

Digital Image Processing (C-7)

Aim & Objective of DIP

To study the fundamental concepts in DIP and discuss the operation of DIP and its components.

Objectives -

1. Learn digital image fundamentals.
2. Expose simple image processing techniques.
3. Familiar with image compression and Segmentation techniques.
4. Represent image in form of features.

At the end of the course, the student should be able to

1. Discuss digital image fundamentals.
2. Apply image enhancement and restoration techniques.
3. Use image compression and segmentation techniques.
4. Represent features of images.

Applications:- Area monitoring, health care monitoring, Environmental/Earth sensing, Air pollution monitoring,

Forest fire detection, Landslide detection, Water quality monitoring and Natural disaster monitoring and prevention, Signal transmission and detection etc.

Industry Connectivity:-
CISCO, IBM, SIEMENS, NORTEL

Unit 1: - Introduction to Image Processing

Digital Image processing refers to preprocessing images by means of a digital computer. Digital image is composed of a finite number of elements, each of which :

- particular location
- value.

- Pixel

In image processing input and outputs are images, but in computer vision, input is scene and after perception its output is image.

In image processing three type of process are

- (1) Image pre-processing (reduce noise, I/p - Image contrast enhancement, image sharpening) O/p - Image
- (2) Segmentation (partitioning an image into regions or objects)

I/p - Image but
O/p → features extracted from images.

- (3) Recognized object

→ Widely used application of image processing is
Adobe Photoshop.

What is an Image?

(3)

It is a two dimensional signals defined by $f(x, y)$, where x & y are coordinate horizontally and vertically.

255	0	32
0	255	77
80	80	255

→ pixel value at co-ordinate (intensity)
(0-255)

Relationship b/w Digital Image & Signal

Signal is a mathematical func, convey some information, like voice.

→ Capturing of an image - using sunlight, object camera to create digital image.

Computer Vision: Input is an image and output is information.

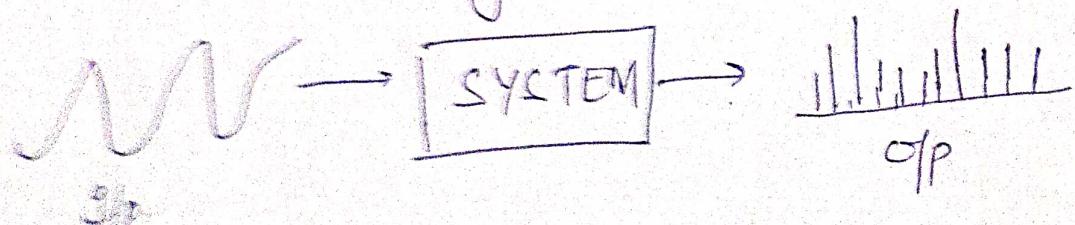
e.g. Biometric: input is face/ finger print and op is attendance or your data

Computer Graphics: Input is object model and op is image.

e.g. Object rendering: Generate image using object

System :-

It convert the analog signal into digital form

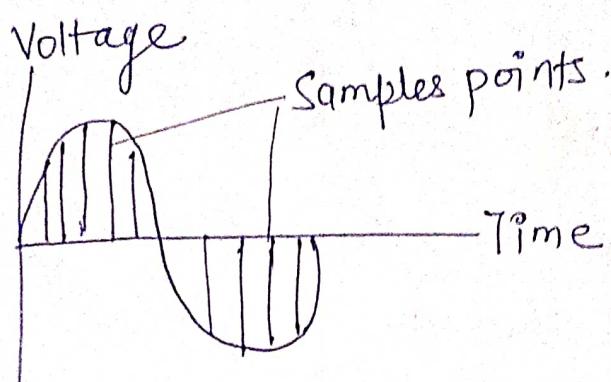


Conversion of Analog signals to digital signal

The two main concept are

- (1) Sampling — digitization of x-ordinate value.
- (2) Quantization — digitization of amplitude value

Sampling :-



Sampling — take sample, over x-axis. or conversion of x-axis to digital is done under sampling

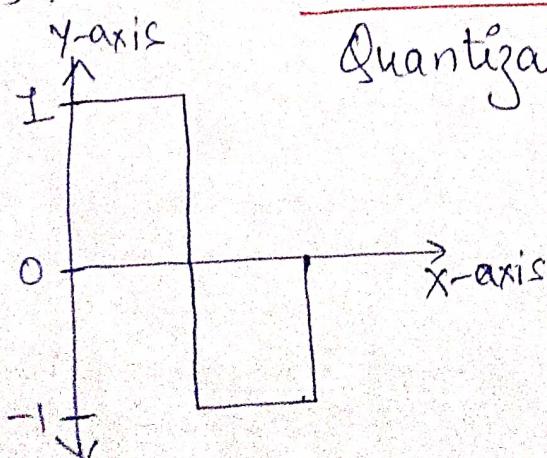
If range of value on x-axis are less then increase the sample of values. \rightarrow up sampling and its viceversa is also known as down sampling.

Quantization :-

divide into quanta (partitions)

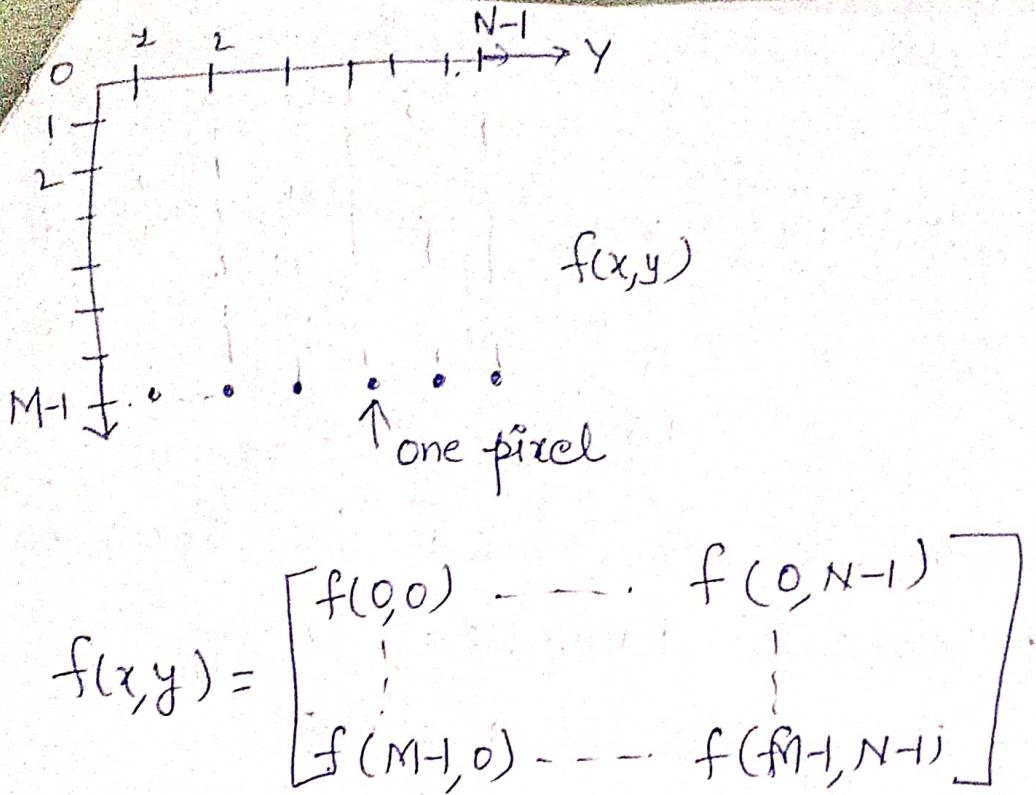
$$\text{eq: } y = \sin(x)$$

Quantization is done on y-variable. It is done on y-axis



Quantization (-1, 0, +1)

(6)



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

If Sampling freq. is increased \rightarrow off image have pleasing effect.

Q1:- What is meant by image sensing? Explain the construction and operation of various image acquisition devices.

Q2:- Write the difference b/w analog and digital signal (in term of analysis, representation, accuracy, storage, subject to noise, etc.)

