UNIT - I

DIGITAL IMAGE FUNDAMENTALS

Introduction

An image is a 2D representation of 3D scene. It is defined as a 2D function represented as $f(x,y)$ where x and y are spatial coordinates.

Digital image : Representation of the original image by a discrete set of data points.

Digital image processing : Processing digital images by means of a digital computer.

Pixel : denotes the elements of a digital image. These elements are called image elements/pels

Intensity / Brightness : Maps the number of pixels at each gray level present in the image.

In an 8 bit gray scale image, the pixel value ranges from 0 to 255. where 0 represents black and 255 for white.

eg: $\begin{bmatrix} 170 & 138 & 85 & 255 \\ 0 & 136 & 17 & 68 \\ 17 & 221 & 119 & 255 \\ 85 & 117 & 140 & 240 \end{bmatrix}$

Analog to digital conversion

Analog image → Sampling → Quantisation. → digital image

To create digital image, the continuous sensed data should be converted into digital form. This involves two processes.

i) Sampling

ii) Quantisation

Sampling

The sampling rate determines the spatial resolution of the digitized image, while the quantisation level determines the number of grey levels in the digitized image.

Quantisation

The transition between continuous values of the image function and its digital equivalent is called quantization.

Advantages of digital images

* Ability to enhance the quality of images.
* used in the field of medicine to improve the accuracy of diagnosis.
* The processing of images is faster and cost-effective.
* copying an image is easy.
* when the image is in digital form, reproduction of image is faster and cheaper.
* used in creative applications such as video editing, graphic design etc.

Disadvantages of digital images

* Complexity - requires lot of mathematical knowledge and expertise.
* The equipments and software used for processing images can be expensive.
* Time consuming process.
* The memory required to store and process good quality images is very high.
* Quality of the image may degrades, if not done properly.
* Ethics.

Q) Why do we need image processing?

* Improvement of pictorial information for human perception.

* Enhance the quality of image for a better look

Applications of image processing.

i) Medicine.

Techniques like image segmentation and pattern recognition is used in digital mammography to identify tumours.

ii) Remote sensing.

It uses remote observation to make useful inference about a target.

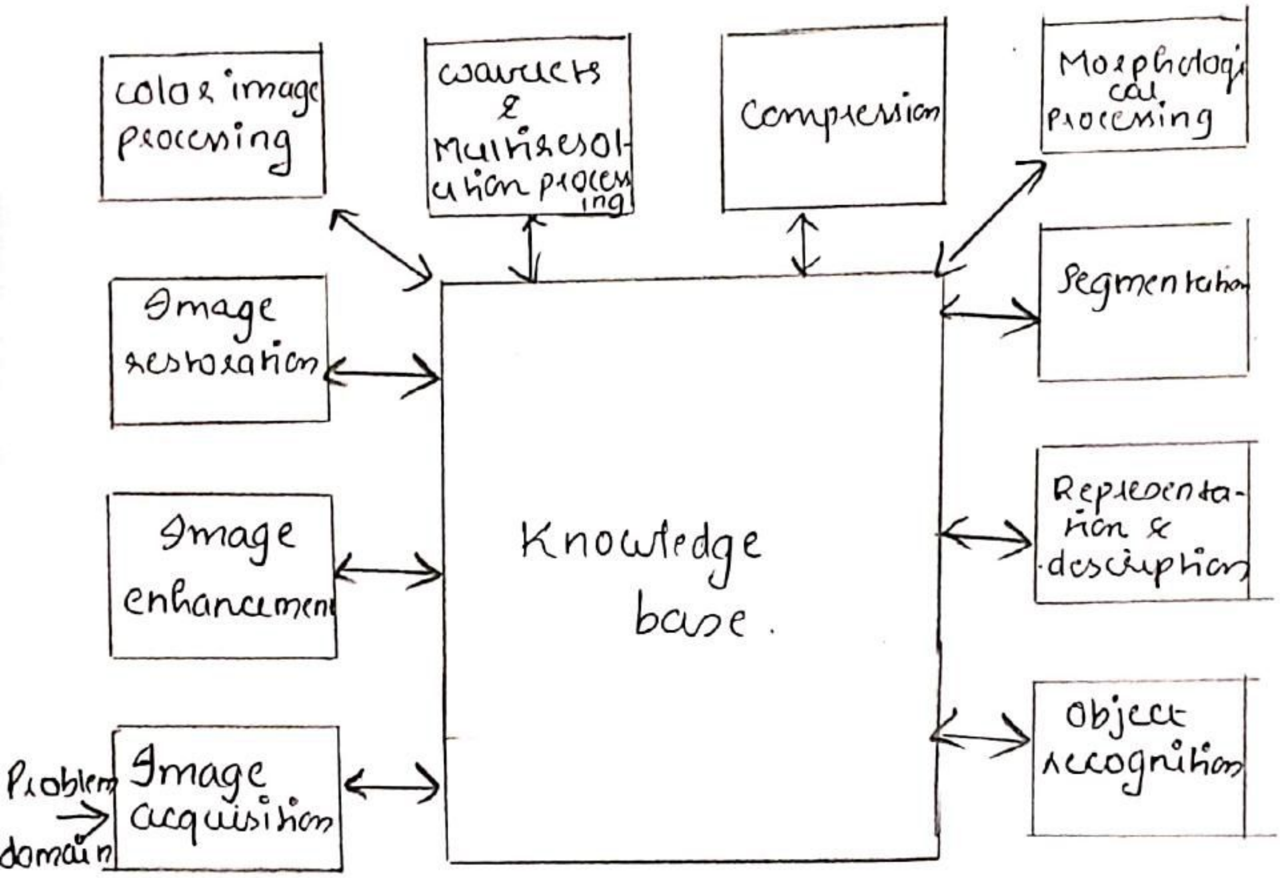
iii) Communications.

Advancement in the technology helps that information can be easily transmitted through the internet.

iv) Automatives

The latest development in the automative sector is "Night vision system". It helps to identify obstacles during night time to avoid accidents.

Fundamental steps in image processing



1) Image acquisition.

It is the first process in digital image processing. Acquisition could be as simple as being given an image that is already in digital form. Generally the acquisition stage involves preprocessing such as scaling.

2) Image enhancement.

After acquisition, based on the application it's required to enhance certain features present in the image.

Image enhancement can be classified into two categories.

i) Spatial domain

ii) frequency domain.

Spatial domain: It is a direct manipulation of pixel on the image

Frequency domain: Techniques based on modifying the fourier transform of an image.

3) Image restoration

Restore image that is not looking good due to some distortion (Noise). It is an objective process ie it deals with mathematical models.

4) Color image processing

It is the area that has been gaining importance because of the significant increase in the use of digital images. color is used as the basis for extracting features of interest in an image.

5) Wavelets and Multiresolution processing

Wavelets is the basis for representing images in various degrees of resolution. It is used for image data compression and for pyramidal representation.

6) Compression

It deals with the techniques for reducing the storage space required to save an image.

Also reduces bandwidth required for transmission of image.

7) Morphological processing

It deals with tools for extracting image components that are useful in the representation and description of shape. In this stages, the input and output are images.

8) Image segmentation.

In this stage, the images are converted into small segments, so that it's easy to extract more accurate information present in the image.

If the segments are properly classified i.e. Autonomous then the representation and description of image will be accurate. If it is rugged segmentation, then the result will not be accurate.

9) Representation and description.

It follows the output of a segmentation stage.

Representation makes a decision whether the data should be represented as a boundary or as a complete region. There are 2 types of representation

i) Boundary representation: focus on external shape characteristics such as corners and inflection.

ii) Regional representation: focus on internal properties such as textures/shape.

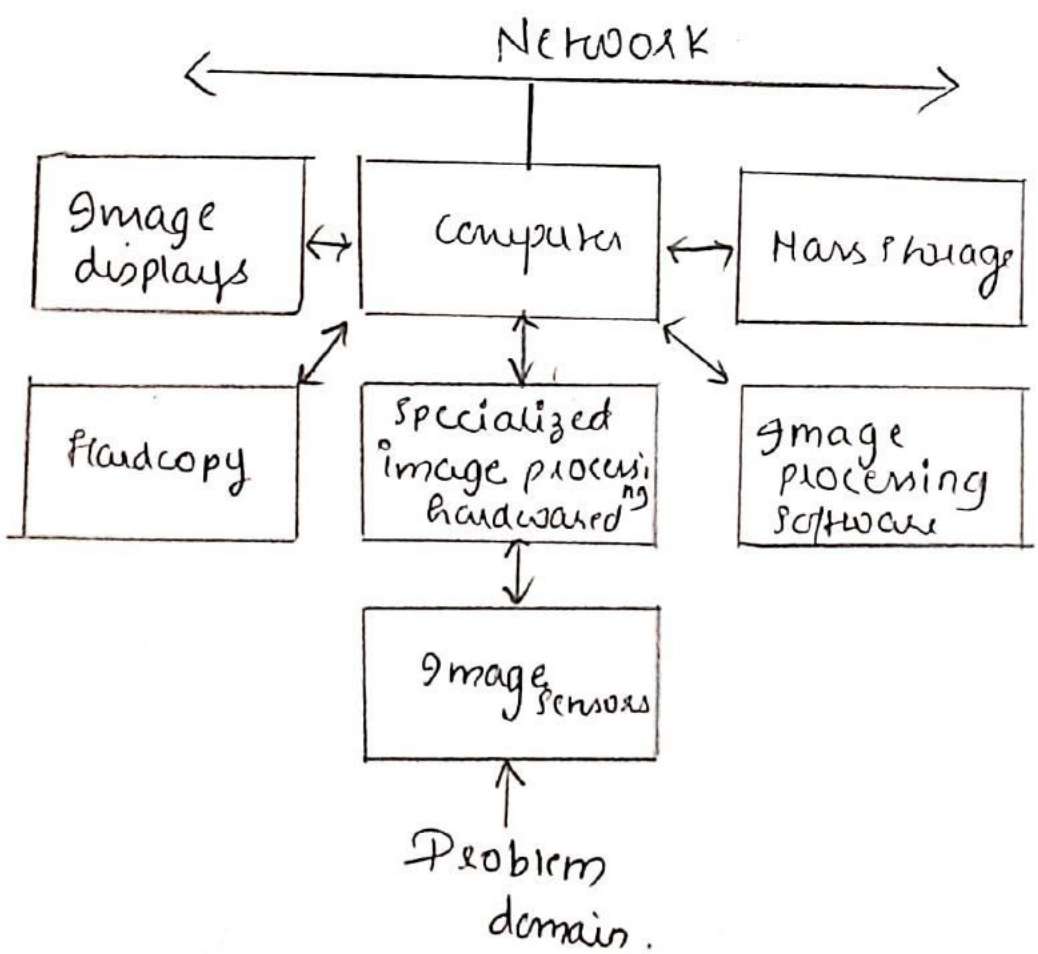
10) Recognition:

It is a process that acquires label to an object based on its descriptors.

11) Knowledge base:

Software that helps the user for proper image enhancement, restoration, segmentation or compression techniques.

Components of digital image processing

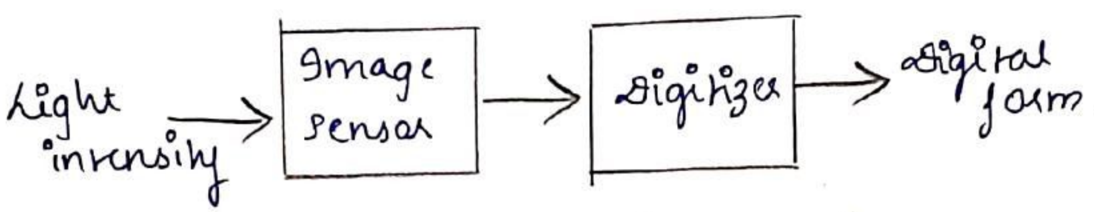


- Sensors

Two elements are required for sensing to acquire digital images.

* The physical device that is sensitive to the energy radiated by the object

* The digitizer is the ~~object~~ device used for converting the o/p into digital form.



Speciali image processing hardware

It consist of hardware and digitizer, which performs primitive operations such as ALU (Arithmetic logic unit) that performs arithmetic and logical operations. This type of systems are sometimes called front end subsystem.

Computer

It is a general purpose computer ranges from a PC to a supercomputer.

Software

It consist of specialised modules that performs specific tasks. It has the facility to write code using specialised modules. The image processing software uses programming language such as MATLAB.

Mass storage

It is a must in image processing applications

The digital storage for image processing application consist of 3 categories

i) Short-term storage - computer memory.

ii) on-line storage - Magnetic disks/optical media.

iii) Archival storage - Jukeboxes.

Storage is measured in bytes, Kb, Mb, Gb, Tb

Image displays

It is used mainly in color TV monitors. The outputs of image and graphics display cards are the integral part of the computer system.

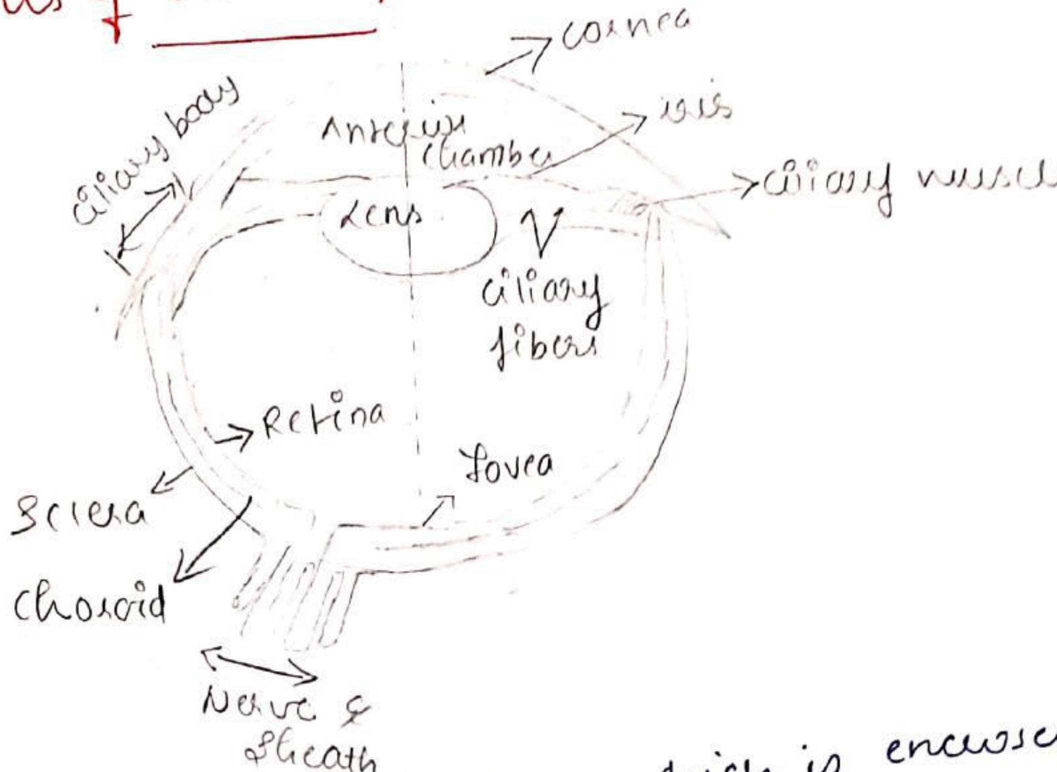
Hardcopy

It is used for recording images includes laser printers, film cameras, heat sensitive devices etc.

Networking

It is almost a default function in computer system. because of the large amount of data inherent in image processing applications.

Elements of visual perception



The eye is nearly a sphere, which is enclosed by 3 membranes : cornea and sclera
: choroid
: retina.

cornea and sclera

* The outer layer, which is tough, transparent tissue that covers the surface of the eye.
* sclera : opaque membrane that encloses the remaining part.

choroid

* It lies below the sclera
* contains a network of blood vessels that serves major nutrition to the eye.
* The choroid is divided into ciliary body and the iris diaphragm.

* The iris diaphragm contracts or expands to control the amount of light that enters the eye.

Lens

* It is made up of concentric layers of fibrous cells.
* It contains 60 - 70% of water, about 6% fat and more protein.
* The lens is colored by a slightly yellow pigmentation that increases with age.

Retina

* The inner most membrane.
* When the eye is properly focused, light from an object outside the eye is imaged on the retina.

Light receptors.

Pattern vision is afforded by the distribution of discrete light receptors. The receptors are classified into two categories

i) cones ii) Rods.

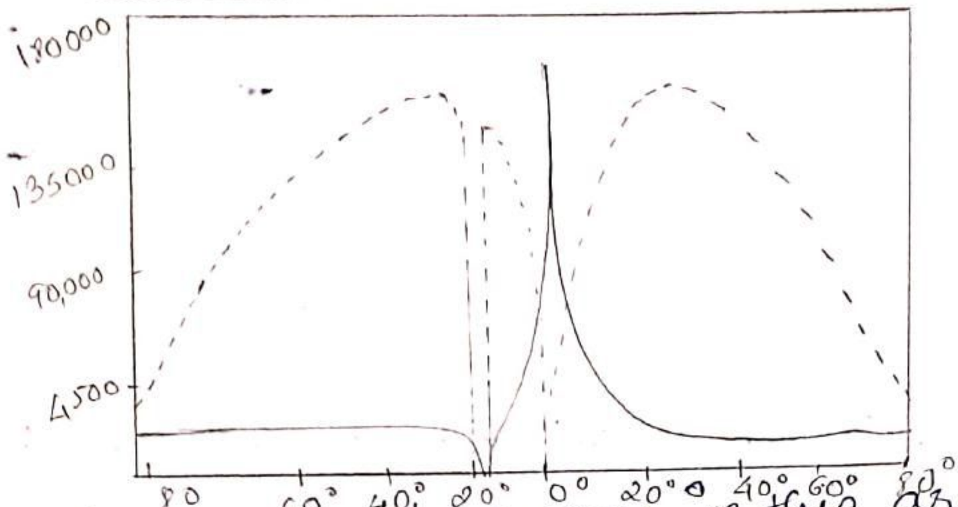
Cones

* Located primarily in the central portion of the retina called fovea.
* Humans can resolve finer details with these cones.
* Cone vision is called photopic / bright light vision

Rods

* Rods are larger in number and have larger area of distribution.
* They are not involved in any color vision and sensitive to low levels of illumination.

