$$f(x,y) = f(x+N, y) = f(x, y+N) = f(x+N, y+N)$$

$$F(u,v) = \frac{1}{N^2}\sum_{x,y=0}^{N-1}\sum f(x,y)\, e^{-j\frac{2\pi}{N}(ux+vy)}$$

$$F(u+N, v+N) = F(u,v) \cdot e^{-j\frac{2\pi}{N}(xN + yN)}$$



$$= F(u,v) \cdot \underbrace{e^{-j\frac{2\pi}{N}(x+y)}}_{1}$$

$$= F(u,v)$$

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

$$\sin^2 x = \frac{2}{e}$$

Conjugate

$f(x,y) \longrightarrow$ Real

$$F(u,v) = F^{*}(-u,-v)$$

$$\boxed{|F(u,v)| = |F(-u,-v)|}$$

$$\cancel{\cos x = \frac{e^{ix} - e^{-ix}}{2i}}$$

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}$$

## Rotation property

$x = r\cos\theta \qquad y = r\sin\theta$

$u = \underline{w\cos\phi} \qquad v = \underline{w\sin\phi}$

$$f(x,y) \Rightarrow f(r,\theta) \qquad F(u,v) = F(w,\phi)$$

$$\boxed{f(r,\theta+\phi) \Rightarrow F(w,\phi+\theta_0)}$$

## Distributivity and Scaling property:

$$\mathcal{F}\{f_1(x,y)+f_2(x,y)\} = \mathcal{F}(f_1(x,y)) + \mathcal{F}(f_2(x,y))$$

Note

$$\mathcal{F}(f_1(x,y)\cdot f_2(x,y)) \neq \mathcal{F}(f_1(x,y)) \cdot \mathcal{F}(f_2(x,y))$$

Note: Distributive property only for addition operation.

## Scaling

scalar Quantity $a$ & $b$

$$\mathcal{F}(a f(x,y)) \Longleftrightarrow a F(u,v).$$

$$\mathcal{F}(f(ax,by)) = \frac{1}{|ab|} F\left(\frac{u}{a},\frac{v}{b}\right)$$

---

Average:

$$\mathcal{F}(f(x,y)) = \frac{1}{N^2} \sum_{x,y=0}^{N-1}\sum f(x,y)$$

$$F(0,0) = \frac{1}{N} \sum_{x,y=0}^{N-1} \sum f(x,y)$$

$$\bar{F}(u,v) = \frac{1}{N} F(0,0)$$

Convolution of

$$\boxed{\begin{aligned} f(x) \cdot g(x) &\Leftrightarrow F(u) * G(u) \\ f(x) * g(x) &\Leftrightarrow F(u) \cdot G(u) \end{aligned}}$$

Correlation

$$f(x,y) \circ g(x,y) \Leftrightarrow F^*(u,v) \cdot G(u,v)$$

$$f^*(x,y) \cdot g(x,y) \Leftrightarrow F(u,v) \circ G(u,v)$$

the FFT algorithm developed using successive doubling method.

Implementation Fast Fourier Transform:

$$F(u,v) = \frac{1}{N^2} \sum_{x,y=0}^{N-1} \sum f(x,y) \, e^{-j\frac{2\pi}{N}(ux+vy)} \quad — N^4$$

$$F(u) = \frac{1}{N} \sum_{x=0}^{N} f(x) \, e^{-j\frac{2\pi}{N}ux} \quad — \cancel{N^2}$$

$$= \frac{1}{N} \sum_{x=0}^{N-1} f(x) \cdot W_N^{ux}$$

$$\boxed{W_N = e^{-j\frac{2\pi}{N}}}$$

$N = 2^n$

$N = 2M$

$$= \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) \, W_{2M}^{ux}$$

$$= \frac{1}{2}\left[\frac{1}{M}\sum_{x=0}^{M-1} f(2x)\, W_{2M}^{u(2x)} + \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)\cdot W_{2M}^{u(2x+1)}\right]$$

$$= \frac{1}{2}\left[\underbrace{\frac{1}{M}\sum_{x=0}^{M-1} f(2x)\cdot W_{2M}^{ux}}_{F_{even}(u)} + \underbrace{\frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)\cdot W_{2M}^{2ux}\cdot W_{2M}^{u}}_{F_{odd}(u)}\right]$$

$$W_{2M}^{u(2x)} = e^{-j\frac{2\pi}{2M}\cdot u2x} = e^{-j\frac{2\pi}{M}\cdot ux} = W_M^{ux}$$

$=$

$u = 0, \dots M-1$

$$F(u) = \frac{1}{2}\left[F_{even}(u) + F_{odd}(u)\cdot W_{2M}^{u}\right]$$

<u>Note:</u> $W_M^{u+M} = W_M^{u}$ and $W_{2M}^{u+M} = -W_{2M}^{u}$

$$e^{-j\frac{2\pi}{M}(u+M)}$$
$$W_M^{u}$$
$$W_{2M}^{u+M} = e^{-j\frac{2\pi}{2M}(u+M)}$$

$\Rightarrow$ $$F(u+M) = \frac{1}{2}\left[F_{even}(u) - F_{odd}(u)\cdot W_{2M}^{u}\right]$$

$M, \dots 2M-1$

$N^2 \quad \frac{N^2}{4} \quad \frac{N^2}{2}$

$O(N\log N)$

$m(n) = 2m(n-1) + 2^{n-1}$

$a(n) = 2a(n-1) + 2^{n}$

**Inverse FFT:**

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-j\frac{2\pi}{N}ux}$$

$$f(x) = \sum_{u=0}^{N-1} F(u)\cdot e^{j\frac{2\pi}{N}ux}$$

1. Find out DFT coefficients of a digital image $f(x,y)$ of size $N \times N$ where $f(x,y) = 1$ for all values of $x$ and $y$.

2. ~~Consider the sample values of a 1D~~ signal ~~as given below. Find out the DFT~~

Ans:

$$F(u,v) = \frac{1}{N} \sum_{x,y=0}^{N-1} \sum f(x,y)\, e^{-j\frac{2\pi}{N}(ux+vy)}$$

$$= \frac{1}{N} \sum_{x=0}^{N-1} e^{-j\frac{2\pi}{N}ux} \sum_{y=0}^{N-1} e^{-j\frac{2\pi}{N}vy}$$

$$\sum_{y=0}^{N-1} e^{-j\frac{2\pi}{N}vy} = 1 + e^{-j\frac{2\pi}{N}v} + e^{-j\frac{2\pi}{N}2v} + e^{-j\frac{2\pi}{N}3v} + \cdots + N$$

$$= \frac{a\left(1 - e^{-j\frac{2\pi}{N}.N.v}\right)}{1 - e^{-j\frac{2\pi}{N}.v}} \qquad \left[ a + ar^1 + ar^2 + \cdots + ar^n = \frac{a(1-r^m)}{1-r} \right.$$

$$= \begin{cases} 1 & v = 0 \\ 0 & v \neq 0 \end{cases}$$

$$F(u,v) = \left.\begin{matrix} 1 \\ 0 \end{matrix}\right\} \quad \begin{matrix} u, v = 0 \\ u, v \neq 0 \end{matrix}$$

2. Consider the sample values of a 1D Signal as given below. Find out the DFT coefficients and also show that Inverse DFT Produces the original sample values

$x_0 = 0.5 \quad x_1 = 0.75 \quad x_2 = 1.0 \quad x_3 = 1.25$

Sol

$$F(0) = \frac{1}{4}\sum_{x=0}^{3} f(x)\, e^{-j\frac{2\pi}{4}\cdot u x} = \frac{1}{4}\Sigma$$

$$= \frac{1}{4}\left[0.5 \times e^{0}\right] \cancel{+ 0.75 \times e^{-j\frac{2\pi}{4}\cdot1\cdot1} + 1.0 \times e^{-j\frac{2\pi}{4}\cdot2\cdot2} + 1.25 \times e^{-j\frac{2\pi}{4}\cdot3}}$$

810 / 360 / 360 / 720

$$\cancel{\frac{1}{4}[0.5 + 0.4]} = 0.125$$

$$\cancel{f(x) = \sum F(u)\, e^{-j\frac{2\pi}{4}\cdot\ldots}}$$

$$= 0.125 \times 1 = 0.125$$

$$F(u) = \frac{1}{4}\left[0.5 \times 1 + 0.75 \times e^{-j90^\circ} + 1.00 \times e^{-j0^\circ} + 1.25 \times e^{-j\ldots}\right]$$

$$= \frac{1}{4}\left[0.5 +\right.$$

$e^{i\theta} = \cos\theta + i\sin\theta$

$e^{-i\theta} = \cos\theta + i\sin\theta$

Other Separable Transforms:

1D DFT $f(x)$

$$T(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\, e^{-j\frac{2\pi}{N}ux}$$

$$\frac{1}{N} e^{-j\frac{2\pi}{N}ux} = g(x,u)$$

$$f(x) = \frac{1}{N}\sum_{u=0}^{N-1} T(u)\, e^{j\frac{2\pi}{N}ux}.$$

$$h(x,u) = \frac{1}{N} e^{j\frac{2\pi}{N}ux}$$

$$f(x) = \sum_{u=0}^{N-1} T(u)\, h(x,u)$$

Similarly, 2D transforms

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

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

where $g(x,y,u,v)$, $h(x,y,u,v)$ are forward and inverse transform kernels.

Forward kernel is said to be separable if

$$g(x,y,u,v) = g_1(x,u)\, g_2(y,v)$$

Symmetric

$$g(x,y,u,v) = g_1(x,u)\, g_1(y,v)$$

The Inverse Fourier kernel also is Separable and Symmetric.

A transform with a separable kernel can be computed in two steps, each requiring a 1D transform. First, taking the 1D transform along each row of $f(x,y)$ yeilds.

$$T(x,v) = \sum_{y=0}^{N-1} f(x,y)\, g_2(y,v)$$

$$T(u,v) = \sum_{x=0}^{N-1} T(x,v)\, g_1(x,u).$$

If the kernel $g(x,y,u,v)$ is separable and symmtric, also may be expressed in matrix from

$$T = AFA$$

$F \rightarrow$ image of $N \times N$

$A \rightarrow$ symmtric transform matrix with elements $a_{ij} = g_1(i,j)$

$T \rightarrow$ result

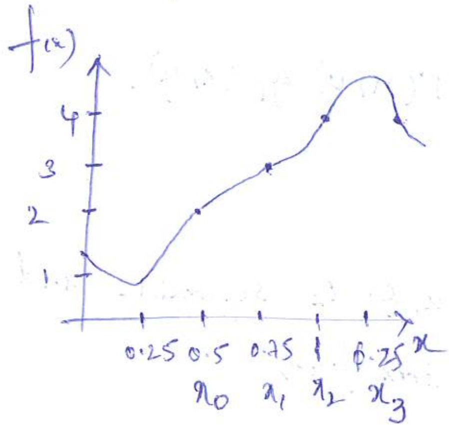
| If $B = A^{-1}$ | If $B \neq A^{-1}$ |
|---|---|
| $BTB = BAFAB$ | $\hat{F} = BAFAB$ |
| $F = BTB$ | |

Questions

1. Find out DFT coefficients of a digital Image $f(x,y)$ of size $N * N$ where $f(x,y) = 1$ for all values of $x$ and $y$.

2. Consider the sample values of a 1D signal as given below. Find out the DFT coefficients and also show that Inverse procedure the original sample values.

$x_0 = 0.5$, $x_1 = 0.75$, $x_2 = 1.0$, $x_3 = 1.25$



gital Image

el for all

1D signal

PT coefficients

re the Original

5

1. What is the time complexity of Fast Fourier Transform?

$O(N \log_2 N)$

2. Show that the DFT and its inverse are periodic functions.

4. Find out the Fourier coefficients for the following of 1D signal using FFT technique. Verify that the results obtained using FFT technique is same as that using direct implementation of DFT.

$f(0)=3,\ f(1)=2,\ f(2)=5,\ f(3)=4$

~~Descrite cosine transform~~

## Descrite cosine Transform!

$$g(x,y,u,v) = \alpha(u).\alpha(v).\cos\left[\frac{(2x+1)u\pi}{2N}\right].\cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$= h(u,v,x,y)$$

$$\cos x = \frac{e^{ix}+e^{-ix}}{2}$$

$$\alpha(u) = \begin{cases} \sqrt{1/N} & u=0 \\ \sqrt{2/N} & u=1,2,\dots N-1 \end{cases}$$

**Forward DCT**

$$C(u,v) = \alpha(u)\alpha(v).\sum_{x,y=0}^{N-1}\sum \left\{ f(x,y).\cos\left[\frac{(2x+1)\pi u}{2N}\right].\cos\left[\frac{(2y+1)\pi v}{2N}\right]\right\}$$

**Inverse DCT**

$$f(x,y) = \sum_{u,v=0}^{N-1}\sum \alpha(u)\alpha(v)\,C(u,v).\cos\left[\frac{(2x+1)\pi u}{2N}\right].\cos\left[\frac{(2y+1)\pi v}{2N}\right].$$

---

## Energy Compaction property:

Energy level are mostly located near to the origin this property called as "Energy compaction". ~~prop~~.

## Walsh Transform

**1D signal**

$$g(x,u) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_{n-1-i}(u)}$$

$N \rightarrow$ Number of samples ✓

$n$ – no. of bits $x/u$ need to represent

$b_k(z) \rightarrow k^{th}$ bit in the digital / binary representation of $z$.

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \cdot \prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_{n-1-i}(u)}$$

$$h(x,u) = \prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_{n-1-i}(u)}$$

$$f(x) = \cancel{\tfrac{1}{N}} \sum_{u=0}^{N-1} W(u) \prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_{n-1-i}(u)}$$

**2D signal**

$$g(x,y,u,v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{\{b_i(x) \cdot b_{n-1-i}(u) + b_i(y)\, b_{n-1-i}(v)\}}$$

$$h(x,y,u,v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{\{b_i(x)\, b_{n-1-i}(u) + b_i(y)\, b_{n-1-i}(v)\}}$$

$$W(u,v) = \frac{1}{N} \sum_{x,y=0}^{N-1}\sum f(x,y) \prod_{i=0}^{n-1} (-1)^{\{b_i(x)\, b_{n-1-i}(u) + b_i(y) \cdot b_{n-1-i}(v)\}}$$

$$f(x,y) = \frac{1}{N} \sum_{u,v=0}^{N-1}\sum W(u,v) \prod_{i=0}^{n-1} (-1)^{\{b_i(x) \cdot b_{n-1-i}(u) + b_i(y) \cdot b_{n-1-i}(v)\}}$$

**Satisfy:** Energy compaction.

nary representation of

$(u)$

$_{n-1-i}(u)$

$(u) + b_i(y)\, b_{n-1-i}(u)\}$

$+ b_i(y)\, b_{n-1-i}(u)\}$

$_{-i}(u) + b_0(y) \cdot b_{n-1-i}(u)\}$

$_{-1-i}(u) + b_i(y) \cdot b_{n-1-i}(u)\}$

$$W(u) = \frac{1}{2}\left[W(u)_{even} + W_{odd}(u)\right]$$

$$W(u+M) = \frac{1}{2}\left[W(u)_{even} + W_{odd}(u)\right]$$

$$\begin{cases} u = 0, \dots N/2 - 1 \\ \cancel{M = 0, \dots N/2} \quad M = N/2 \end{cases}$$

1.1

## 1D Walsh Transform kernel for N=8

| u \ x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 3 | 2 | 7 | 6 | 4 | 5 |
| 0 | + | + | + | + | + | + | + | + |
| 1 | + | + | + | + | − | − | − | − |
| 2 | + | + | − | − | + | + | − | − |
| 3 | + | + | − | − | − | − | + | + |
| 4 | + | + | + | − | + | − | + | − |
| 5 | + | − | + | − | − | + | − | + |
| 6 | + | − | − | + | + | − | − | + |
| 7 | + | − | − | + | − | + | + | − |

$(-1)^{0.0}$

$(-1)$

N=4

| u \ x | 0 (00) | 1 (01) | 2 (10) | 3 (11) |
|---|---|---|---|---|
| 00 0 | + | + | + | + |
| 01 1 | + | + | + | + |
| 10 2 | + | + | + | − |
| 11 3 | + | − | − | + |

$(-1)^{0.1} \cdot (-1)^{0.0}$

$(-1)^{0.0} \cdot (-1)^{0.0}$

$(-1)^{0.1} (-1)^{0.1}$

$(-1)^{0.0} \cdot (-1)^{1.0}$

$(-1)^{1.0} (-1)^{0.0}$

$(-1)^{1.0} (-1)^{1.0}$

$(-1)^{0.0} \cdot (-1)^{1.0}$

$(-1)^{0.0} (-1)^{1.0} = -1$

$(-1)^{0.1} (-1)^{1.1} = -$

$(-1)^{1.0} (-1)^{0.1} =$

$(-1)^{1.1} (-1)^{0.1} = -1$

00

$(-1)^{0.0} \quad (-1)^{0.1} \cdot (-1)^{1}$

| u \ x | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 00 0 | + | + | + | + |
| 01 1 | + | + | − | − |
| 2 | + | | | |
| 3 | + | | | |

# Hadamard Transform

0·0+0·0+1·1

0 0 + 0·0 + 0·0

1D

$$g(x,u) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} b_i(x)\, b_i(u)}$$

$$H(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) (-1)^{\sum_{i=0}^{n-1} b_i(x)\, b_i(u)}$$

$$h(x,u) = (-1)^{\sum_{i=0}^{n-1} b_i(x)\, b_i(u)}$$

$$f(x) = \sum_{u=0}^{N-1} H(u) \cdot (-1)^{\sum_{i=0}^{n-1} b_i(x)\, b_i(u)}$$

2D

$$g(x,y,u,v) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} \left( b_i(x)\, b_i(u) + b_i(y)\, b_i(v) \right)}$$

$$h(x,y,u,v) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} \left( b_i(x)\, b_i(u) + b_i(y)\, b_i(u) \right)}$$

$$H(u,v) = \frac{1}{N} \sum_{x,y=0}^{N-1}\sum f(x,y) (-1)^{\sum_{i=0}^{n-1} \left( b_i(x)\, b_i(u) + b_i(y)\, b_i(u) \right)}$$

1D ~~Transform~~ Hadamard Matrix ( N = 8)

| | 0 | 7 | 3 | 4 | 1 | 6 | 2 | 5 |
|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 0 | + | + | + | + | + | + | + | + |
| 1 | + | - | + | - | + | - | + | - |
| 2 | + | + | - | - | + | + | - | - |
| 3 | + | - | - | + | + | - | - | + |
| 4 | + | + | + | + | - | - | - | - |
| 5 | + | - | + | - | - | + | - | + |
| 6 | + | + | - | - | - | - | + | + |
| 7 | + | - | - | + | - | + | + | - |

$N = 2$

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_{2N} = \begin{bmatrix} H_N & H_N \\ H_N & H_{-N} \end{bmatrix}$$

Number of sign change in Hadamard matrix is not in order.