Examples of fields that we studied till now

Electromagnetic spectrum

(nuclear medicine)

(Wavelength)	Gamma ray	X-ray	UV	Infrared	Microwave	Radio
freq:-	10^{12}	10^{10}	10^8	10^6	10^4	10^2
	10^{20}	10^{18}	10^{16}	10^{14}	10^{12}	10^4

Electromagnetic waves can be thought of as stream of particles, each particle is moving with the speed of light. Each particle carries a bundle of energy. This bundle of energy is called a photon.

Q. Why do we need to analyse all that other stuff in EM spectrum?

Applications of DIP :-

- Image sharpening and restoration - process image
- Medical field
- Remote sensing
- Transmission and encoding vice versa, Image retrieval and image recognition.
- Machine/Robot Vision
- Color processing
- Pattern recognition
- Video processing
- Microscopic Imaging etc.

Medical field :-

- 1) Gamma ray imaging
- 2) PET scan
- 3) X-ray Imaging
- 4) Medical CT
- 5) UV imaging

Fundamental steps in Digital Image Processing

- (1) Image acquisition
- (2) Image Enhancement
- (3) Image restoration
- (4) Image Compression
- (5) Image Segmentation
- (6) Image representation & description

1. Image Acquisition:-

It is the first step of image processing. It involve capturing of image using cameras and apply preprocessing and scaling on image.

2. Image Enhancement:-

It is simple and most appealing area of DP. Using this step interesting features of an image are highlighted. highlights

e.g. Improve the contrast of image because image looks better.

3. Image Restoration:-

It improve the appearance of an image. It is used to apply mathematical & probabilistic models of image degradation

Enhancement - is subjective and

Restoration - is objective

like color model or wavelets of image

4. Image Compression:-

- Used to reduce the size of Image (for storage or bandwidth (for transmit))

5. Image Segmentation:-

Segmentation method is used to divide the image into subparts to get the Region of Interest (ROI).

6. Image Representation and description:-

They use segmented image, like edges of the objects differentiating all objects or boundary of the region (separating one image region from other regions)

According to the image representation features are extracted from the image. This feature extraction is called description of the image.

Recognition is applied on the extracted features and identify the object based on its description.

—X—

Components of an Image Processing System

① Image Sensors

To capture a image is called image sensors. It require two elements

(1) Physical device - It's sensitive to the energy radiated by the object

(2) Digitizer - It convert the output of physical sensing device into digital form.

(2) Specialized image processing hardware

This hardware consist of digitizer plus ALU that perform arithmetic and logical operations.

(3) Computer

It's a image processing system and it's a general-purpose computer and can range from a PC to supercomputer.

(4) Software

The software for image processing system consist of specialized modules that perform specific task.

Software packages are used to write code using specialized modules.

(5) Mass Storage

Mass storage capability is needed if the image is not compressed.

Three principal categories.

- Computer memory, frame buffer are used for short storage during processing.
Frame buffer are special board that store more images and accessed rapidly at video rates.
- On-line storage for relatively fast recall
e.g. magnetic disk or optical media.
- Archival storage characterized by frequent access e.g.: magnetic tapes and optical disks.

6. Image Display:-

Display are color TV monitors. These monitors are driven by the off of image and graphics display cards.

7. Hard copy

Record image and print them. It include laser printer, film cameras, heat sensitive device, inkjet printers and digital units.

8. Networking

Used for the transmission of images. It include optical fiber and other broad band technologies.

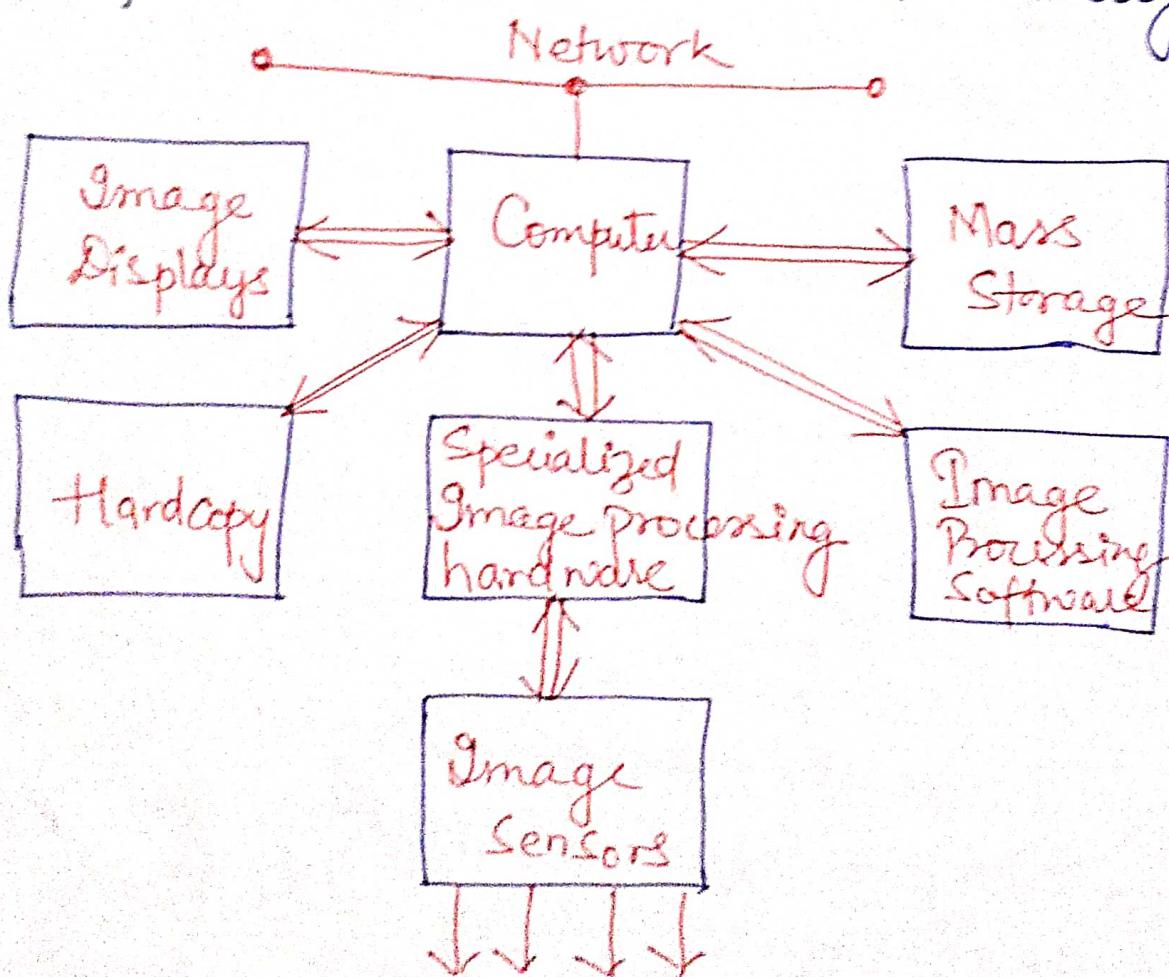
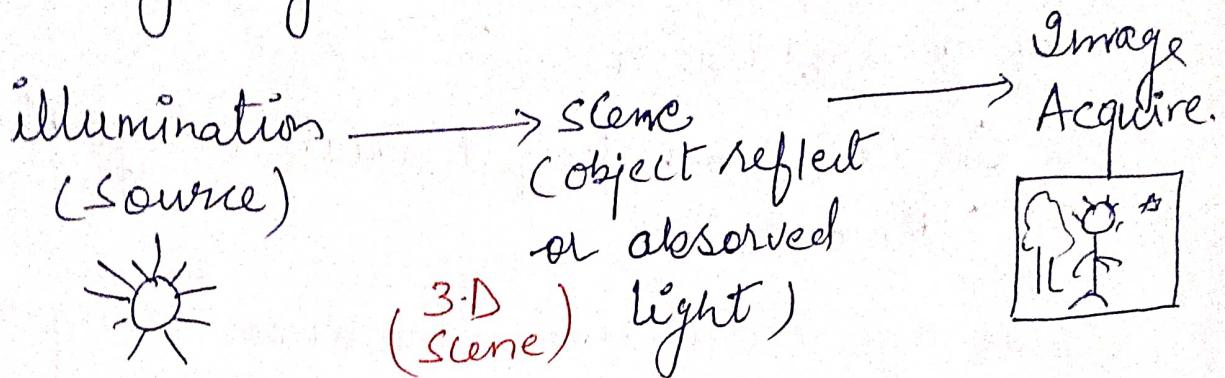