Rods and cones distribution.



The absence of receptors in this area is called blind spot.

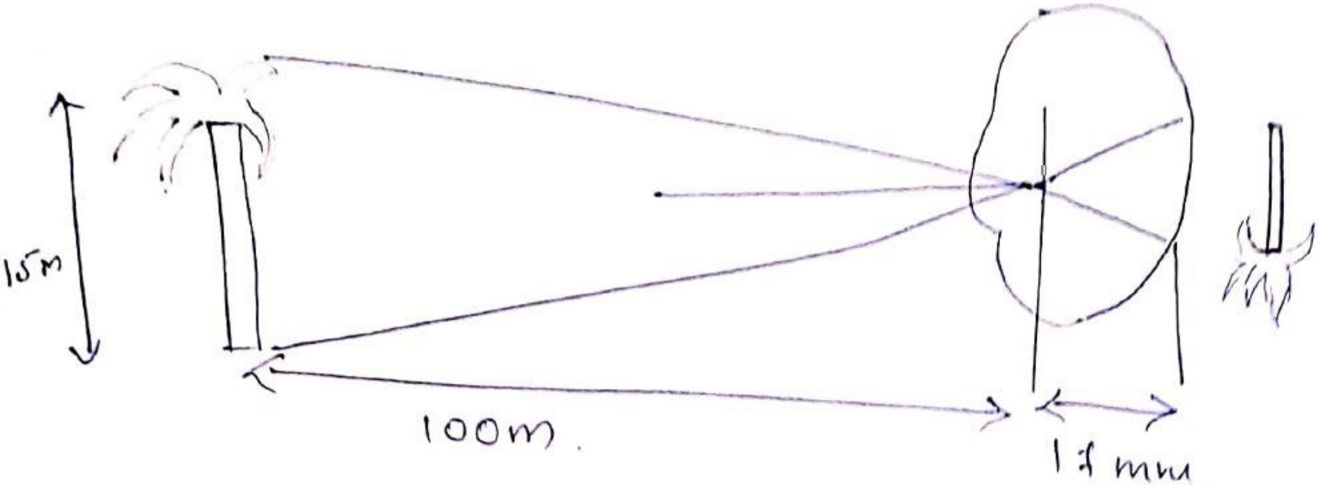
* Symmetrical distribution except the blind spot.
* Density is measured in degrees.
* Cones are most dense in the center of the retina whereas rods density decreases to the extreme periphery of the retina.

Image formation in the eye

* The principal difference between the lens of the eyes and an ordinary optical lens is former flexible.

* The shape of the lens is controlled by the tension in the fibers.

* The distance between the center of the lens and retina is called focal length. [Varies from 17mm to 14mm].



The height of the object in the retinal image is 'h' in mm

$$\frac{15}{100} = \frac{h}{17}$$

$$h = \frac{15 \times 17}{100} = 2.55mm$$

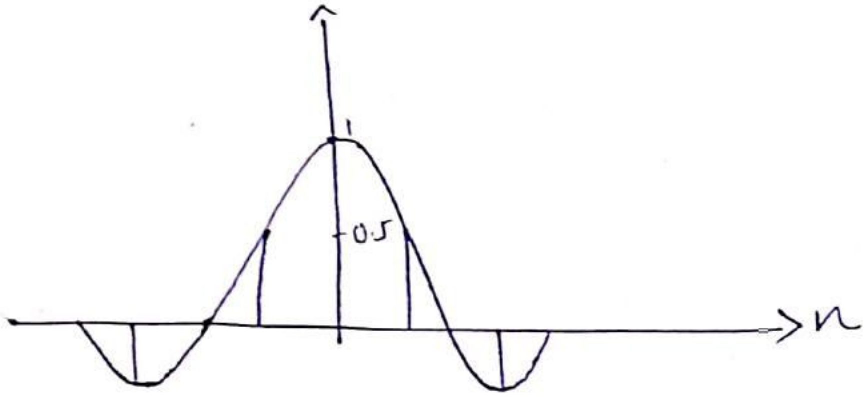
Mach band effect

* It describes an effect where the human brain subconsciously increases the contrast between two surfaces with different luminance.

* The visual appearance of the each strip is darker towards the left side than right.

* Machbanding is due to the lateral inhibition of the receptors in the eye.

* It can be used to estimate the impulse response of the visual system.



The negative lobes represents the visual phenomenon known as lateral inhibition.

Image Sensing and Acquisition

Image acquisition

* The process of capturing real world images and storing them into a computer.
* Films are not used in digital cameras.
* Instead, uses charge-coupled device | CMOS device as image sensors to convert light into electrical charges.

Image sensor

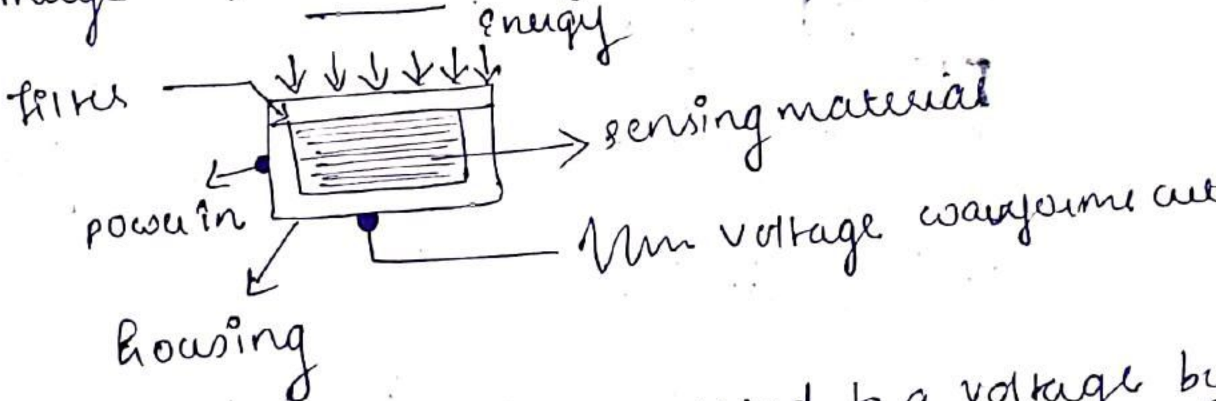
* It is a 2D array of light sensitive element that converts photons to electrons.
* All the digital camera uses either CCD or CMOS image sensor.
* A solid state image sensor consist of many discrete photo sensing element.
* The photo sensing element converts incoming photons into electrical charges and integrates these charges into a charges packet.
* The charge packet is then transferred through transport mechanism.
* The types of photo sensitive elements used in solid state images are photodiodes, MOS capacitors, Schottky barrier diodes and photoconductive layers.

Illumination and Reflection

* The images are generated by the combination of "illumination" and "reflection" of energy from the source.
* Depending upon the nature of the source, the illumination energy is reflected from or transmitted

- To transform illumination energy into digital images three sensor arrangements are used
    * Single Imaging sensor
    * Line sensor
    * Array sensor.

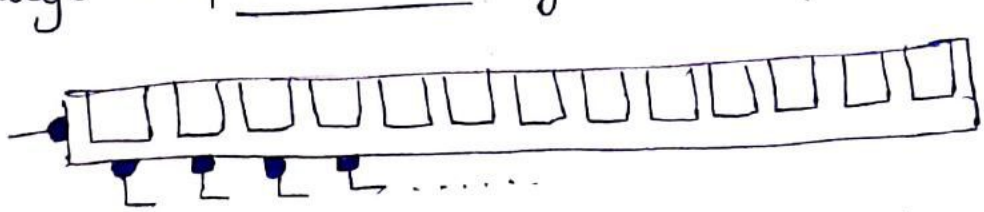
Image acquisition using single sensor



* Incoming energy is converted to a voltage by the combination of i/p electric power and sensor material.
* The response of the sensor is the o/p voltage waveform which has to be digitized.
* Filters are used to improve selectivity.

    eg: photodiode.

Image acquisition using sensor strips.

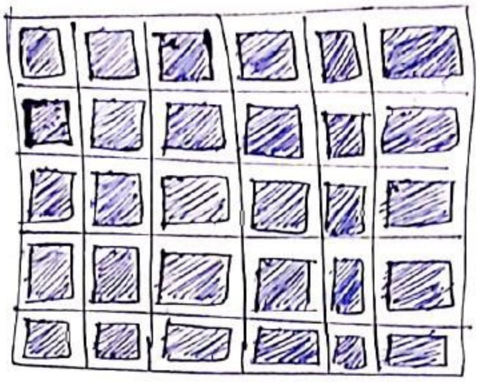


* Sensor strip has in line arrangement of sensors.
* Provides imaging elements in one direction.

* Motion perpendicular to the strip provides imaging in the other direction thereby completing the 2D image.

* Sensing devices with 4,000 or more in line sensors are possible.

Image acquisition using sensor arrays



Array sensors used in ~~~~ ordinary digital camera.

* Imaging system collects the incoming energy from an illumination source and focuses it onto a image.
* The front end of the imaging system is a lens.
* Then sensory array coincide with the focal plane produces output proportional to the intensity of light received at each sensor.
* The output is then digitized by another section of the imaging system.

- Simple image formation model

* Images are denoted using 2D functions of $f(x,y)$
* The value of a f is a +ve scalar quantity.
* $f(x,y)$ must be nonzero and finite.

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

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

→ Illumination component:

The amount of source illumination incident on the scene being viewed, denoted by $i(x,y)$.

→ Reflection component:

The amount of illumination reflected by the objects in the scene, denoted by $r(x,y)$.

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

where,

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

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

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

Sampling and Quantization

To create digital image, the continuous sensed data should be converted into digital form.

Digital images can be represented in two different ways.

Method 1

The complete M×N digital images can be written in a matrix form.

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

Each element of this matrix array is called an image element / picture element / pixel / pel.

Method 2

It is more advantageous to use a more traditional matrix notation to denote a digital image and its elements.

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

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

- The digitization process requires decision about values for M, N and L. The M and N should be positive integers. Due to processing, storage and sampling the no: of gray levels typically is an integer power of 2

$$L = 2^K$$

Dynamic range

The range of values spanned by the gray scale [0, L-1] is called dynamic range.

$$\text{Dynamic range} = \frac{\text{Maximum measurable intensity}}{\text{Minimum detectable intensity}}$$

* Upper limit is determined by saturation
* Lower limit is determined by noise.
* If the dynamic range is high, image will have high contrast.
* If the dynamic range is low, image will be dull.

No: of bits required, $b = M \times N \times K$

where M and N are no: of rows and columns.

If $M = N$

$$\boxed{b = N^2 K}$$

Spatial and intensity resolution

Spatial resolution

* Sampling is the principal factor for determining the spatial resolution.
* It measures the detailed information present in the image.
* It is also depends upon the no: of pixels.

Intensity resolution

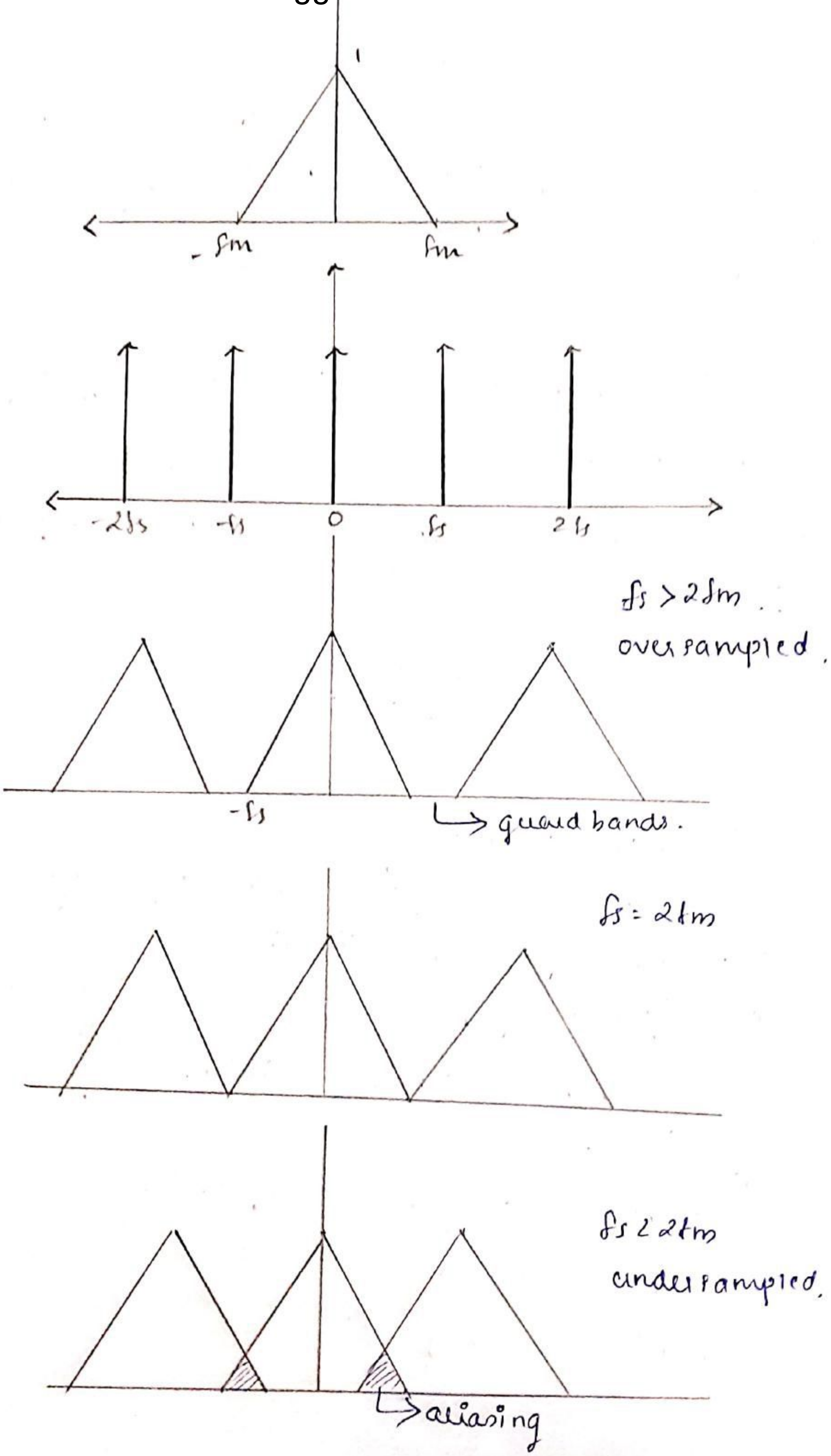
* It measures the smallest discernible change in intensity.
* Depends upon the no: of gray levels.
* Highly subjective process.

Aliasing and Moire patterns

Shannon sampling theorem states that the signal should be sampled at a rate equal to or greater than twice its frequency.

$$f_s \geq 2f_m$$

where f_s is the sampled frequency and f_m is the maximum frequency of the signal.


1
-fm
fm
-2fs
-fs
0
fs
2fs
fs > 2fm
oversampled.
-fs
guard bands.
fs = 2fm
fs < 2fm
undersampled.
aliasing

How to reduce aliasing?

To reduce the aliasing effect on image is to reduce the high frequency components by blurring the image prior to sampling.

Moire patterns

* Moire patterns are used to view the effect of aliased frequencies on a sampled image.
* Occurs while violating the Shannon sampling theorem.

Zooming and shrinking of digital images

Zooming

* Viewed as oversampling.

There are two steps needed for zooming, they are

i) creation of new pixel locations.
ii) Assignment of gray levels to those new locations.

Zooming can be implemented ~~~~ using three different methods.

→ Nearest neighbor interpolation
→ Pixel Replication
→ Bilinear interpolation.

→ Nearest neighbor interpolation

* Basic tool used in tasks as zooming, shrinking, rotating and geometric corrections.

* Interpolation is the process of using known data to estimate values of unknown locations.

Steps:

* Consider an image of size 500 * 500 pixels and we want to enlarge it to 1.5 times 750 * 750 pixels.

* One of the easiest way to visualize zooming is to create an imaginary 750 * 750 grid over the original image and then shrink it so that it fits exactly over a smaller image.

* Compare the grid to the original image.

* Gray level assignment is performed by assigning the gray level of the closest pixel in the original image to the new pixel in the grid.

→ Pixel Replication

* Special case of nearest neighbor interpolation.

* Applicable when the size of an image is to be increased to an integer number of times like double, triple, quadruple etc.

* To duplicate the size of an image, we can duplicate each column. This doubles the image size in the horizontal direction.

* Based on the duplication of pixels for the required number of times to achieve the desired size.

→ Bilinear Interpolation

The four nearest neighbors to estimate the intensity.

- Bilinear transformation is the most preferred method of gray level assignment.
- This method overcomes the drawback of nearest neighbor interpolation.
- It uses four nearest neighbours of a point.

Let (x', y') represents the co-ordinates of a point in the zoomed image and $v(x', y')$ is the gray level assigned to that image.

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

where a, b, c, d coefficients using the 4 nearest neighbor of points x', y':

Shrinking

It can be performed by two methods.

method 1

Row-column deletion method - the required size of the image is obtained by deleting the additional rows and columns.