Walsh-Hadamard Transforms

$$g(x,u) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} b_i(x)\, P_i(u)}$$

$P_0(u) = b_{n-1}(u)$

$P_1(u) = b_{n-1}(u) + b_{n-2}(u)$

$P_2(u) = b_{n-2}(u) + b_{n-3}(u)$

$\vdots$

$P_{n-1}(u) = b_1(u) + b_0(u)$

1D Transform Modified Hadamard matrix ($N=8$)

0 u→ 1 v 3 4 5 6 7

| x ↓ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | + | + | + | + | + | + | + | + |
| 1 | + | + | + | + | − | − | − | − |
| 2 | + | + | − | − | − | − | + | + |
| 3 | + | + | − | − | + | + | − | − |
| 4 | + | + | − | + | + | − | − | + |
| 5 | + | − | − | + | − | + | + | − |
| 6 | + | − | + | + | − | + | − | + |
| 7 | + | − | + | − | + | + | + | − |

1. Which property of DCT makes it so popular for image compression application:

Energy compaction (Energy levels are mostly located near to the origin)

2. What is the transformation kernel for walsh transform?

forward kernel $\rightarrow g(x, y, u, v) = \frac{1}{N}\prod_{i=0}^{n-1}(-1)^{\{b_i(x)\, b_{n-1-i}(u) + b_i(y)\, b_{n-1-i}(v)\}}$

Inverse kernel $\rightarrow h(x, y, u, v) = \frac{1}{N}\prod_{i=0}^{n-1}(-1)^{\{b_i(x)\, b_{n-1-i}(u) + b_i(y)\, b_{n-1-i}(v)\}}$

3. what is the transformation kernel for Hadamard transform?

forward kernel $\quad g(x, y, u, v) = \frac{1}{N}(-1)^{\sum_{i=0}^{n-1}\{b_i(x)\, b_i(u) + b_i(y)\, b_i(v)\}}$

Inverse kernel $\quad h(x, y, u, v) = \frac{1}{N}(-1)^{\sum_{i=0}^{n-1}\{b_i(x)\, b_i(u) + b_i(y)\, b_i(v)\}}$

4. Explain the significance of modified Hadamard transform?

Ans

To order the Number of sign changes in Hadamard transform.

5. Find out the walsh transform coefficients for the following samples of a 1D signal.

$$f(0) = 3, \quad f(1) = 2, \quad f(2) = 5, \quad f(3) = 4$$

Ans

$$N = 4$$

$$W(u) = \frac{1}{\cancel{N}4} \sum_{x=0}^{4} f(x) \prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_{n-1-i}(u)}$$

1D ~~walsh~~ Transform Walsh Matrix $(N=4)$

| $x \backslash u$ | | 00 / 0 | 01 / 1 | 10 / 2 | 11 / 3 |
|---|---|---|---|---|---|
| 00 | 0 | + | + | + | + |
| 01 | 1 | + | + | − | − |
| 10 | 2 | + | − | + | + |
| 11 | 3 | + | − | + | + |

$(-1)^{0\cdot1} \cdot (-1)^{1\cdot0}$

$(-1)^{0\cdot0} (-1)^{1\cdot1} = -1$

$$W(0) = \frac{1}{4}\left[f(0)\cdot1 + f(1)\cdot1 + f(2)\cdot1 + f(3)\cdot1\right]$$

$$= \frac{1}{4}[3 + 2 + 5 + 4]$$

$$= \frac{14}{4} = 3.25$$

$$W(1) = \frac{1}{4}\left[f(0) + f(1) - f(2) - f(3)\right]$$

$$= \frac{1}{4}[3 + 2 - 5 - 4] = -1$$

$$W(2) = \frac{1}{4}\left[f(0) - f(1) + f(2) - f(3)\right]$$

$$= \frac{1}{4}[3 - 2 + 5 - 4] = \frac{2}{4} = 0.5$$

$$W(3) = \frac{1}{4}\left[f(0) - f(1) - f(2) + f(3)\right] = \frac{1}{4}[3 - 2 - 5 + 4]$$

$$= 0.$$

# Image Enhancement

The principal objective of enhancement is to process an image so that the result is more suitable than the original image for a specific application.

1) Spatial Domain

2) Frequency Domain

**Spatial Domain:**

The term spatial domain refers to the image plane itself, and approaches in this category are based on direct manipulation of pixels in an image.

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

**Frequency Domain:**

The term frequency domain refers to the Fourier transform coefficients, and approaches in this category are based on indirect manipulation of pixels in an image.

$$g(x,y) = \mathcal{F}^{-1}[\mathcal{F}(f(x,y))]$$

# Histogram Processing

- The histogram of a digital image with gray levels in the range $[0, L-1]$ is a discrete function $h(r_k) = n_k$

$r_k \rightarrow k^{th}$ gray level

$n_k \rightarrow$ number of pixels in the image with gray level $r_k$

**Normalized histogram:**

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

- An estimate of the probability of occurrence of gray level $r_k$:

$$P(r_k) = n_k / n$$

$n \rightarrow$ total number of pixels in image.

**Note:** Dark image that the components of the histogram are concentrated on the low (dark) side of the gray scale.

Bright image — Higher side

low contrast image — middle

High contrast image — equally distributed.

| Type of Image | Pixel distribution |
| --- | --- |
| Dark Image | Lower side |
| Bright image | Higher side |
| Low contrast | Middle |

41.

# Histogram Equalization

Histogram equalization is a technique for adjusting image intensities to enhance contrast.

Let $f(x,y)$ — $m,n$ — $0, L-1$

$P$ — denote Normalized histogram of $f$

$$P(r_k) = \frac{\text{Number of pixels with intensity } r_k}{\text{total Number of pixels.}}$$

$$S_k = T(r_k) = \text{Round}\left((L-1)\sum_{j=0}^{K} P_r(r_j)\right)$$

Histogram equalization of image $f$ is $g$

$$g(x,y) = \text{floor}\left((L-1)\sum_{r_k=0}^{\cancel{f(x,y)}\; L-1} P(r_k)\right)$$

This is equivalent to transforming the pixel intensities, $k$ of $f$ by function

$$T(k) = \text{floor}\left((L-1)\sum_{r_k=0}^{\cancel{f(x,y)}\; L-1} P(r_k)\right)$$

If $f(x,y)$ ~~are~~ continuous function $X, Y$

$$Y = T(X) = (L-1)\int_0^X P_X(x)\, dx$$

$P_X \rightarrow$ PDF of $f$

$T \rightarrow$ CDF of $X$

$T$ is invertible

$$X = T^{-1}(Y)$$

$$T(X) = (L-1)\int_0^{T^{-1}(y)} P_x(x)\,dx = Y$$

$$\frac{d}{dy}\left(\int_0^{y} P_y(z)\,dz\right) = P_y(y) = P_x(T^{-1}(y))$$

Let consider an continuous function

$$0 \le r \le 1$$

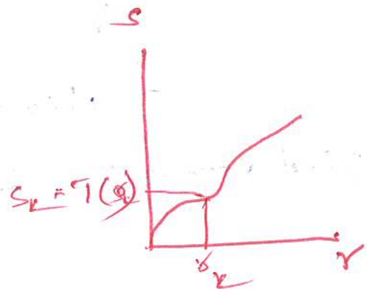
Pixel value – $[0, L-1]$

$$s = T(r) \qquad 0 \le r \le 1$$

$T(r) \rightarrow$ Transformation function & Inverse transform

(a) $T(r)$ is a single-valued and monotonically increasing function in the interval $0 \le r \le 1$

(b) $0 \le T(r) \le 1$ for $0 \le r \le 1$




$$r = T^{-1}(s) \qquad 0 \le s \le 1$$

$$P_s(s) = P_r(r)\left|\frac{ds}{dr}\right|$$

$$s = T(r) = \int_0^{r} P_r(w)\,dw$$

CDF

$$\frac{ds}{dr} = \frac{d\,T(r)}{dr}$$

$$= \frac{d}{dr}\left[\int_0^{r} P_r(w)\,dw\right]$$

$$= P_r(r)$$

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

$$= 1 \qquad 0 \le s \le 1$$

$$s_k = T(r_k) = \sum_{j=0}^{k} P_r(r_j)$$

Ex Perform Histogram Equalization on the following 8x8 image the gray level distribution is given in the below

90%

64

| Graylevel | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| No. of pixels | 8 | 10 | 10 | 2 | 12 | ~~16~~ 11 | 9 | 2 |

64

Total No. pixels = 64

| $r_k$ | $n_k$ | PDF $P_r(r_k)$ | CDF $C_m$ | $(L-1)*C_m$ | Round.off | No. pixels |
|---|---|---|---|---|---|---|
| 0 | 8 | 0.125 | 0.125 | 0.875 | 1 | 8 |
| 1 | 10 | 0.15625 | 0.28125 | 1.96 | 2 | 10 |
| 2 | 10 | 0.15625 | 0.34375 | 3.06 | 3 | 10 |
| 3 | 2 | 0.03125 | 0.46785 | 3. | 3 | 2 |
| 4 | 12 | 0.1875 | 0.65625 | 5 | 5 | 12 |
| 5 | 16 | 0.25 | 0.90625 | 6 | 6 | 9 |
| 6 | 9 | 0.0625 | 0.96875 | 7 | 7 | |
| 7 | 2 | 0.03125 | 1 | 7 | 7 | 2 |
| | 69 | | | | 34 | 53 |

Exercise

The histogram of the whole image value r ranging from 0 to 7 is given by

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| h | 2 | 4 | 5 | 2 | 3 | 3 | 3 | 3 |

Use histogram equalization and arrive at the mapping of old r to new r and plot the corresponding new r values.

# Image Enhancement

* Processing an Image to enhance certain features of the image.

* The result is more suitable than the original image for certain specific application.

Ex:
- Best technique for enhancement of x-ray image may not be the best for enhancement of microscopic images.

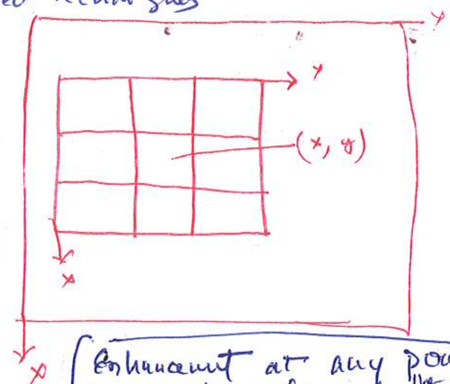
Image Enhancement

Direct manipulation of pixels — Spatial { Point processing, Mask processing, Histogram Based techniques

Frequency — Modify DFT | DCT | DWT coefficients of an image.

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

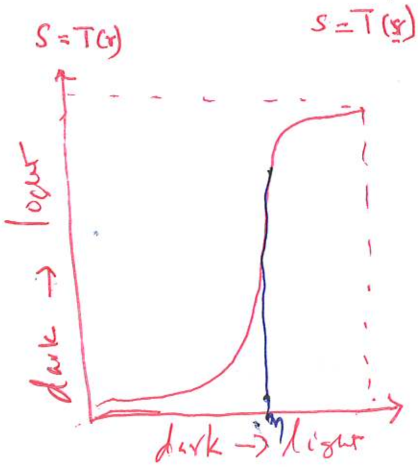
$g(x,y) = T[f(x,y)]$ ↳ neighborhood



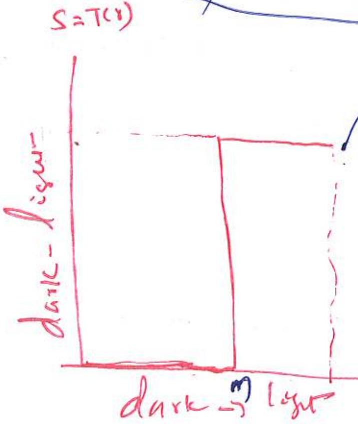

Point processing:

Enhancement at any point in an image depends only on the gray level at that point; techniques in this category often referred to as point processing

$S = T(r)$




Construct stretching function




Thresholding function

## Mask processing

- brightening
- sharpening
- edge darking

| | | |
|---|---|---|
| $w_{-1,-1}$ | $w_{-1,0}$ | $w_{-1,1}$ |
| $w_{0,-1}$ | $w_{0,0}$ | $w_{0,1}$ |
| $w_{1,-1}$ | $w_{1,0}$ | $w_{1,1}$ |

3×3 Mask

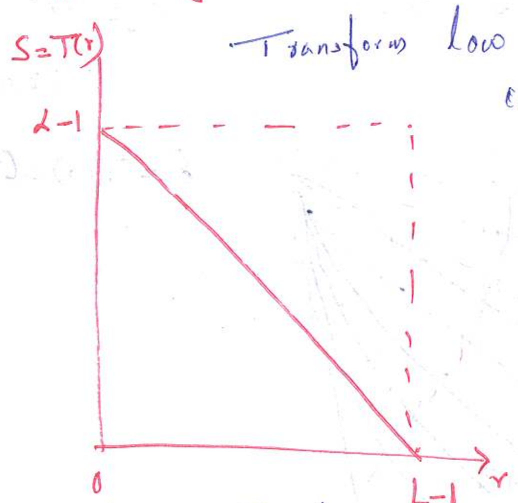
$$g(x,y) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} w_{i,j} \cdot f(x+i, y+j)$$

Gray level Trasform :-

Image Negative

- x-ray
- MRI scan
- CT scan

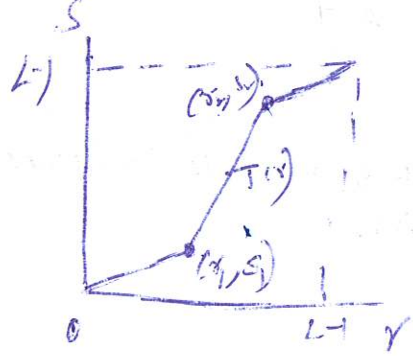
Transform low pixel ~~intensity~~ intensity to high in processed image and vise verse.