Image Acquisition:-

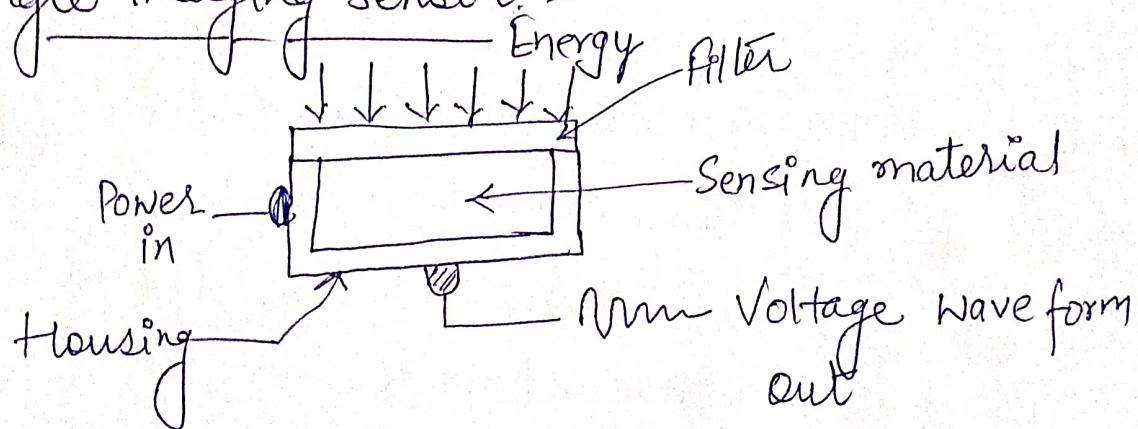
Image acquisition and sensing depend on the "illumination" source and reflected or absorption of energy from that source by the elements of the "Scene" being imaged



e.g.: 1) light reflected from rock

2) X-ray passes through human body

Single Imaging Sensor:-

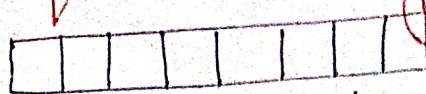


(1) Image Acquisition Using a Single Sensor

Photodiode is the single sensor and constructed by silicon material and its C/P voltage wavelength is proportional to light.

In order to generate 2D image using a single sensor, there has to be relative displacement in both X and Y directions b/w the sensor and the area to be imaged. e.g.: In scanner

(2) Image Acquisition using Sensor strips:-

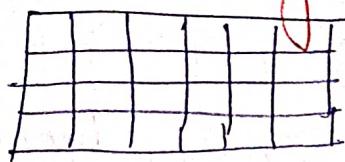


By using On-line sensor, two dimension images can be captured at a time.

e.g.: - 1) Sensor strip on aircraft can capture the image of geographical area.

2) In CAT (computerized axial tomography)
MRI and PET

(3) Image Acquisition using sensor arrays:-



Used in CCD (charge-coupled device) array, which can be manufactured with a broad-range of sensing properties and can be packaged in rugged array of 4000×4000 elements or more.

→ Image is a two-dimensional func $f(x, y)$

• It must be nonzero and finite so

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

• $f(x, y)$ is characterized by two components

(1) illumination of source $i(x, y)$

(2) amount of illumination reflected $r(x, y)$

$$\text{so, } f(x, y) = i(x, y) r(x, y)$$

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

$$0 < r(x, y) < 1$$

(total absorption) (total reflected)

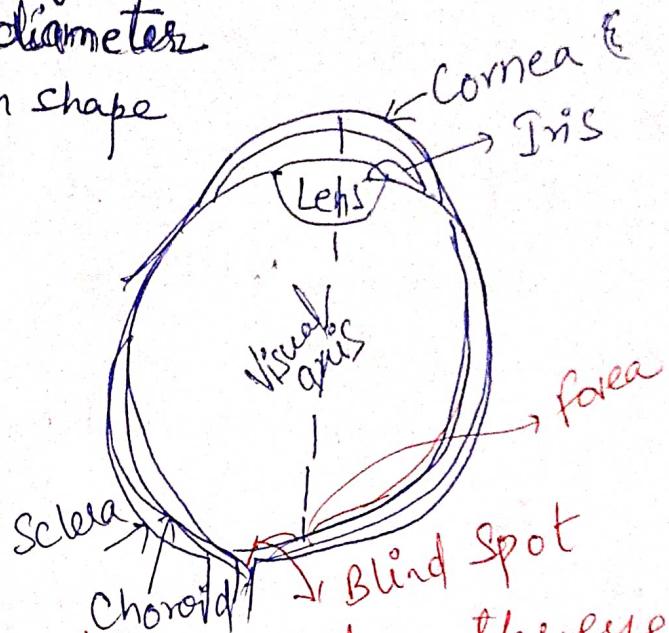
Elements of Visual Perception

(2)

① Human Eye & Its Structure

- 20mm diameter

- Sphere in shape



Three membranes enclose the eye

① Cornea & Sclera (Outer cover)

② Choroid

③ retina

→ tough, transparent tissue that covers anterior surface of the eye

Sclera: - The Sclera is continuation of cornea and also covers remain portion of the optic globe.

Choroid: - contains m/p of blood vessels. The major source of nutrition to the eye is divided into ciliary body and iris diaphragm. The Iris diaphragm is used to control the amount of light enter into the eye.

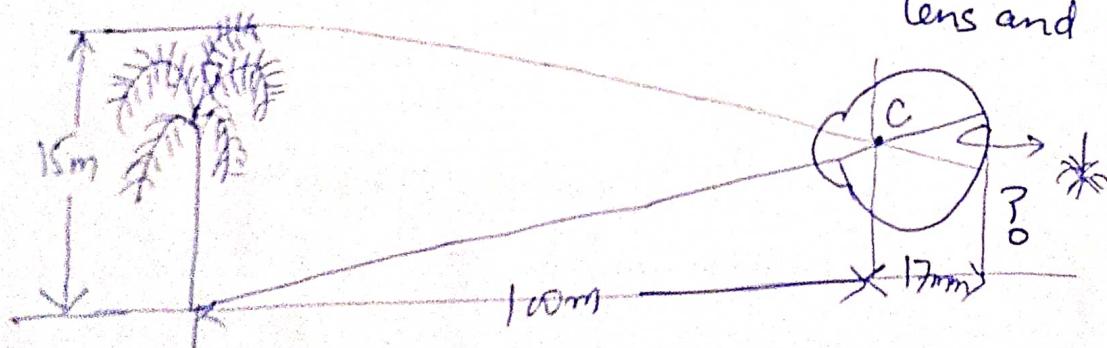
Lens: - The lens is made up of fibrous cell. It contains 60 to 70% water about 4% fat and more protein. It is colored with a slightly yellow pigmentation.

Inner most membrane of the eye is the retina.
Retina - The retina consist of two type of light receptors - cones and rods.

Cones - There are 6 to 7 million no. of cones in each eye. The cones are located at the central portion of the eye called ^{retina} fovea and it highly sensitive to color. Each cone connected to its nerve end. The cone vision is called photopic or bright-light vision.

Rods - There are 45 to 100 million distributed over retinal surface. The rods are connected to a single nerve, hence not involved in color vision. These are sensitive to low levels of illumination. Rod vision is called scotopic or dim-light vision.

