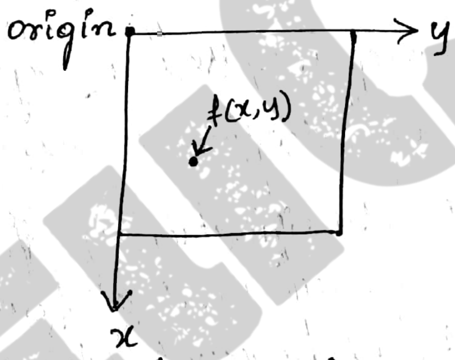
MODULE 1

UNIT-1 : Digital Image fundmentals

* Digital Image processing refers to processing of a two dimensional picture by a digital Computer.

* A monochrome image or an image is a two dimensional intensity function $f(x,y)$, where x and y denote spatial co-ordinates and the value of $f(x,y)$ is proportional to the brightness (or gray level) of an image at (x,y) point.



* A digital image is an image $f(x,y)$ that has been discretized both in spatial co-ordinates and brightness.

* A digital image can be considered as a matrix whose rows and columns indices refer to a point in the image and corresponding matrix element value ~~intens~~ identifies the gray level at that point.

* Matrix elements of digital image matrix are called as image elements, picture elements, pixels or pels.

Types of Image processing

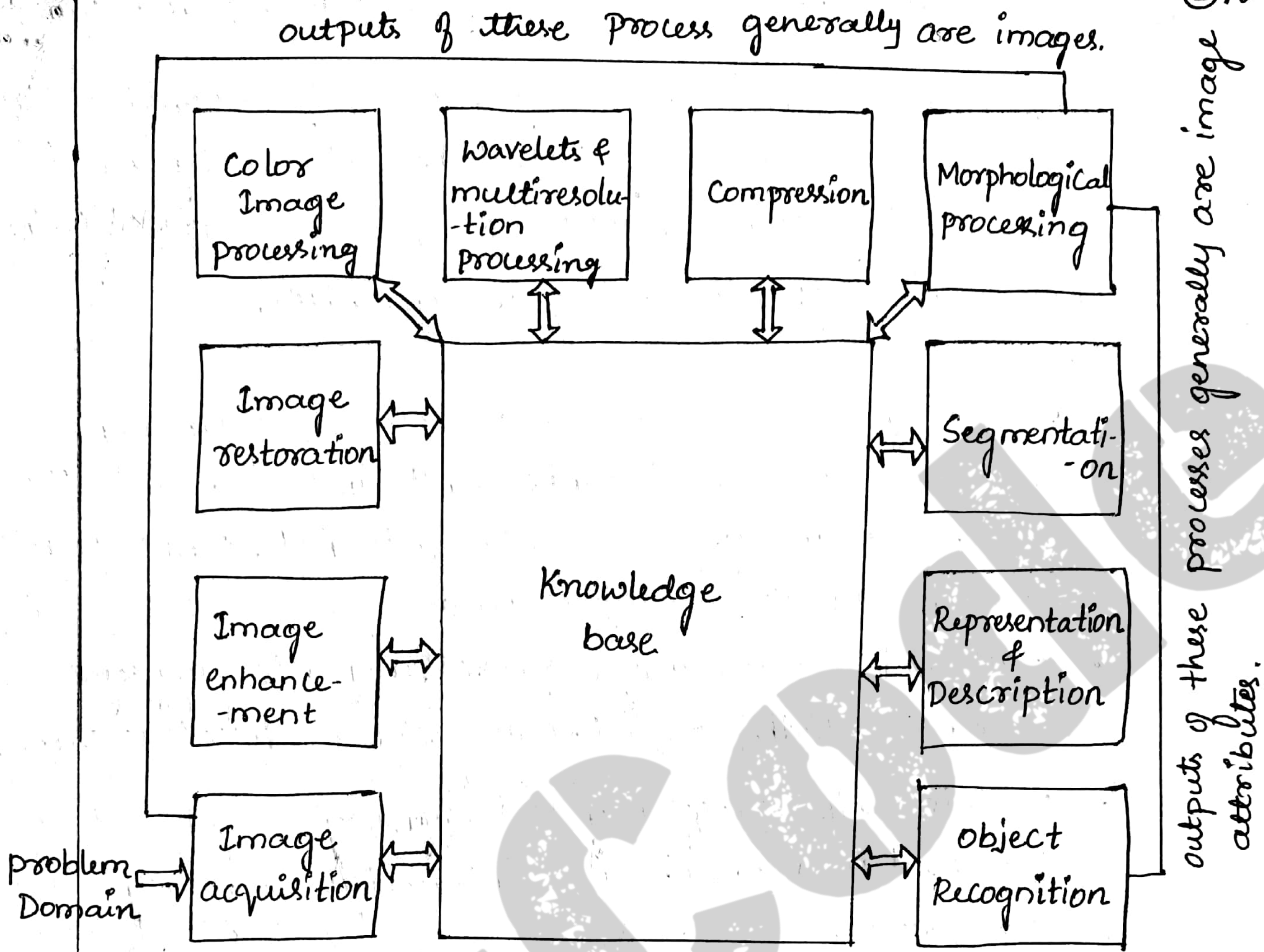
* Low level processing includes noise reduction, contrast enhancement and image sharpening. Input and output of low level processing are images.
* Mid level processing includes segmentation process which is the process of partitioning an image into regions or subdivisions. Here inputs are images but outputs are attributes extracted from images like edges, contours and identifying of individual objects.
* High level processing involves image analysis. (understating of a collection of recognized objects)

Eg:- In automated classification of fruits and vegetables:- Images of all kind of fruit and vegetables are acquired, preprocessed (low level), segmented and described (mid level) in a form suitable for computer processing and finally fruits and vegetables are recognized (high level).

*** Fundamental Steps in Digital Image processing

Problem Domain : Images or data should be taken according to the requirements or applications.

②.a



* <u>Image Acquisition</u> :- This is the first step in image processing. Aquiring an image requires
  - → Imaging sensor eg:- Camera
  - → Capability to digitize the signal produced by the sensor i.e. Analog to digital Converter.
  - → This block also involves preprocessing of the images such as scaling.

* <u>Image enhancement</u> :- This step involves processing of the acquired image so that it looks better than the original image. Eg:- Contrast enhancement. This step is very subjective part of image processing because this can be used with one type of image (say Photographic image) that may not be

applicable to other type of images (eg: biomedical images).

* Image restoration:- This is a process of reconstruction or recovering of an image that has been degraded by using some prior knowledge of the degradation phenomenon. Eg:- Removal of image blur.

* Color image processing:- This area is gaining importance due to significant increase in the use of digital images over the internet. This can be divided to two major areas
  - → Full color processing:- processing the images acquired with a full-color sensor like color TV camera.
  - → pseudo-color processing:- Assigning a color to a particular monochrome intensity or range of intensities.