$S = T(r) = \underline{L-1-r}$

Piecewise - Linear Transformation :-

Constrast stretching



$s_1 \leq s_2$

$r_1 \leq r_2$

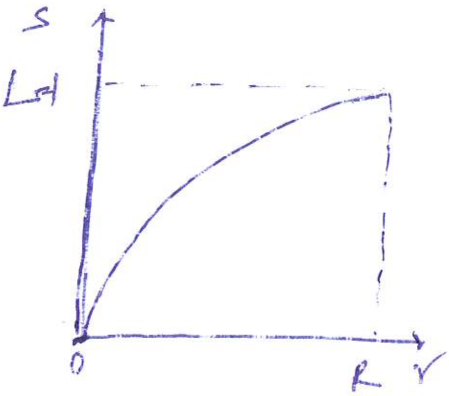
- poor (or) illumination
- object itself dark
- Dynamic range of the Sensor ✓

grey

1 - 255

$s_1$ $s_2$

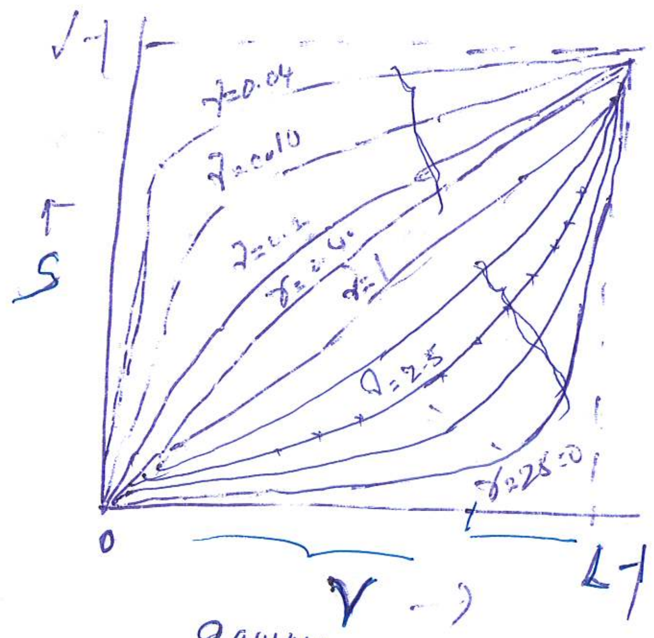
1) Dynamic Range compression (or) Log Transform.




Limited

$s = T(r) = C \log(1 + |r|)$

Gamma Transform
(or) power Law Transforms (Image capture, Image display, printing)




$s = T(r) = C r^{\gamma}$

gamma

The process used to correct this powerlaw response phenomena is called gamma correction.

Transform.

$c\log(1+|r|)$

...age capture

...age display

printing

$= T(r) = c r^{\gamma}$

...aw response

1. What meant by image enhancement?

2. What are the different types of image enhancement techniques?

3. What is the transformation function to create image negatives?

4. For what type of images, negative transformation is useful?

5. A captured image appears very dark because of wrong lens aperture setting. Which enhancement technique is appropriate to enhance such image?

1. What is a histogram?

~~2. Give~~

Ans. An Image histogram is a graphical representation of the tonal distribution in a digital image.

(or)

It plots the number of pixels for each tonal value.

2. Give the transformation function of histogram equalization technique.

Ans.

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

$$= \sum_{i=0}^{k} \frac{n_i}{n} \qquad \left[\because P_r(r_i) = \frac{n_i}{n}\right]$$

3. What should be the nature of the histogram of a histogram equalized image?

Ans. In histogram, pixels distributed across the all tonal values.

(or)

Uniform distribution of pixels in all tonals.

~~4.~~

4. Suppose a digital image is subjected to histogram equalization. What effect will a second pass equalization have over the equalized image?

Ans In idle case no effect.

5. What is histogram specification technique?

Ans. To convert the image so that it has a particular histogram that can be arbitrarily specified. Such mapping function can be found on three steps

1) Equalize the histograms of input image
2) Equalize the specified histogram 3) Relate the two equalized histogra

6.

to

Ans

ion of the tonal

tonal value.

m equalization

$(r_i) = \frac{n_0}{n}$ ]

m of a histogram

tonal values.

onals.

to histogram
ss equalization

nique?

rticular histogram
function can be

image

3> Relate the two equalized histograms.

6. What conditions must be satisfied by the target histogram to be used in histogram specification technique?

Ans

$P_z(z) \neq 0$ ~~for every possible value of~~ $\forall z$

| $r$ | $n_k$ | $n_k/n$ |
| --- | --- | --- |
| 0 | 729 | 0.19 |
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 7 | | |

$s_0 = (L-1) \sum_{i=0}^{0} p_r(r_i)$

$= (L-1) \times 0.19$

$= 7 \times 0.19 = 1.33$

$s_0 = 7 \times \frac{729}{64 \times 64}$

$729 = \frac{1.33 \times 64 \times 64}{7}$

$s_0 = 729 \to 1.33$

$s_{k=0,7}$

$s_0 = \sum_{i=0}^{0} p_r(r_i)$ $\quad n_k/n$

$= 0.19$

$s_1 = \sum_{i=0}^{1} \left( p_r(r_0) + p_r(r_1) \right)$

$s_7 = 1$

Histogram specification / Matching: —

It is the transformation of an image so that its histogram matches a specific histogram.

Procedure 1. Obtain histogram of a specific image

2. Calculate $P_r(r_k) = \frac{n_k}{n} \longrightarrow$ from the given image

PDF $r_k$ to CDF

$\boxed{P_z(z_k) \Rightarrow \text{Target histogram}}$

PDF

$S_k = T(r_k) = \sum_{i=0}^{k} \frac{n_i}{n} = \sum_{i=0}^{k} P_r(r_i)$ CDF

Calculate CDF to $G = P_z(z)$

3. $Z_k \Rightarrow V_k = G(z_k) = \sum_{i=0}^{k} P_z(z_i) \cong S_k$

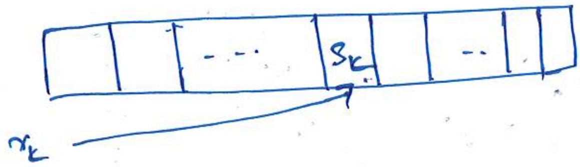
4. Equate $Z_k = G^{-1}(S_k)$

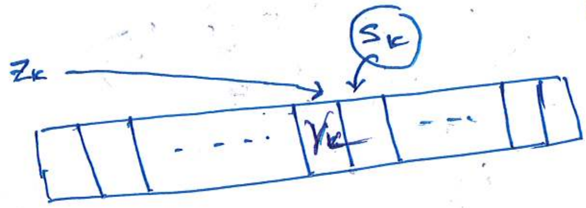
$= G^{-1}[T(r_k)]$ [Always Analytical solution not possible]

Iterative approach

Def

The method used to generate a processed image that has a specific histogram is called histogram matching.





$V_k = G_*(z_k) = S_k$

$G(z_k) - S_k = 0$

$\hat{z} \rightarrow$ Assume value

$G(\hat{z}_k) - S_k \geq 0$

$\hat{z} \rightarrow$ start from zero

it is the transform
ation of an image so
that its histogram
matches a specific
histogram.

DF

solution not possible]

ethod used to
e a processed
that has a
ic histogram
called histogram
ching.

## Mask Processing

- A subimage called Mask (or) filter (or) kernel.

- Linear Smoothing filter
- Median filter (nonlinear)
- Sharpening filter

## Smoothing filter (or) Averaging Filter (Lowpass filter)

→ Smoothing filters are used for blurring and Noise reduction.

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \text{Box Filter}$$

- The Smoothing linear filter is simply the Average of pixel contained in the neighborhood of the filter mask.

$$g(x,y) = \frac{1}{9} \sum_{i=-1}^{1} \sum_{j=-1}^{1} f(x+i, y+j)$$

- It is also called Averaging filters (or) also referred to low pass filters.

* Sharp Edge become blur (or) bluring effect

## Weighted average

$$\frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$g(x,y) = \frac{\sum_{i=-1}^{1} \sum_{j=-1}^{1} w_{i,j} * f(x+i, y+j)}{\sum_{i=-1}^{1} \sum_{j=-1}^{1} w_{i,j}}$$

Mask size = M x N

$M = 2a$   $N = 2b$

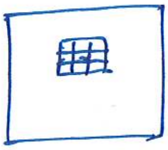
$$g(x,y) = \frac{\sum_{i=-a}^{a} \sum_{j=-a}^{a} w_{i,j} f(x+i, y+j)}{\sum_{i=-a}^{a} \sum_{j=-a}^{a} w_{i,j}}$$

→ The values of every pixels in an image by the average of the gray levels in neighborhood defined by the filter mask, this process results in an image with reduced "sharp" transitions in gray levels.

* Reduced bluring effect

Nonlinear filter (Median filter) 

→ Order statistic filter



Non-linear filter.

→ The output is based on ordering the pixels contained in the image area encompassed by the filter, and then replacing the value of the center pixel value determined by ranking result.

Eg

| | | |
|---|---|---|
| 100 | 85 | 98 |
| 99 | 105 | 102 |
| 90 | 101 | 108 |

$f(x,y)$ ⇒

85
90
98
99 } lessing
100
101
102
105
108 } greater than.

| | | |
|---|---|---|
| x | x | x |
| x | 100 | x |
| x | x | x |

→ Median filter are effective on presence of salt-and-pepper Noise,

Sharpening spatial filter:

First order derivation
Second Order derivation

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$$\frac{df}{dx} = f(x+1) - f(x) \rightarrow 1 \text{ order}$$

$$\frac{d^2f}{dx^2} = f(x-1) + f(x+1) - 2f(x) \rightarrow 2 \text{ order}$$

Laplacian operator

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

...statistic filter(s)

...mean filter,

...ar based on the pixels contained ...image area encompa... the filter, and then the value of the center pixel value determined by ranking result.

...dian filter are ...ive on presence of ...-and-pepper ...noise,

order

...,y)

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

$$\nabla^2 f = \frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2}$$

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

Laplacian Mask

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

| | | |
|---|---|---|
| -1 | -1 | -1 |
| -1 | 8 | -1 |
| -1 | -1 | -1 |

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f \\ f(x,y) + \nabla^2 f \end{cases}$$

Laplacian Mask is negative

Laplacian mask is positive

Composite Mask

| | | |
|---|---|---|
| 0 | -1 | 0 |
| -1 | 5 | -1 |
| 0 | -1 | 0 |

| | | |
|---|---|---|
| -1 | -1 | -1 |
| -1 | 9 | -1 |
| -1 | -1 | -1 |

Unsharp Masking

Get sharpen image from blurred image of original image.

$$\boxed{f_s(x,y) = f(x,y) - \bar{f}(x,y)}$$

$f(x,y) \to$ original image

$\bar{f}(x,y) \to$ blurred $f(x,y)$

$f_s(x,y) \to$ sharped image

Highboost Filtering

It is Enhance form of Unsharp masking.

$A \geq 1$

$$f_{hb}(x,y) = A f(x,y) - \bar{f}(x,y)$$

$$= (A-1) f(x,y) + f(x,y) - \bar{f}(x,y)$$

$$f_{hb}(x,y) = (A-1) f(x,y) + f_s(x,y)$$

$$f_{hb}(x,y) = \begin{cases} A f(x,y) - \nabla^2 f(x,y) & \text{Negative} \\ A f(x,y) + \nabla^2 f(x,y) & \text{positive} \end{cases}$$

| | | |
|---|---|---|
| 0 | -1 | 0 |
| -1 | A+4 | -1 |
| 0 | -1 | 0 |

| | | |
|---|---|---|
| -1 | -1 | -1 |
| -1 | A+8 | -1 |
| -1 | -1 | -1 |

High boost Mask

Note:
If A=0, it is Laplacian operator
If A ≥ 2, by increasing A value contrast of image will increase.

Gradient Operator:

$$\vec{\nabla} f = \begin{bmatrix} \frac{df}{dx} \\ \frac{df}{dy} \end{bmatrix}$$

The gradient of a digital image $f(x,y)$ is defined as first order derivative in two dimensional column vector.

Magnitude:

$$\nabla f = |\vec{\nabla} f| = \left[ \left(\frac{df}{dx}\right)^2 + \left(\frac{df}{dy}\right)^2 \right]$$

| | | |
|---|---|---|
| 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 |

$$\simeq \left|\frac{df}{dx}\right| + \left|\frac{df}{dy}\right| \quad \left[ f(x-1,y) + f(x,y) \right] +$$

$$\frac{df}{dx} \simeq \left[ f(x+1, y-1) + f(x+1, y+1) + 2f(x+1, y) \right] - \left[ f(x-1, y-1) + f(x-1, y+1) + 2f(x-1, y) \right]$$

Robert's

| | | |
|---|---|---|
| $z_1$ | $z_2$ | $z_3$ |
| $z_4$ | $z_5$ | $z_6$ |
| $z_7$ | $z_8$ | $z_9$ |

$G_x = (z_8 - z_5)$

$G_y = (z_6 - z_5)$

(or)

$G_x = (z_9 - z_5)$

$G_y = (z_8 - z_6)$

| | |
|---|---|
| -1 | 0 |
| 0 | 1 |

| | |
|---|---|
| 0 | -1 |
| 1 | 0 |

Sobel operator $\frac{df}{dx}$

| y / x | -1 | 0 | 1 |
|---|---|---|---|
| -1 | -1 | -2 | -1 |
| 0 | 0 | 0 | 0 |
| 1 | 1 | 2 | 1 |

$$\frac{df}{dy} \simeq \left[ f(x-1, y+1) + f(x+1, y+1) + 2f(x, y+1) \right] - \left[ f(x-1, y-1) + f(x+1, y-1) + 2f(x, y+1) \right]$$

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

Sobel operator

$$\nabla f = \left| (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \right| + \left| (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \right|$$

Note :

If $A = 0$, it is Laplacian operator

If $A \geq 2$, by increasing $A$ value contrast of image will increase.

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

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

$2f(x+1, y)] +$

$2f(x-1, y)]$

Robert's

| | | |
|---|---|---|
| | | $z_3$ |
| | | $z_6$ |
| | | $z_9$ |

$G_x = (z_8 - z_5)$

$G_y = (z_6 - z_5)$

(or)

$G_x = (z_9 - z_5)$

$G_y = (z_8 - z_6)$

| | |
|---|---|
| 0 | -1 |
| 1 | 0 |

$2f(x, y+1)] -$

$2f(x, y+1)]$

$(z_7 + 2z_8 + z_9)$

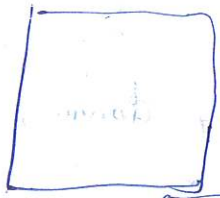
$z_1 + 2z_2 + z_3)$

$z_3 + 2z_6 + z_9) -$

$(z_1 + 2z_4 + z_7)$

## Mask Processing (Convolution)

Localization of Histogram



$512 \times 512$

# Frequency Domain Image Enhancement

Apply Image enhancement technique such filtering on frequency domain coefficients.

Basics of filtering in frequency Domain:

1. Multiply the input image by $(-1)^{x+y}$

$$f'(x,y) = f(x,y) * (-1)^{x+y}$$

2. Compute $T(u,v)$ frequency domain coefficients.

$$T(u,v) = T(f'(x,y))$$

Example

$$F(u,v) = \mathcal{F}(f'(x,y))$$

3. Multiply $T(u,v)$ by filter function

$$G(u,v) = H(u,v) * F(u,v)$$ [Zero-phase shift filters]

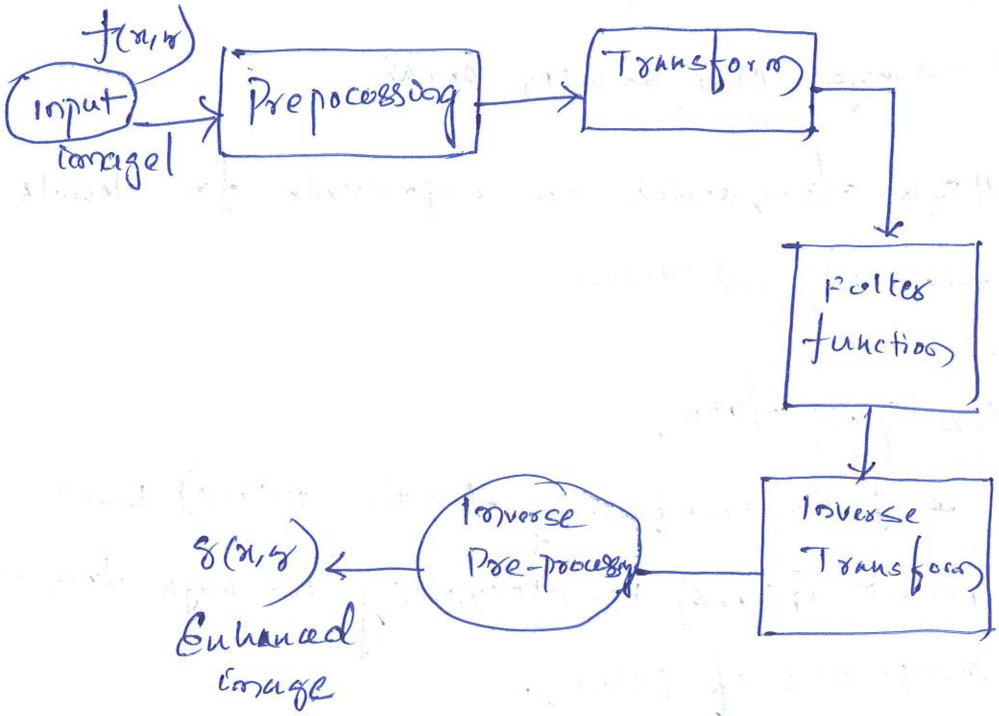
4. Compute the inverse transform of $G(u,v)$.

$$f''(x,y) = T^{-1}(G(u,v))$$

5. Obtain real part of 4

6. Multiply $(-1)^{x+y}$ in 5.

$$g(x,y) = f'''(x,y) = (-1)^{x+y} \times Real(f''(x,y))$$




## Some Basic filter

1) Low pass filter

2) High Pass filter

~~3) Butterworth filter~~

1) Low Pass Filter:

A filter that attenuates high frequencies while "passing" low frequencies is called a low pass filter.