Method 2

- This is similar to grid analogy for zooming.

* Keep an imaginary grip of required size over the original image.
* Expand the grid to fix to fit over the original image.
* Perform gray level assignment.
* Shrink the grid back to its original size which gives the shrinked image of required size.

Relationships between pixels.

Neighbors of a pixel

* 4 - Neighbors, $N_4(P)$
* Diagonal Neighbors, $N_D(P)$
* 8- Neighbors, $N_8(P)$

$N_4(P)$

* Any pixel $P(x,y)$ has 2 vertical and 2 horizontal neighbors given by $(x+1,y)$, $(x-1,y)$, $(x,y+1)$, $(x,y-1)$.
* This set of pixels are called the 4 neighbors of P, and is denoted by $N_4(P)$.
* Each of them is at a unit distance from P.

$N_D(P)$

	P	
x		x

* This set of pixels, called 4 neighbors and denoted by $N_D(P)$.
* Four diagonal neighbors of P have coordinates.

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

* Each of them are at euclidean distance of 1.414 from P.

$N_8(P)$

* $N_4(P)$ and $N_D(P)$ together are called 8 neighbors of P, denoted by $N_8(P)$.

x	x	x
x	P	x
x	x	x

* $N_8 = N_4 \cup N_D$.

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

Adjacency

Two pixels are connected if they are neighbors and their gray levels satisfy some specified criterion of similarity.

Types of adjacency

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

* 8-adjacency: Two pixels p and q with values from V are 8-adjacent if q is in the set N8(P).

* m-adjacency (mixed adjacency).

Two pixels p and q with values from V are m-adjacent if:

→ q is in N4(P)

OR

→ q is in ND(P) and the set N4(P) ∩ N4(q) has no pixel whose values are from V.

Connectivity

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

Region and Boundary

Let R be the region of the image connected to a set. If R_1 and R_2 are said to be adjacent if their union forms a connected set. If the regions are not adjacent, they are said to be disjoint.

The boundary (border/contour) of a region R is the set of point in the region that have one or more neighbors that are not in R.

Distance measures

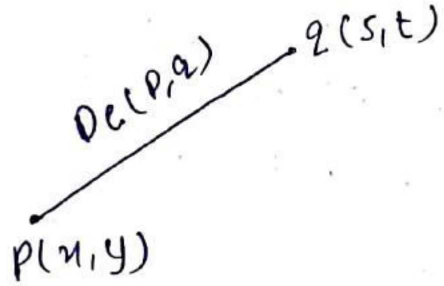
For pixels P, q and z with coordinates (x, y), (s, t) and (v, w) respectively, D is a distance function

Properties

* $D(P,q) \geq 0$ $[D(P,q) = 0,$ if $P = q]$
* $D(P,q) = D(q,P)$
* $D(P,z) \leq D(P,q) + D(q,z)$

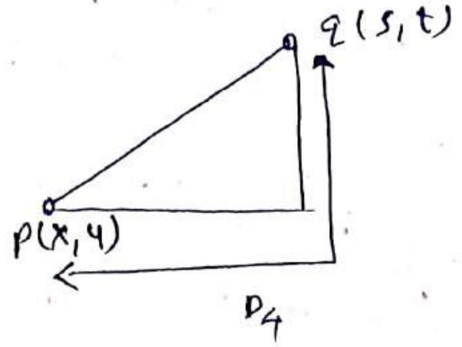
The euclidean distance b/w P and q is defined as

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



The city block distance b/w P and q is defined as

$$\boxed{D_4(P,q) = |x-s| + |y-t|}$$



Color image processing

* useful in object detection, extraction, segmentation etc.
* Human eye can easily distinguish shades of color than shades of grey.

Basics

* Light is an electromagnetic wave having a wide range of frequencies but we can see only the visible spectrum [400nm to 700nm].

Radiance: It is the total amount of energy that flows from the light source, measured in watts (W).

Luminance: Measures the amount of energy an observer perceives from a light source, measured in lumens (lm).

Brightness: It is a subjective nature that varies from human to human.

Primary colors are

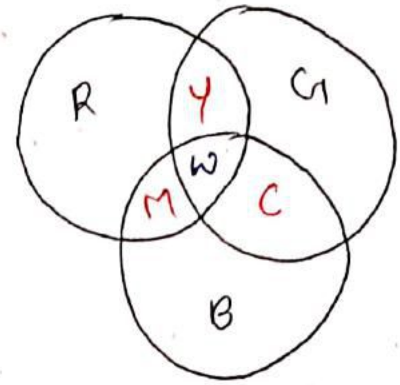
Red (R), Green (G), Blue (B)

Secondary colors are:

G + B = cyan.

R + G = Yellow

R + B = Magenta.



Chromaticity

Hue + saturation = chromaticity

* The chromaticity diagram represents the position of different colors in an image.
* The amount of red, green and blue needed to form any particular color are called the tristimulus.

A color is then specified by its trichromatic coefficients defined as

$$x = \frac{X}{X+Y+Z} = \frac{R}{R+G+B}.$$

$$y = \frac{Y}{X+Y+Z} = \frac{G}{R+G+B}$$

$$z = \frac{Z}{X+Y+Z} = \frac{B}{R+G+B}.$$

$$x + y + z = 1$$

Color Models

* The purpose of color model is to faciliate the specification of colours in some standards.
* Also known as color space / color system.
* Provides a standard way to specify particular color in defining a 3D coordinate system.

In terms of DIP, the hardware oriented models are classified into three types.

→ RGB [Red, Green, Blue]

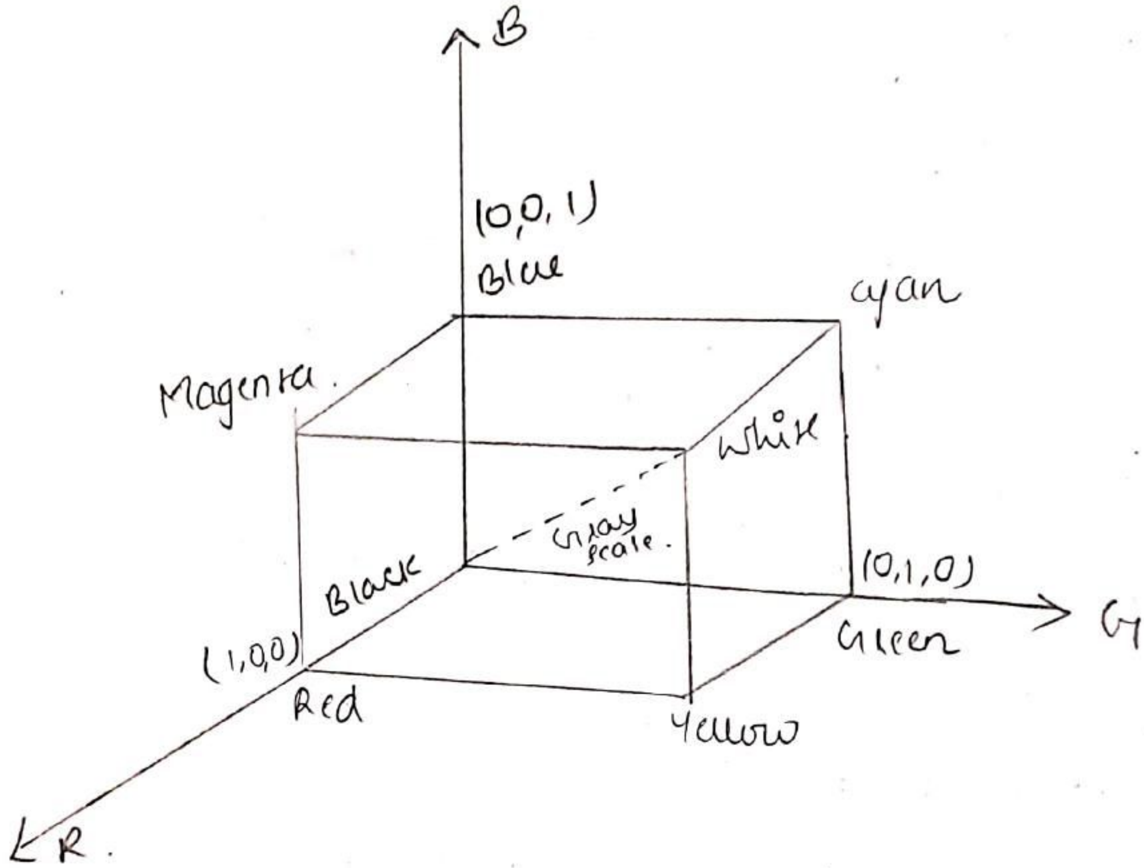
→ CMY [Cyan, Magenta, Yellow]

→ HSI [Hue, Saturation, Intensity]

RGB color model

* Additive color model, with the combination of primary colours to produce secondary colours such as cyan, Magenta & Yellow.
* This model is based on cartesian coordinate system
* The color subspace is the cube in which RGB values are in 3 corners and their secondary colours are in other three corner, black is at the origin and white is the corner farthest from the

origin.



(R, G, B) ⟶ (8 bits, 8 bits, 8 bits)

The no: of bits used to represent each pixel in RGB space is called pixel depth.

Applications:

* color monitors
* Color video cameras.

Advantages

* It is an easy process and therefore it is an ideal tool for image color generation.
* Changing to other models such as CMY is straight forward.

Limitations

* It is device dependent.
* It is not perceptually uniform.
* It is difficult to relate this model to color appearance because its basis is to device signals and not display luminance values.

HSI color model

H - Hue - attributes that describe a pure color

S - Saturation - purity of color

I - Intensity - Achromatic notation of intensity.

To find intensity:

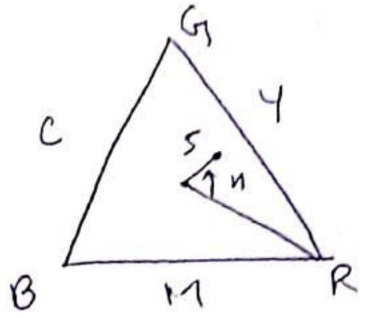
* It is a most useful descriptor of monochromatic images.
* HSI model decouples the intensity component from the color carrying information in a color image.
* Thus HSI model is an ideal tool for developing image processing algorithms based on color descriptions that are natural to humans.

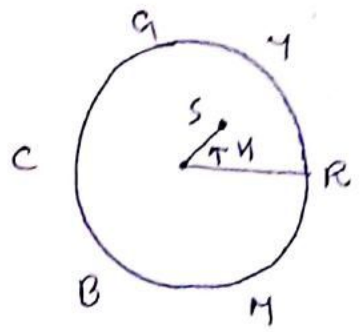
To find saturation:

The saturation of a color increases as a function of distance from the intensity axis.

To find hue:

For a given RGB point, hue can be determined.





Converting colors from RGB to HSI

Hue,

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

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

Saturation,

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

Intensity,

$$I = \frac{1}{3}\left[R+G+B\right]$$

Converting color from HSI to RGB

Hue to its original range of $[0°, 360°]$.

a) Red - Green [RG] , $0° \le H < 120°$

b) Green - Blue [GB] , $120° \le H < 240°$

c) Blue - Red (BR), $240° \le H \le 360°$

RG sector

$$B = I[1-S]$$

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

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

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

GB sector

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

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

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

BR sector

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

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

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

Manipulating HSI components images

a) To change the Hue or Individual color.

- First hue values in the region to be modified are changed.
- The new H image along with the unchanged S and I images is converted back into RGB model.

- The resulting RGB image will have the modified color in the desired region.

b) To change the saturation.

- The saturation value of the desired region is changed in the S component image and H and I images are kept unchanged.
- The new S image and unchanged H and I images are converted back into RGB space.

c) To change average intensity.

- The target region's intensity is changed in the I image where H and S images are unchanged.
- Then the H, S and new I images are converted into RGB ~~spacing~~ ~~colour~~ space.

Two-Dimensional Fourier Transform

The 2D Fourier transform of $f(x,y)$ is represented as $F(u,v)$

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

Inverse

$$f(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int F(u,v)\, e^{j2\pi(ux+vy)}\, du\,dv.$$

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

Amplitude Spectrum.

$$\boxed{|F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)}}$$

- phase spectrum,

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

Power spectrum,

$$\boxed{P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v)}$$

2D Transform - DFT

If $f(x,y)$ is an $M \times N$ array, the DFT is

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

where $u = 0, 1, 2, \ldots M-1$
$v = 0, 1, \ldots N-1$

Inverse DFT ie IDFT is

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

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

If the images are sampled in a square array,

$M = N$.

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

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

Properties of DFT

* Spatial frequencies
* uniqueness
* Separability.
* Rotation

$$f(\pm x, y) \longrightarrow F(\pm u, v)$$

* Linearity

$$a_1 f_1(x,y) + a_2 f_2(x,y) \longrightarrow a_1 F_1(u,v) + a_2 F_2(u,v)$$

* Conjugation

$$f^*(x,y) \rightarrow F^*(-u,-v)$$

* Seperability

$$f_1(x) f_2(y) \longrightarrow F_1(u) \cdot F_2(v)$$

* Scaling

$$f(ax, by) \longrightarrow \frac{F(u/a, v/b)}{|ab|}$$

* Shifting

$$f(x \pm x_0, y + y_0) \longrightarrow F(u,v) \cdot e^{\pm j 2\pi (x_0 u + y_0 v)}$$

Need for transform

* For mathematical convenience
* To extract more relevant information.

Discrete Cosine Transform (DCT)

This transform is very similar to fourier transform. The DCT is defined as,

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

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

The inverse DCT is defined as,

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

Properties

* Real transform
* It has excellent energy compaction properties
* There are fast algorithm to compute DCT.
* It has been widely used in image compression standards.

UNIT - II

IMAGE ENHANCEMENT

Image ~~en~~ enhancement refers to sharpening of image features such as edges, boundaries / contrast.

Examples of image enhancement:

* Noise reduction
* Graylevel and contrast manipulation.
* Filtering
* Pseudocoloring.

Image enhancement can be classified into two categories.

* Spatial domain methods
* Frequency domain methods.

Spatial Domain

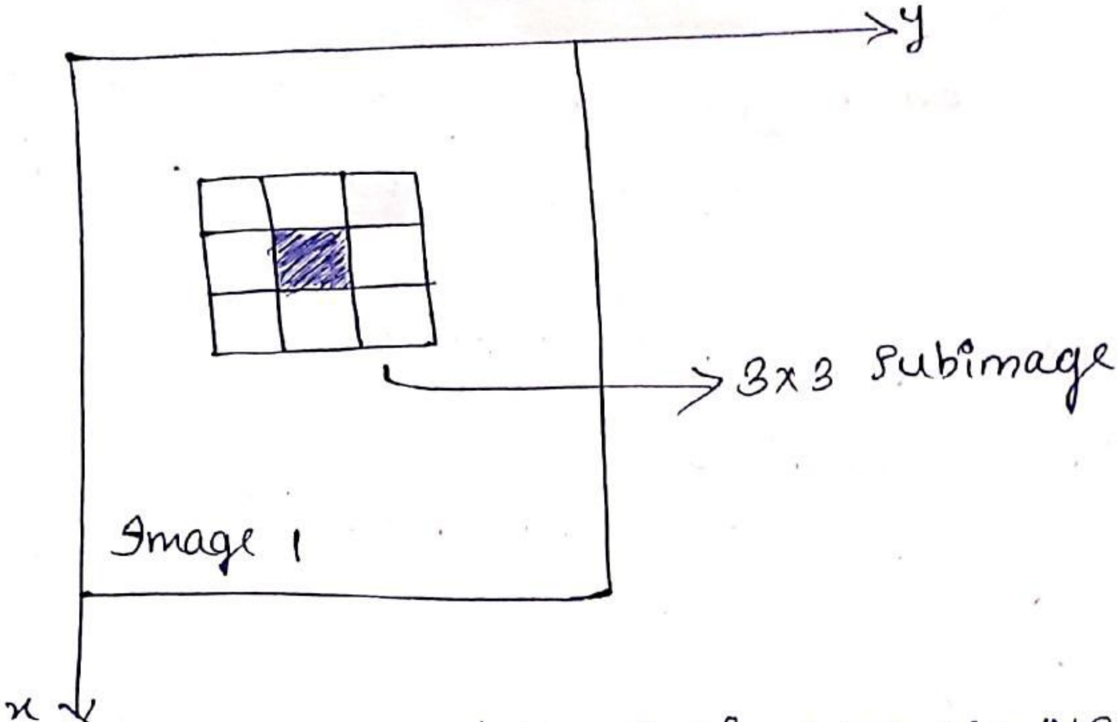
* Operates directly on the pixels.
* More efficient communication and requires less processing resources to implement.

The spatial domain will be denoted by

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

where, $f(x,y)$: Input image and

$g(x,y)$ = processed image.



The centre of the subimages are moved pixel by pixel in the image. Then the operator T is applied for each pixel to obtain the output g.

* These subimages are also known as 'filter', 'mask', 'Kernel', or 'window'.

There are two types of spatial filtering they are

i) Linear filtering

ii) Non linear filtering.

Linear filtering

For spatial filtering, the ~~spatial~~ response is given by a sum of products of the filter coefficients and the corresponding image pixels in the area spanned by the filter mask.

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

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

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

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

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

$$\boxed{R = w^T z}$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Non linear spatial filtering

* Same as that of linear spatial filtering which uses ~~filtering~~ filter bank
* Doesn't use the coefficients in the sum of product form.
* The main application is noise reduction.