* Wavelets and multiresolution processing:- Wavelets are the foundation for representing images in various degrees of resolution. They are also widely used for image compression and noise elimination.

* Compression:- This step involves reducing the storage space required and also the bandwidth required for transmitting. The image storage technology has improved significantly and BW capability has also improved a lot due to the use of optical fibre but still image compression is an important step and the standard (file-name.jpg) widely used.
[JPEG - Joint Photographic Expert Group]

* Morphological processing :- This step involves extracting image components useful in the representation and description of the shape.

* Segmentation :- partitioning or subdividing the image into its constituent parts or objects is called as segmentation. A good segmentation procedure helps in successful identification of objects in an image and a weak segmentation procedure may lead to failure. Therefore segmentation process should be very accurate.

* Representation and description :- The output of segmentation process is raw data. This data should be properly represented in order to help further processing. There are two types of representation

→ Boundary representation :- Helps in studying external shape of an object.

→ Regional representation :- Helps in studying internal properties like texture.

→ In some applications both types of representation is used.

Description or feature selection deals with extracting attributes from representation which leads to getting some quantitative information of interest or helps in differentiating one class of objects from other.

* <u>object recognition</u> :- It is a process of assigning a label to an object (eg: car, table etc) based on the description.

* <u>Knowledge base</u> :- Knowledge about a problem domain is placed in the form of knowledge base. Knowledge base can be as complex as a list of all possible defects in a material or it can be as simple as giving information about region in the image from which useful information can be obtained.

<u>NOTE</u>:-

→ Knowledge base controls interaction between modules (or blocks)

→ Result of processing can be viewed at the output of any block.

→ All the processing steps may not be needed for all the applications.

→ All the steps may not require complex interaction interation between each other.

***

<u>Components (or Elements) of Digital Image processing</u>

* Basic components of general-purpose system used for digital image processing is shown in the figure ②

* <u>Image Sensors</u>:-

Two Elements are required for acquiring digital images.

→ A physical device that is sensitive to a band of electromagnetic energy spectrum, like X Rays, visible, Infrared etc. and produces an electrical signal proportional to energy level sensed.

→ A digitizer for converting the device output to digital form.

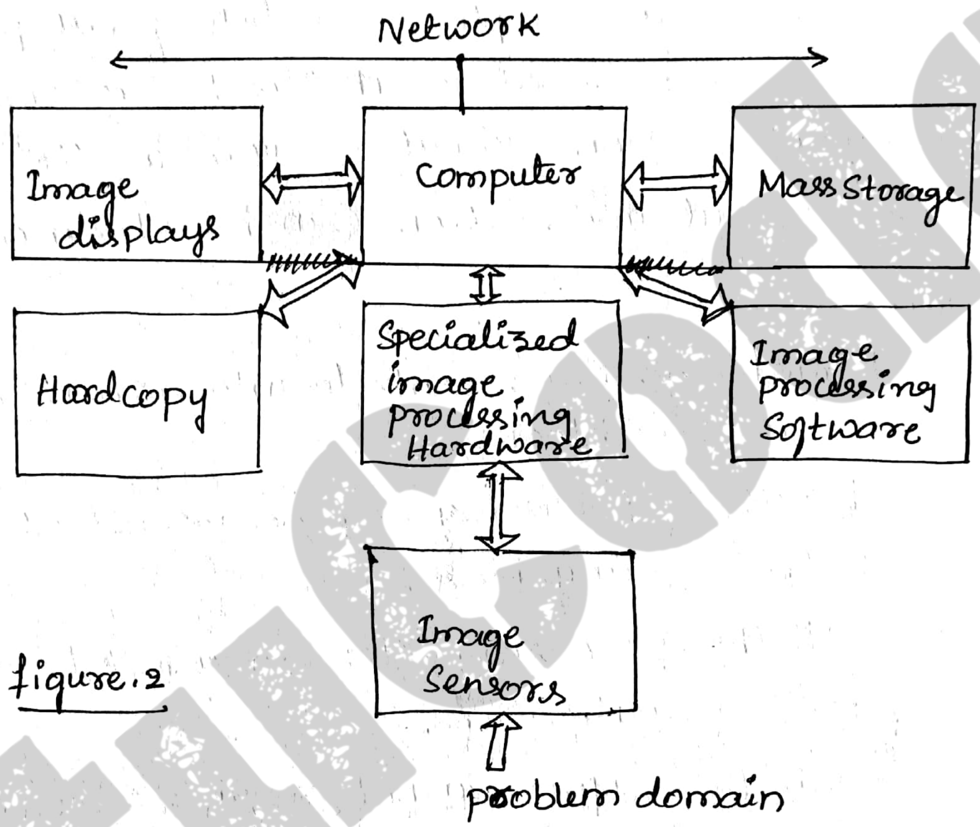


figure.2

* Specialized Image processing hardware:-

This consists of digitizer and a hardware that performs some operations eg:- it may be arithmatic logic unit (ALU) which performs arithematic and logical operations on the entire image. one such operation is averaging of the image as soon as they are digitized in order to reduce the noise. This hardware is some-times called as front end subsystem. Important characteristic of this device is its speed of operation.

(4)

* <u>Computer</u> :- General purpose computer is used in image processing system. It can be ranged from personal computer (PC) to super computer. Sometimes specially designed computers are also used in order to get required level of performance.

* <u>Software</u> :- Image processing uses specialised softwares for performing specific tasks. Some well designed softwares allows the user to write codes. eg:- Matlab, C,

- Astra Image → PC based scientific image processing
  CREAM → used for biological images.
  GRASS → Geographic Resource Analysis Support System.
  OCTAVE → open source → UNIX platform
  free image → open source for project developers.

* <u>Mass Storage</u> :- Mass storage capability is the most important part of image processing applns.

An image of size 1024 x 1024 pixels requires one megabyte of storage space. If ~~a~~ number of images are used then storage becomes a problem.

There are three types of storage used in digital image processing applications.

i) Short term ~~processing~~ storage - used during processing
ii) on-line storage - for fast (frequently used)
iii) Archival storage - for storing images which are not frequently used.

Computer memory is used as short term storage. Another type of storage device called frame buffers that can store one or more images and can be accessed rapidly are also used for short term storage. Frame buffers are placed in specialized image process-ing unit of the block diagram.

on line storage are usually magnetic disk (or) optical storage devices. Archival storage requires large storage space so magnetic tapes or optical disk are used.

* Image displays:-

Flat screen TV monitors are used for displaying images. Monitors are driven by output images and graphics display cards that are integral part of the computer system.