2) High pass filter

A filter that attenuates low frequencies while "passing" high frequencies is called a "high pass filter."

Note:

Low frequencies in the Fourier transform are represent the general graylevel appearance of an image over smooth areas.

High frequencies are responsible for details such edge and Noise.

---

Low pass filter:

A filter transforming function $H(u,v)$ that result $G(u,v)$ by attenuating the high frequency components of $F(u,v)$.

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

$F(u,v) \rightarrow$ Fourier transform function

$H(u,v) \rightarrow$ transform function

$G(u,v) \rightarrow$ resultant

Types of low pass filters.

1) Ideal Lowpass filter (ILPF)
2) Butterworth Lowpass filter (BLPF)
3) Gaussian Lowpass filter (GLPF)

## Comparison between Spatial Domain & frequency Domain filters:

→ Most fundamental relationship between spatial and frequency domain filters are "Convolution theorem".

$$\begin{cases} f(x,y)\,h(x,y) \Leftrightarrow F(u,v) * H(u,v) \\ f(x,y) * h(x,y) \Leftrightarrow F(u,v) \cdot H(u,v) \end{cases}$$

— Displayment

→ Notch filter (frequency domain)

$$H(u,v) = \begin{cases} 0 & \text{if } (u,v) = (M/2, N/2) \\ 1 & \text{otherwise} \end{cases}$$

Mask

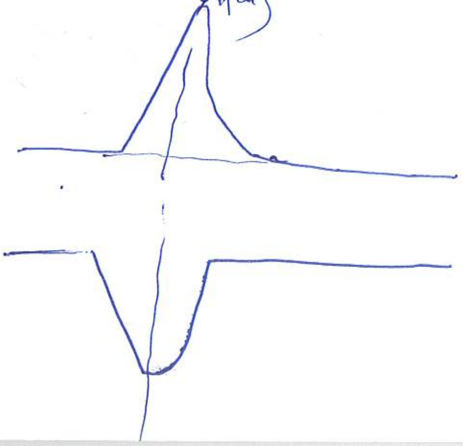
→ Example: Gaussian filter

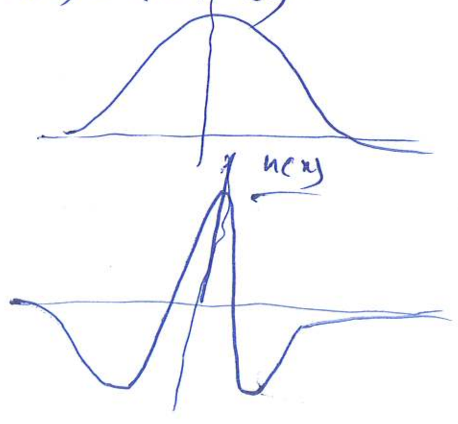
Frequency

$H(u) = A e^{-u^2/2\sigma^2}$

Spatial

$h(x) = \sqrt{2\pi}\,\sigma A e^{-2\pi^2\sigma^2 x^2}$





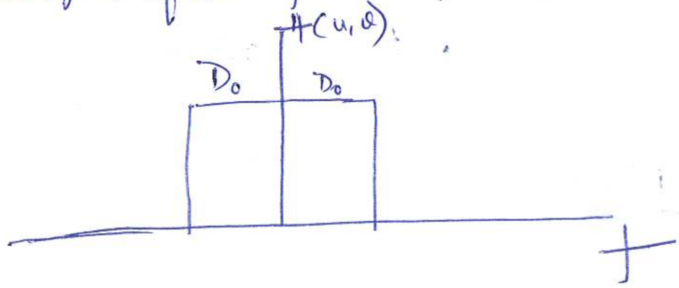
## Ideal Lowpass Filters:

A filter that "cuts off" all high-frequency components of Fourier transform that are at a distance greater than a specific distance $D_0$ from the origin of the transform. Such a filter is called a two-dimensional (2D) Ideal LowPass filter (ILPF).

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

where $D_0$ is a specified nonnegative quantity

$D(u,v)$ is the distance from point $(u,v)$ to the origin of the frequency rectangle.




Let

Image $f(x,y)$ of size $M \times N$.

Origin of the frequency $(u,v) = (M/2, N/2)$

Distance between origin and any point:

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

cut off frequency:

standard cutoff frequency: $P_T$ Calculation

total image power $P_T$:

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

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

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

If the transform has been centered, a circle of radius r with origin at the center of the frequency rectangle encloses $\alpha$ percent of the power,

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

---

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

from convolution theorem

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

Note:

$h(x,y)$ has two major distinctive characteristics:

i) Dominant component at the origin ~~and~~

ii) Concentric, circular components about the center.

The center component is responsible for blurring.

Concentric components are responsible for ringing characteristic of ILPL.

Note

⇒ Radius of the center component and the number of concentric circles per unit distance are inversely proportional to cut off frequency.

Note: Ringing artifacts are artifacts that appears spurious signals near edges.

## Butterworth Lowpass Filters:—

Objective:

Show that it is possible to achieve blurring with little or no ringing.

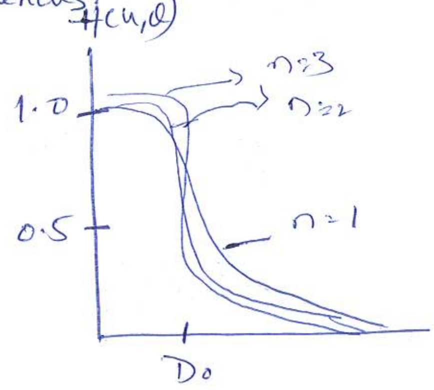
The transform function of a Butterworth lowpass filter (BLPF) of order n, with cutoff frequency at distance $D_0$ from origin, defined as

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

$$D(u,v) = |\cancel{F(u,v)}|^2$$
$$= \left[(u-M/2)^2 + (v-N/2)^2\right]^{1/2}$$

Difference:

Unlike the ILPF, the BLPF transform function does not have a sharp discontinuity that establishes a clear cutoff between passed and filter frequencies




Note:

- A Butterworth filter of order n=1, has no ringing.

- Generally, Ring behaviour is directly proportional to order n.

Note:

A Butterworth filter order n=20 exhibits the characteristics of ILPF.

BLPF of order n=2 is a good compromise between effective low pass filtering and acceptable ring characteristic.

## Gaussian Lowpass Filters:—

The transform function of a 2D Gaussian filter,

defined as

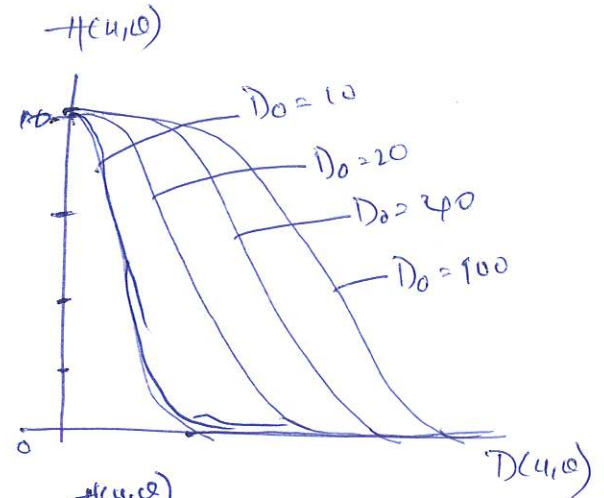
$$H(u,v) = e^{-D^2(u,v)/2\sigma^2}$$

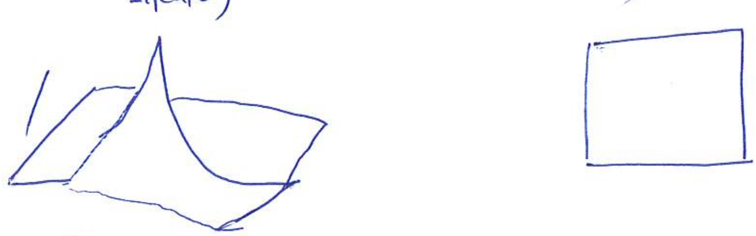
$D(u,v) \rightarrow$ distance from origin

$\sigma$ — is measure of spread of Gaussian curve

$\sigma = D_0$

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




Perspective plot of GLPF

Note:

- An Image can be blurred by attenuating the high-frequency components of it Fourier transform.
- Because edge and other abrupt changes in grey levels are associate with ~~edges~~ high-frequency components.

## High Pass Filters:

Image sharpening can be achieved in frequency domain by a high-pass filtering process, which attenuates the low-frequency components without disturbing high frequency information in frequency domain in Fourier transform.

A high pass filter in terms of low pass filter, and defined as

$$H_{hp}(u,v) = 1 - H_{lp}(u,v)$$

## Ideal High pass filter:

A Ideal high pass filter is defined as

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

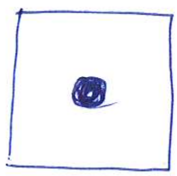
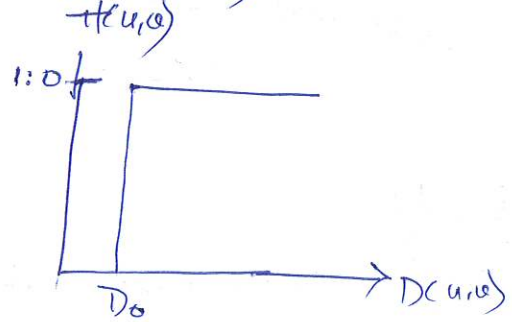


Image of IDHF




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

$D_0$ – cut off frequency

**Note:**

* Ringing effect

# Image Compression

Image compression addresses the problem of reducing the amount of data required to represent a digital image, by removal of redundant data.

## Applications

Televideo conferencing
remote sensing
document and medical imaging
facsimile transmission (FAX)

control of remotely piloted vehicles in military.
space, and
hazardous waste control application

## Categories

1) Information preserving (legal (or) Medical records)
2) Lossy (which provide higher levels of data reduction but result in a less than perfect reproduction of original image).
(FAX, Broadcaste television, videoconferency)

## Fundamentals

The term data compression refers to the process of reducing the amount of data required to represent a given quantity of information.

## Data Redundancy:

It contains data that either provide no relevant information (or) simply restate that which is already known. called as data redundancy.

If $n_1$ and $n_2$ denote the number of information carrying units in two data sets that represent the same information, the relative data redundancy $R_D$ of the first data set can be defined as

$$R_D = 1 - \frac{1}{C_R}$$

where $C_R$ commonly called the compression ratio

$$C_R = \frac{n_1}{n_2}$$

If $n_2 = n_1$, $C_R = 1 \Rightarrow R_D = 0$; No Redundant data

If $n_2 << n_1$, $C_R \simeq n_1$; $R_D \approx 1$; Significant compression & highly redundant

If $n_2 >> n_1$, $C_R \approx 0$; $R_D = -\infty$ Second data set consists much more ~~information~~ data than the original.

Compression ratio 10 (10:1)

$R_D = 0.9$ 90% of first data set is redundant.

In digital image compression, three basic data redundancies can be identified and exploited:

1) Coding redundancy
2) Interpixel redundancy
3) psychovisual redundancy

## Code Redundancy :- (Histogram Modification)

Discrete Random Variable $r_k \rightarrow [0,1]$

$$P_r(r_k) = \frac{n_k}{n} \; ; \quad k = 0, 1, 2, \ldots L-1$$

where $L$ is the number of gray levels,

$n_k$ is number of times that the $k^{th}$ gray level appears in the image

$n$ is the total number of pixels in the image

If the number of bits used to represent each value of $r_k$ is $l(r_k)$, the Average number of bits required to represent each pixel is

$$L_{avg} = \sum_{k=0}^{L-1} l(r_k) \cdot P_r(r_k)$$

to code an $M \times N$ Image $MNL_{avg}$.

**Example**

An 8-level image has the gray level distribution shown below. If a natural 3-bit binary code is used to represent the 8 possible gray levels, $L_{avg}$ is 3 bits,

Find the $L_{avg}$ and $R_D$.

Variable Length Coding Example

| $r_k$ | $P_r(r_k)$ | code1 | $l_1(r_k)$ | code2 | $l_2(r_k)$ |
|---|---|---|---|---|---|
| $r_0 = 0$ | 0.19 | 000 | 3 | 11 | 2 |
| $r_1 = 1/7$ | 0.25 | 001 | 3 | 01 | 2 |
| $r_2 = 2/7$ | 0.21 | 010 | 3 | 10 | 2 |
| $r_3 = 3/7$ | 0.16 | 011 | 3 | 001 | 3 |
| $r_4 = 4/7$ | 0.08 | 100 | 3 | 0001 | 4 |
| $r_5 = 5/7$ | 0.06 | 101 | 3 | 00001 | 5 |

ldification)

vels, $r_k$

$k^{th}$ gray level

s in the image

epresent each

umber of bits

MN $L_{avg}$.

distribution shows

used to represent

| ) | code2 | $l_2(r_k)$ |
|---|---|---|
| | 11 | 2 |
| | 01 | 2 |
| | 10 | 2 |
| | 001 | 3 |
| | 0001 | 4 |
| | 00001 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| $r_6 = 6/7$ | 0.03 | 110 | 3 | 000001 | 6 |
| $r_7 = 1$ | 0.02 | 111 | 3 | 000000 | 6 |

$$L_{avg_1} = \sum_{k=0}^{L-1} l(r_k) \cdot P_r(r_k)$$

$$= 3$$

$$L_{avg_2} = 2\cdot(0.19) + 2\cdot(0.25) + 2(0.21) + 3(0.16) + 4(0.08) + 5(0.06) + 6(0.03) + 6(0.02)$$

$$= 2.7 \text{ bits}$$

$$C_R = 3/2.7 = 1.11$$

$$R_D = 1 - \frac{1}{C_R} = 1 - \frac{1}{1.11}$$

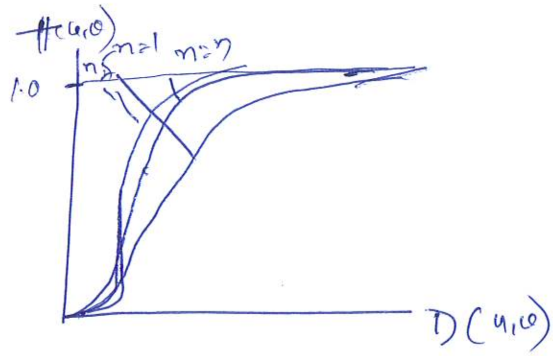
$$= 0.099$$

Butterworth High Pass Filter

- A Butterworth high pass filter (BHPF) of order n and cutoff frequency $D_0$ from origin, is defined as

transforming function of

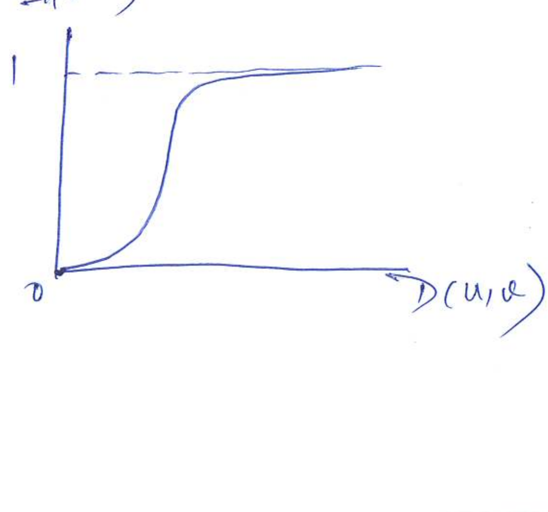
$$H(u,v) = \frac{1}{1+\left[D_0/D(u,v)\right]^{2n}}$$




Gaussian High pass Filter

- The transfer function of the Gaussian highpass filter (GHPF) with cut off frequency $D_0$, is defined as

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



## The Laplacian in the frequency domain:—

Enhancement by

The Laplacian ~~filter~~ of frequency domain, as defined as

$H(u,v)$ $\mathcal{F}[\nabla^2 f(x,y)]$. where $\nabla^2 f(x,y)$ is Laplacian operator and $\mathcal{F}$ is forward fourier transform.

$$\mathcal{F}[\nabla^2 f(x,y)] = \mathcal{F}\left[\frac{d^2 f(x,y)}{dx^2} + \frac{d^2 f(x,y)}{dy^2}\right].$$

$$\boxed{\therefore \mathcal{F}\left[\frac{d^n f(x)}{dx^n}\right] = (ju)^n F(u)}$$

$$\left[\because \mathcal{F}[f(x,y)] = [F(u,v)]\right.$$

$$= \iint_{x\,y} f(x,y) \cdot e^{-j\frac{2\pi}{N}(ux+vy)}$$

$$= (ju)^2 F(u,v) + (jv)^2 F(u,v)$$

$$= -(u^2+v^2) F(u,v)$$

From that laplacian transforming function

$$H(u,v) = -(u^2+v^2)$$

By shifting the center to $(M/2, N/2)$

$$H(u,v) = -\left[(u-M/2)^2 + (v-N/2)^2\right]$$

From Laplacian transformed image.