Gray level Transformation function, $T(r)$

The transformation function,

$$s = T(r).$$

Based on T(r) form, there are two techniques which are used for contrast enhancement. They are

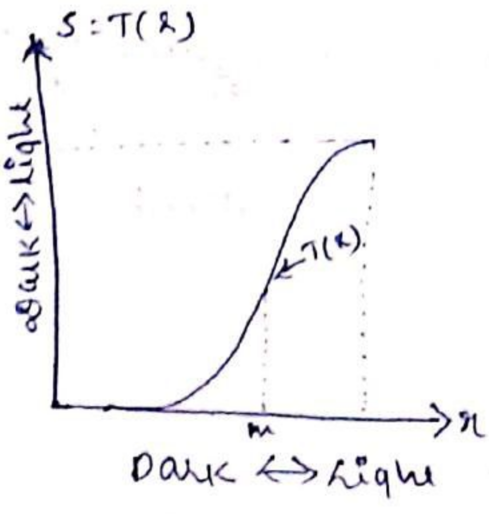
* Contrast stretching.
* Thresholding.

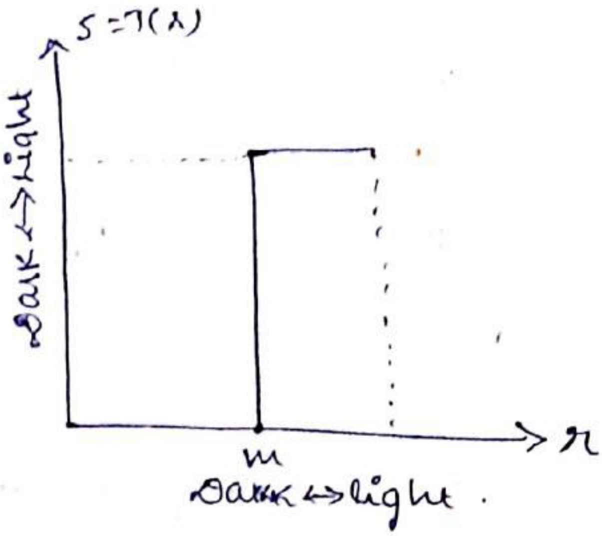
Contrast Stretching

The effect of this transformation to produce an image of higher contrast than the original by darkening the levels below m and brightening the level above m are compressed by the transformation function in the original image. This technique is known as contrast stretching.

Thresholding

* T(r) produces a two-level (binary) image. A mapping of this form is called a thresholding function.





Mask processing / filtering

The values of the mask coefficients determine the nature of the process, such as image sharpening. The enhancement technique based on this type of approach are called mask processing / filtering.

Gray level transformations

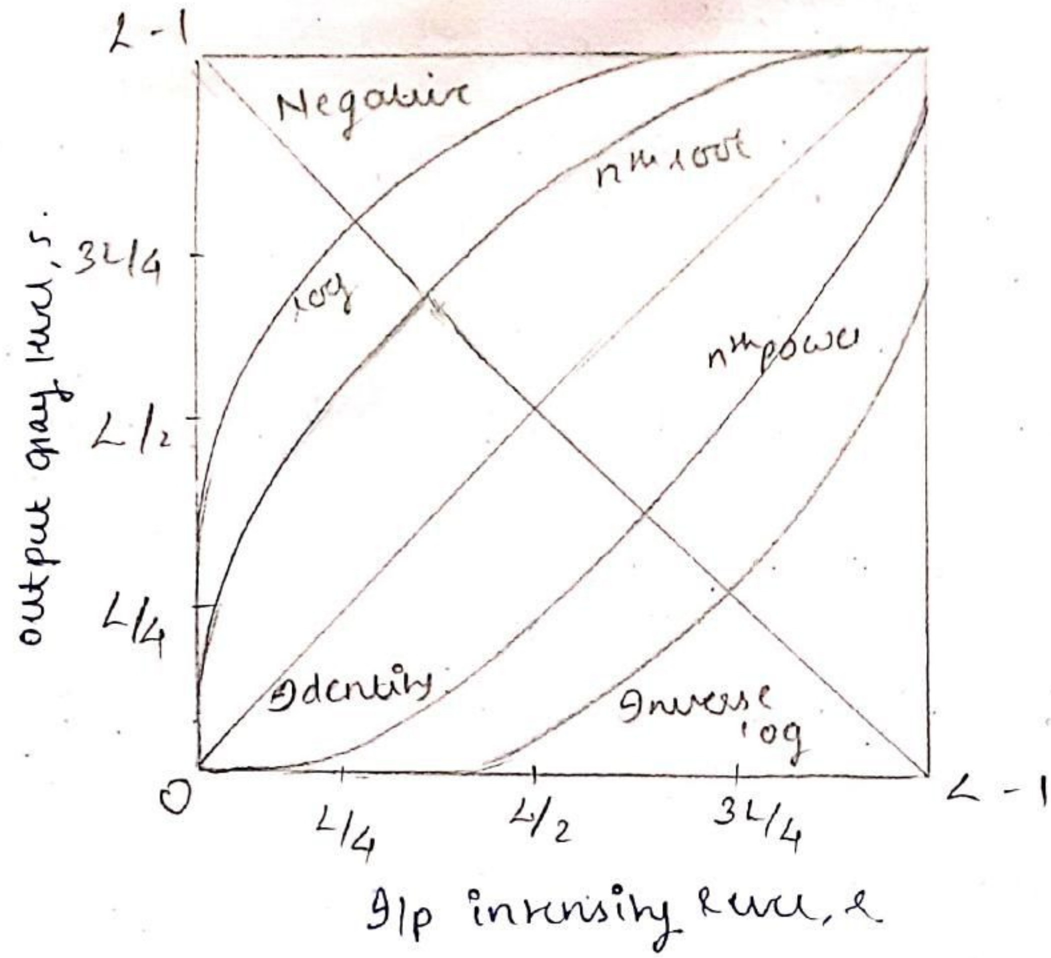
The pixel values r and s are related by

$$s = T(r)$$

where T is a transformation that maps a pixel value r into a pixel value s.

There are 3 basic types of functions that are used frequently used in image enhancement.

- Linear functions.
  - Negative transformation.
  - Identity transformation.
- Logarithmic functions.
  - Log transformation.
  - Inverse log transformation.
- Power law function.
  - n^{th} power transformation.
  - n^{th} root transformation.



Negative Transformations.

* Image negatives.
* The inverse transformation reverse light and dark.
* ~~Image~~ A negative image is obtained by subtracting each pixel from the maximum pixel value.

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

Identity function

* Output intensities are identical to input intensities.
* This function doesn't have an effect on an image, the graph only for completeness.

$$\boxed{s = r}$$

Log transformations.

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

where c is a constant and it is assumed that $r \geq 0$.

* The opposite of this applies for inverse log transform

* This transform is used to expand values of dark pixels and compress values of bright pixels.

Power law (Gamma) Transformation

The nth power and nth root is defined as

$$\boxed{s = cr^{\gamma}}$$

where c and γ are positive constants.

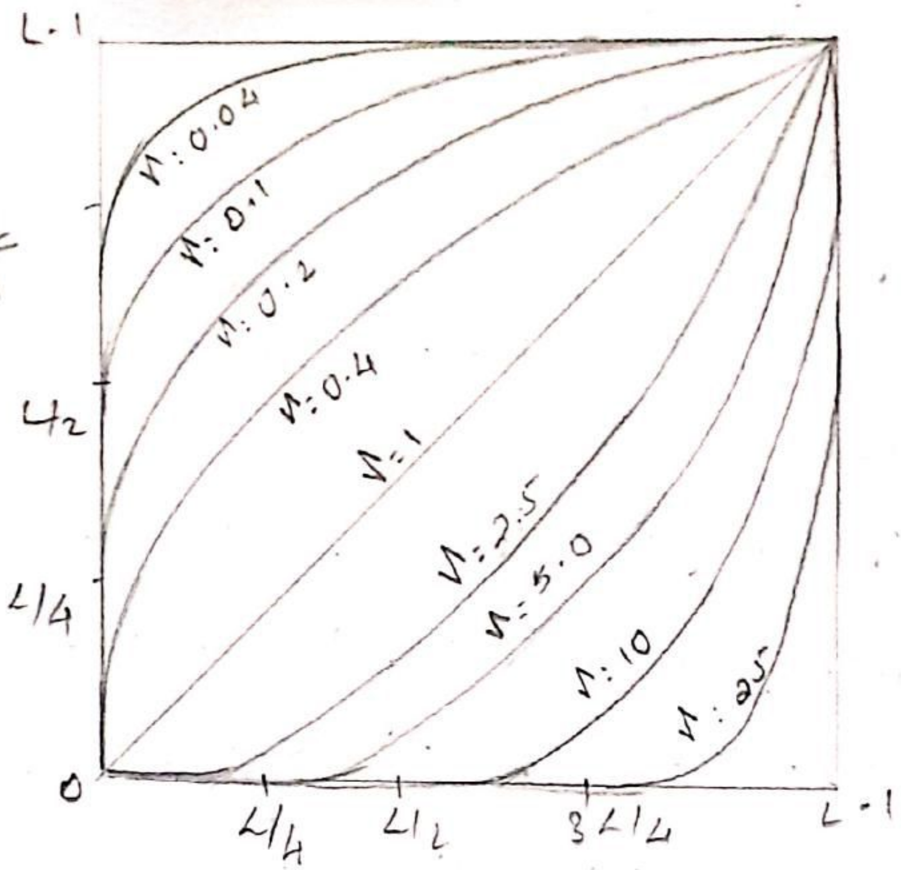
For an offset

$$\boxed{s = c[r+\varepsilon]^{\gamma}}$$

This transformation is called gamma correction.

For an identity transformation.

$$\boxed{c = \gamma = 1}$$



Piecewise linear transformation function

Inorder to use the mathematical function, some defined transforms can be used.

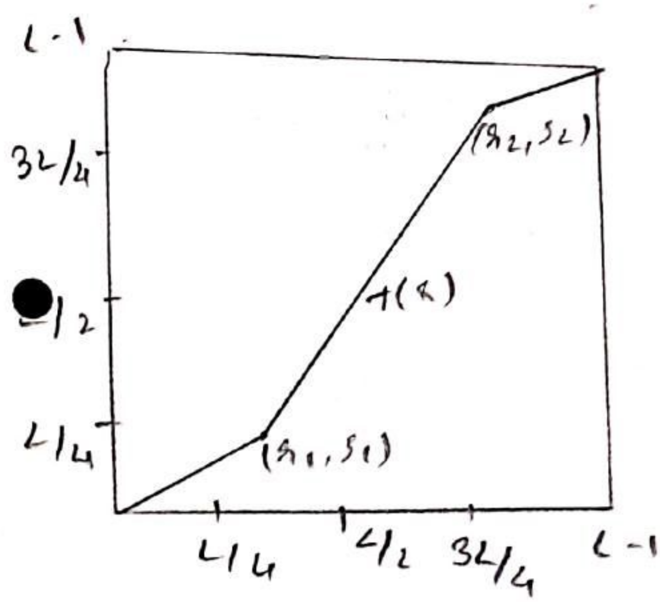
* The main advantage of this is some important transformations can be formulated using piecewise linear transformation function.
* The disadvantage behind this is, it requires more user input than previously used transformation.

The piecewise transformation can be classified into 3 types.

→ Contrast stretching
→ Gray level slicing
→ Bit-plane slicing.

Contrast stretching

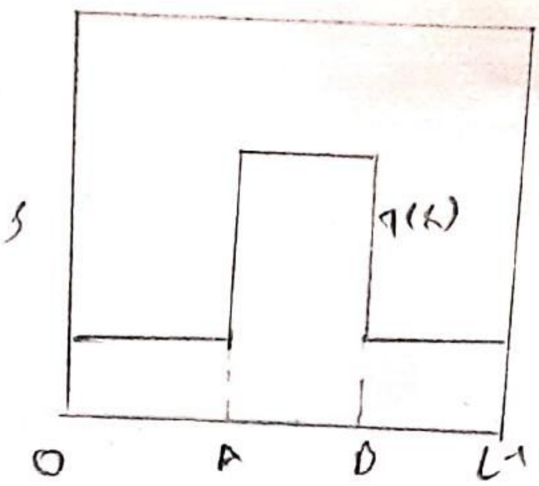
* One of the simplest piecewise linear function.
* It is used to enhance the low contrast images.
* Low contrast images results:
  → poor illumination.
  → ~~wrong~~ setting of lens aperture during image acquisition.



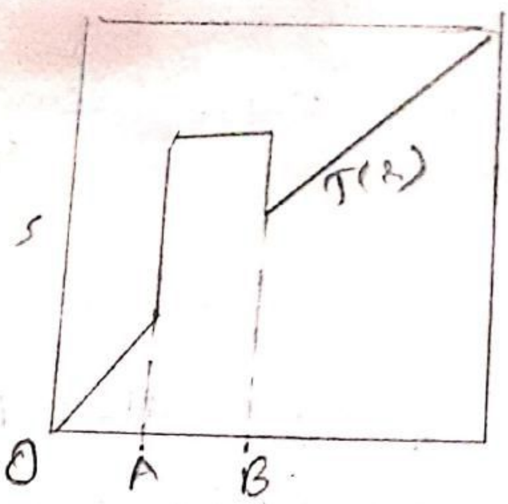
* If $r_1 = s_1$ and $r_2 = s_2$, the transformation is a linear function that produces changes in gray levels.
* If $r_1 = r_2$, $s_1 = 0$ and $s_2 = L-1$ the transformation becomes a thresholding function.

Intensity level slicing

* It is used to highlight a specific range of gray levels in a given image.



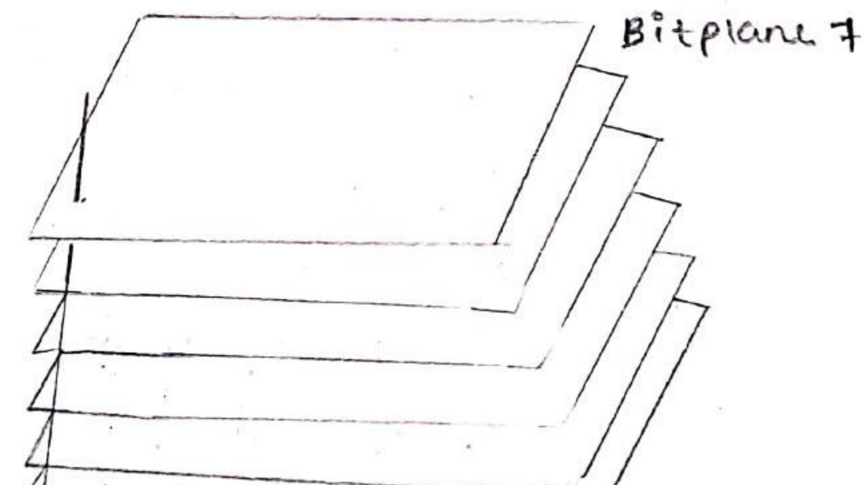
Binary gray level slicing



Gray level slicing - 2nd approach

Bit plane slicing

The intensity of each pixel in a 256 gray scale image is composed of 8 bits.



bitplane representation.

* Converting a gray level image to binary image
* Representing an image with fewer bits and compressing.

- Enhancing the image by focusing.

*

Histogram Processing

* The histogram of a digital image with intensity levels in the range $(0, L-1)$ is a discrete function $h(r_K) = n_K$, where r_K is the K^{th} intensity value and n_K is the no: of pixels in the image with intensity r_K.

* Normalise a histogram by dividing each of its components by the total no: of pixels in the image denoted by the product MN. where M and N are the rows and column dimensions of the image.

* A normalised histogram is given by

$$P(r_K) = \frac{n_K}{MN}, \text{ for } K = 0,1,2, \dots L-1$$

n_K : no: of pixels in the K^{th} gray level.

* Histogram manipulation can be used for image enhancement.

* Histogram processing for intensity transforms 4 basic intensity characteristics

dark, light, low contrast & high contrast.

* The horizontal axis of each histogram.

Image Histogram

* Image histogram is a type of histogram that acts as a graphical representation of the of the tonal distribution in a digital image.
* plots the number of pixels for each tonal value.
* Horizontal axis of graph represents tonal variations and vertical axis represents the number of pixels in that particular tone.
* Histogram of an image represents relative frequency of occurence of various gray level.

a) Histogram equalization / Histogram Linearization

Histogram equalization is a process of automatically determining a transformation function which produces an output image with ~~an~~ a uniform histogram.

Histogram equalization is a process for increasing the contrast in an image by spreading the histogram out to be approximately uniformly distributed.

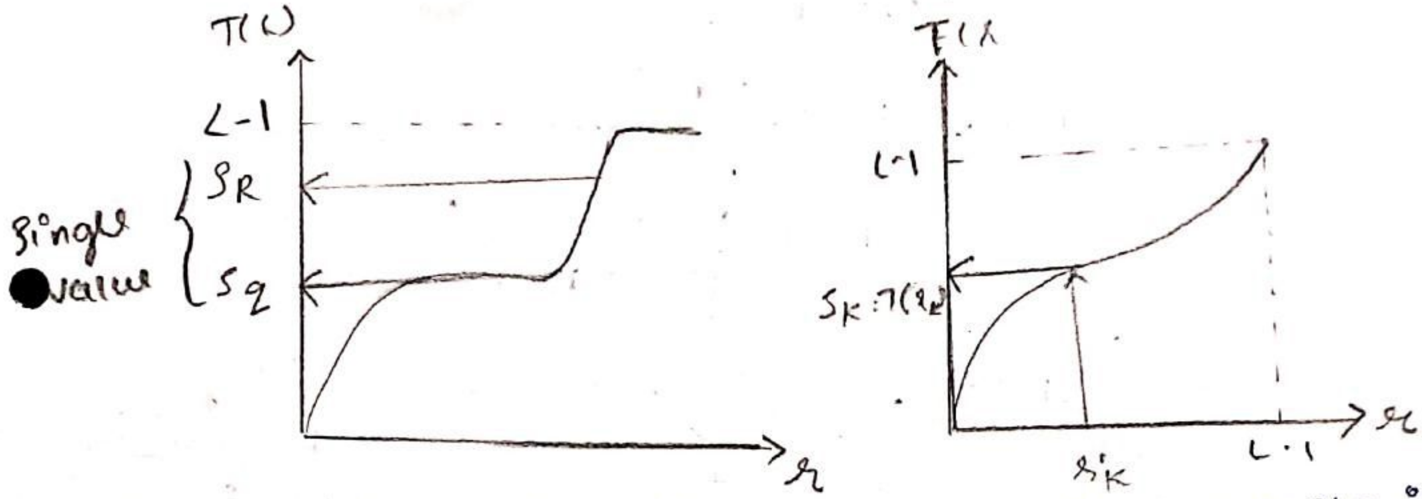
Assume that r has been normalized to the interval $[0,1]$ with $r=0$ representing black and $r=1$ representing white.