* Hard Copy:-

Laser printers, film cameras, inkjet devices, heat sensitive devices and digital units like CD-ROM disks and optical disks are used for obtaining hardcopy of the image. Films give best resolution

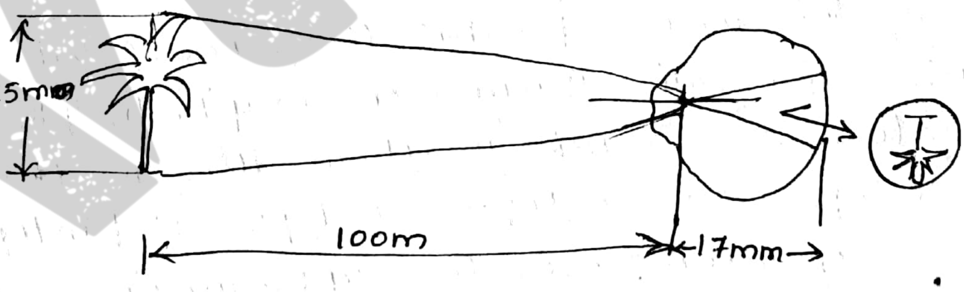
* Networking:-

Networking is always used in computers now-a-days. Bandwidth is the most important consideration while transmitting images. This is because large amount of data is associated with image processing applications. The use of optical fibre and other broad band technologies have helped in overcoming bandwidth problem.

Image formation in the Eye:-

In an ordinary photographic camera, the lens has a fixed focal length, and focusing at various distances is achieved by varying the distance between the lens and the imaging plane, where the film is located. In the human eye, the converse is true; the distance between the lens and the imaging region (the retina) is fixed, and the focal length needed to achieve proper focus is obtained by varying the shape of the lens. The fibers in the ciliary body accomplish this, flattening or thickening the lens for distant or near objects, respectively.

The distance between the center of the lens and the retina along the visual axis is approximately 17mm. The range of focal lengths is approximately 14mm to 17mm, the latter taking place when the eye is relaxed and focused at distances greater than about 3m.



The above figure illustrates how to obtain dimensions of an image formed on the retina

Eq:- $\frac{15}{100} = \frac{h}{17}$ h denote height of that object in the retina image.

$h = 2.5mm$

Brightness Adaptation and Discrimination

The range of light intensity levels ~~is an important~~ to which the human visual system can adapt is enormous - on the order of 10^{10} - from the scotopic threshold to the glare limit. Experimental evidence indicates that subjective brightness (intensity as perceived by the human visual system) is a logarithmic function of the light intensity incident on the eye. Fig 1.0, a plot of light intensity versus subjective brightness, illustrates this characteristics.

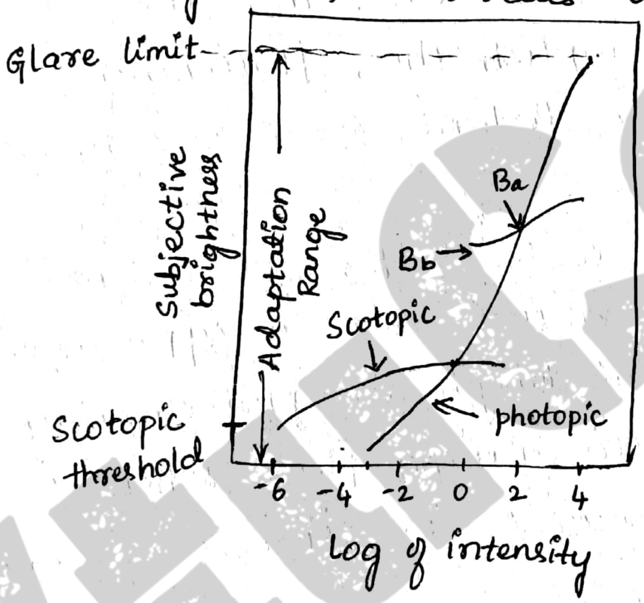


fig 1.0

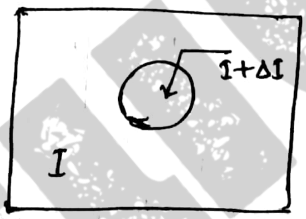
The long solid curve represents the range of intensities to which the visual system can adapt. In photopic vision alone, the range is about 10^6. The transition from ~~photopic~~ to ~~scotopic~~ scotopic to photopic vision is gradual over the approximate range from 0.001 and 0.1 millilambert (-3 to -1 mL in the log scale), as the double branches of the adaptation curve in this range show.

The phenomenon in which large variation by changing the overall sensitivity is called brightness apaptation. The total range of distinct intensity levels the eye can discriminate simultaneously is rather small when compared with the total adaptation range. The current sensitivity level of the visual system is called the brightness adaptation level.

The ability of the eye to discriminate between changes in light intensity at specific adaptation level.

eg:- Illuminated light source intensity I and increme-nt of illumination (added) ΔI, in the form of a short duration flash ~~to~~ that appears as a circle in the center of the uniformly illuminated field as shown in figure 2.

when



fig② Basic Experimental setup used to char brightness discrimination.

The quantity $\frac{\Delta I_c}{I}$, where ΔI_c is the increment of illumination discriminable 50% of the time with background illumination I, is called the Weber's ratio. A small value of $\frac{\Delta I_c}{I}$ means that a small percentage change in intensity is discriminable. This represents "good" brightness discrimination. Conversely, a large value of $\frac{\Delta I_c}{I}$ means that a large percentage change in intensity is required. (Poor brightness discrimination)

Two phenomenon clearly demonstrates that perceived brightness is not a simple function of intensity. The first is based on the fact that the visual system tends to undershoot or overshoot around the boundary of regions of different intensities. Below figure shows a example of this phenomenon.

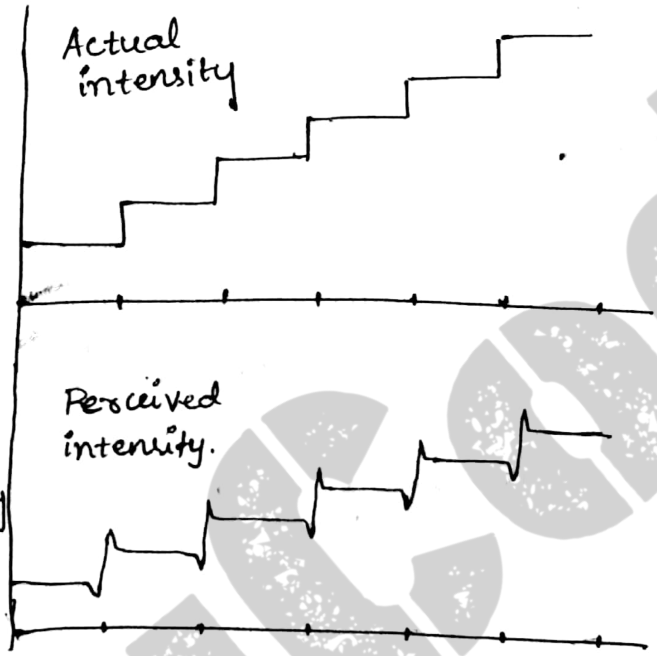


Illustration of the Mach band effect perceived intensity is not a simple function of actual intensity

Although the intensity of the stripes is constant, we actually perceive a brightness pattern that is strongly scalloped near the boundaries. These seemingly scallope-d bands are called Mach bands after Ernst Mach, who first described the phenomenon 1865.