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

$$\mathcal{F}[\nabla^2 f(x,y)] = -[(u-M/2)^2 + (v-N/2)^2] \cdot F(u,v)$$

$$\nabla^2 f(x,y) = \mathcal{F}^{-1}\{-[(u-M/2)^2 + (v-N/2)^2] \cdot F(u,v)\}$$

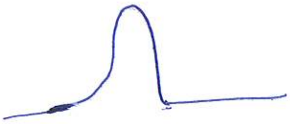
$$= \mathcal{F}[(u-M/2)^2 + (v-N/2)^2] \cdot f(x,y)$$

Laplacian Enhancement

Enhanced image in spatial domain:

$$g(x,y) = f(x,y) + \nabla^2 f(x,y)$$

$$= \mathcal{F}^{-1}\{F(u,v) + [(u-M/2)^2 + (v-N/2)^2] f(x,y)\}$$



$$= \mathcal{F}^{-1}\{F(u,v) - [((u-M/2)^2 + (v-N/2)^2] F(u,v)\}$$

$$G(u,v) = \mathcal{F}^{-1}\{(1 - [(u-M/2)^2 + (v-N/2)^2]] F(u,v)\}$$

$$G(u,v) = \boxed{[1 - d^2(u,v)]} F(u,v)$$

Note 1: Difference b/w Laplacian frequency domain filter and spatial domain filter?

Note 2: Difference b/w Laplacian filter and high pass filters.

## Unsharp Masking (High boost filter and High Frequency Emphasis filtering) :—

Generating a sharp image by subtracting from an image a blurred version of itself called Unsharp Masking. From frequency domain terminology,

A high pass filter $f_{hb}$ is obtained using

$$f_{hp}(x,y) = f(x,y) - f_{lp}(x,y)$$

$$\left[\because H_{hp}(u,v) = 1 - H_{lp}(u,v)\right]$$

$$F(u,v).H(u,v) = F(u,v) - F(u,v).H_{lp}(u,v)$$

$$\mathcal{F}^{-1}\left[F(u,v).H_{hp}(u,v)\right] = \mathcal{F}^{-1}\left[F(u,v) - F(u,v).H_{lp}(u,v)\right]$$

$$f_{hp}(x,y) = f(x,y) - f_{lp}(x,y)$$

$$f_{lp}(x,y) = f(x,y) - f_{hp}(x,y)$$

High boost filter

$$f_{hb}(x,y) = Af(x,y) - \bar{f}(x,y)$$

$$f_{hb}(x,y) = A f(x,y) - f_{lp}(x,y)$$

[∵ low pass filter has blurry effect]

$$f_{hb}(x,y) = (A-1) f(x,y) + f(x,y) - f_{lp}(x,y)$$

$$= (A-1) f(x,y) + f_{hp}(x,y)$$

Frequency Domain

$$H_{hb}(u,v).F(u,v) = (A-1) F(u,v) + H_{hp}(u,v).F(u,v)$$

$$H_{hb}(u,v) = (A+1) + H_{hp}(u,v)$$

Note: $A=1$ the $H_{hb} = H_{hp}(u,v)$, $A=0$ $H_{hb} = -H_{lb}(u,v)$

Homomorphic Filtering :—

The illumination-reflectance model ~~intrado~~ used to develop a frequency domain prodedure for improving the appearance of an image by simultaneous gray-level range compression and contrast enhancement.

An image $f(x,y)$ can be expressed as product of illumination and reflectance components

$$f(x,y) = i(x,y)\, r(x,y) \quad \text{—} (1)$$

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

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

$$= \log_e [i(x,y)] + \log_e (r(x,y)) \quad \text{—} (1)$$

$$\mathcal{F}[\log_e (f(x,y))] = \mathcal{F}[\log_e (i(x,y))] + \mathcal{F}[\log_e (r(x,y))]$$

$$H(u,v) \cdot F_{\ln}(u,v) = H(u,v) I_{\ln}(u,v) + H(u,v) R_{\ln}(u,v)$$

$$G_{\ln}(u,v) = I'_{\ln}(u,v) + R'_{\ln}(u,v)$$

~~subtractive~~ used to
... improving the
... gray-level range
... as product of
... s
... ①
... exp → g(x,y)
... $\cdot r(x,y)]$
... $\log_e(r(x,y))]$ — ①
... $(x,y))] + \mathcal{F}[\log_e(r(x,y))]$
... $(u,v)\, R_{ln}(u,v)$
... $_{ln}(u,v)$

$$\mathcal{F}^{-1}[G_{ln}(u,v)] = \mathcal{F}^{-1}[I'_{ln}(u,v)] + \mathcal{F}^{-1}[R'_{ln}(u,v)]$$

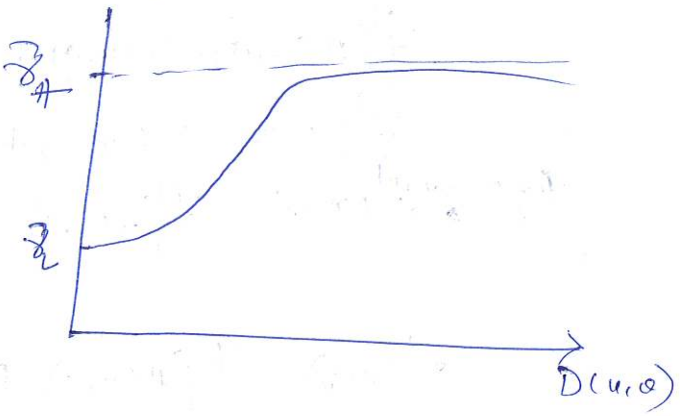
$$g'_{ln}(x,y) = i'_{ln}(x,y) + r'_{ln}(x,y)$$

$$e^{g'_{ln}(x,y)} = e^{i'_{ln}(x,y)} \cdot e^{r'_{ln}(x,y)}$$

$$g(x,y) = \cancel{e^{i'_{ln}(x,y)}}\; i'(x,y) \cdot r'(x,y)$$
$$i_0(x,y) \cdot r_0(x,y)$$

$$\boxed{H(u,v) = (\gamma_H - \gamma_L)\left[1 - e^{-c(D^2(u,v)/D_0^2)}\right] + \gamma_L}$$

$H(u,v)$




**Note:**

Similar to High pass filter

$\gamma_L = 0.5$ $\gamma_H \geq 2.0$

$$F_f(u,v) = F_i(u,v) + F_r(u,v) \quad \text{--- (3)}$$

Multiply transforming function $H(u,v)$

$$H(u,v) \cdot F_f(u,v) = H(u,v)\, F_i(u,v) + H(u,v)\, F_r(u,v)$$

Assume $S(u,v) = H(u,v) \cdot F_f(u,v)$

$$S(u,v) = H(u,v) \cdot F_i(u,v) + H(u,v) \cdot F_r(u,v) \quad \text{--- (3)}$$

Spatial domain function $S(u,v)$ is $s(x,y)$

$$s(x,y) = \mathcal{F}^{-1}[S(u,v)]$$

$$= \mathcal{F}^{-1}[H(u,v) \cdot F_i(u,v) + H(u,v) \cdot F_r(u,v)]$$

$$= \mathcal{F}^{-1}[H(u,v) \cdot F_i(u,v)] + \mathcal{F}^{-1}[H(u,v)\, F_r(u,v)] \quad \text{--- (4)}$$

By Assuming $i'(x,y) = \mathcal{F}^{-1}[H(u,v) \cdot F_i(u,v)]$

$$r'(x,y) = \mathcal{F}^{-1}[H(u,v) \cdot F_r(u,v)]$$

By substituting above equation into (4)

$$s(x,y) = i'(x,y) + r'(x,y)$$

$$e^{s(x,y)} = e^{i'(x,y) + r'(x,y)}$$

$$= i_0(x,y) \cdot r_0(x,y)$$

By Assuming

$$i_0(x,y) = e^{i'(x,y)}$$

$$r_0(x,y) = e^{r'(x,y)}$$

Gaussian High pass filter

$$H(u,v) = (\gamma_H - \gamma_L)\left(1 - e^{-c(D^2(u,v)/D_0^2)}\right) + \gamma_L$$

C → Constant to control the sharpness of the slope the filter function.

$\gamma_L$ \$ $\gamma_H$ → transmission constant

$\gamma_H$ → High pass

$\gamma_L$ – Low pass

# Color Image Processing

→ Color image processing is divided into two major areas;

- full-color processing (A color TV, color scanner sensor)
- Pseudo color processing (A color to a particular monochrome intensity)

## Color Fundamentals:

→ In 1666, Sir Isaac Newton discovered light has a color.

Then, color spectrum divided into six broad regions:

violet, blue, green, yellow, orange and red.

→ The colors that human perceive in an object are determined by the nature of the light reflected from object.

→ Characterization of light is central to the science of color.

Chromatic [ 400nm to 700nm ] → visible spectrum

achromatic [ Rest ]

→ Three basic quantities are used to describe quality of light source:

Radiance (W)

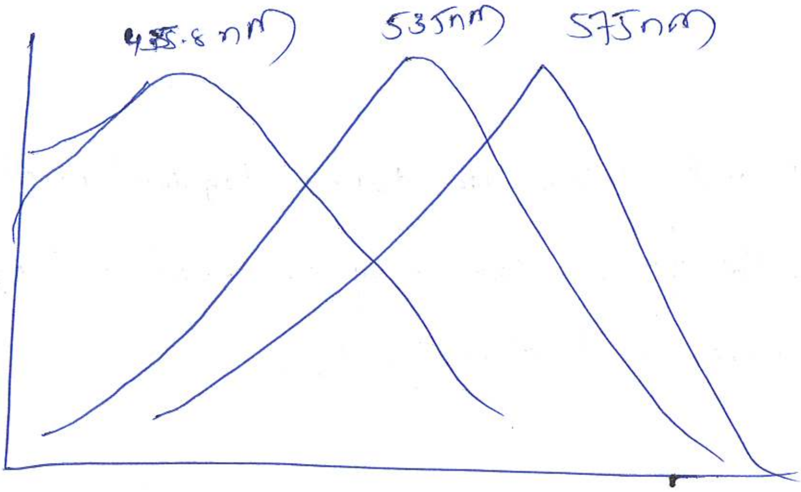
Luminance (lm)

brightness

→ Primary colors Red (R) — 700nm

Green (G) — 546.1nm

Blue (B) — 435.8nm



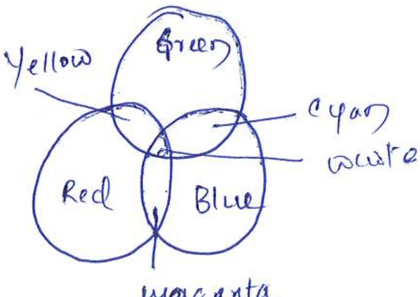

→ secondary colors of light

magenta — (R+B)

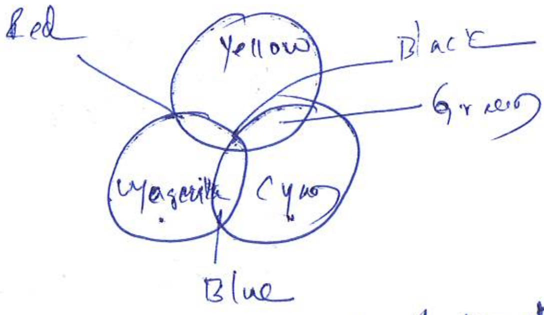
cyan (G+B)

yellow (R+G)

→ Difference b/w primary colors of light and primary colors of pigment.




Primary colors of light




Pigment color of light

→ The characteristic generally used to distinguish one color from another are brightness, $\underset{\downarrow \text{ dominant wavelength}}{\underline{\text{hue}}}$, and $\underset{\downarrow \text{ amount white}}{\underline{\text{Saturation}}}$.

→ Hue and Saturation taken togather are called chromaticity, and, therefore, a color may be characterized by its brightness and chromaticity.

→ The amounts of red, green and blue needed to form any particular color are called the tristimulus values. denoted by $(X, Y, Z)$

→ A color is then specified by its trichromatic coefficients

defined

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z}$$

$$\boxed{x+y+z=1}$$

## Color Models

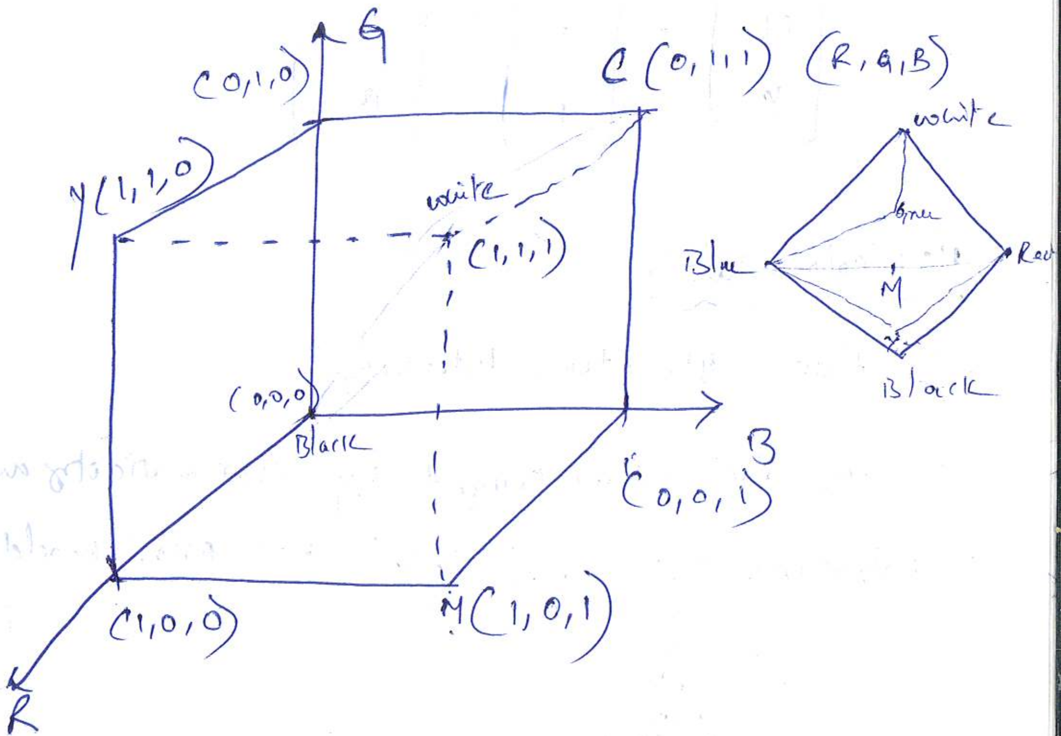
→ The purpose of a color model is to facilitate the specification of colors in some standard form.

Color model
- RGB
- CMY/CMYK
- HSI

## RGB Color Model

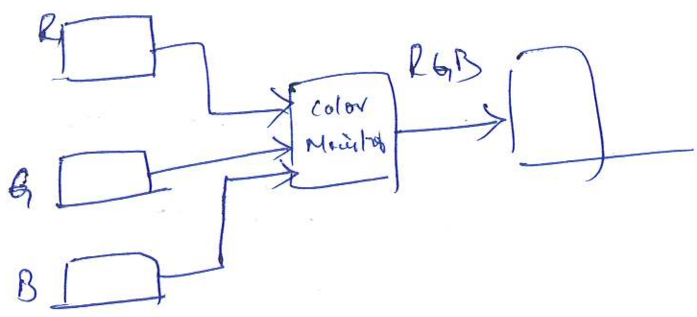
→ In the RGB model, each color appears in its primary spectral components of red, green, and blue.

→ This is based on cartesian coordinate system.




→ $(2^8)^3$ color

→ size of the image $\underline{M \times N \times 3 \times 8}$ bits.




CMY and CMYK Model

→ In the CMY model, each color appears in its pigment processing color such as: Cyan (G+B), Magenta (B+R), Yellow (R+G)

→ This is based on cartesian coordinate system.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

HSI color Model:

→ Hue Saturation Intensity

→ A color characterized by chromaticity and brightness. This model called HSI color model.