$\boxed{s = T(r)}$, $0 \leq r \leq L-1$

Inverse transformation

The inverse transformation for s back to r is denoted

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



a) Monotonically increasing function

b) strictly monotonically increasing function

* The intensity levels in an image as random variable in the interval [0, L-1].
* A fundamental descriptor of random variable is its probability distribution function (pdf).
* Let $P_r(r)$ and $P_s(s)$ denote the PDFs of r and s, P_s and P_r are different functions.

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

Transformation function as

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

ω is a dummy variable of integration

To find $P_s(s)$;

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

$$= (L-1)\frac{d}{dr}\left[\int_0^r P_r(\omega)\,d\omega\right]$$

$$= (L-1)\,P_r(r)$$

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

$$= P_r(r)\left|\frac{1}{(L-1)P_r(r)}\right|$$

$$\boxed{P_s(s) = \frac{1}{L-1}}\quad , 0 \le s \le L-1 \text{ ; uniform probability density function}$$

Q) Suppose that the continuous intensity values in an image have the PDF.

$$P_r(r) = \begin{cases} \frac{2r}{(L-1)^2}, & \text{for } 0 \le r \le L-1 \\ 0, & \text{o/w} \end{cases}$$

$$s = T(r) = L-1\int_0^r P_r(\omega)\,d\omega = \frac{2}{L-1}\int_0^r \omega\,d\omega$$

$$= \frac{2}{L-1}\left[\frac{\omega^2}{2}\right]_0^r$$

$$s = \frac{r^2}{L-1}$$

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

$$= \frac{2r}{(L-1)^2}\left[\left[\frac{ds}{dr}\right]^{-1}\right]$$

$$= \frac{2r}{(L-1)^2}\cdot\left[\frac{d}{dr}\cdot\frac{r^2}{L-1}\right]^{-1}$$

$$= \frac{2r}{(L-1)^2}\cdot\left[\frac{2r}{L-1}\right]^{-1}$$

$$= \frac{\cancel{2r}}{(L-1)^2}\cdot\frac{L-1}{\cancel{2r}}$$

$$= \frac{1}{L-1}$$

The result is a uniform PDF.

Histogram Matching

* Histogram equalisation automatically determines a transfer function that seeks to produce an output image that has a uniform histogram.

* When automatic enhancement is desired, this is good approach because the results from this technique are predictable and method is simple to implement.

* In particular, it is useful sometimes to be able to specify the shape of the histogram that the processed image have.

* The method used to generate a processed image that has a specified histogram is called histogram matching or histogram specification.

* Histogram ~~~~ matching is the transformation of an so that its histogram matches a specified histogram.

Let s be a random variable with the property.

$$s = T(r) = L-1 \int_0^r P_r(w)\,dw \quad \text{——} ①$$

we define a random variable z with the property

$$G(z) = (L-1) \int_0^z P_z(t)\,dt = s \quad \text{——} ②$$

t is a dummy variable of integration.

① = ②

$G(z) = T(r)$

$$z = G^{-1}[T(r)] = \underline{\underline{G^{-1}(s)}}$$

Steps for finding histogram equalisation

1) compute the running sum of histogram values. The running sum of histogram values is otherwise known as cumulative frequency distribution.
2) Divide the running sum in step 1 by the total number of pixels.

3) Multiply the result obtained in step 2 by the maximum gray level value.

4) Mapping of gray level by a one to one correspondance.

Q) Obtain the histogram equalization of image.

$$\begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 5 & 5 & 5 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix}$$

Step 1 - Step 4

Gray levels	0	1	2	3	4	5	6	7.
No: of pixels	0	0	0	6	14	5	0	0
Running sum	0	0	0	6	20	25	25	25
$\frac{\text{Running sum}}{\text{Total no: of pixels}}$	$\frac{0}{25}$	$\frac{0}{25}$	$\frac{0}{25}$	$\frac{6}{25}$	$\frac{20}{25}$	$\frac{25}{25}$	$\frac{25}{25}$	$\frac{25}{25}$
Multiply the result by max. gray level	$\frac{0}{25}\times 7$	$\frac{0}{25}\times 7$	$\frac{0}{25}\times 7$	$\frac{6}{25}\times 7$ 1.68	$\frac{20}{25}\times 7$ 5.6	$\frac{25}{25}\times 7$	$\frac{25}{25}\times 7$	$\frac{25}{25}\times 7$
	0	0	0	2	6	7	7	7

$$\begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 5 & 5 & 5 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix} \xrightarrow[\text{equalisation}]{\text{Histogram}} \begin{bmatrix} 6 & 6 & 6 & 6 & 6 \\ 2 & 6 & 7 & 6 & 2 \\ 2 & 7 & 7 & 7 & 2 \\ 2 & 6 & 7 & 6 & 2 \\ 6 & 6 & 6 & 6 & 6 \end{bmatrix}$$

original image

Histogram equalization image.

Q) Obtain histogram equalised image for the following 8 bit image segment size 5x5.

$$\begin{bmatrix} 200 & 200 & 200 & 180 & 240 \\ 180 & 180 & 180 & 180 & 190 \\ 190 & 190 & 190 & 190 & 180 \\ 190 & 200 & 220 & 220 & 240 \\ 230 & 180 & 190 & 210 & 230 \end{bmatrix}$$

~~Gray level~~

Pixels	Occurrences of probability	cdf
180	7/25	7/25
190	7/25	14/25
200	4/25	18/25
210	1/25	19/25
220	2/25	21/25
230	2/25	23/25
240	2/25	25/25

8 bit image $\rightarrow 2^8 = 256$.

- The pixel values ranges from 0 to 255.

$$\frac{7}{25} \times 255 = 71 \qquad \frac{21}{25} \times 255 = 214$$

$$\frac{14}{25} \times 255 = 143 \qquad \frac{23}{25} \times 255 = 235$$

$$\frac{18}{25} \times 255 = 184 \qquad \frac{25}{25} \times 255 = 255$$

$$\frac{19}{25} \times 255 = 194$$

-

$$\begin{bmatrix} 200 & 200 & 200 & 180 & 240 \\ 180 & 180 & 180 & 180 & 190 \\ 190 & 190 & 190 & 190 & 180 \\ 190 & 200 & 220 & 220 & 240 \\ 230 & 180 & 190 & 210 & 230 \end{bmatrix} \xrightarrow[\text{equalisation}]{\text{histogram}} \begin{bmatrix} 184 & 184 & 184 & 71 & 255 \\ 71 & 71 & 71 & 71 & 143 \\ 143 & 143 & 143 & 143 & 71 \\ 143 & 184 & 214 & 184 & 255 \\ 235 & 71 & 143 & 194 & 235 \end{bmatrix}$$

Basics of Spatial filtering - Smoothing.

- Smoothing filters are used for blurring and for noise reduction.
  - * Blurring is used in preprocessing steps.
  - * Noise reduction can be accomplished by blurring with a linear filter and also by non linear filtering.

There are 2 ways of smoothing spatial filters.

→ Smoothing linear filters.

→ Order statistics filters / Non linear filter.

Smoothing linear filters

* Linear spatial filters is simply the average of the pixels contained in the neighborhood of the filter mask.
* These filters are also called "averaging filters".
* The idea is replacing the value of every pixel in an image by the average of the gray levels in the neighborhood defined by the filter mask.
* It is also called low pass filters.

Application

→ Noise reduction

→ Smoothing of false contours

Two 3×3 smoothing linear filters.

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(a) Standard average

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

(b) Weighted average.

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

The spatial averaging filters in which all coefficients are equal is called box filter.

Process of median filter

* crop region of neighborhood
* Sort the values of the pixel in the neighbourhood.
* determine their median.
* Assign that value to the corresponding pixel in the filtered image.

Q) Compute the median value of the masked pixel using 3×3 mask.

$$\begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix}$$

Step 1

The pixel values are arranged in ascending order.

1 1 2 2 3 4 5 6 7

The median value of the ordered pixel.

~~1~~ ~~1~~ ~~2~~ ~~2~~ 3 ~~4~~ ~~5~~ ~~6~~ ~~7~~.

Thus the median value is 3.

Replace 5 by 3.

$$\begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 3 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix}$$

Original image data. After median filter.

* The second mask called weighted average is used to indicate that pixels are multiplied by different coefficients.

* Sum of all the coefficients in the mask is equal to 16.

* The general implementation for filtering MXN image with a weighted average filter of size mxn is given by.

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

Order statistics filter [Non linear filter]

* It is also known as median filters.
* Median filter replaces the value of a pixel by the median of the intensity values in the neighbourhood of that pixel.
* It is very effective in the presence of impulse noise also called as "salt and pepper noise" because of its appearance as white and black dots superimposed on an image.

Sharpening Spatial filtering

* The principal objective of sharpening is to highlight fine details in an image or to enhance detail that has been blurred.
* The image blurring is accomplished in the spatial domain by pixel averaging in a neighborhood.
* Since averaging is analogous to integration, sharpening could be accomplished by spatial differentiation.

First derivative	Second derivative
* Must be zero in areas of constant intensity	* Must be zero in constant areas.
* Must be nonzero at the onset of a intensity step/ramp	* Must be nonzero at the onset and end of a intensity step/ramp.
* Must be nonzero along ramps.	* Must be zero along ramps of constant slope.
* Defined as $\frac{\partial f}{\partial x} = f(x+1) - f(x)$	* Defined as $\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$

Unsharp Masking

* One of the typical technique used for edge enhancement.

* A process to sharpen images consist of subtracting a blurred version of an image from the image itself. This process is called unsharp masking.

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

Steps:

1] Blur the original image.

2] Subtract the result obtained in step 1 from original image.

3] Multiply the result obtained in step 2 by weighted fraction.

4] Add the result obtained in step 3 to the original image.

High boost filtering

* High boost filtering is also a technique used for sharpening.

* Adding and subtracting with the gain factor.

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

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

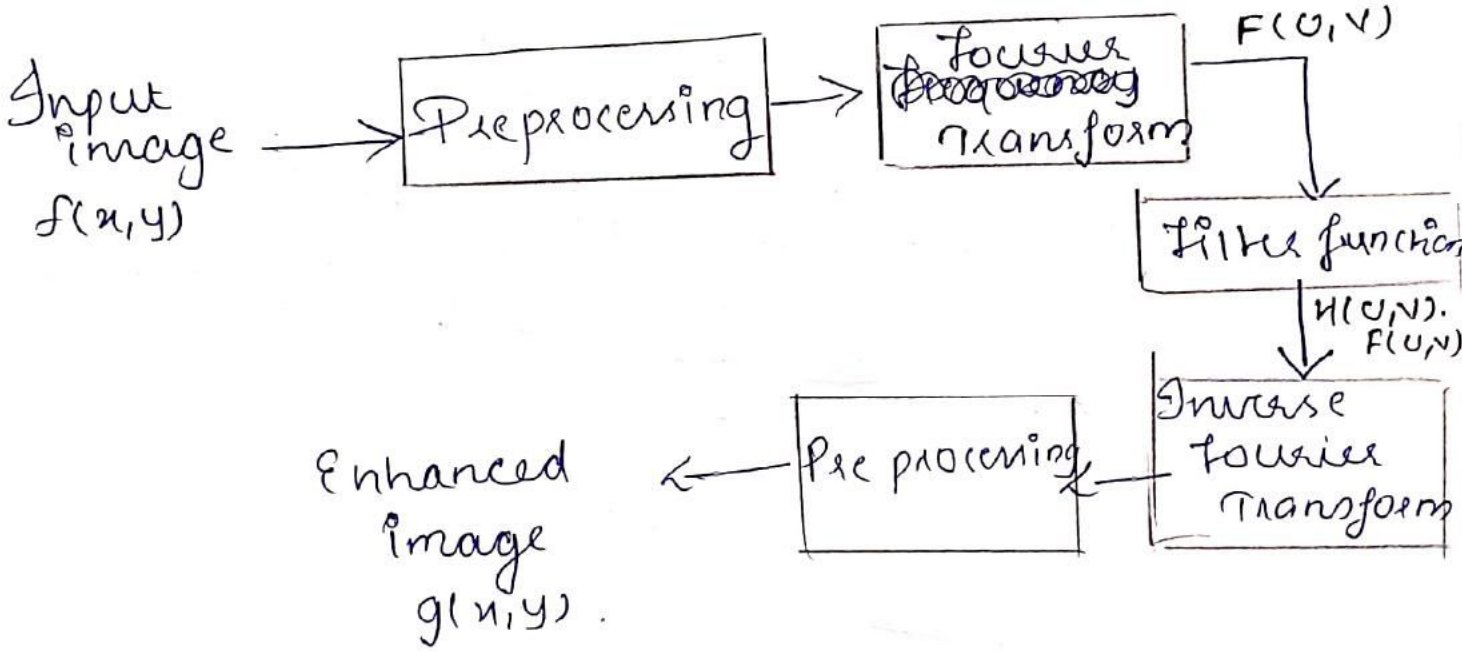
$$= (A-1)f(x,y) + f_s(x,y)$$

Frequency Domain

Frequency: Rate of repetition of some periodic event.

The selective enhancement / suppression of frequency components is termed as fourier filtering or frequency domain filtering.

Filtering in frequency domain



* Low frequency in the transform are related to slowly varying intensity in an image.
* High frequency is caused by sharp transitions in intensity such as edges and noise.
* Filter $H(U,V)$ that attenuates high frequencies while passing low frequencies would blur an image.

* Multiply all values of $s(x,y)$ by the filter function

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

Relationship between filtering in spatial and frequency domain.

The discrete convolution of two functions $f(x,y)$ and $h(x,y)$ of size $M \times N$ is defined as

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

Let $F(u,v)$ and $H(u,v)$ denote the fourier transform of $f(x,y)$ and $h(x,y)$.

$$f(x,y) * h(x,y) \longrightarrow F(u,v) \cdot H(u,v)$$

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

$$= \frac{1}{MN}$$

$$f(x,y) = \delta(x,y)$$

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

$$= \frac{1}{MN} h(x,y)$$

$$f(x,y) * h(x,y) \longrightarrow F(u,v) \cdot H(u,v)$$

$$\delta(x,y) * h(x,y) \longrightarrow \mathcal{F}[\delta(u,v)] \cdot H(u,v)$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\frac{1}{MN} h(x,y) \qquad \frac{1}{MN} H(u,v)$$

$H(U,V)$ denote frequency domain

$$H(U) = Ae^{-u^2/2\sigma^2}$$

where σ^2 is the standard deviation of the gaussian curve.

The corresponding filter in spatial domain

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

There are several standard LPF [Low Pass Filters].

* Ideal Lowpass filter
* Butterworth Lowpass filter.
* Gaussian Lowpass filter.

Ideal Lowpass filters [ILPS]

* The simplest lowpass filter is a filter that cut off all high frequency components of the fourier transform that are at a distance greater than a specified distance D_0 from the origin of the transform

$$H(U,V) = \begin{cases} 1 & D(U,V) \le D_0 \\ 0 & D(U,V) > D_0. \end{cases}$$

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

For an ILPF cross section, the point of transition between $H(U,V) = 1$ and $H(U,V) = 0$ is called the cut off frequency. The sharp cut off frequencies of an ILPF cannot be realized with electronic components

Image power

Oneway to establish a set of standard cut off frequency loci is to compute circles that encloses specified amount of the ~~xxxxx~~ total image power P_T:

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

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

~~Ideal low pass filter~~

Butterworth Lowpass filters (BLPFs) with order n.

* Doesn't have sharp discontinuity
* No clear cut off between passed and filter frequencies.
* The transfer function of a BLPF of order n and with cut off frequencies at a distance D_0.

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

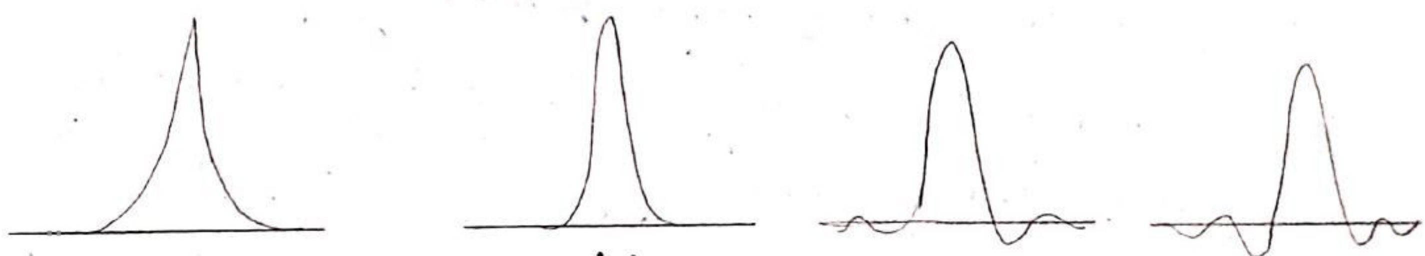
* The BLPF transfer function doesn't have a sharp discontinuity that gives a clear cut off between passed and filtered frequencies.