The colors that humans perceive in an object are determined by the nature of the light reflected from the object. A body that reflects light relatively balanced in all visible wavelengths appears white to the observer.

Light without color is called monochromatic or achromatic light. The only feature of monochromatic light is intensity. The intensity of monochromatic light perceived to vary from black to grays & finally to white, The term gray level is used commonly to denote monochromatic intensity.

Chromatic (color) light spans the electromagnetic energy spectrum from approximately 0.43 to 0.79 µm.

The three basic quantities of chromatic light are,

a) Radiance: It is the total amount of energy that flows from the light source. (watts)

b) Luminance: It gives a measure of the amount of energy that (~~flows from the light source~~) an observer can perceive from a light source. (unit - Lumens (Lm))

c) Brightness: It is a subjective descriptor of light perception that is practically impossible to measure.

Image Sensing and Acquisition.

The images which of our interest are generated by the Combination of an "Illumination" Source & the reflection or absorption of energy. from that Source by the elements of the object or Scene being imaged.

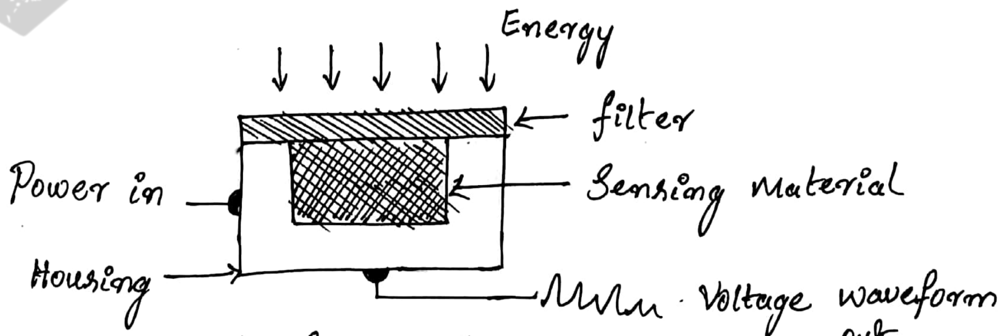
Reflection ⟶ Eg: Light reflected from a planar Surface.

Transmitted through ⟶ X-rays pass through a patient's body for the Purpose of generating a diagnostic X-ray film.

There are 3 basic Sensors. used to transform illumination energy into digital images.

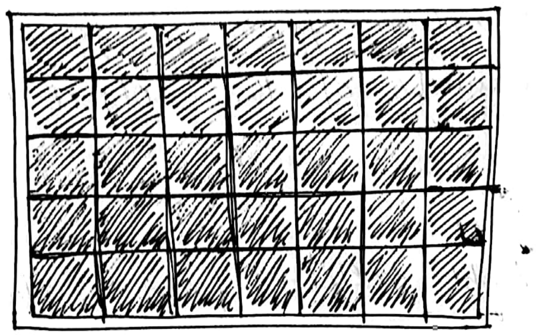
a) Single imaging Sensor
b) Line Sensor
c) Array Sensor.

Process : Incoming energy is transformed into a voltage by the combination of input electrical power & Sensor material that is responsive to the particular type of energy being detected. output voltage waveform is the response of the sensors.



a) Single Sensor.

b) Line Sensor



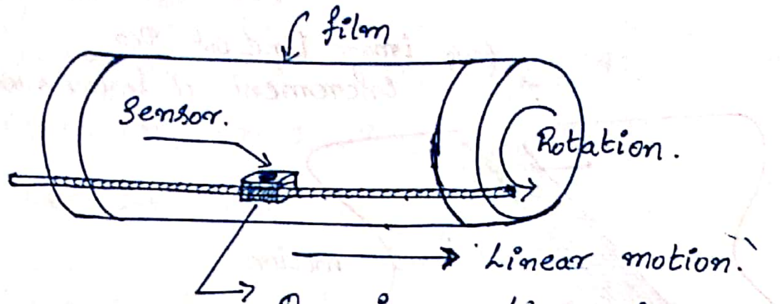
c) Array Sensor

Image acquisition using a Single Sensor.

Example : Photo diodes, LED's etc

A green LED o/p will be stronger for green light than for other components in the visible spectrum. This is accomplished by means of a green (pass) filter in front of the LED (Sensor).

To generate a 2D image using a single sensor, there has to be relative displacement in both horizontal (x) & vertical (y) directions between the sensor & the area to be imaged.



A film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. A single sensor is mounted on a lead screw the provides motion in the perpendicular direction.

Advantages

1) Less expensive (In expensive)
2) Provides high resolution images

Disadvantage.

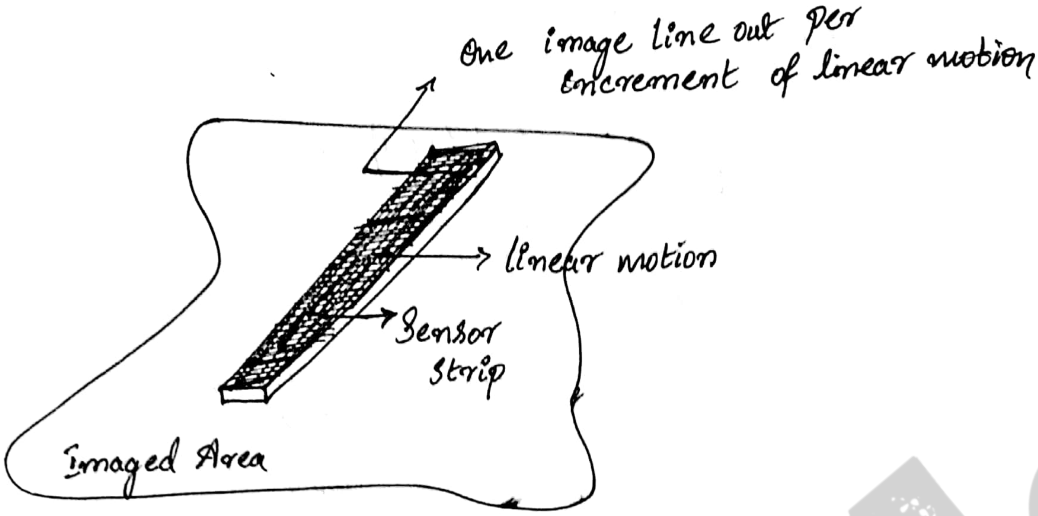
Slow process.