Image formation in the eye :-



17 mm = center of lens and retina

$$\frac{100 \text{ m}}{15 \text{ m}} = \frac{17 \text{ mm}}{x}$$
$$\Rightarrow x = \frac{17 \times 15}{100}$$
$$\Rightarrow x = 2.55 \text{ mm}$$

$$\begin{array}{r} 17 \times 15 \\ \hline 255 \end{array}$$

Color image Representations

③

Types of Images →

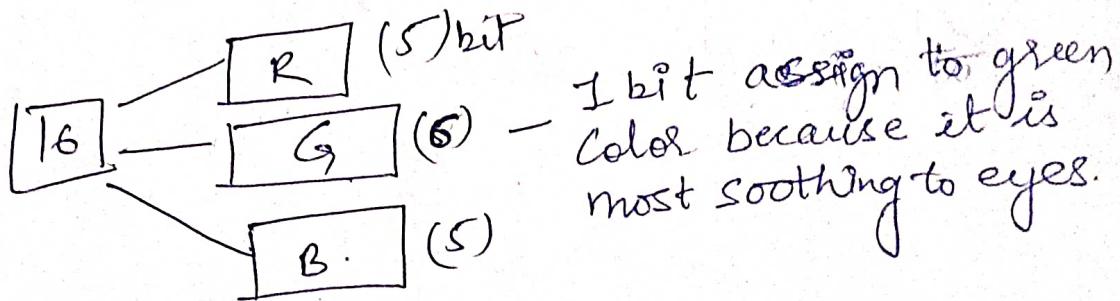
① Binary Image: → 0 & 1 — Monochrome

② 8 bit color format: Black White
 $2^8 = 256$ colors

0 - - 127 - - 255
Black ↓ gray ↑ White
Color color color

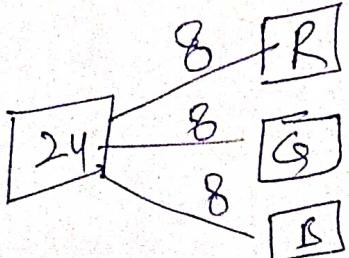
③ 16-bit color

$2^{16} = 65,536$ — High color format



④ 24-bit color

True color



Black (0,0,0)

White (255, 255, 255)

Red (255, 0, 0)

Green (0, 255, 0)

Blue (0, 0, 255)

Grey (128, 128, 128)

(5) CMYK Color Model

C - cyan (0, 255, 255)
 (255, 0, 255) M - magenta. in printers

Y - yellow (255, 255, 0)

K - black.

two carters

CMYK Black.

Conversion

RGB → to Hex

(1) Take color (255, 255, 255) white

(2) Take first portion 255

(3) Divide by 16

$$\begin{array}{r} 255 \\ 16 \overline{) 95} \\ \hline 80 \\ \hline 15 \end{array}$$

↓ P
↓ P

FF

(4) apply same process for remaining two parts

FF FF FF Ans

Hex to RGB

(1) Take hex no FFFF

(2) Break it into three part

$$(1111)(1111) = 256$$

Answer (255, 255, 255)



Poornima

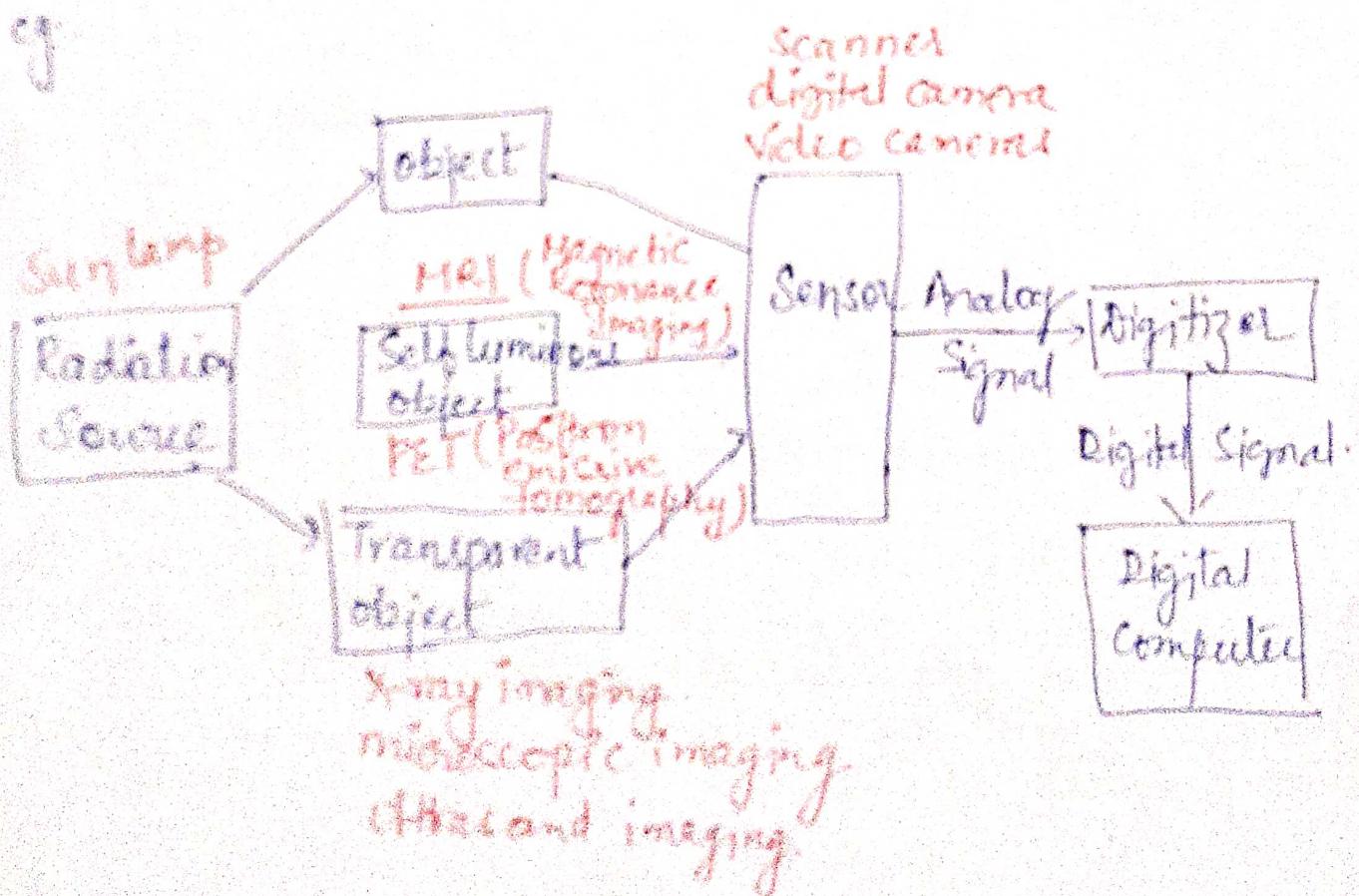
COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

0

PAGE NO.

- Image is a two dimensional (2D) signal $f(x, y)$ where the value of $f(x, y)$ represents amplitude or intensity of the image.
- So, first we have to store the natural pictures in the form of image. Natural pictures are captured with the help of camera phone etc. So we convert it into digital form (Sampling & Quantization).



- Our challenge is to acquire the image for p.

Three type of processing

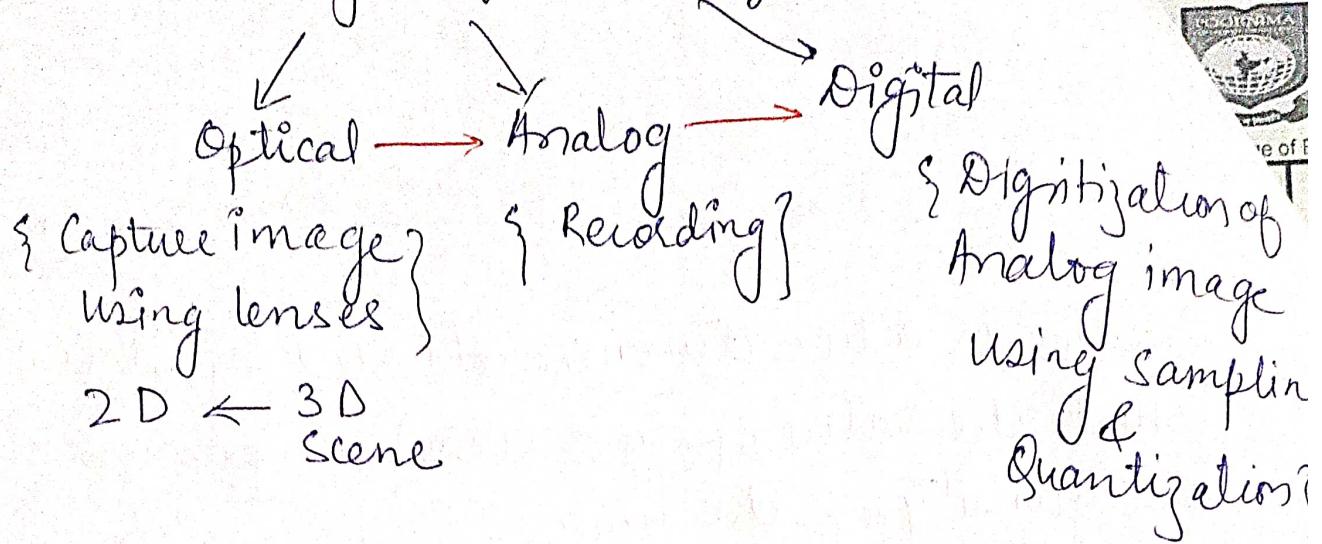
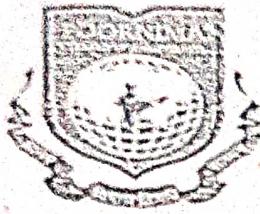


Image processing & Related fields.