## Converting colors from RGB to HSI:

Given an image in RGB color format, the H component of each RGB pixel is obtained using the equation

$$H = \begin{cases} \theta & \text{if } B \le G \\ 360 - \theta & \text{if } B > G \end{cases}$$

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

Saturation component given by

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

Intensity component

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

## Converting colors from HSI to RGB:

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

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

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

1) Define orthogonality:

Transform an 1D signal into a series summation of orthogonal basic vector.

2) Define connected component

S is a connected component of for each pixel pairs $(x_1, y_1) \in S$ and $(x_2, y_2) \in S$ there is a path passing through ~~4~~- neighbours (or) 8- neighbours.

3) Specify the elements of DIP system:

Preprocessing
Image Enhancement
Image Transform
Image classification & Analysis

4) Define weber ratio:

Weber fraction is known in color sciences as contrast sensitivity.

Contrast ratio refers to the ratio of intensities between the most intense (brightness) and less intense (darkest) element of the scene.

5) What is mean by mach band effect?

It exaggerates the contrast between edges of the slightly differing shades of gray, as soon as they contact one another, by triggering edge-detection in the human visual system.

6. Describes the elements of visual perception.

Cornea
Iris
Lens
Sclera
choroid
Fovea
Retina
Optical nerve

7. Consider the two image subsets $S_1$ and $S_2$ shown below

| | $S_1$ | | | | $S_2$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |

For $V = \{1\}$, determine whether $S_1$ and $S_2$ are (a) 4-connected

(b) 8-connected

(c) m-connected.

(a) 4-connected

$\{P_4 = P_6$

$\{P_6 = P_3$

| | $S_1$ | | | | $S_2$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | $P_1$ | $P_1$ | $\cancel{7}$0 |
| $P_2$ | 0 | 0 | $P_3$ | 0 | 0 | $P_4$ | 0 | 0 | $P_5$ |
| $P_2$ | 0 | 0 | $P_3$ | 0 | $P_6$ | $P_4$ | 0 | 0 | 0 |
| 0 | 0 | $\cancel{P_4}$ | $P_3$ | $P_3$ | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | $\cancel{P_6}$ | $P_3$ | $P_3$ | 0 | 0 | $\cancel{P_8}$ | $\cancel{P_8}$ | $\cancel{P_8}$ |

ception.

hown below

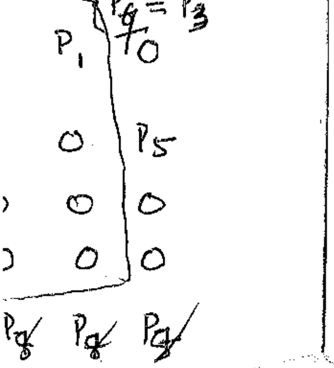
are (a) 4-Connected

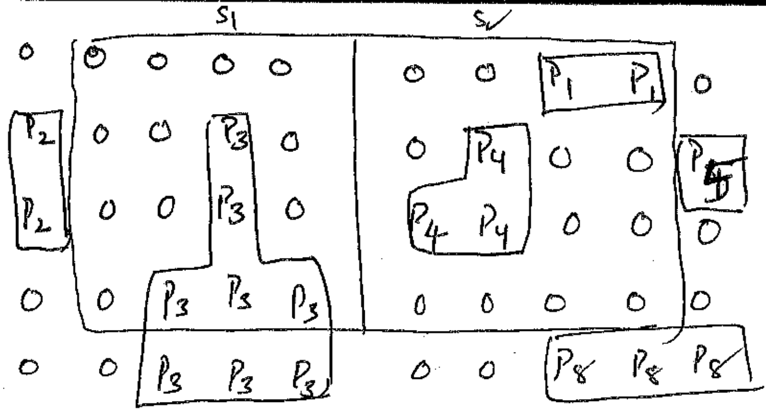
(b) 8-Connected

(c) m-connected.

$\{P_4 = P_6$

$\{P_9 = P_3$





$P_1 = C_1$

$P_2 : C_2$

$P_3 : C_3$

$P_4 : C_4$

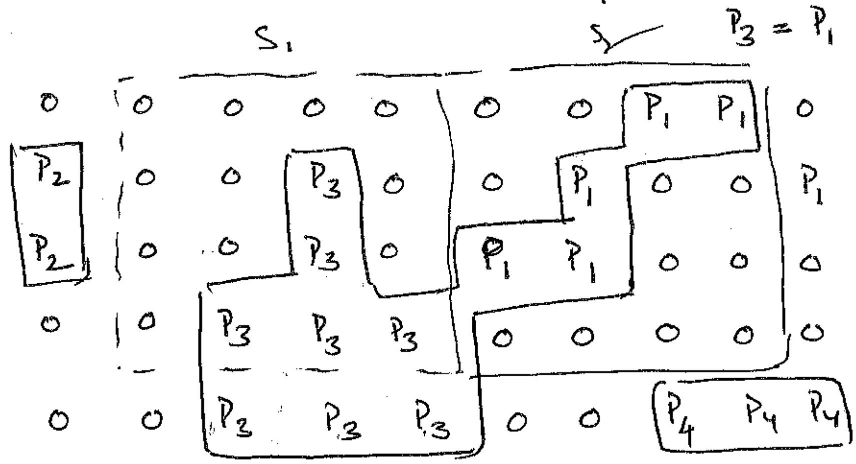
$P_5 : C_5$

$P_8 : C_6$

(b) 8 - Connected

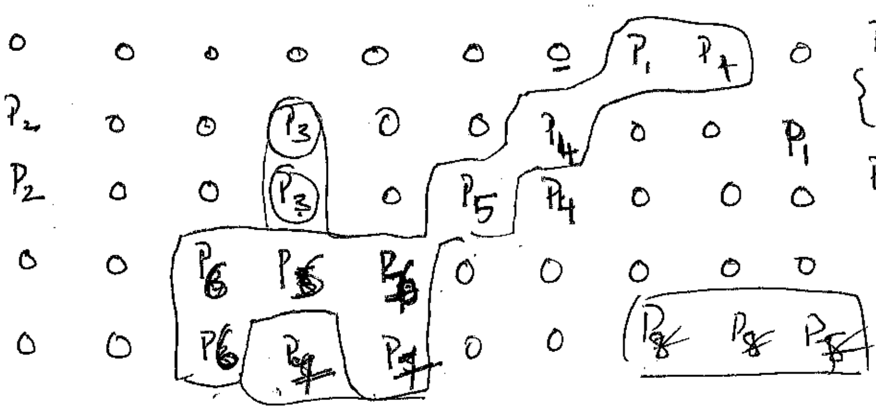
$N_4(P) \cup N_D(P)$

$P_3 = P_1$



(c) M - connected

$\underline{N_4(P) \cap N_4(q) = \phi}$



$P_1 = P_3 = P_5$

$\{P_4 = P_5\}$

$P_3 = P_6$

$P_7 = P_6$

8) Consider an image segment shown

(a) Let $V = \{0,1\}$ and compute the length of the shortest 4-, 8- and m- path between p and q. If a particular path does not exist between these two points, explain why?

(b) Repeat for $V = \{1,2\}$

| | | | | |
|---|---|---|---|---|
| | 3 | 1 | 2 | 1 (q) |
| | 2 | 2 | 0 | 2 |
| | 1 | 2 | 1 | 1 |
| (p) | 1 | 0 | 1 | 2 |

Sol.

$V = \{0,1\}$

(a) 4 - connected path

$P_4 = P_3$

| | | | | | |
|---|---|---|---|---|---|
| | 3 | $P_1$ | 2 | $P_2$ | q |
| | 2 | 2 | $P_3$ | 2 | |
| | $P_{4/3}$ | 2 | $P_3$ | $P_3$ | |
| p | $P_{4/3}$ | $P_{4/3}$ | $P_3$ | 2 | |

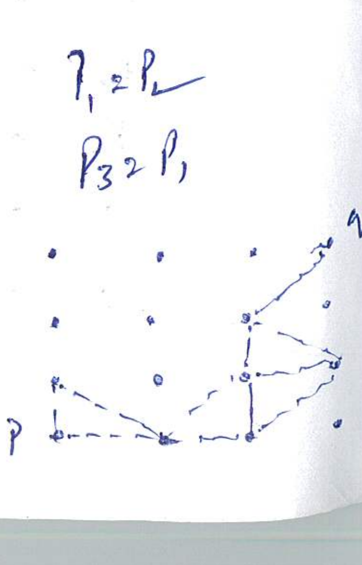
points p, q are not connected. So, it don't have a path.

8 - connected path

$P_1 = P_2$

$P_3 = P_1$

| | | | | | |
|---|---|---|---|---|---|
| | 3 | $P_1$ | 2 | $P_{2/1}$ | q |
| | 2 | 2 | $P_1$ | 2 | |
| | $P_{3/1}$ | 2 | $P_1$ | $P_1$ | |
| p | $P_{3/1}$ | $P_{3/1}$ | $P_1$ | 2 | |



— There exist multiple paths. shortest 8-path ~~is 4~~ length is 4.

M- connected path

| | | | |
|---|---|---|---|
| 3 | $P_1$ | 2 | $P_{2/1}$ |
| 2 | 2 | $P_1$ | 2 |
| $P_{3/1}$ | 2 | $P_1$ | $P_1$ |
| $P_{3/1}$ | $P_{3/1}$ | $P_{3/1}$ | 2 |

$P_1 = P_2$

$P_3 = P_1$

M- Connected

— There exist Multiple paths. shortest m- path length is 5.

(b) $V = \{1, 2\}$

$P_2 = P_1$

4 - path

| | | | | |
|---|---|---|---|---|
| | 3 | $P_1$ | $P_1$ | $P_1$ q |
| | $P_{2/1}$ | $P_1$ | $\boxed{0}$ | $P_1$ |
| | $P_{2/1}$ | $P_1$ | $P_1$ | $P_1$ |
| p | $P_{2/1}$ | 0 | $P_1$ | $P_1$ |

— There exist multiple paths. shortest 4-path length is 6.

8 - path

| | | | |
|---|---|---|---|
| 3 | $P_1$ | $P_2$ | $P_1$ |
| $P_1$ | $P_1$ | $\boxed{0}$ | $P_1$ |
| $P_1$ | $P_1$ | $P_1$ | $P_1$ |
| $P_1$ | 0 | $P_1$ | $P_1$ |

— There exist multiple paths. shortest 8-path length is 4.

M-path

| | | | |
|---|---|---|---|
| 3 | $P_1$ | $P_1$ | $P_1$ |
| $P_1$ | $P_1$ | $\boxed{0}$ | $P_{2}/1$ |
| $P_1$ | $P_{3}/1$ | $P_{3}/1$ | $P_{3}/1$ |
| $P_1$ | 0 | $P_{4}/1$ | $P_{4}/1$ |

$P_1 = P_2 = P_3$

$P_3 = P_4$

— There exist multiple paths. shortest M-path length is 6.

9) A common measure of transmission for digital data is the baud rate, defined as the number of bits transmitted per second. Generally, transmission is accomplished in packets consists of a start bit, a byte of information, and a stop bit. Using these facts, answer the following:

(a) How many minutes would it takes to transmit a 1024×1024 image with 256 gray levels using a 56 Baud modem?

(b) What would the time be at 750 baud, a representative speed of a phone DSL connection?

Ans

(a)

$$T = (1024 \times 1024) \times (8+2) / 56 \times 10^3$$

Time = Data / speed

$$T = 187.24 \text{ bit/sec}$$

(b)

$$T = \frac{(1024 \times 1024) \times (8+2)}{750 \times 10^3}$$

$$T = 13.98 \text{ bits per Sec}$$

imageDatastore

Splot

2016a

Bagoffeatures

Approaches for Deep learning

1). Train a Deep NeuralNetwork from scratch

(CNN)

* 2). Fine tune a pretrained model (transfer learning)

cration

nsfer learning)

# Fundamental of Image Processing

## Image processing Applications

1) Computerized photography (e.g photoshop, Maya Animation s/w)

2) Space Image Processing (e.g telescope image, interplanetary probe images)

3) Medical/Biological Image Processing (e.g X-ray, CT scan, MRI scan, Angeogram)

4) Automatic character recognization (e.g license plate recognization)

5) Biometric (e.g finger print/face/iris)

6) Remote Sensing: (arial and satellite images interpretation)

7) Industrial Application (e.g Product Inspection/ sorting)

## What is Digital Image Processing?

An Image may be defined as a two-dimensional function, $f(x,y)$, where $x$ and $y$ are spatial (plane) Coordinates, and the amplitude of $f$ at any pair of coordinates $(x,y)$ is called the intensity (or) gray level of the image at that point. When $x, y$ and amplitude values of $f$ are all finite, discrete quantities, we call the image a digital image.

$$f(x,y) = \underbrace{i(x,y)}_{\text{illumination}} \cdot \underbrace{r(x,y)}_{\text{reflectance}}$$

## History

— News paper Industry

- 1920 London to Newyork pictures are transmitted across atlantic. It reduced time taking to transmit from more weeks to 3 hours.
- special printing equipment coded from transmission and reconstruction.

Initial Problems

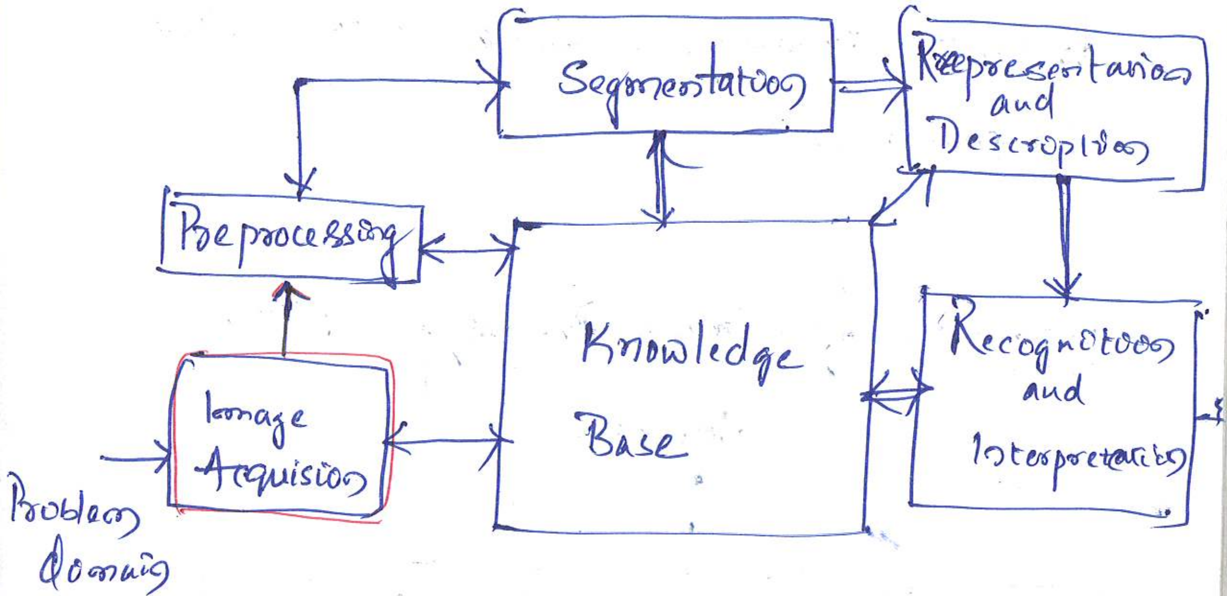
i) Visual Quality (selection of printing procedure and intensity distribution)

5 levels

1921 — 5 gray levels

— 15 gray levels 1929

## Fundamental Steps in digital Image Processing:




- Image Acquision: An image sensor and capability to digitize the signal produced by the sensors.

- Preprocessing: Enhances the image quality, filtering, Contrast enhancement etc.