Other examples are, Microdensitometers :- A flat bed with the sensor moving in two linear directions.

Image Acquisition using Sensor strips.

In this arrangement, the strip provides imaging elements in one direction, Motion perpendicular to the strip provides imaging in the other direction.

This type of sensors are made up of 4000 or more in line sensors.



Applications:

a) In airborne imaging applications, in which imaging system is mounted on an aircraft that flies at a constant altitude & speed over the geographical area to be imaged.

An imaging sensor strip that responds to various bands of the electromagnetic spectrum are mounted perpendicular to the direction of flight.

The imaging strip gives one line of an image at a time, & the motion of the strip completes the other dimension of a two-dimensional image.

b) In medical & industrial imaging, sensor strip will be mounted in a ring configuration to obtain cross sectional images of 3-D objects.

A rotating X-ray source provides illumination & the sensors opposite to the source collect the X-ray energy that passes through the object.

Other Applications are:

→ Computerized axial tomography [CAT]

→ Magnetic Resonance Imaging [MRI]

→ Positron Emission Tomography. [PET]

Image acquisition using Sensor arrays.

Elementary sensors are arranged in the form of a 2D-array, leads to sensor arrays.

Eg: * Electromagnetic & ultrasonic sensing devices.

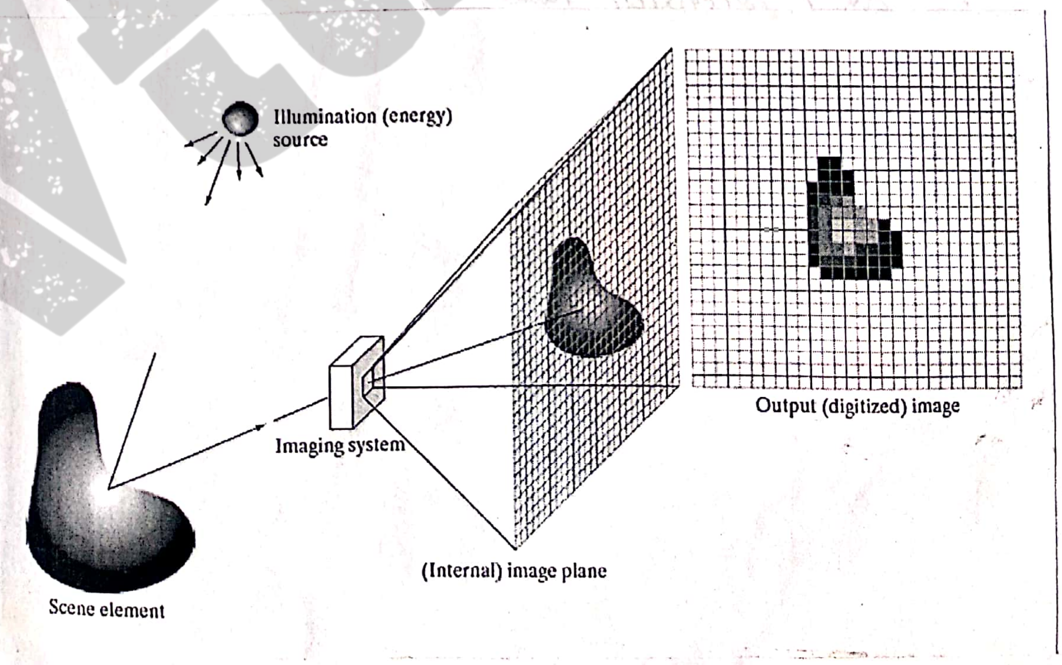
* Digital cameras [CCD Arrays typically]

Array sensors typically have 4000 x 4000 elements or more.

Advantages: * Motion of sensor is not necessary.

* Noise reduction is achieved by integrating i/p light signal with respect to time.

Digital image acquisition process



* Illumination Source projects light energy on a Scene element (object to be imaged). Some part of energy being reflected from the Scene element.

* Imaging System collects the reflected energy from the object & focus it onto an image plane.

* The Sensor array, which is coincident with the focal plane of lens, produces o/p proportional to the integral of the light received at each Sensor.

* Analog & digital circuitry of imaging System produces analog & digital images (Signals) respectively.

A Simple Image formation Model

An image is denoted by a 2 dimensional function of the form $f(x,y)$. The amplitude of 'f' at (x,y) is a positive Scalar quantity whose value is determined by the Source of image.

i.e Pixel intensity value depends on Source of image. in other words, intensity values are proportional to energy radiated by the Source.

$f(x,y)$ must be non zero & finite.

i.e $0 < f(x,y) < \infty$ [Pixel values are non zero & finite]

The function $f(x,y)$ is characterized by

a) Illumination $\longrightarrow i(x,y)$

b) Reflectance $\longrightarrow r(x,y)$

* The amount of the Source illumination incident on the scene being viewed is Called <u>Illumination</u>.

* The amount of the Source illumination reflected back by the objects in the Scene. is Called. <u>Reflectance</u>.

$f(x,y)$ Can be represented as,

$$f(x,y) = i(x,y)\, r(x,y) \longrightarrow (1)$$

where, $0 < i(x,y) < \infty \longrightarrow (2)$

& $0 < r(x,y) < 1 \longrightarrow (3)$

if $r(x,y) = 0$, it Corresponds to total absorption.

& $r(x,y) = 1$, it Corresponds to total reflectance.

<u>Eg</u>: Let the intensity of a monochrome image at any Coordinates (x_0, y_0) be denoted by,

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

In an image, 'L', the min.^{m} pixel intensity is denoted by L_{min} & max^m pixel intensity by L_{max}.

$\therefore \quad L_{min} \leq l \leq L_{max} \longrightarrow (4)$

$L_{min} \longrightarrow$ to be positive
$L_{max} \longrightarrow$ finite.
} Requirements.

$$L_{min} = i_{min}\, r_{min}.$$

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

$\therefore$ The interval $[L_{min}, L_{max}]$ is Called the <u>gray scale</u>.

if $l = L_{min} = 0$, is considered as black.

& $l = L_{max} = L-1$, is considered as white.

∴ A gray scale or black & white image can have the range of intensities $[0, L-1]$.

Image Sampling and Quantization.

To create a digital image, we need to convert the continuous sensed data into digital form. it requires two processes. a) Sampling.
b) Quantization

An image may be continuous with respect to the (x, y) & also in amplitude. To convert it into digital form we have to sample the function in both. co-ordinates & in amplitudes.