- In Image Processing - Computer Graphics - Input is value & op is image & op is image.
- Input is
- Signal Processing { deal with processing of a signal }
 - Machine Vision (MV) { one-dimensional signal }
MV is to interpret the image and to extract its physical, geometrical or topological properties }
- Computer Vision:-
- Input is an image & op is info.
eg: Biometric m/c.



Poornima COLLEGE OF ENGINEERING

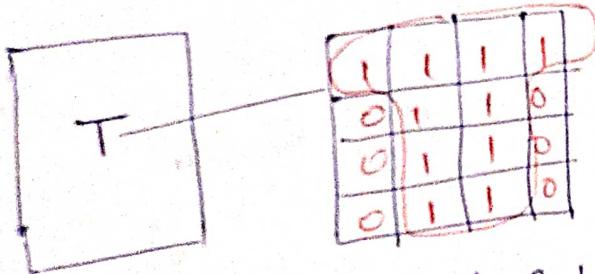
DETAILED LECTURE NOTES

(3)
PAGE NO.

Digital-Image representation.

An image is a 2D signal in spatial domain with coordinates $x \& y \rightarrow f(x, y)$

- Medical images : MRI (Magnetic Resonance Image)
CT (Computerized Tomography)
are 3D images $f(x, y, z)$



So, $f(x, y)$ is divided into X rows & Y columns.

$$x = 0, \dots, X-1$$

$$y = 0, \dots, Y-1$$

pixel \rightarrow intersection of row & column

pel \downarrow "picture element"

Value at $f(x, y)$ is called Intensity of image.

\rightarrow No of rows \rightarrow vertical resolution

\rightarrow No of columns \rightarrow horizontal resolution

e.g: $256 \times 256, 512 \times 512$

Resolution of image depend on two parameters

→ pixel intensity value & no. of pixels

e.g. for Binary Image value is 0 or 1 → So 1 bit is sufficient.

No. of bits necessary to encode array → bit depth.

e.g. for Gray image values = 0, 255 → $2^8 \rightarrow 8$ bits.

~~Resolution~~ = $\frac{\text{Size of image}}{\text{No. of rows} \times \text{No. of columns} \times \text{Bit Depth}}$

Q: What is the storage requirement for

1024×1024 binary image.

$$\begin{aligned}\text{Ans} = \text{Size} &= 1024 \times 1024 \times 1 \text{ bit} \\ &= \frac{1024 \times 1024}{8} \text{ bytes} \\ &= \frac{128 \times 1024}{1024} \text{ KBytes} \\ &= 128 \text{ KB.}\end{aligned}$$

Q: What is storage requirement for

1024×1024 24-bit color image.

$$\begin{aligned}\text{Ans} - \text{Size} &= \frac{1024 \times 1024 \times 24}{8} \text{ bytes} \\ &= \frac{1024 \times 1024 \times 3}{1024} \text{ kb} \\ &= 3072 \text{ kb.}\end{aligned}$$



POORNIMA

COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

PAGE NO. ①

- Q. If a 375×300 Gray-scale Image needs to be sent across the channel of capacity 28 kbps , then how much transmission time is required.

Solu:-

Gray image - 8 bits used

$$\text{Size} = \frac{375 \times 300 \times 8}{8 \times 10^3} \text{ KB} = \frac{375 \times 300}{1000} \text{ KB}$$

$$\text{Time} = \frac{375 \times 300 \times 8}{1000 \times 28} \text{ sec}$$

$$\begin{aligned} \text{Time} &= \frac{375}{280} \text{ sec} = \frac{375}{280} \text{ sec} \\ &= \frac{37.5}{280} \text{ sec} = \frac{37.5}{280} \text{ sec} \\ &= \frac{1125}{5600} \text{ sec} = \frac{1125}{5600} \text{ sec} \\ &= \frac{1125}{5600} \text{ sec} = 4.2 \text{ sec.} \end{aligned}$$

$$\begin{array}{r} 375 \\ \times 280 \\ \hline 1125 \\ 1125 \\ \hline 105000 \end{array}$$

- Q:- Given Gray -image of size 5 inches by 6 inches Scanned at the rate of 300 dpi , find

(1) Bits to store image

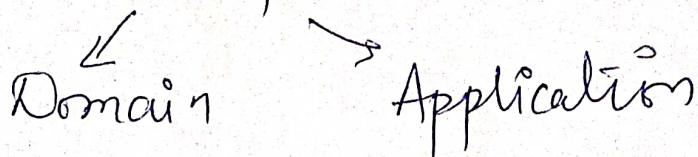
(2) Time required to transmit at 28 kbps .

Ans (1) $= 5 \times 300 \times 6 \times 300 \times 8 = 21600000 \text{ bits}$

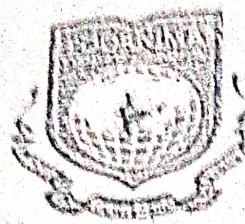
(2) $\text{Time} = \frac{\text{No. of bits in image}}{\text{Transmission Speed}} = 771.43 \text{ sec}$

Domain Specific Images:-

Images can be classified on the basis off



- ① Range Images: → Encountered in Computer Vision.
 - ✓ Pixel value denotes the distance b/w object & camera
 - ✓ Also known as depth images.
- ② Multispectral Images: → Remote Sensing Application
 - ✓ Taken on different bands of visible or infrared regions. of electromagnetic wave
 - ✓ Bands like RGB, ultraviolet, electromagnetic



POORNIMA

COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

PAGE NO. 6

Image Processing Applications:

Electromagnetic Domain \rightarrow visible light

Type of radiation	Freq Range	Wavelength	Image processing
① Radio Waves Remote sensing (Radar System)	$10^5 - 10^{10}$	$> 10^7$	AM/FM Radio Radar image
② Microwaves (SAR) \rightarrow Radar system	$10^{10} - 10^{12}$	$10^9 - 10^6$	
③ Infrared Infrared camera by military	$10^{12} - 10^{14}$	$10^9 - 1000$	Thermal imaging
		$10^9 - 1000$	Optical
④ Visible light photography, telescopes	$4 - 7.5 \times 10^{14}$	$7000 - 4000$	Optical
⑤ Ultraviolet Biomedical imaging	$10^{15} - 10^{17}$	$4000 - 10$	
⑥ X-ray X-ray	$10^{17} - 10^{20}$	$10 - 0.1$	Medical & Industrial
⑦ Gamma rays Nuclear medicine	$10^{20} - 10^{24}$	< 0.1	Medical

RGB to Gray

Two methods

(1) Average method

(2) Weighted method or luminosity method.

Average Method :-

$$\text{Gray image} = (R + G + B)/3$$

Outcome is not as expected, because three different colors have different wavelength, so we have to take average according to their contributions.

Weighted method or luminosity method :-

- Red color has more wavelength of all three.
- Green less wavelength but also soothing effect to eyes.

Decrease contribution of Red color and increase contribution of green color.

$$\text{New gray image} = (0.3 \times R) + (0.59 \times G) + (0.11 \times B)$$

mean 30% of Red

59% of Green

11% of Blue

Sampling & Quantization relationship with Pixels

- If we are doing sampling only, the signal is not converted into digital format.
- Quantization is also needed.

if 5 rows and 5 column

$\text{Pixel (total no)} = \text{total no. of rows} \times \text{total no. of columns}$

$$25 = 5 \times 5$$

Total pixel = 25

In CCD

- no of pixels equal to the no of pixels.

Sampling → up Sampling (Over Sampling) or

down sampling

Zooming

Q: What is the difference b/w Sampling & Zooming?