Butterworth Lowpass filters (BLPFs) with order n

* Doesn't have sharp discontinuity
* No clear cut off between passed and filtered frequencies.

The transfer function of a BLPF of order n with cut off frequency at a distance D_0 from the origin is defined as

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



(a) (b) (c) (d)

Spatial representation of BLPFs order 1, 2, 5, and 20.

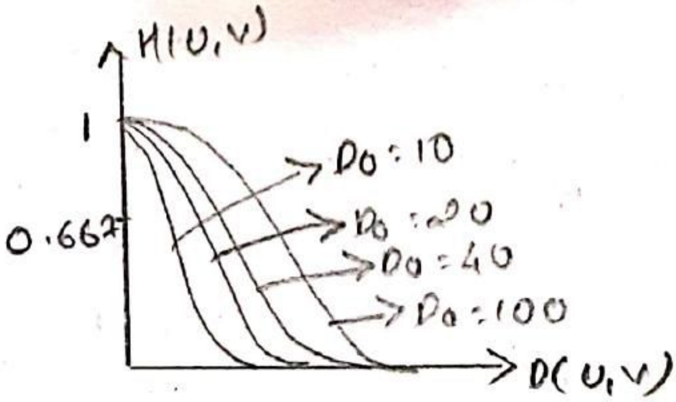
Gaussian Lowpass filters (GLPFs)

* Doesn't have sharp discontinuity.
* No clear cut off between passed and filtered frequencies.
* $H(u,v) = 0.607$ when $D(u,v) = D_0$

Gaussian LPFs of two dimension

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

where $D(U,V)$ is the distance from the centre of the frequencies and cut off frequency.



Sharpening Frequency Domain Filters

* Image sharpening can be achieved by a HPF process, which attenuates the low frequency components without disturbing high frequency information.
* Smoothing & blurring achieved by attenuating high frequency content of fourier transform.
* Edge can be enhanced by attenuating low frequency content.
* Achieved by high pass filtering.

Ideal High pass filter

A highpass filter is obtained from a given LPF using this equation.

$$H_{hp}(U,V) = 1 - H_{lp}(U,V)$$

A 2D ideal HPF is defined as

$$H(U,V) = \begin{cases} 1, & D(U,V) \le D_0 \\ 0, & D(U,V) > D_0 \end{cases}$$

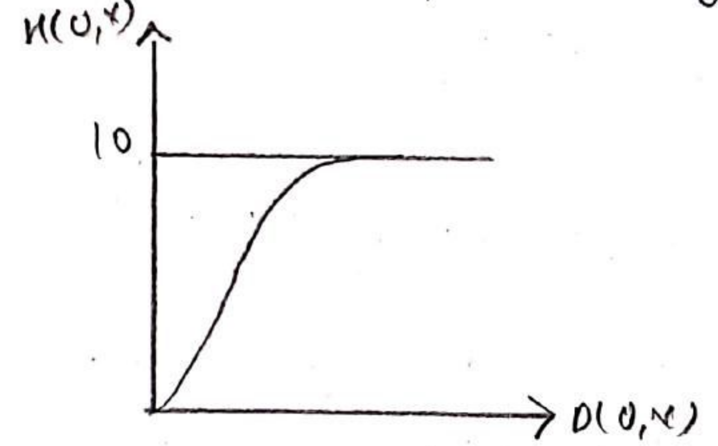
where D_0 is the cut off frequency.

- Butterworth Highpass filter

A 2D butterworth highpass filter of order n and cut off frequency D_0 is defined as

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

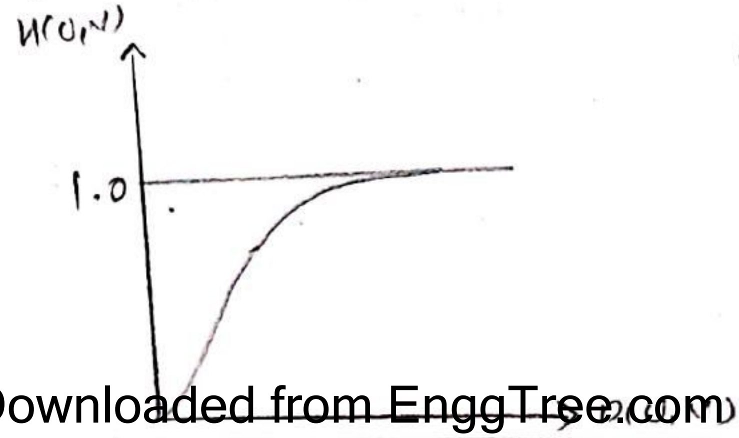
where D_0 is the cut off frequency.

- 

Gaussian Highpass filter

- The transfer function of the gaussian high pass filter with cut off frequency locus at a distance D_0 from the center of the frequency rectangle is given by

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



The Laplacian in the frequency domain

The laplacian can be implemented in the frequency domain using the filter.

$$H(u,v) = -4\pi^2(u^2+v^2)$$

$$H(u,v) = -4\pi^2\left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right]$$

$$H(u,v) = 4\pi^2 D^2(u,v)$$

The laplacian filtered image in the spatial domain is obtained by computing the inverse fourier transform of $H(u,v)F(u,v)$.

$$\nabla^2 f(x,y) = \mathcal{F}^{-1}\{H(u,v)\,F(u,v)\}.$$

where $F(u,v)$ is the DFT of $f(x,y)$.

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

Here $c=-1$ because $H(u,v)$ is negative.

Enhancing the filtered image.

A major drawback in HPF is that the high pass filters eliminate the zero frequency component in the fourier transform of the images. Thus it helps to reduce average background ~~ima~~ intensity of the image to near back.

To ~~con~~ overcome this problem, three processes are used. They are.

i) unsharp masking
ii) High boost filtering
iii) High frequency Emphasis filtering.

i) unsharp masking

unsharp masking is the process of producing a sharp image by subtracting an image and its blurred version.

$$g_{mask}(x,y) = f(x,y) - f_{LP}(x,y)$$

$$f_{LP}(x,y) = \mathcal{F}^{-1}[H_{LP}(u,v) \cdot F(u,v)]$$

where

$H_{LP}(u,v)$ is a LPF.
$F(u,v)$ is the fourier transform of $f(x,y)$

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

ii) High boost filtering

High boost filtering is based on the concept that increasing the contribution made by the original image to the filtered image would enhance the overall result.

$$f_{hb}(x,y) = A f(x,y) - f_{hp}(x,y) \quad A \geq 1$$

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

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

when $A=1$ the highboost filtering is reduces to normal high pass filter.

iii) High frequency emphasis filtering

High frequency emphasis filtering is the process of increasing the contribution made by high frequency components of an image to the overall result enhancement.

There are two steps in this process, they are.

i) Multiply a HP filter function $H_{hp}(u,v)$ by constant b.

ii) Add an offset a so that the frequency term is removed by the filter.

$$H_{hfe}(u,v) = a + b H_{hp}(u,v) \quad a \geq 0, b \geq 0.$$

when $b > 1$, high frequency are given importance as the name states.

when $a = (A-1)$ and $b = 1$ its reduced to high. boost filtering.

The typical ranges of 'a' and 'b' are.

$a = 0.25$ to 0.5

$b = 1.5$ to 2.0.

HOMOMORPHIC FILTERING

An image is modeled as the product of an illumination function and the reflectance function at every point.

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

The illumination component has low frequency components of the image and the reflectance component is associated with high frequency components of the image.

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

Taking fourier transform.

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

consider $z(x,y) = \ln f(x,y)$

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

$$F[z(x,y)] = F\{\ln f(x,y)]$$

$$= F\{\ln i(x,y)\} + F\{\ln r(x,y)].$$

$$\boxed{Z(u,v) = F_i(u,v) + F_r(u,v)}$$

where $F_i(u,v)$ and $F_r(u,v)$ are the fourier transform of $\ln i(x,y)$ and $\ln r(x,y)$ respectively.

If we process $z(u,v)$ by means of a filter function $H(u,v)$ then,

$$S(u,v) = H(u,v) Z(u,v)$$
$$= H(u,v) F_i(u,v) + H(u,v) F_r(u,v)$$

where $S(u,v)$ is the fourier transform of the result.

$$s(x,y) = F^{-1}\{S(u,v)\}.$$
$$= F^{-1}[H(u,v) F_i(u,v)] + F^{-1}[H(u,v) F_r(u,v)]$$

$$i'(x,y) = F^{-1}\{H(u,v) F_i(u,v)\}.$$

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

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

$$g(x,y) = e^{s(x,y)}$$
$$= e^{i'(x,y)} \cdot e^{r'(x,y)}$$
$$= i_0(x,y)\, r_0(x,y).$$

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

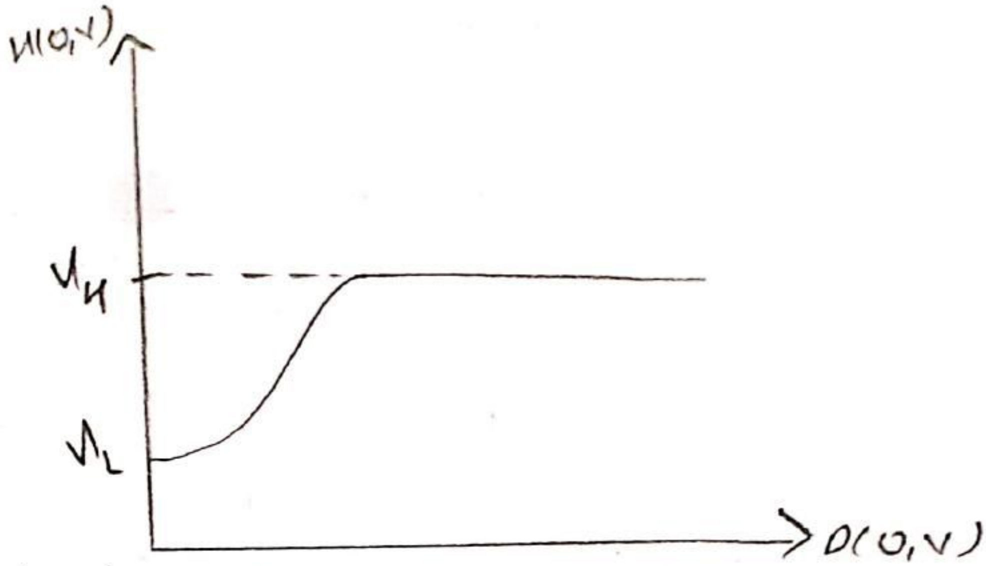
$r_0(x,y) = e^{r'(x,y)}$.

Advantages

* A good deal of control can be gained over the illumination and reflectance components with a homomorphic filter.
* Attenuates the contribution made by the low frequencies.
* Dynamic range compression and contrast enhancement.

The gaussian highpass filter yields the function

$$H(U,V) = [\gamma_H - \gamma_L] \, 1 - e^{-2(D^2(U,V)/D_0^2)}] \, \gamma_L$$



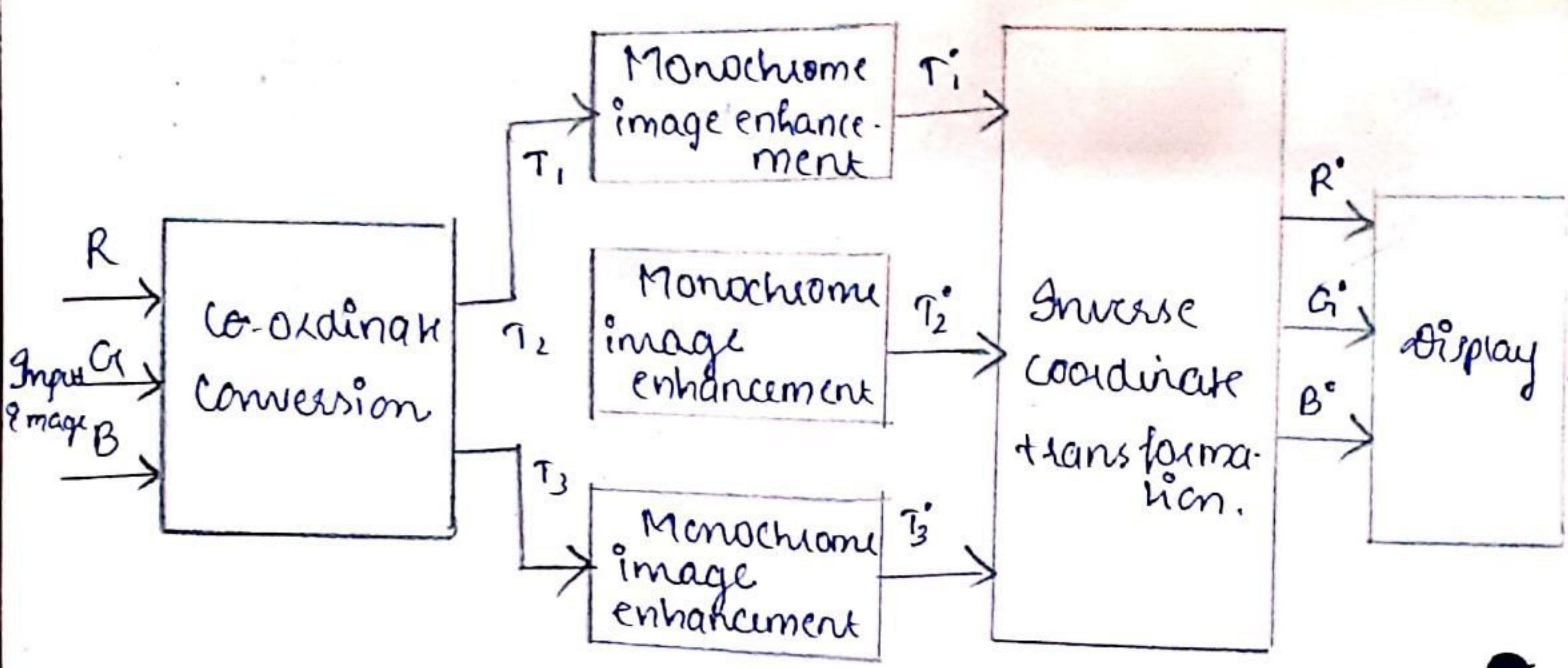
Colour image Enhancement

* The monochrome image enhancement, colour image enhancement may require improvement of colour balance or colour contrast in a colour image

* Enhancement of colour images becomes a more difficult task not only because of the added dimension of the data but also due to the added complexity of colour perception.

* The enhanced image coordinates T_1, T_2, T_3 are inverse transformed to R, G, B for display.

* Since each image plane $T_K(m,n)$, $K = 1, 2, 3$ is enhanced independently



* The enhanced image is represented as T_k'.
* The choice of colour coordinate system T_K, $K = 1, 2, 3$ in which the enhancement algorithms are implemented.

UNIT-III

IMAGE RESTORATION

Image restoration is defined as the process of reconstructing or recovering an image which is in the degraded or distorted state. It attempts to recover an image that has been degraded by using a prior knowledge of the degradation phenomenon.

The main objective of restoration process is to improve the appearance of an image. This is same as enhancement process. It depends only on class or ensemble properties of dataset whereas image enhancement technique are much more image dependent.

The different causes of image degradation are

→ Improper opening and closing of the shutter

→ Atmospheric turbulance

→ Misfocus of lens

→ Relative motion between camera and object which causes motion blur.

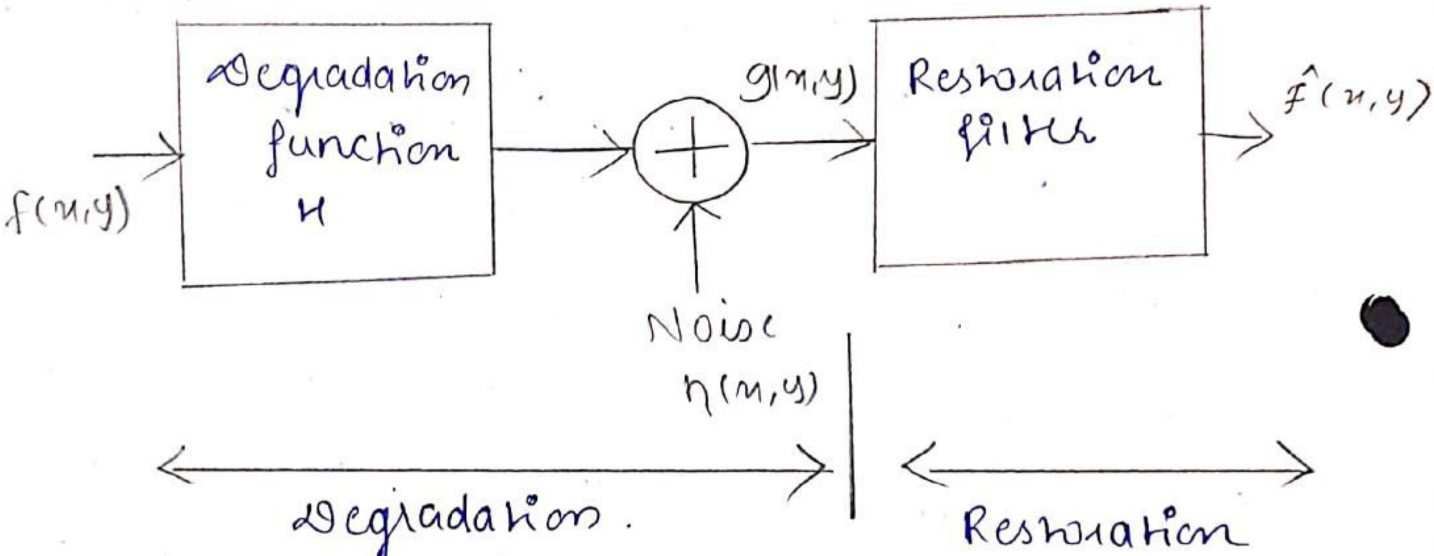
Some examples of restoration process are

* Correction of geometric distortion or nonlinearities due to sensors.