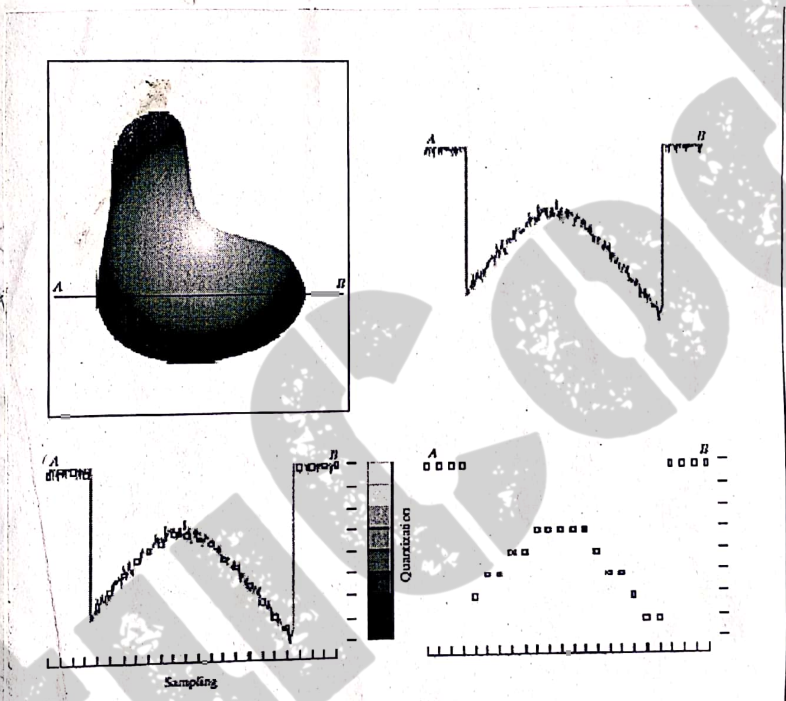
* Digitizing the co-ordinates values is called. Sampling.
* Digitizing the amplitude values is called. Quantization.

eg: Consider a line segment AB in the continuous. image. To sample this function, we take equally spaced samples along line AB. [

The spatial coordinates are indicated in the x-axis. [bottom part]. The samples are shown as white squares.

The set of these discrete locations gives the sampled function. However, the values of the samples still span (vertically) a continuous range of intensity values.

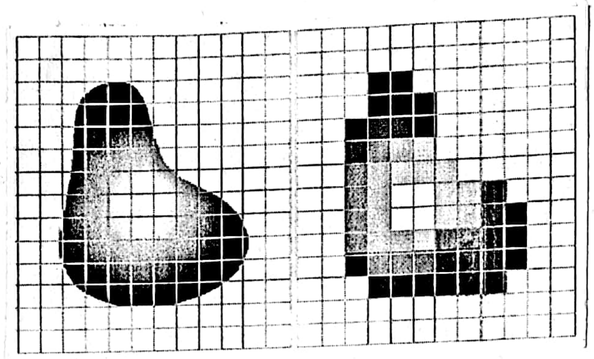
In order to form a digital function, the intensity values also must be converted into (quantized) discrete quantities.



The continuous intensity levels are quantized by assigning one of the eight values to each sample.

Starting at the top of the image and carrying out this procedure line by line produces 2 dimensional digital image.

In addition to the no. of discrete levels used, the accuracy of quantization is dependent on the noise content of the sampled signal.



Clearly, the quality of digital image depends on the no of Samples and discrete intensity levels used in Sampling & quantization.

Representing Digital images.

Let us Consider a Continuous image $f(s,t)$ converted into digital image $f(x,y)$ by sampling & quantization process.

where $(s,t) \longrightarrow$ Continuous co-ordinates
$(x,y) \longrightarrow$ discrete Co-ordinates.

Let $f(x,y)$, Containing M rows & N Columns

$\therefore$ $x = 0\ 1\ 2\ ---\ M-1$
$y = 0\ 1\ 2\ ---\ N-1$

$\therefore$ digital image origin is at $f(0,0)$

The Section of the real plane Spanned by the co-ordinates of an Image is called the Spatial domain. with x & y are Spatial Variables

A 2-D array $f(x,y)$ with M rows & N Columns Can be represented as,

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

· Each element of the matrix is called an image element, picture element, pixel or pel.

Alternate representation: [Traditional Representation]

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} & ---- & a_{0,N-1} \\ a_{1,0} & a_{1,1} & a_{1,2} & ---- & a_{1,N-1} \\ \vdots & & & & \\ a_{M-1,0} & a_{M-1,1} & a_{M-1,2} & ---- & a_{M-1,N-1} \end{bmatrix}$$

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

An image can be represented by a vector (v) of size $MN \times 1$.

Conventional representation of digital image

(Origin)

$f(0,0)$ $f(0,1)$ ------ $f(0,N-1)$

$f(M-1,0)$ $f(M-1,1)$ ---- $f(M-1,N-1)$

Right handed Cartesian Co-ordinate system

Digitization of an image depends on no of intensity levels, let it be "L". due to storage & quantizing hardware considerations, the number of intensity levels typically is an integer power of 2;

ie $L = 2^K$ where $K \rightarrow$ no of bits

Dynamic range: is defined as, the ratio of the maximum measurable intensity to the minimum detectable intensity level in an image.

'upper limit is determined by Saturation

Lower limit by Noise

Contrast: The difference in intensity between the highest & lowest intensity levels in an image.

if dynamic range is high → Good Contrast image

if dynamic range is low → Dull (or) washed-out gray look.

The number of bits (b) required to store a digitized image is.

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

when $M = N$, $b = N^2 K$

Image Interpolation :

"Interpolation is the process of using known data to estimate values at unknown locations."
It is also called as image "Resampling". It is used extensively in tasks such as zooming, shrinking, rotating & geometric corrections.

example : Consider an image of size 500×500, suppose it has to be enlarged. to 750 × 750 pixels [1.5 times]. To achieve zooming, let us create an imaginary grid of 750×750 with the same pixel spacing as the original. Then shrink it so that it fits exactly over the original image.

It is obvious fact that, the pixel spacing in the shrunken 750×750 grid will be less than the original image.

To perform intensity level assignment for any point in the imaginary 750 x 750 grid, we look for its closest pixel in the original image & assign the intensity of that pixel to the new pixel in the 750 x 750 grid.

After assigning intensities to all the points in the overlay grid, we expand it to the original specified size to obtain the zoomed image.

This method is called Nearest Neighbour interpolation. This approach is simple but it produces undesirable artifacts, such as severe distortion of straight edges. hence it is impractical

Bilinear Interpolation :-

In this method we use 4 nearest neighbours to estimate the intensity at a given location.

Let (x, y) be the co-ordinates of the location to which we want to assign an intensity value $v(x, y)$

Then bilinear interpolation is written as,

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

where a, b, c & d are the co-efficients determined from the four equations in the four unknowns that can be written using the four nearest neighbors of point (x, y).

This approach provides better results compared to nearest neighbour method, with a modest increase in computational burden.