Sampling on signals

Zooming on the digital images.

Resolution :- $m \times n$

if resolution high ↑ Quality of Image ↑ high
Mega pixel

$m \times n / 1 \text{ million}$

Size of image = no. of pixels \times [bpp]

bits/pixel

(7)

g.: Image dimension = 2500×3192

$$\text{Resolution} = \frac{2500 \times 3192}{1000000} = 7.98 \approx 8 \text{ mega pixel}$$

$$\begin{array}{r} 2500 \\ \times 3192 \\ \hline 1596 \\ +10000 \\ \hline 798 \end{array}$$

$$\begin{array}{r} 2500 \\ \times 3192 \\ \hline 400 \\ 200 \\ 100 \\ \hline 798 \end{array}$$

Aspect Ratio:-

Another important concept with the pixel resolution is aspect ratio.

$$\text{Aspect ratio} = \frac{\text{width of image}}{\text{length of image}} = \frac{C}{R}$$

Like $1.33:1$, $1.66:1$, $2.00:1$ etc.

Eg:- Aspect ratio = $6:2$

Pixel resolution = 480000

Bits per pixel grayscale image = 8 bpp

find: Number of rows? Size of
Number of column?

$$\text{Ans: } A.R = \frac{C}{R} = \frac{6}{2} \Rightarrow C = \frac{6R}{2} \quad \text{---(1)}$$

$$\text{Pixel resolution} = 480000 = R \times C$$

$$\Rightarrow C = \frac{480000}{R} \quad \text{---(2)}$$

Using (1) & (2)

$$\frac{6R}{2} = \frac{480000}{R}$$

$$R^2 = \frac{480000 \times 2}{6}$$

$$R = \sqrt{2 \times 2 \times 2 \times 2 \times 100 \times 100} = 2 \times 2 \times 100 = 400$$

$$C = \frac{6 \times R}{2}$$

$$= \frac{6 \times 400 \times 200}{2} = 1200$$

Size of image = $r \times c \times \text{bpp}$

$$= 400 \times 1200 \times 8 \text{ bits}$$

$$= \frac{480000 \times 8}{800000} \text{ bytes}$$

$$= \frac{480000}{1024} \text{ KB}$$

$$\approx 48 (\text{kb})$$

Zooming — (Two types of zooming)

(1) Optical Zoom (2) Digital Zoom

(Before taking an particular image)

- Known as pre-processing Zoom
- True Zoom
- Camera lens is physically extended to zoom or magnify an object

(After an image has been captured.)

- Known as processing of image
- pixels get expand
- Quality of image compromised.
- Different algorithms as applied to make zoom of image.

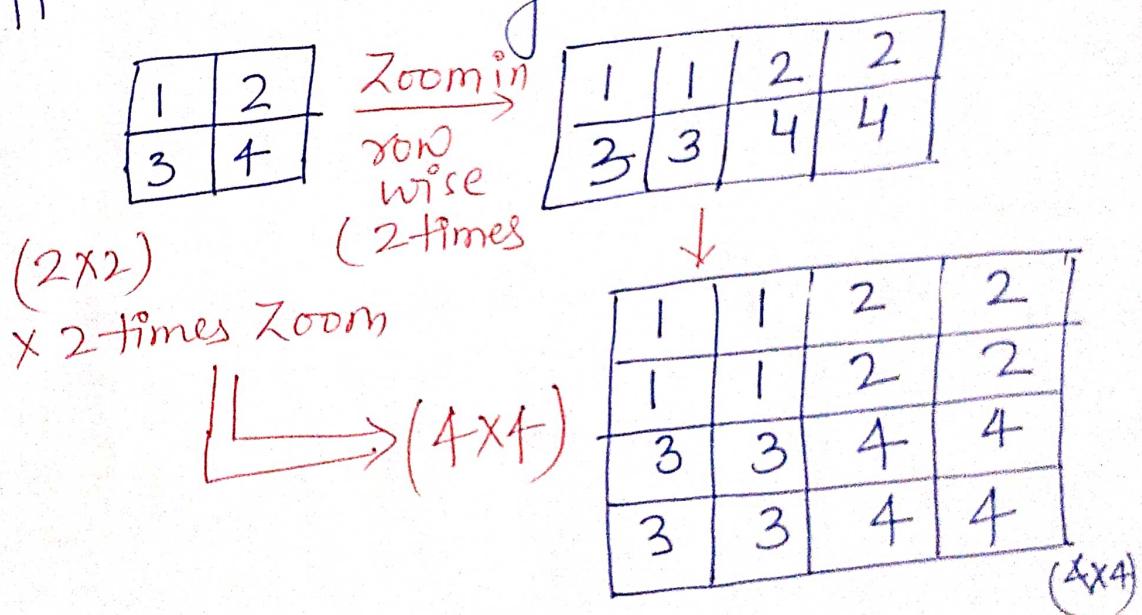
Zooming methods:-

- (1) Pixel replication or (nearest neighbor)
- (2) Zero order hold method (interpolation)
- (3) Zooming k times

(i) Pixel replication:

- As name suggest, just replicate the neighbouring pixels.

- eg.: Suppose we have image 2×2



Advantage: - Very simple

Disadvantage: - Output image get blur and if zooming factor increased, the image got more and more blurred.

(ii) Zero Order hold

It's known as zoom twice.
 (i) Pick two adjacent elements from the rows respectively, add them and divide the result by two
 (ii) Place the result in b/w those two elements.

Eg:- Image (2×2)

1	2
3	4

Row wise zooming

1	1	2
3	3	4

$$(2+1) = 3/2 = 1.5 \approx 1$$

$$(4+3) = 7/2 = 3.5 \approx 3$$

Column wise zooming

1	1	2
2	2	3
3	3	4

$$(1+3) = 4/2 = 2$$

$$(4+2) = 6/2 = 3$$

Size of Image :- $(2(\text{no. of row}) - 1) \times (2(\text{no. of col}) - 1)$

$$= (2 \times 2 - 1) \times (2 \times 2 - 1) = 3 \times 3$$

Advantage :- Does not create blurry image as compare to nearest neighbor interpolation method.

Disadvantage :- It only run on the power of 2. like if we have image of 2×2 and want to zoom

it 6 times then

⑦

$$\text{New image size} = (6 \times 2 - 1) \times (6 \times 2 - 1)$$
$$= 11 \times 11$$

Now after applying this method

$$2 \times 2 \rightarrow 3 \times 3 \rightarrow 5 \times 5 \rightarrow 9 \times 9$$

$$\downarrow$$
$$17 \times 17$$

So, you are not getting the dimensions according to

(iii) k-Times Zooming: —

- One of the perfect zooming algorithm

e.g.:— Steps: (i) Take two adjacent pixel and then subtract smaller from the greater one (OP)

(ii) Divide the output (OP) with zooming factor (k). Now, add result in the smaller value in between the adjacent pixel.

(iii) Repeat step (ii) for place (k-1) values.
Add OP in the current calculated value.

(iv) Repeat all steps for all rows and columns

(v) get zoomed image.

Row wise Zooming

Zoom three times

K=3

15	30	15
30	15	30

$$(i) 15 - 30 = (30-15) = 15/3 = 5 \text{ (OP) same for col 2 and 3}$$

$$30 - 15 = (30-15) = 15/3 = 5 \text{ (OP)}$$

Adjacent pixel

(iii) now add 5 into 15, because 15 is the smaller value.

$k=1$	$k=2$	$k=1$	$k=2$
15	20	25	30
30	20	25	15

Add OP to last value

$$\text{Add OP to smaller value} = 15 + 5 = 20$$

(iv) Repeat same procedure for last two pixel values.