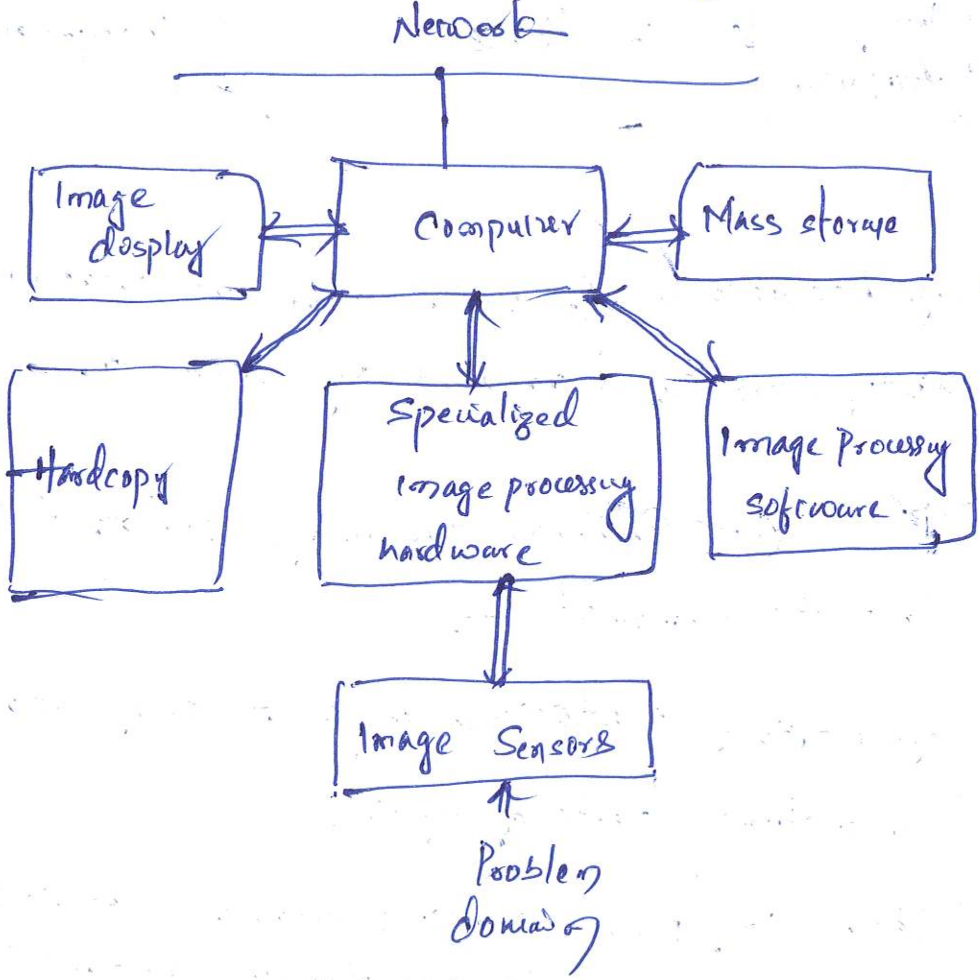
- Segmentation: Partitions an input image into constituent parts of objects.

- Description/Feature selection: extracts description of image objects suitable for further computer processing

- Recognition & Interpretation: Assigning a label to the object information provided by its descriptor. Interpretation assigns meaning to a set of labeled objects.

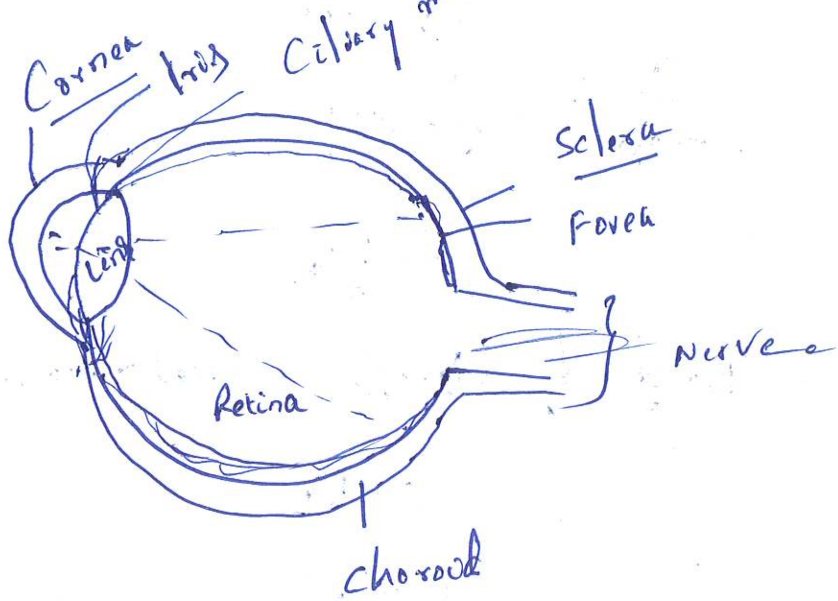
- Knowledge base: Knowledge base helps for efficient processing as well as inter module cooperation.

Components of an Image Processing System:



## Structure of the Human Eye:

3-2-2




## Image formation in Eye


$\frac{15}{100} = \frac{h}{17} \Rightarrow$

Large variation by changes in its overall sensitivity a phenomenon known as brightness adaptation.

$I + \Delta I$.

Weber ratio: $\Delta I_c / I$

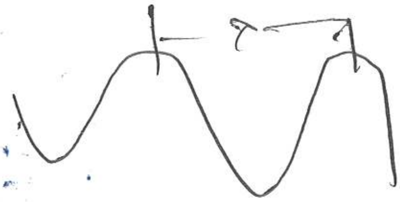
I ~~Light~~ Background light source intensity.

$\Delta I_c$ is ~~increment~~ of ~~illumination~~ d short duration flash

incremental illumination by background illumination, called Weber ratio.

Light and Electromagnetic Spectrum:—

$$\underset{\text{wavelength}}{\lambda} = \frac{c}{\nu} \quad \begin{matrix} \text{— light speed} \\ \text{— frequency} \end{matrix}$$




$$\boxed{E = h\nu}$$

Bundle of energy is called photon.

$$E = h\nu$$

$$E \propto \nu$$

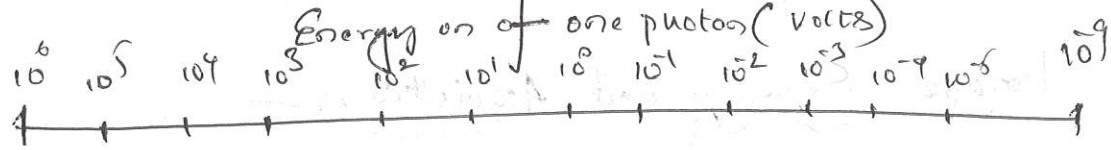
frequency and wavelength

$$\lambda = \frac{c}{\nu}$$

$$\lambda \propto \frac{1}{\nu}$$

shorter wavelength more Energy and penetration.

Note:
Shorter imaging more clarity and long imaging less clarity.




~~Chromatic~~

monochromatic → Light that void color.
Light intensity changes from black to white is usually called the grayscale. monochromatic images are called as gray-scale images.

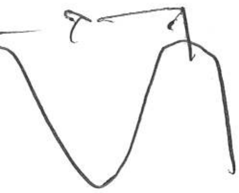
chromatic – light spans the electromagnetic energy spectrum from approximately 0.43 μm to 0.79 μm.

chromatic light describe the following quantities.

Radiance (Total amount of energy flows from light source) W

Luminance (Amount of energy an observer perceives from a light source)

Brightness



rgy adaptation.

e long image

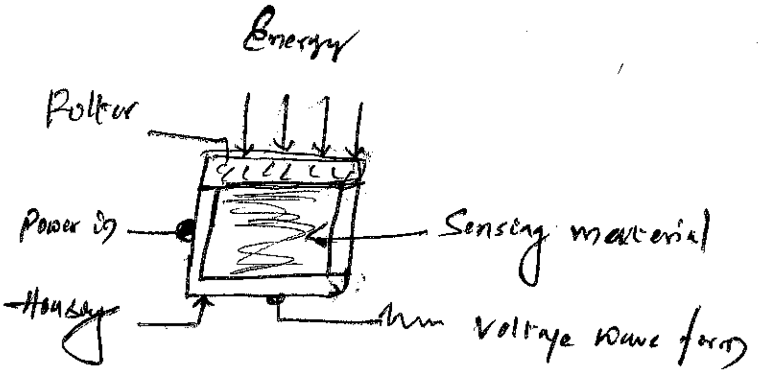
# Image Sensing and Acquisition: —

Image Sensing and Acquisition are used for processing the analog image of physical scenes (or) the interior structure of an object, and convert into digital.

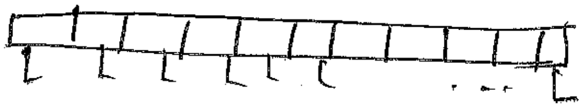
⇒ Image Sensing refers to Sensing an analog image ~~and giving it as input to~~

⇒ Image aquisition include processing, compression, and finally storing of image into digital form.

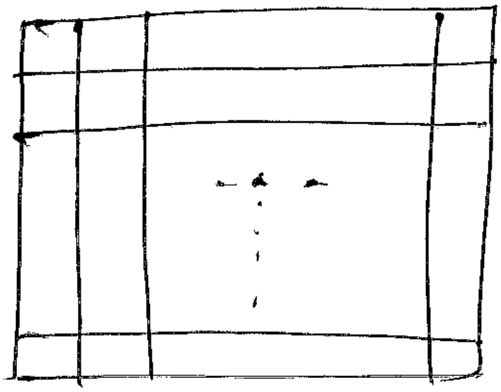
1) Image acquisition using single sensor
2) Image acquisition using sensor strip
3) Image acquisition using sensor array




Sensor



Linear Sensor



Array Sensor

Image Acquisition using Single sensor

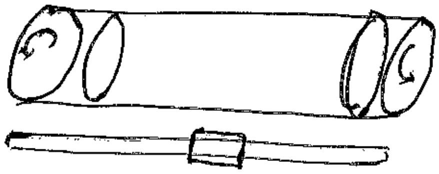
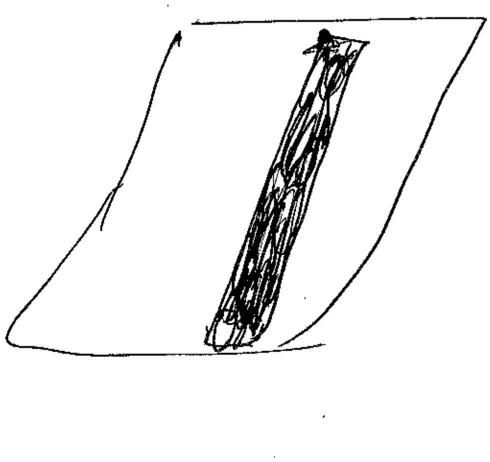
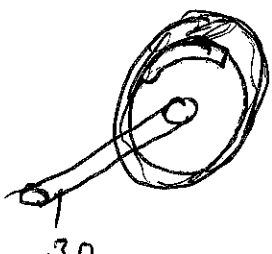


Image Acquisition using sensor strip:

- → Linear sensor strip
- → Circular sensor strip





3D

## Image Acquisition using Sensor Arrays: ③

→ Individual sensors arranged in the form of a 2D array.

→ Numerous electromagnetic and some ultrasonic sensing devices are arranged in an array format.

→ { CCD (Charge Couple Device).
CMOS (complementary metal oxide Semiconductor) }

Sample Image Formation

→ Image

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

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

$$0 < r(x,y) < 1$$

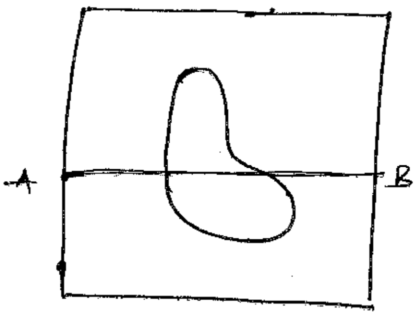
→ Intensity of monochrome image at any coordinate $(x_0, y_0)$ gray level (l)

## Image Sampling and Quantization :—

To Generate digital image from sensed data.

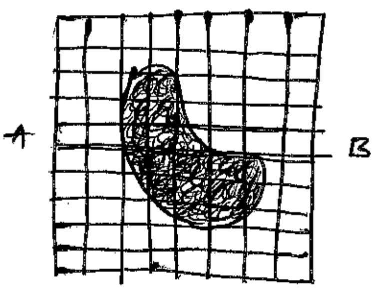
### Basic concept

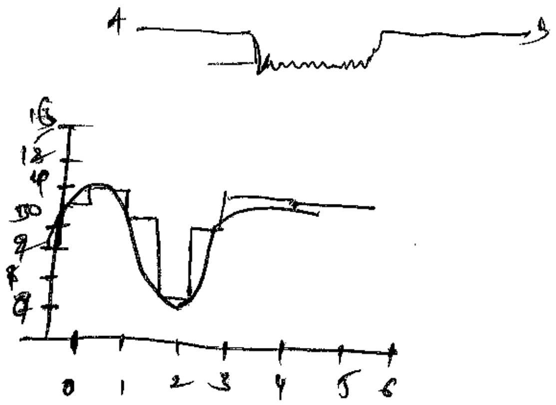
→ Continuous image 'f' that we want to convert to digital form.




→ Sampling: Digitizing the coordinate values is called sampling.

→ Quantization: Digitizing the amplitude values is called quantization.



## Representing Digital image :—

Let $f(s,t)$ represent a continuous image function of two continuous variables. Convert continuous image into digital by sampling and Quantization.

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

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

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

The digitization process requires that decisions be made regarding the values for M, N, and for the number, L of discrete intensity levels.

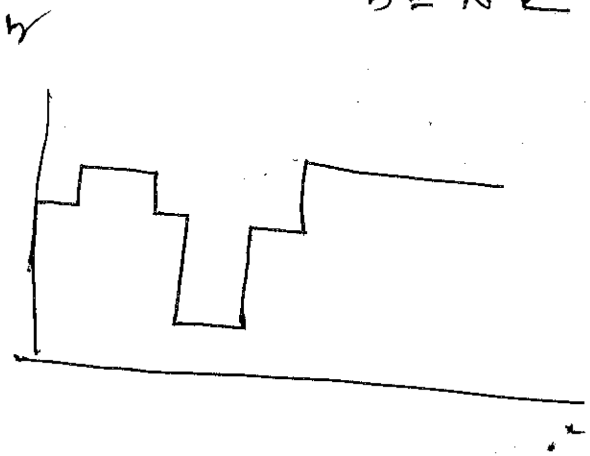
$$\boxed{L = 2^k}$$

$$b = M \times N \times k \qquad N = M$$

Saturation → Upper limit detected intensity

Noise — lower limit detected intensity.

$$b = N^2 k$$



Sensed data.

- to convert

values is called

e values is

4 5 6

Number of storage bits for various values of N and k

| N \ k | 1 (L=2) | | 8 (L=256) |
| --- | --- | --- | --- |
| | | | |

## Spatial and Gray-level Resolution

→ Sampling is the principle factor determining the spatial resolution of an image.

→ Smallest number of ~~discernible~~ Perceptible line pairs per unit distance. Called spatial resolution (or) dpi (or) lines per inch

→ Gray level resolution similarly refers to the smallest ~~discernible~~ Perceptible change in gray level.

## Aliasing and Moire Patterns:

→ Shannon Sampling theorem

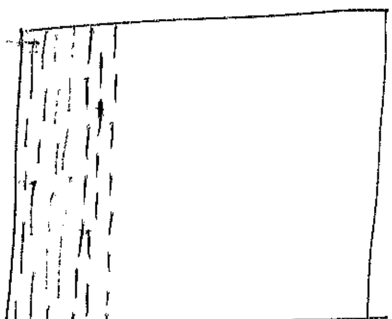
[ If the function is sampled at a rate equal to (or) greater than twice its frequency, it is possible to recover completely the original function from its sample. ]

⇒ If the function is undersampled, then a phenomenon called alliasing.

→ Alliasing corrupts the sampled image.

⇒ The corruption is in the form of additional frequency components being introduced into the sampled function. These frequencies called aliasing frequencies.

⇒ The aliasing frequencies can be seen under some special conditions as Moire patterns.



## Some Basic Relationships Between pixels:—

— A image is denoted by $f(x, y)$.

### Neighbors of a pixel:—

— A pixel $p$ at coordinates $(x, y)$ has four horizontal and vertical neighbors.

| | | |
|---|---|---|
| | $(x-1, y)$ | |
| $(x, y-1)$ | $p\,(x, y)$ | $(x, y+1)$ |
| | $(x+1, y)$ | |

— This set of pixel called 4-neighbors of $P$. denoted as $N_4(P)$. Each pixel has unit distance from $(x, y)$.

— The 4 diagonal neighbors of $P$ have coordinates

| | | |
|---|---|---|
| $(x-1, y-1)$ | | $(x-1, y+1)$ |
| | $p\,(x, y)$ | |
| $(x+1, y-1)$ | | $x+1, y+1)$ |

are denoted by $N_D(P)$. These four diagonal neighbors and 4-Neighbors called 8-Neighbors. denoted as $N_8(P)$.

## Adjacency, Connectivity, Regions and Boundaries:-

Let V be the set of grey-level values used to define adjacency. In binary image, $V=\{1\}$ if we refer adjacency of pixels with value 1.

In grey-scale image, the idea is same, but set V typically contain more elements.

(a) 4-adjacency:- Two pixels p and q with values from V are 4-adjacent, if q is in the set $N_4(p)$.

(b) 8-adjacency:- Two pixels p and q with values from V are 8-adjacent, if q is in the set $N_8(p)$.