* Deblurring of images degraded by the limitations of a sensors or its environments.

* Noise filtering.

Degradation Model



The degradation process is modeled using a degradation function together with an additive noise term, operates on an image $f(x,y)$ to produce a degraded image $g(x,y)$.

The objective of restoration is to obtain an estimate $\hat{f}(x,y)$ of the original image.

Spatial domain representation

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

where $h(x,y)$ is the spatial representation of the degradation function and as in the symbol '*' denotes convolution.

Frequency domain representation

$$G(U,V) = H(U,V)F(U,V) + N(U,V)$$

where the terms in capital letters are the fourier transforms.

Properties

The input-output relationship before the restoration stage is expressed as

$$g(x,y) = H[f(x,y)] + \eta(x,y).$$

$$\eta(x,y) = 0, \quad g(x,y) = H[f(x,y)].$$

i) Additivity property.

If H is linear, the additivity property gives.

$$H[a f_1(x,y) + b f_2(x,y)] = aH[f_1(x,y)] + bH[f_2(x,y)]$$

where a and b are scalars and $f_1(x,y)$ and $f_2(x,y)$ are any two input images.

$$a = b = 1$$

$$H[f_1(x,y) + f_2(x,y)] = H[f_1(x,y)] + H[f_2(x,y)]$$

which is called the property of additivity.

ii) Homogenity property

$$f_2(x,y) = 0$$

$$H[af_1(x,y)] = aH[f_1(x,y)]$$

Which is called the property of homogenity.

Thus a linear operator possesses both the property of additivity and the property of homogenity.

iii) Impulse Response:

An operator having the input-output relationship $g(x,y) = H[f(x,y)]$ is said to be position invariant if.

$$g(x-\alpha, y-\beta) = H[f(x-\alpha, y-\beta)]$$

$f(x,y)$ can be expressed in terms of a continuous impulse function.

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\,\delta(x-\alpha, y-\beta)\,d\alpha\,d\beta.$$

$$g(x,y) = H[f(x,y)] = H\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\,\delta(x-\alpha, y-\beta)\,d\alpha\,d\beta\right]$$

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H[f(\alpha,\beta)\,\delta(x-\alpha, y-\beta)]\,d\alpha\,d\beta.$$

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\,H[\delta(x-\alpha, y-\beta)]\,d\alpha\,d\beta.$$

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\,H[\delta(x-\alpha, y-\beta)]\,d\alpha\,d\beta$$

The term

$$h(x,\alpha,y,\beta) = H[\delta(x-\alpha, y-\beta)]$$

is called the impulse response of H.

If $\eta(x,y) = 0$, then $h(x,\alpha,y,\beta)$ is the response of H to an impulse at the coordinates (x,y). If the impulse becomes a point of light and $h(x,\alpha,y,\beta)$ is commonly referred to as the point spread function ~~PUSF~~ (PSF).

Then,

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\, h(x,\alpha,y,\beta)\, d\alpha\, d\beta.$$

which is called the superposition (Fredholm) integral of first kind.

If H is position invariant,

$$H[\delta(x-\alpha, y-\beta)] = h(x-\alpha, y-\beta)$$

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\, h(x-\alpha, y-\beta)\, d\alpha\, d\beta.$$

This expression is called the convolution integral.

Additive noise

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\, h(x,\alpha,y,\beta)\, d\alpha\, d\beta + \eta(x,y)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\, h(x-\alpha, y-\beta)\, d\alpha\, d\beta + \eta(x,y)$$

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

based on convolution theorem, the expression in the frequency domain is

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

Noise models

The principal sources of noise in digital images arise during image acquisition and transmission additivity and the property of homogeneity.

* Imaging sensors can be affected by ambient conditions.

* Inference can be added to an image during transmission.

Noise from sensors
- → Electronic circuits
- → Light level
- → Sensor temperature.

Noise from environment
- → Lightening
- → Atmospheric disturbance
- → Other strong electric/magnetic signals.

The noisy image is modelled as follows

$$g(x,y) = f(x,y) + \eta(x,y).$$

where $f(x,y)$ is the original image pixel, $\eta(x,y)$ is the noise term and $g(x,y)$ is the resulting noise pixel.

Noise categories based on distributions.

There are many different models for the image noise term $\eta(x,y)$.

a) Gaussian

b) Rayleigh

c) Erlang

d) Exponential

e) Uniform

f) Impulse (salt and pepper noise)

i) Gaussian Noise

Probability density function (PDF)

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

where z = gray level (Gaussian Random variable)

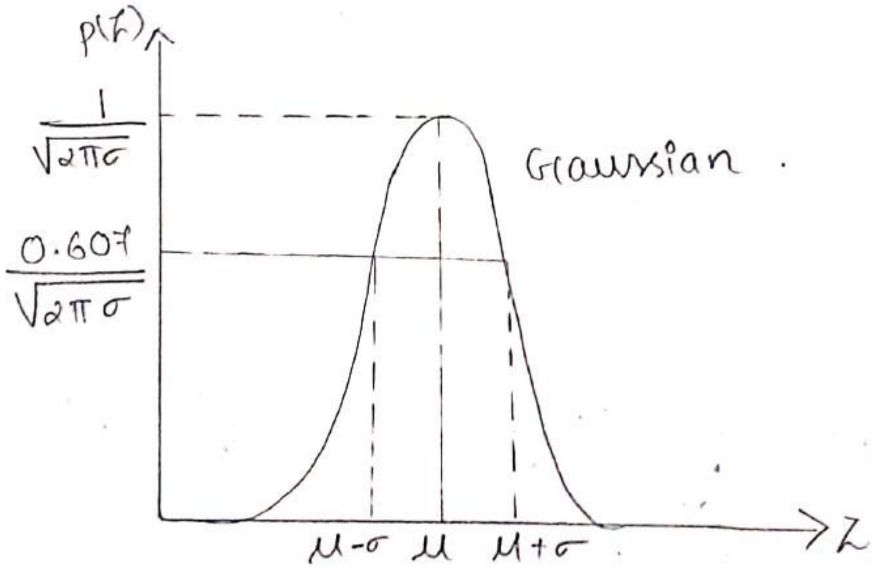
μ = mean of average value of z.

σ = Standard deviation of z.

σ^2 = variance of z.

→ 70% of z in $[\mu-\sigma, \mu+\sigma]$

→ 90% of z in $[\mu-2\sigma, \mu+2\sigma]$



PDF of gaussian noise.

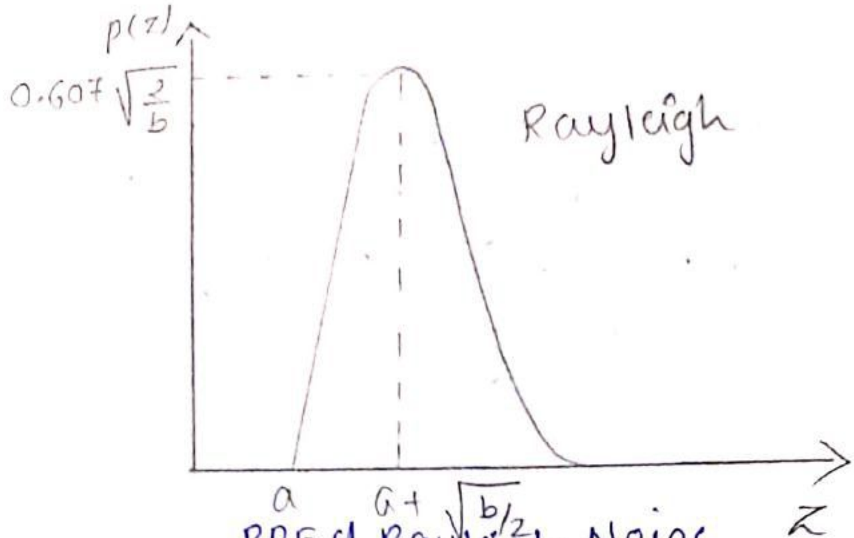
ii) Rayleigh Noise

This type of noise is mostly present in range images.

Ranges images are used in many remote sensing applications where the pixel value indicates the distance between the object and the camera system.

The PDF of Rayleigh noise is given by

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



PDF of Rayleigh Noise

The mean and variance of this density are given by

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

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

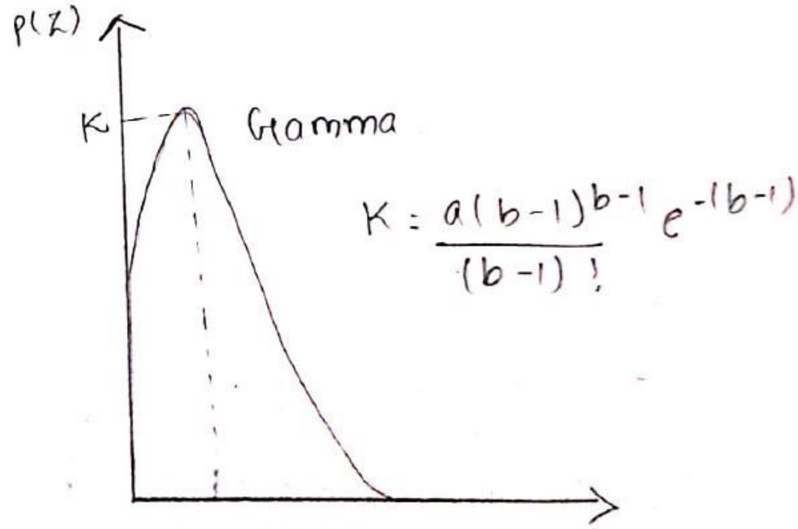
iii) Erlang [Gamma] Noise

The PDF of Erlang noise is given by.

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

$$\mu = \frac{b}{a}$$

$$\sigma^2 = \frac{b}{a^2}.$$

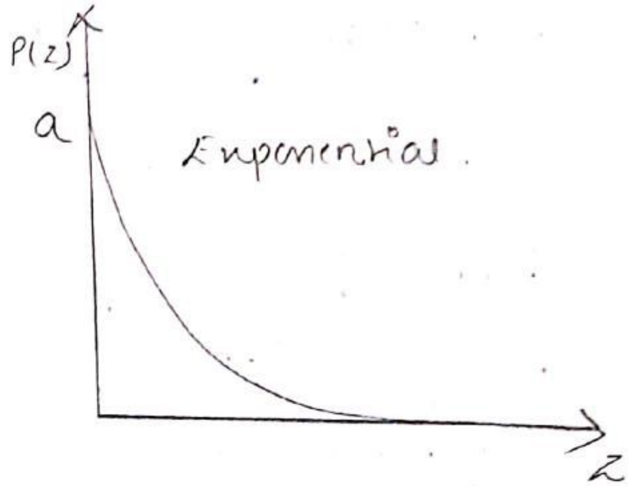


PDF of Erlang Noise

iv) Exponential Noise

The PDF of exponential noise is given by

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



Where $a > 0$.

$$\mu = \frac{1}{a}$$

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

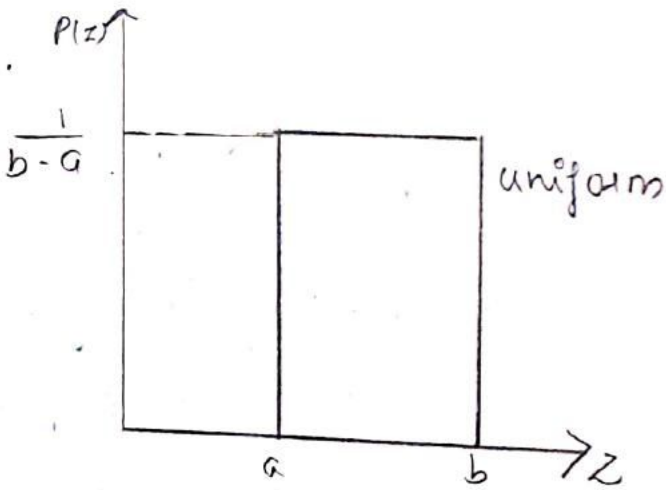
v) Uniform Noise

The PDF of uniform noise is given by.

$$p(z) = \begin{cases} \frac{1}{b-a} & , a \leq z \leq b \\ 0 & , o/w. \end{cases}$$

$$\mu = \frac{a+b}{2}$$

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



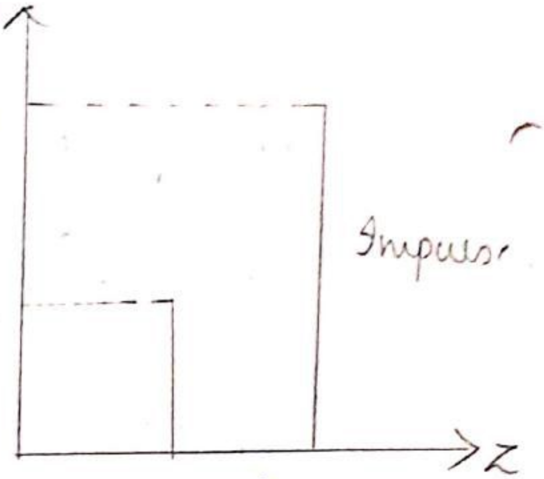
PDF of uniform noise.

vi) Impulse (salt and pepper) Noise.

The PDF of bipolar impulse noise is given by.

$$p(z) = \begin{cases} P_a & , z = a \\ P_b & , z = b \\ 0 & , o/w. \end{cases}$$

If $b > a$, intensity b will appear as light dot in the image.



PDF of impulse noise

- bipolar if $P_a \neq 0$, $P_b \neq 0$
- unipolar if one of P_a and P_b is zero.
- Noise looks like salt and pepper granules if $P_a \simeq P_b$.
- Negative / positive - scaling is often necessary to form digital images.

Estimation of noise parameters

↳ Periodic Noise

↳ Random noise with unknown PDFs.

<u>Case 1</u> : Imaging system is available.

* Capture images of 'flat' environment.

<u>Case 2</u> : Noisy images available.

* Take a strip from constant area
* Draw the histogram and observe it.
* Measure the mean and variance.

Mean, $\mu = \sum_{z_i \in S} z_i \, p(z_i)$

variance, $\sigma^2 = \sum_{z_i \in S} (z_i - \mu)^2 p(z_i)$.

Mean filters

The mean filter is a simple spatial filter that replaces the center value in the window with the average (mean) of all pixel values in the window.

There are different kinds of mean filters all of which exhibit slightly different behaviour.

- Geometric Mean.
- Harmonic Mean
- Contraharmonic mean.

Consider an example of mean filtering of a single 3x3 window of values is shown below.

5	3	6
2	1	9
8	4	7

$$\text{Mean}, \mu = \frac{5+3+6+2+1+9+8+4+7}{9}$$

$$= \frac{45}{9} = 5$$

Replace 1 by the mean of all nine values to 5.

5	3	6
2	5	9
8	4	7

Geometric Mean filter

An image restored using a geometric mean filter is given by,

$$\hat{f}(x,y) = \left[\prod_{(s,t)\in S_{xy}} g(s,t)\right]^{1/mn}.$$

Each restored pixel is given by the product of the pixels in the subimage window raised to the power $\frac{1}{mn}$. A geometric mean filter achieves smoothing.

It is well suited for random noise like Gaussian or uniform noise.

Harmonic mean filter

The harmonic mean filtering operation is given by.

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

The harmonic mean filter works wells for salt noise, but fails for pepper noise. It does well also with other types of noise like Gaussian noise.

Contraharmonic mean filter

The contraharmonic mean filter yields a restored image based on the expression

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

where Q is called the order of the filter.

→ For +ve values of Q, the filter eliminates pepper noise.

→ For -ve values of Q, it eliminates salt noise.

→ It cannot do both simultaneously.

→ Contra harmonic filter reduces to arithmetic mean filter if Q=0, and to the harmonic mean filter, Q = -1.

Disadvantages

* It must be known whether the noise is dark or light inorder to select the proper sign for Q.
* The results of choosing the wrong sign for Q can be disastrous.

Order statistics filter

Order statistics filters are spatial filters whose response is based on ordering the values of the pixels contained on the image area encompassed by the filter.

Suitable for unipolar or bipolar noise (salt and pepper noise).

→ Median filter.

→ Max/Min filter.

→ Midpoint filter.

→ Alpha-trimmed mean filter

i) Median Filter

The best known order-statistics filter is the median filter, which as its name implies replaces the value of a pixel by the median of the intensity levels in the neighborhood of that pixel.

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

* The value of pixel at (x,y) is included in the computation of the median.
* Median filters are quite popular because for certain types of random noise, they provide excellent noise reduction capabilities.
* It is particularly effective in the presence of both bipolar and unipolar impulse noise.

Properties

1] A median filter smoothens additive white noise.
2] A median filter doesn't degrade edge.
3] A median filter is effective in removing impulses.

Problem 1

A 4x4 image is given by

$$\begin{bmatrix} 3 & 2 & 1 & 4 \\ 5 & 2 & 6 & 3 \\ 7 & 9 & 1 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

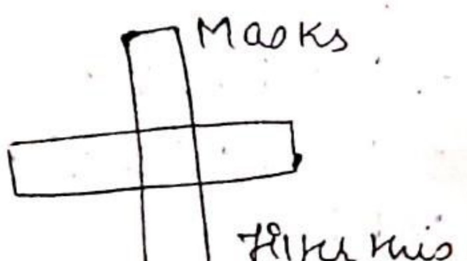


image using a median filter. Assume replicate padding.