So, Row wise Zooming is present

15	20	25	30	20	25	15
30	20	25	15	20	25	30

After inserting the values sort the inserted values in ascending order according to smaller value. Start sorting from smaller value

So, New table

15	20	25	30	25	20	15
30	25	20	15	20	25	30

Column Wise Zooming:-

(i) Calculate the different b/w adjacent pixel value

$$(30-15) \quad (25-20) \quad (25-20) \quad (30-15) \quad (25-20) \quad (25-20) \quad (30-15)$$

1	↓	↓	↓	15/3	↓	↓	↓
15/3	5/3	5/3	5/3	15/3	5/3	5/3	15/3
OP	5	1	1	5	1	1	5

	15	20	25	30	25	20	15
K=1	20	21	21	25	21	21	20
K=2	25	22	22	25	22	22	25
	30	25	20	15	20	25	30

New image size:-

Old image size = 2×3

New image size = 4×7

So

$$\begin{aligned}
 \text{formula} &= (K(\text{no. of row}-1)+1) \times \\
 &\quad (K(\text{no. of column}-1)+1) \\
 &= (3 \times (2-1)+1) \times (3 \times (3-1)+1) \\
 &= (3 \times 1+1) \times (3 \times 2+1) \\
 &= \boxed{4 \times 7}
 \end{aligned}$$

Advantage:- It gives improved result and comprises the power of the two algorithms.

Disadvantage:- In the end, sort the pixel value, which increase the cost of computation.

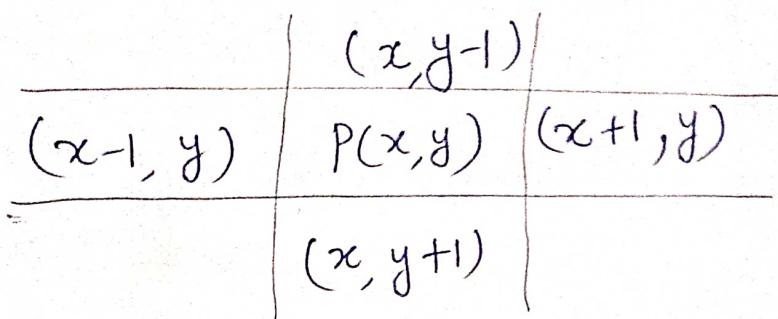
ic Relationship b/w Pixels:-

①

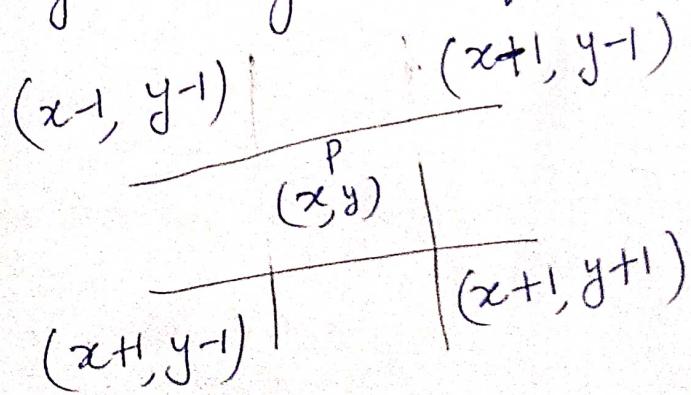
An image is denoted by $f(x, y)$. To represent a pixel, we are using term p and q.

① Neighbors of a Pixel:-

A pixel p at co-ordinates (x, y) has four horizontal and vertical neighbors.



- This set of pixels, called the 4-neighbors of p, is denoted by $N_4(p)$.
- Each pixel is a unit distance from (x, y) .
- If four diagonal neighbors of p have coordinate



- So, denoted by $N_D(p)$
- These four and four points together are called 8-neighbors of p, denoted by $N_8(p)$.

② Adjacency, Connectivity, Regions and Boundaries

Let V be the set of intensity values used to define adjacency. In a binary image $V = \{1\}$, if adjacency of pixels with value 1. for gray image V could be $\{0, 255\}$

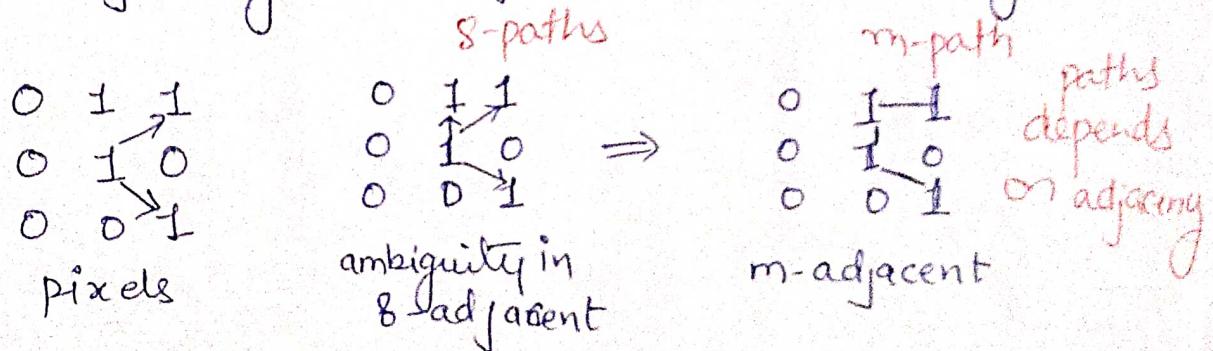
Three type of adjacency:-

(1) 4-adjacency:- Two pixels p and q , with values from V are 4-adjacent if q , is in the set $N_4(p)$.

(2) 8-adjacency:- Two pixels p and q , with values from V are 8-adjacent if q , is in the set $N_8(p)$.

(3) Mixed m-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 $N_8(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V .

• Mixed adjacency is modified of 8-adjacency



• A (digital) Path from pixel p with coordinates (x, y) to pixel q , with coordinate (s, t) is a sequence of distinct pixel with coordinates where $(x_0, y_0) = (x, y)$ & $(x_n, y_n) = (s, t)$

Closed Path. if $(x_0, y_0) = (x_n, y_n)$ $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ { n is the length of path }

S represents a subset of pixels in an image. (2)
 Two pixels p and q are said to be connected in S if there exist a path b/w them consisting entirely of pixels in S . (connected set)

for 1 1 1
 4-adjacency 1 0 1
 these regions 0 1 0
 are not connected 0 0 1
 Set 8-adjacency {Connected Set}

Two regions R_i and R_j are said to be adjacent if their union forms a connected set.

Regions that are not adjacent are said to be disjoint.

Boundaries:-

e.g:-

Suppose image

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

have K regions

0 1 (1) 1 0
 0 1 1 1 0

R_k .

0 0 0 0 0

) $K = 1, 2, 3, \dots, K$

R_u = union of all K regions - foreground

$(R_u)^c$ = complement of $R_u \rightarrow$ background.
 Complement

- The boundary (contour) of a region R is the set of points that are adjacent to points in the complement of R .
 OR The border of a region is the set of pixels in the region that have at least one background neighbor.
- See in above example - for 4-adjacent

Distance Measures:-

for pixels p, q and z with co-ordinates
 (x, y) (s, t) (v, w) respectively, D is a
 distance func or metric if following properties satisfied.

- (i) $D(p, q) \geq 0$ ($D(p, q) = 0$ iff $p = q$),
- (ii) $D(p, q) = D(q, p)$ and
- (iii) $D(p, z) \leq D(p, q) + D(q, z)$, this is called the
Property of triangular inequality.

① The Euclidean distance b/w p and q is defined

as

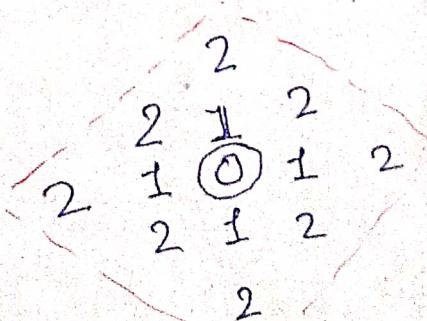
$$D_e(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$$

② D_4 distance (or city-block distance) b/w p and q

$$D_4(p, q) = |x-s| + |y-t|$$

So, the pixels having a D_4 distance from (x, y) less than
or equal to some value r & form a diamond
centered at (x, y)

e.g.: - D_4 distance ≤ 2 , (x, y) - center point



If $D_4 = 1 \rightarrow 4$ neighbours
of (x, y)

D_8 : distance (chessboard distance) b/w p and q

$$D_8(p, q) = \max(|x_p - x_q|, |y_p - y_q|)$$

So, the pixels with D_8 distance from (x, y) , less than or equal to some value r , form a square centered at (x, y) .