Bicubic interpolation :

This method uses 16 nearest neighbors to estimate the intensity at a given location. (x, y).

$$v(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij}\, x^i y^j$$

where 16 co-efficients are determined from the 16 equations in 16 unknowns that can be written using the sixteen nearest neighbor of point (x,y).

Spatial & Intensity Resolution :

Spatial resolution is a measure of the smallest discernible detail in an image. It can be stated as "line pairs per unit distance," & "dots (pixels) per unit distance"

Suppose that we construct a chart with alternating black & white vertical lines, each of width W units. $(W<1)$ The width of a line pair is thus $2W$, & there are $\frac{1}{2}W$ line pairs per unit distance.

Example: if the width of a line is 0.1mm, there are 5 line pairs per unit distance (mm).

Dots per unit distance is a measure of image resolution used commonly in the printing & publishing industry. In U.S it is expressed as dots per inch (dpi)

generally, newspapers are printed with a resolution of 75 dpi

magazines at 133 dpi; glossy brochures at 175 dpi

Intensity Resolution.

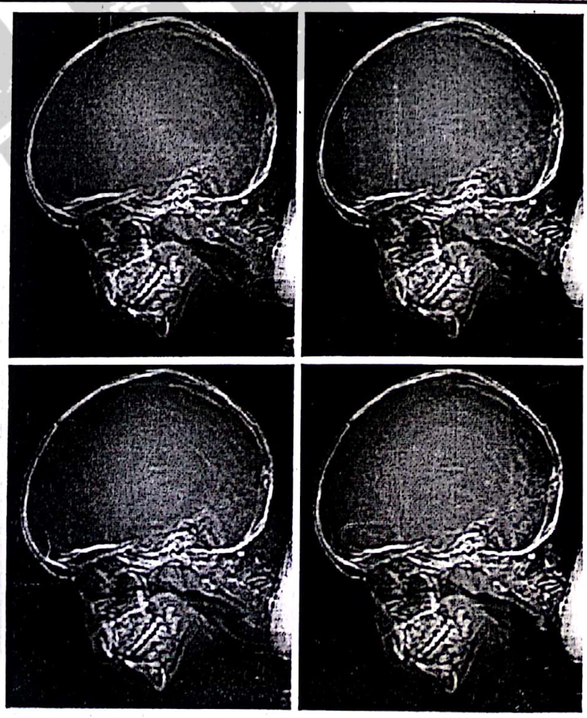
It refers to the smallest discernible change in intensity level. Based on hardware considerations, the number of intensity levels usually is an integer power of two.

Most common number is 8-bits, 16 bits & rarely 32 bits. In practice it refers to the number of bits used to quantize intensity as the intensity Resolution

eg: An image whose intensity is quantized into 256 levels has 8-bits of intensity resolution.

An 8 bit system quantizes intensity in fixed increments of 1/256 units of intensity amplitude.

lower intensity resolution results in false contouring & checker boards. Images of size 256X256 pixels with 64 intensity levels & printed on a size format of 5x5 cm have lowest acceptable spatial & intensity resolution.



a b
c d

FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

e f
g h

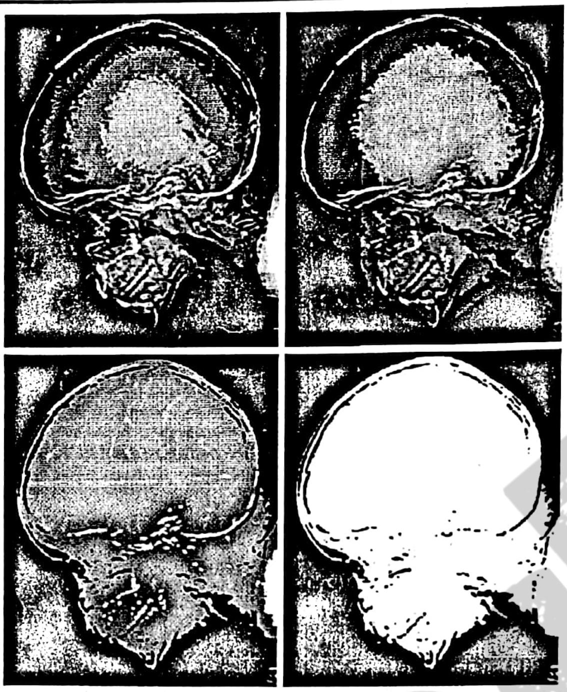


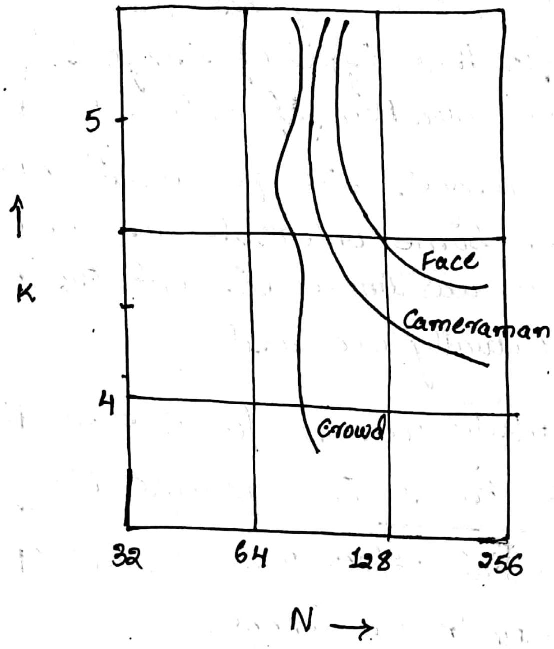
FIGURE 2.21 *(Continued)* (e)-(g) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

- Effects of Varying N & K Simultaneously.

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

The woman's face is representative of an image with relatively little detail; the picture of the Cameraman contains an intermediate amount of detail & the Crowd picture contains by comparison, a large amount of detail.



Isopreference Curve.

The Subjective quality of all the three images have been analysed & results were summerized in the form of "isopreference curves." in the NK-plane.

Each point in the NK plane represents an image having values of N & K equal to the co-ordinates of that point. Points lying on an isopreference curve correspond to images of equal subjective quality.

From the above isopreference curve we can observe that, The curves tend to shift up & right, it means larger values for N & K, which implies better picture quality.

Isopreference curves tend to become more vertical as the detail in the image increases. This result suggests that for images with a large amount of detail only a few intensity levels may be needed.

ie Crowd image requires few intensity levels, because it has large amount of detail. hence the corresponding isopreference curve is nearly vertical.

This indicates that, for a fixed value of N, the perceived quality for this type of image is nearly independent of the number of intensity levels used.

In other two images, the perceived quality remained same in some intervals in which the number of samples was increased. but the number of intensity levels actually decreased.

The most likely reason for this result is that a decrease in k tends to increase the apparent contrast, a visual effect that humans often perceive as improved quality in an image.