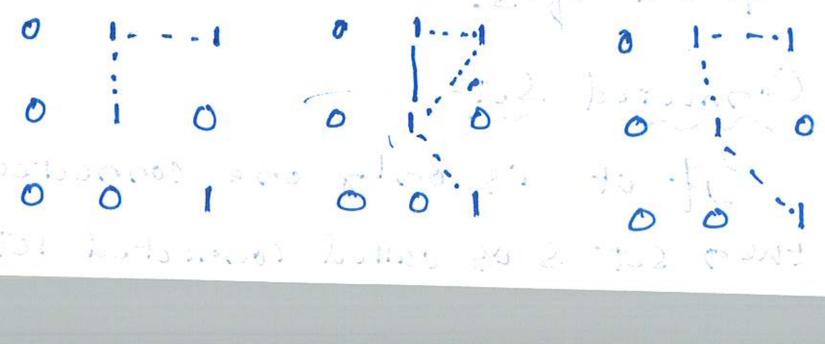
(c) m-adjacency (mixed adjacency):- Two pixels p and q with values from V are m-adjacent if

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

(ii) q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$, has no pixels whose values are from V.

Note:

Mixed adjacency is modified version of 8-adjacency. It is used to remove ambiguities that often arise when 8-adjacency is used.



## Path:

A path from pixel p with coordinates $(x, y)$ to pixel q with coordinates $(s, t)$ is a sequence of distinct pixels with coordinates.

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

where $(x, y) = (x_0, y_0)$ and $(s, t) = (x_n, y_n)$ and pixels $(x_i, y_i)$ and $(x_{i-1}, y_{i-1})$ are adjacent for $1 \leq i \leq n$.

This case length of path is n.

Note: Let S is a subset of pixels in an image.

## Connected

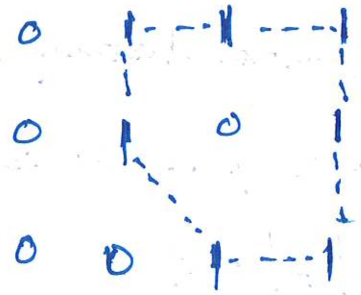
Two pixels p and q are said to be connected in S, if there exists a path between them ~~……~~ with all pixels in S.

## Connected Component

For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S.

## Connected Set

If it is only one connected component, then set S is called connected set.



## Region

Let $R$ be a subset of pixels in image. We call $R$ a region of the image, if $R$ is a connected set.

## Boundary

— The boundary or border of a region $R$ is the set of pixels in the region that have one (or) more neighbors that are are not in $R$.

## Distance Measures:

For pixels $p$, $q$ and $z$ with coordinates $(x, y)$, $(s, t)$ and $(v, w)$, respectively, $D$ is a distance function (or) metric of

(a) $D(p, q) \geq 0$ $(D(p, q) = 0$ iff $p = q)$

(b) $D(p, q) = D(q, p)$

(c) $D(p, z) \leq D(p, q) + D(q, z)$

(i) Euclidean distance between $\underset{(x,y)}{p}$ and $\underset{(s,t)}{q}$ is defined as

$$D_e(p, q) = \left[(x-s)^2 + (y-t)^2\right]$$

(ii) $D_4$ distance (city-block distance):

- Distance between $p$ and $q$ is defined as

$$D_4(p,q) = |x-s| + |y-t|$$

Note: ~~$D_4$~~ The pixels with $D_4 = 1$ are the 4-neighbors of $(x,y)$.

(iii) $D_8$ distance: (chessboard distance): —

- Distance between $p$ and $q$ is defined as

$$D_8(p,q) = \max(|x-s|, |y-t|).$$

Note

- The pixels with $D_8 = 1$ are the 8-neighbors of $(x,y)$.

Note:

$D_4$ and $D_8$ distances between $p$ and $q$ are independent of any paths that might exist between the points.

## Image Operations on a pixel Basis:—

→ Image operations are carried out between corresponding pixels in the two images.

→ Arithmetic and logical operations can take place on images.

## Linear and Non-Linear operations:—

Let $H$ be an operator whose input and output are images. $H$ is said to be linear operator, if for any two images $f$ and $g$ and any two scalars $a$ and $b$.

$$H(af + bg) = aH(f) + bH(g)$$

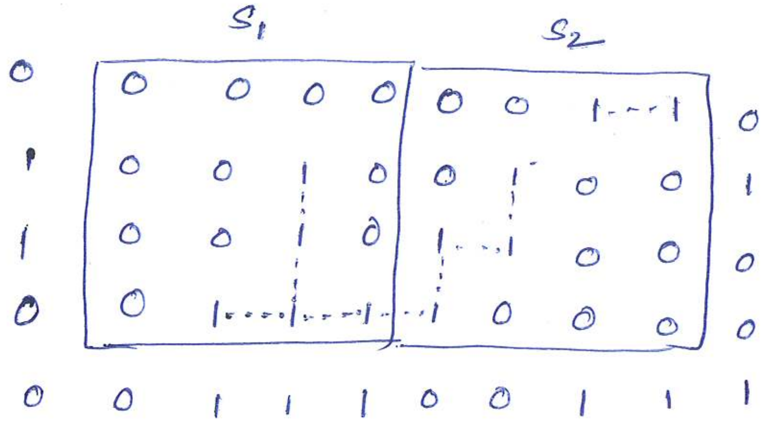
Otherwise, operator called as non-linear.

## Exercise Problems:

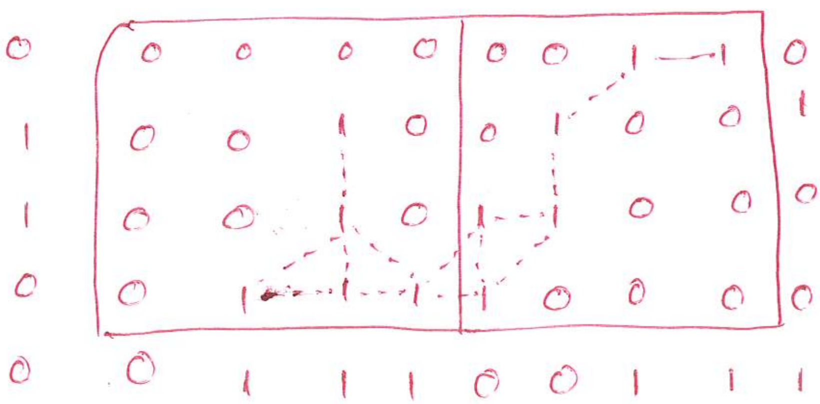
1) Consider the two images subsets, $S_1$ and $S_2$, shown in the following figure. For $V = \{1\}$, determine whether these two subsets are (a) 4-adjacent (b) 8-adjacent (c) m-adjacent

| | $S_1$ | | | | $S_2$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |

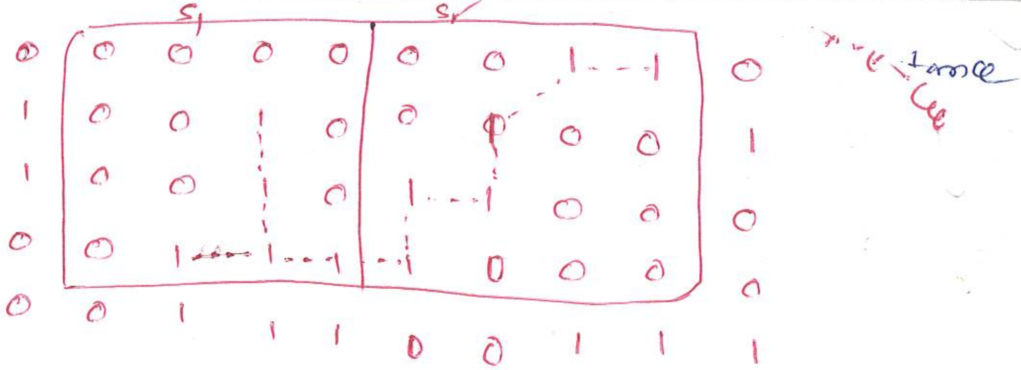
(a)




(b)



(c)



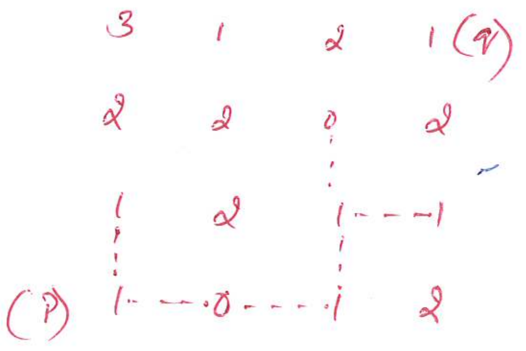
2) Consider the image segment shown

(a) Let $V=\{0,1\}$ and compute the lengths of the shortest 4- 8- and m-path between p and q. If a particular path does not exist between these two points, explain why.

(b) Repeat for $V=\{1,2\}$

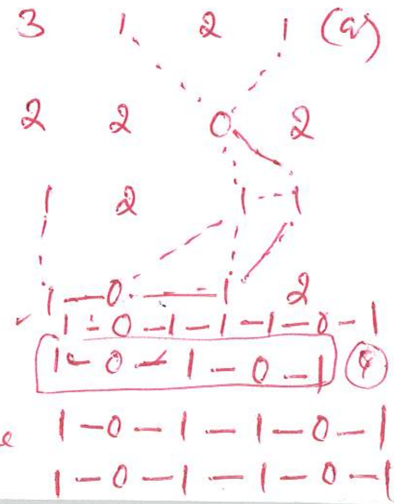
| | | | | |
|---|---|---|---|---|
| | 3 | 1 | 2 | 1 (q) |
| | 2 | 2 | 0 | 2 |
| | 1 | 2 | 1 | 1 |
| (p) | 1 | 0 | 1 | 2 |

(a) 4-path

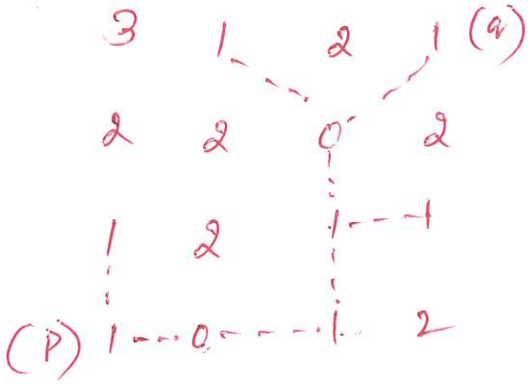


No path. Because q does not have pixel in the set V.

8-path

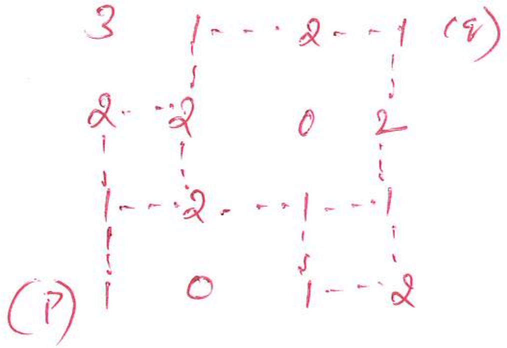


8-path



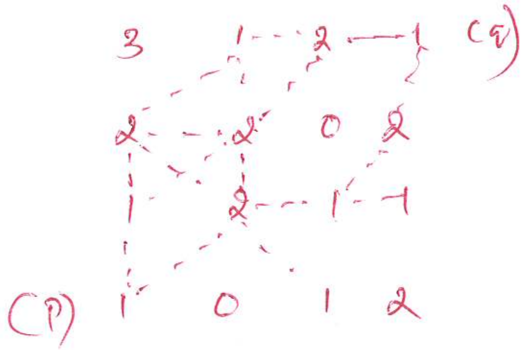
5

(b) 4-path



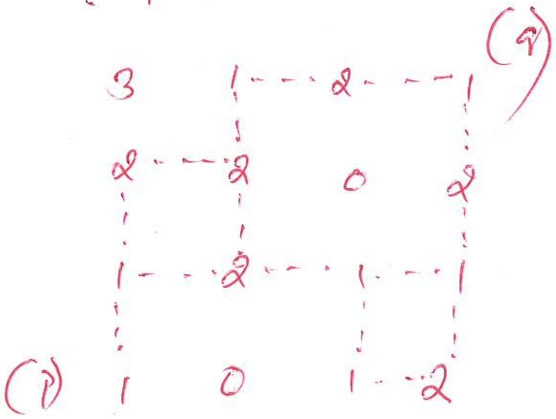
(6)

8-path



(4)

m-path



3) (a) Give the condotions under which the $D_4$ distance between two points p and q is equal to the shortest 4-path between these points.

(b) Is this path unique?

Sol

$p - (x, y)$
$q - (s, t)$

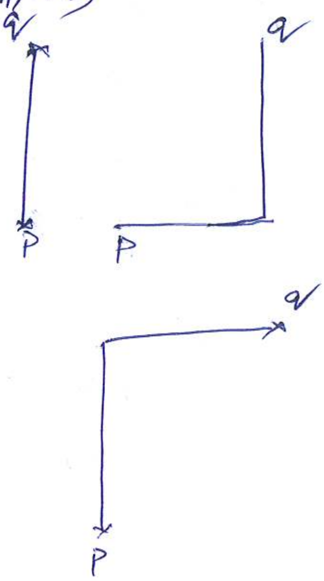
(a) $D_4$ distance $= |x-s| + |y-t|$

4-path $- (x_0, y_0), (x_1, y_1) \dots (x_n, y_n)$

$(x, y) = (x_0, y_0)$ $(s, t) = (x_n, y_n)$

distance $= n$

The following condotions



$D_4$ and 4-path shortest distance are same.

(b) No, it is depend on V.

(4) Repeat with $D_8$ distance.

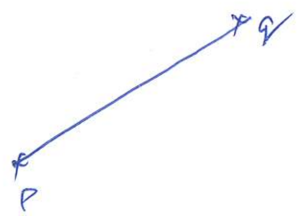
(a) Give the condotions under which the $D_8$ distance b/w two points p and q is equal to the shortest 8-path between these points.

(b) Is this path is unique? Yes.

Sol

$p - (x, y)$

$q - (s, t)$



$D_8$ distance $- \max(|x-s|, |y-t|)$

$n = \max(|x-s|, |y-t|)$, 8-path $(x_0, y_0), (x_1, y_1), \dots (x_n, y_n)$

$(x, y) = (x_0, y_0)$ $(s, t) = (x_n, y_n)$

If p and q are in same diagonal then ~~both distance are~~ distance $= n$

## Mathematical tools used in Digital Image Processing:

Image Array vs Matrix operation:—

Let consider 2x2 Image

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

Image Array product—

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

Matrix product given by

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

## Arithmetic Operations:—

Four arithmetic operations are denoted by

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

$$s(x,y) = f(x,y) - g(x,y)$$

$$s(x,y) = f(x,y) * g(x,y)$$

$$s(x,y) = f(x,y) \div g(x,y)$$

## Application of Image Geometry: (cont)

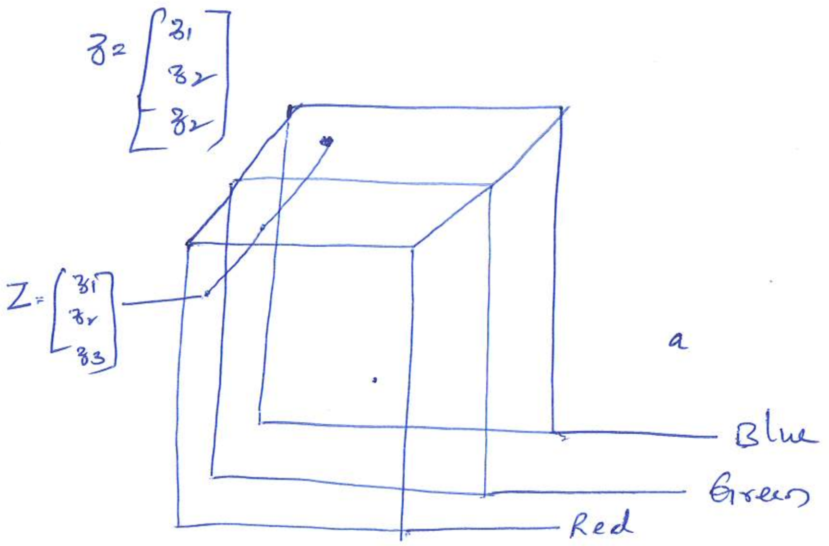
### Vector and Matrix Operations:—

Multispectral image processing is a typical area in which vector and matrix operations are used commonly.

**Ex**

RGB image

RGB image has three components, which organized in the form of a column vector.




$$D(z, a) = \left[(z_1 - a_1)^2 + (z_2 - a_2)^2 + (z_3 - a_3)^2\right]^{1/2}$$

# Image Transforms:-

$$f(x,y) \longrightarrow \boxed{\text{Transform}} \xrightarrow{T(u,v)} \boxed{\text{Operation } R} \xrightarrow{R[T(u,v)]} \boxed{\text{Inverse Transform}} \longrightarrow g(x,y)$$

spatial domain — spatial domain

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

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