e.g. D_8 distance ≤ 2 (x, y) — center point

1	2	2	2	2	2	
2	1	1	1	1	2	
2	1	0	1	1	2	
2	1	1	1	1	2	
2	2	2	2	2	2	

The pixel $D_8=1$ are 8 neighbors of (x, y)

Example:-

$$p(2, 2)$$

$$q = (1, 1)$$

1	2	3
2	q	p
3	.	.

$$\text{Euclidean distance} = \sqrt{(2-1)^2 + (2-1)^2}$$
$$(1-2)^2 + (1-2)^2 = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}$$

$$D_4 = |1-2| + |1-2| = 2$$

$$D_8 = \max(|1-2|, |1-2|) = 1$$

Example :-
Use D_4 distance

$$\text{Pixel A: } |2-2| + |2-2| = 1$$

$$\text{Pixel B: } |3-2| + |2-2| = 1$$

$$\text{Pixel C: } |2-2| + |3-2| = 1$$

$$\text{Pixel D: } |1-2| + |2-2| = 1$$

1	2	3
2	a	p
3	b	c

$$a = (2, 1)$$

$$b = (3, 2)$$

$$c = (2, 3)$$

$$d = (1, 2)$$

$$p = (2, 2)$$

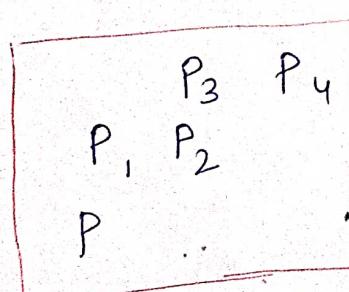
Do:- D_8 distance

Distance Measures:-

- D_m distance: is defined as the shortest m -path b/w the points
- In this case, the distance b/w two pixels will depends on the values of the pixels along the path, as well as the values of their neighbors.

e.g:-

Consider the following arrangement of pixels



Assume P_1, P_2 and P_4

have value 1 and

P_1 and P_3 can have value 0 or 1

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

length 1
4(P, P_1, P_2, P_3, P_4)

Compute
 D_m b/w
 P and P_4 ?

Case 1:-
 $\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}$
if $P_1 = 0$ & $P_3 = 0$

The length of shortest path
 m -path is $2(P, P_1, P_4)$

$$V = \{1\}$$

Case 2:
 $\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}$
if $P_1 = 1$ and $P_3 = 0$

Then length of the
shortest path
 $3(P, P_1, P_2, P_4)$

Case 3:
if $P_1 = 0$ & $P_3 = 1$

$\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}$
Same 3(P, P_1, P_2, P_4)



POORNIMA

COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

PAGE NO ①

Resampling →

- During the scaling of an image, the number of pixels should be increased to retain image quality.
- Pixels are reduced for better compression.
- This increase and decrease of the no. of pixels is called resampling. It is of two forms

- (1) Downsampling (subsampling)
- (2) Upsampling

→ Spatial resolution technique in which image is scaled down by half, by reducing the sampling rate.
- This is done by replacement of a group of pixels by an arbitrarily chosen pixel value. (pixel chosen randomly or top-left pixel with in the neighborhood).

$$P = \begin{array}{|c|c|c|c|} \hline & 3 & 3 & 3 \\ \hline 9 & 9 & 9 & 9 \\ \hline 3 & 3 & 3 & 3 \\ \hline 9 & 9 & 9 & 9 \\ \hline \end{array}$$

$$\Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \quad \text{Single-pixel selection}$$

$$\text{Or } = \frac{3+9+3+9}{4} = 6$$

$$P = \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix}$$

Upsampling:-

Upsampling can be done using replication or interpolation.

Replication:- Zero-order hold process, where each pixel

along the scan line is repeated once. Then the scan line is repeated.

Interpolation:- Linear Interpolation is equivalent to fitting a straight line by taking the average along the rows and columns.

$$f = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$



$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 3 & 3 \\ 1 & 1 & 3 & 3 \end{bmatrix}$$

$$f = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$



$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

↑
for column
Zero Pinterlaced

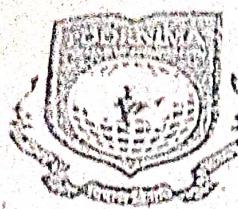
↑
for column
Interpolation (avg of column)

$$\begin{bmatrix} 2 & 1.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 1.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 2 & 1.5 & 1 & 0.5 \\ 1.5 & 1.75 & 2 & 1 \\ 1 & 2 & 3 & 1.5 \\ 0.5 & 1 & 1.5 & 0.75 \end{bmatrix}$$

↑
for row



POORNIMA

COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

PAGE NO. 2

Q: Assume a 1D image $F = [23, 45, 60]$. Perform upsampling using zero hold interpolation and linear interpolation.

Ans:

$$\text{Zero hold } F = [23 \ 0 \ 45 \ 0 \ 60 \ 0]$$
$$= [23 \ 23 \ 45 \ 45 \ 60 \ 60]$$

$$\text{Linear } F = [23 \ 24 \ 45 \ 53 \ 60 \ 30]$$

Q: Apply linear interpolation

$$F = \begin{bmatrix} 0 & 7 \\ 3 & 15 \end{bmatrix}$$

Ans: -

$$\begin{bmatrix} 0 & 3.5 & ? & 3.5 \\ 1.5 & 6.25 & 11 & 15.5 \\ 3 & 9 & 15 & 17.5 \\ 1.5 & 4.5 & 7.5 & 13.75 \end{bmatrix}$$

B

Q: Let $V = \{0, 1\}$ compute D_e , D_4 , D_8 and D_m d.

b/w two pixels p and q.

$$P(3,0)$$

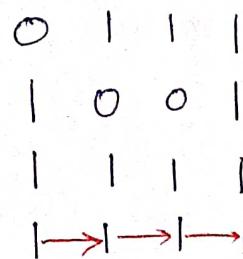
$$q = (2,3)$$

$$D_e = \sqrt{(3-2)^2 + (0-3)^2} \\ = \sqrt{1^2 + 9} = \sqrt{10}$$

$$D_4 = |3-2| + |0-3| \\ = 4$$

$$D_8 = \max(|3-2|, |0-3|) \\ = 3$$

$$D_m = \text{if } V = \{1\}$$



4 Path length

$$P = (0,0) \\ q = (4,4)$$

Q: - Compute D_4 D_8 D_m

$$(1) V = [2,3]$$

$$(2) V = [2,6]$$

D_4 , D_8 don't depend on V value

only D_m depends on V value

$$D_4 = |4-0| + |4-0| \\ = 8$$

$$D_8 = \max(|4-0|, |4-0|)$$

$$= 4$$

$D_m = \text{not possible (no path)}$
 $V = [2,3]$

	0	1	2	3	4
0	0	1	1	1	1
1	1	0	0	0	1
2	1	1	1	1	1
3	1	1	1	1	1

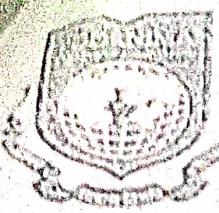
↓ P

P	0	1	2	3	4
0	2	3	2	6	1
1	6	2	3	6	2
2	5	3	2	3	5
3	2	4	3	5	2
4	4	5	2	3	6

q

$$V = [2,6]$$

$D_m = \text{No path.}$



POORNIMA

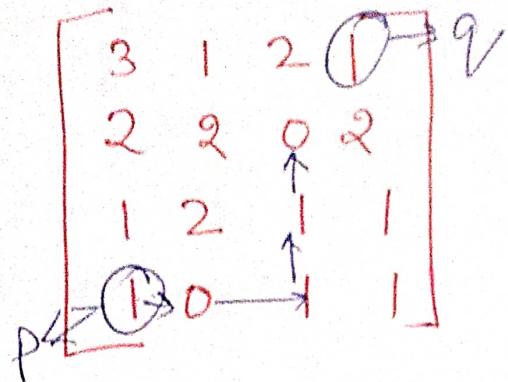
COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

PAGE NO.

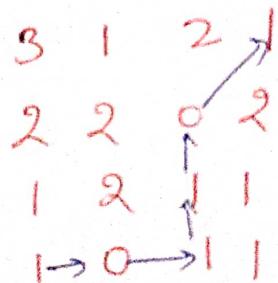
Q) $V = \{0, 1, 3\}$

find δ, λ, β



1) 4-adjacent
not possible

2) 8-adjacent



3) m-path