Some Basic relationship between Pixels:

In an image $f(x,y)$ let us consider an arbitrary pixel 'p' at co-ordinates (x,y), it has four horizontal & vertical neighbors whose co-ordinates are given by.

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

This set of pixels, called the 4-neighbors of 'p' denoted by $N_4(P)$

Each pixel is a unit distance from (x,y) $(x, y+1)$

Diagonal neighbors $[N_D(P)]$

$f(x-1,y) \leftarrow f(x,y) \rightarrow f(x+1,y)$

$f(x, y-1)$

The co-ordinates are,

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

8-neighbor : $[N_8(p)]$

The 4-neighbors together with diagonal neighbors are called 8-neighbors.

Note: Some of the neighbor locations in $N_D(p)$ & $N_8(p)$ fall outside the image, if (x,y) is on the border of the image. in Such Circumstances zero padding will comes into picture.

Adjacency, Connectivity, Regions & Boundaries.

Let V be the set of intensity values used to define adjacency. In a binary image, $V=\{1\}$ if we are referring to adjacency of pixels with value 1.

In case of gray scale image, intensity values ranging from 0 to 255, hence set V could be any subset of these 256 values.

There are 3 types of adjacency:

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

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

3) m-adjacency: (Mixed adjacency). Two pixels p & q with values from V are m-adjacent if.

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

b) q is in $N_D(p)$ & the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

Mixed adjacency is a modification of 8-adjacency. it eliminates the ~~ab~~ ambiguities that often arise when 8-adjacency is used.

Note: In m-adjacency, first priority is given to 4-neighbor,

Digital path

A digital path (Curve) from pixel 'p' with co-ordinates (x, y) to pixel 'q' with co-ordinates (s, t) is a sequence of distinct pixels with co-ordinates.

$(x_0, y_0), (x_1, y_1), ----, (x_n, y_n)$

where $(x_0, y_0) = (x, y)$ & $(x_n, y_n) = (s, t)$

& pixels (x_i, y_i) & (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$, where n is the length of the path.

* If $(x_0, y_0) = (x_n, y_n)$, the path is closed path.
* we can define 4, 8, ~~or~~ m-paths depending on the type of adjacency specified.

Let "S" represent a subset of pixels in an image. Two pixels p & q are said to be connected in "S" if there exist a path between them consisting entirely of pixels in 'S'.

For any pixel 'p' in 'S'. the set of pixels that are connected to it in S is called a Connected Component of 'S'. If it has only one connected component, then set 'S' is called a Connected Set.

Let "R" be a subset of pixels in an image. We call 'R' a <u>region</u> of the image if R is a connected set. Two regions R_i & R_j are said to be adjacent if their union forms a connected set.

Regions that are not adjacent are said to be <u>disjoint</u>.

Suppose that an image contains K disjoint regions R_k, $k=1,2,\ldots K$, none of which touches the image border. Let R_u denote the union of all the K regions, & let $(R_u)^c$ denote its complement.

We call all the points in R_u the foreground, & all the points in $(R_u)^c$ the background of the image.

<u>Boundary</u> (border or contour)

The boundary of a region R is the set of points that are adjacent to points in the complement of "R"

In other words, the border of a region is the set of pixels in the region that have atleast one background neighbor.

<u>Note</u>: If R happens to be an entire image, then its boundary is defined as the set of pixels in the first & last rows & columns of the image.

The boundary of a finite region forms a closed path & is thus a "global" concept. But edges are formed from pixels with derivative values that exceed a preset threshold. Thus, the idea of an edge is a "local" that is based on a measure of intensity level discontinuity at a point.

Distance measures.

For pixels p, q & z, with co-ordinates (x, y), (s, t) & (v, w) respectively, D is a distance function or metric if.

a) $D(p,q) \geq 0$ $\left(D(p,q) = 0 \text{ if } p=q\right)$

b) $D(p,q) = D(q, p)$

c) $D(p,z) \leq D(p,q) + D(q,z)$

Euclidean distance between p & q is defined as

$$D_e(p,q) = \left[(x-s)^2 + (y-t)^2\right]^{1/2}$$

For this distance measure, the pixels having a distance less than or equal to some value r from (x, y) are the points contained in a disk of radius r centered at (x,y).

D4 distance [City-block distance] between p & q is defined as,

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

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)

here $r=2$

for D_4 distance ≤ 2

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

The pixels with $D_4 = 1$ are the 4-neighbors of (x,y).

D_8 distance [Chess board distance] between p & q

is defined as.

$$D_8(p,q) = \max(|x-s|, |y-t|)$$

The pixels with D_8 distance from $(x,y) \leq r$ form a Square Centered at (x,y)

D_8 distance ≤ 2

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

The pixels with $D_8 = 1$ are the 8-neighbors of (x,y)

D_m distance : Note that the D_4 & D_8 distances between p & q are independent of any paths that might exist between the points. because these distances involve only the co-ordinates of the points. If we elect to m-adjacency. The D_m distance ~~between~~ two points is defined as the shortest m-path between points. D_m ~~will~~ depend on the values of the pixels along the path, as well as the values of their neighbor.

eg

$$\begin{matrix} & P_3 & P_4 \\ P_1 & P_2 & \\ P & & \end{matrix}$$

let $V = \{1\}$

if

* if $P_1 = P_3 = 0$, then D_m distance betn P & P_4 is $\underline{2}$ Path[$P\,P_2\,P_4$]
* if $P_1 = 1$, then $D_m = 3$, path $\rightarrow P P_1 P_2 P_4$
* if $P_3 = 1$ & $P_1 = 0$ then $D_m = 3$, path $\rightarrow P P_2 P_3 P_4$
* if $P_1 = P_3 = 1$, then $D_m = 4$, path $\rightarrow P P_1 P_2 P_3 P_4$

Arithmetic operations

Arithmetic operations between images are array operations.

$$s(x,y) = f(x,y) + g(x,y) \longrightarrow \text{Addition}$$
$$d(x,y) = f(x,y) - g(x,y) \longrightarrow \text{Subtraction}$$

$$P(x,y) = f(x,y) \times g(x,y) \longrightarrow \text{multiplication}$$

$$v(x,y) = f(x,y) \div g(x,y) \longrightarrow \text{division.}$$

where $x = 0, 1, 2, \ldots M-1$ & $y = 0, 1, 2, \ldots N-1$

Logical operations

when dealing with binary images, foreground (1) & background (0). Then if we define regions (objects) as being composed of foreground pixels,

while processing binary images, it is common to refer to union, intersection & complement as the OR, AND & NOT logical operations.

Consider two regions A & B composed of foreground pixels.

* The OR of these two regions is the set of elements (coordinates) belonging either to A or B or to Both.
* The AND operation is the set of elements that are common to A & B.
* The NOT operation of a set A is the set of elements not in A.