Step 1: To perform replicate padding.

$$\begin{bmatrix} 3 & 3 & 2 & 1 & 4 & 4 \\ 3 & 3 & 2 & 1 & 4 & 4 \\ 5 & 5 & 2 & 6 & 3 & 3 \\ 7 & 7 & 9 & 1 & 4 & 4 \\ 2 & 2 & 4 & 6 & 8 & 8 \\ 2 & 2 & 4 & 6 & 8 & 8 \end{bmatrix}$$

Step 2: compute median using given masks.

The median of the pixel 3 marked

$$\begin{bmatrix} 3 & 3 & 2 & 1 & 4 & 4 \\ 3 & 3 & 2 & 1 & 4 & 4 \\ 5 & 5 & 2 & 6 & 3 & 3 \\ 2 & 2 & 4 & 6 & 8 & 8 \\ 2 & 2 & 4 & 6 & 8 & 8 \end{bmatrix}$$

3, 3, 2, 2, 2

The four neighbors including the found pixels are 3, 3, 3, 2 and 5.

The pixels are then arranged in the decending order as 5, 3, 3, 3, 2

3 2 2 2 3 1

5 3 $\underline{3}$ 3 2

6 4 2 1 1

The median value is 3.

Q) What is the value of the marked pixel after 5×5 median pixel.

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 5 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 0 & 0 & 1 & 2 \\ 5 & 1 & 2 & 3 & 1 \\ 4 & 3 & 1 & 2 & 0 \end{bmatrix}$$

Step 1

Arranging the pixels in ascending order.

0 0 0 1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 4 4 5 5

Step 2

$\cancel{0}$ $\cancel{0}$ $\cancel{0}$ $\cancel{1}$ $\cancel{1}$ $\cancel{1}$ $\cancel{1}$ $\cancel{1}$ $\cancel{1}$ 1 $\cancel{2}$ $\cancel{2}$ 2 2 $\cancel{2}$ $\cancel{2}$ $\cancel{3}$ $\cancel{3}$ $\cancel{3}$ $\cancel{3}$ $\cancel{4}$ $\cancel{4}$ $\cancel{5}$ $\cancel{5}$

The median value is found to be 2. Hence the marked pixel 0 is replaced by 2.

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 5 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 0 & 2 & 1 & 2 \\ 5 & 1 & 2 & 3 & 1 \\ 4 & 3 & 1 & 2 & 0 \end{bmatrix}$$

Problem 2:

Compute the median value of the marked pixel using 3x3 mask.

$$\begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix}$$

Step 1

Pixel values are arranged in ascending order

1 1 2 2 3 4 5 6 7

Step 2

The median value of the ordered pixel.

1 1 2 2 3 4 5 6 7

The median value is computed to be 3.

$$\begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 5 & 7 \\ 2 & 3 & 6 \\ 3 & 2 & 1 \end{bmatrix}$$

Problem 3

What is the value of the marked pixel after a 5x5 median filter.

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 5 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 0 & 0 & 1 & 2 \\ 5 & 1 & 2 & 3 & 1 \\ 4 & 3 & 1 & 2 & 0 \end{bmatrix}$$

Step 1

Arranging the pixels in ascending order.

0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 5, 5

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 5 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 0 & 2 & 1 & 2 \\ 5 & 1 & 2 & 3 & 1 \\ 4 & 3 & 1 & 2 & 0 \end{bmatrix}$$

ii) Max and Min filters.

* Median filter is most used in image processing.
* using 100th percentile results in the so called max filter given by.

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

The maximum filter selects the largest value within of pixel values.

The 0th percentile filter is the min filter:

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

12, 13, 14, 15, 17, 18, 19, 20, 21.

Max filter : 21 - select largest value within pixel values.

Min filter : 12 - select smallest value within pixel values.

iii) Midpoint filter.

The midpoint filter simply computes the midpoint between maximum and minimum values in the area encompassed by the filter.

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

Problem 1

consider the pixel.

$\{12, 13, 14, 15, 17, 18, 19, 20, 21\}$

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

$$= \frac{1}{2}\left[21+12\right]$$

$$= \frac{33}{2} = 16.5$$

works well for gaussian noise or uniform noise.

(iv) Alpha-trimmed mean filter

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

→ Able to delete $d/2$ lowest and $d/2$ highest gray levels.

Let $g_r(s,t)$ represent the remaining $mn-d$ pixels. A filter formed by averaging these remaining pixel is called an alpha trimmed mean filter.

* $d=0$, Arithmetic Mean filter
* $d = mn-1$, the filter becomes median filter
* Suitable for situation involving types of noise.

Adaptive Filters

The behaviour of adaptive filters changes depending on the characteristics of the image inside the filter region.

* Improved performance Vs increased complexity.
* Example: Adaptive local noise reduction filter.

i) Adaptive, local noise reduction filter.

* The simplest statistical measures of a random variable are its mean and variance.

→ Mean of a random variable : Measure of average ~~bright~~ gray level in some region.

→ variance of a random variable : Measure of average contrast in the region.

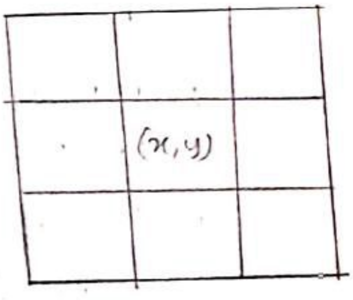
Response based on four quantities

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

σ_η^2 : variance of the noise

m_L : local mean in S_{xy}

σ_L^2 : local variance in S_{xy}.



S_{xy}

* If $\sigma_\eta^2 = 0$, $\hat{f}(x,y) = g(x,y) = f(x,y)$

* If $\sigma_L^2 >> \sigma_\eta^2$, $\hat{f}(x,y) \approx g(x,y)$

* If $\sigma_L^2 \approx \sigma_\eta^2$, $\hat{f}(x,y)$ = arithmetic mean value in S_{xy}.

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

* when $\sigma_L^2 \geq \sigma_\eta^2$, $\hat{f}(x,y) \geq 0$

* when $\sigma_L^2 < \sigma_\eta^2$, $\hat{f}(x,y)$ may be negative.

ii) Adaptive Median Filtering

* The median filter performs relatively well on impulse noise.
* The adaptive median filter can handle much more spatially dense impulse noise.
* The filter size changes depending on the characteristics of the image.

The main objective of adaptive median filtering are

- Removing salt and pepper noise
- Reducing distortion.
- Remove impulse noise
- To provide smoothing of other noise
- Reduce distortion such as excessing thinning or thickening of object boundaries.

Periodic Noise Reduction [Frequency Domain Filtering]

* Frequency domain techniques in Fourier domain are most effective at removing periodic noise.

Four types of selective filters are used for periodic noise reduction.

- Band Reject filters
- Bandpass filters
- Notch filters.
- Optimum notch filtering.

i) Band Reject filters

Removing periodic noise ~~from~~ form an image involves removing a particular range of frequencies from that image. It can be used for this purpose.

An ideal band reject filter is given as.

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

$D(u,v)$ represents the distance from the origin of the centered frequency rectangle.

$W \rightarrow$ width of the band

$D_0 \rightarrow$ Radial center of the band.

ii) Bandpass filters

* Performs opposite operation of a band reject filter.

$$H_{BP}(u,v) = 1 - H_{BR}(u,v)$$

iii) Notch filters

* Rejects / passes frequencies in predefined neighborhoods about a center frequency.

* Notch pass filter passes the frequencies contained in the notch areas

$$H_{NP}(u,v) = 1 - H_{NR}(u,v)$$

iv) Optimum Notch filtering

* When several interference patterns are present, filtering may remove much image information.

First filter out noise interference by placing a notch pass filter $H(u,v)$ at the location of each spike.

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

where $G(u,v)$ is the fourier transform of the corrupted image.

The corresponding spatial domain representation.

$$\eta(x,y) = F^{-1}\{N(u,v)\}.$$

$$= F^{-1}\left[H(u,v) \cdot G(u,v)\right]$$

Then subtract a weighted portion of $\eta(x,y)$ from the image $g(x,y)$ to obtain restored image.

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

$w(x,y)$ is called weighting or modulation function $\hat{f}(x,y)$ over a specified neighbourhood of every point (x,y).

Inverse Filtering

The simplest method to restore images degraded by a degradation function H is direct inverse filtering.

Inverse filtering is defined as the process of recovering the input of a system from its output.

Let an image f(x,y) degraded by a degradation function H. The obtained degraded image is denoted as g(x,y). The inverse filtering divides the transform of the degraded image G(u,v) by the degradation function H(u,v).

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

$F(u,v) \rightarrow$ Transform of the original image

$\hat{F}(u,v) \rightarrow$ Approximation of $F(u,v)$.

In the presence of noise, the degraded image is given as

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

Applying Fourier transform

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

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

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

Not able to restore the image exactly with knowledge of $H(u,v)$ because $H(u,v)$ is unknown

If $H(u,v)$ has values turns to zero, then $\frac{N(u,v)}{H(u,v)}$ will dominate $F(u,v)$. This indicates the poor performance of the system which results in bad approximation of the original function. This is known as "zero or small value problem".

* Inverse filtering is very sensitive to noise and has no provision to handle noise.
* Blurring corresponds to lowpass filtering and inverse filtering corresponds to high pass filters.

Disadvantages

* Highly sensitive to noise
* It is not always possible obtain an inverse.

Application

* Removal of blur caused by uniform linear motion.

Pseudo inverse filter

Removes the problem at zero (or near zero) frequencies, but still ampifies noise at other frequencies where the blurring filter response is not zero but small.

A pseudo inverse filter is defined as

$$= \begin{cases} \frac{1}{H} & , H > \varepsilon \\ \varepsilon & , H \leq \varepsilon \end{cases}$$

The value of ε affects the restored image.

Weiner Filtering [Minimum Mean Square Error Filtering]

* For the restoration of an image, this method considers the degradation function as well as the statistical properties of noise.

* It is used to improve direct inverse filtering because it has a provision to handle the noise.
* Weiner filter is an optimal filter
* It's purpose is to reduce the amount of noise in an image.
* It not only restores the image, but also removes noise by image smoothing.

"The main objective of weiner filtering is to approximate the original image in such a way that the mean square error between original and approximated images is minimized."

Mean square error:

$$e^2 = E\{(f-\hat{f})^2\}$$

 (u,v) is unknown

where $E(x)$ - expected value of vector x.

f - uncorrupted / original image.

$\hat{f}$ - approximation of f.

There some assumptions are made to perform least mean square error filtering.

* The image and the noise are un-correlated

* Either image or noise has zero mean.

* The approximated gray levels are a linear function of the degraded gray levels.

Frequency domain which satisfies the minimum error function is given by.

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

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

Multiplying and dividing by $H(u,v)$

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

for u,v ranges from 0 to $N-1$.

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

$H(u,v)$ is the transform of the degradation function.

$H^*(u,v) \rightarrow$ complex conjugate of $H(u,v)$

$S_\eta(u,v) \rightarrow |N(u,v)|^2 \Rightarrow$ power spectrum of noise

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

If $N(u,v) = 0$.

$$\hat{F}(u,v) = \frac{1}{H(u,v)}\left[\frac{|H(u,v)|^2}{|H(u,v)|^2 + 0}\right] G(u,v)$$

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

Thus if noise is zero, weiner filter reduces to the inverse filter.

Advantages

* It has no small or zero value problem
* The results obtained in weiner filtering are more closer to the original image than the inverse filtering.

Disadvantages

* It requires the power spectrum of the undegraded image and noise to be known, which makes the implentation more difficult.
* Weiner filter is based on minimizing a statistical criterion.

* The weiner filter requires a prior knowledge of the power spectral density of original image.

Signal to Noise ratio

The signal to noise ratio can be approximated using the equation

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

*Low noise gives high SNR and high noise gives low SNR.

Mean square Error

The MSE is defined as

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

The SNR in spatial domain is defined as

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

1

UNIT-IV

IMAGE SEGMENTATION

Segmentation subdivides an image into its constituent regions or objects.

Region of Interest (ROI) vary with applications.
eg: If we want to analyse the tumour (mass) in an X-ray image, then the tumour in the image is the ROI.

The segmentation methods or algorithms are of two types
* Discontinuity
* Similarity.

Methods based on Discontinuity

This approach is to partitioning an image based on abrupt changes in intensity such as edges in an image.

Methods based on Similarities

This kind of methods partitioning an image into similar region depending on a specified criteria.

eg: Thresholding, Region growing, Region splitting and merging

Applications

* Industrial inspection.
* Autonomous target acquisition.

Edge Detection

There are three type of gray level discontinuities known as

i) Points
ii) Lines
iii) Edges.

The general 3x3 mask is represented as

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

i) Point Detection

$$|R| \geq T$$

where T is a nonnegative threshold

R is the sum of products of the coefficients with gray levels contained in the region.

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

ii) Line Detection

* It is little more ~~more~~ complex than point detection.

* In line detection, 4 types of masks are used to get the responses for the directions vertical, horizontal, +45°, -45° respectively.

-1	-1	-1
2	2	2
-1	-1	-1

Horizontal

-1	-1	2
-1	2	-1
2	-1	-1

+45°

-1	2	-1
-1	2	-1
-1	2	-1

vertical

2	-1	-1
-1	2	-1
-1	-1	2

-45°.

* Horizontal mask

It responds more strongly to lines oriented horizontally with one pixel thickness.

* vertical mask

This mask will have strong respond to vertical lines.

* +45° mask

This one responds best to lines in the +45° direction.

* -45° mask

It has strong responds to lines oriented at -45° direction.

Method of detection

Case i

To detect all line in the image.

If $R_1 > R_j$, for $j = 2,3,4$ the point is said to be related with a horizontal line.

If $R_2 > R_j$, for $j = 1,3,4$ the point is said to be related with a vertical line.

Case ii

To detect lines in a specified direction.

The mask for the particular direction is run through the image and its output is thresholded ie its output is greater than a threshold, T it is considered as a line point.

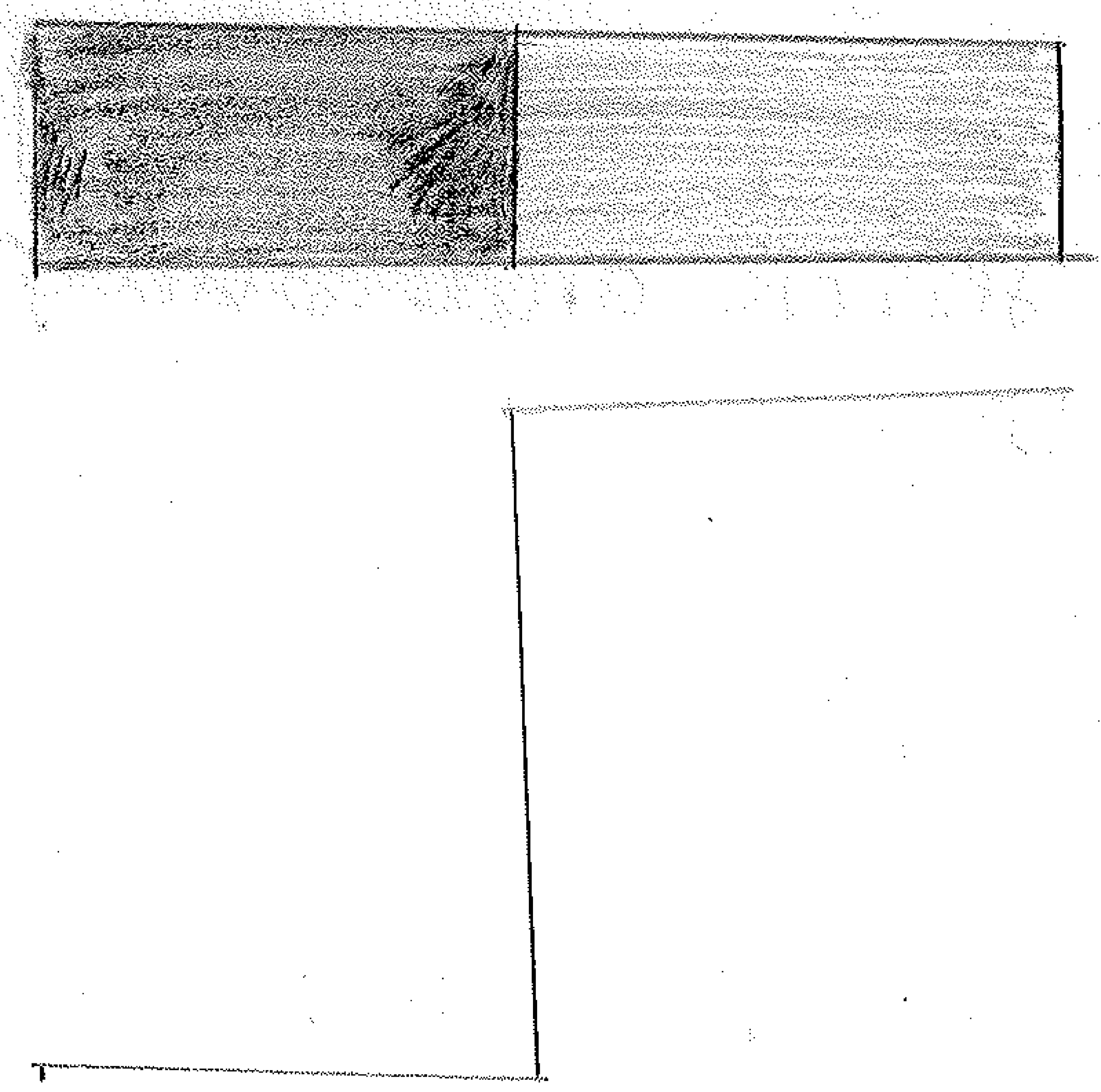
iii) Edge detection

* This approach is mostly used for segmenting images based on local changes in intensity.

Edge models are classified according to their intensity profiles.

→ Ideal Digital Edge model.

A step edge involves a transition between two intensity levels occurring ideally over the distance of one pixel.



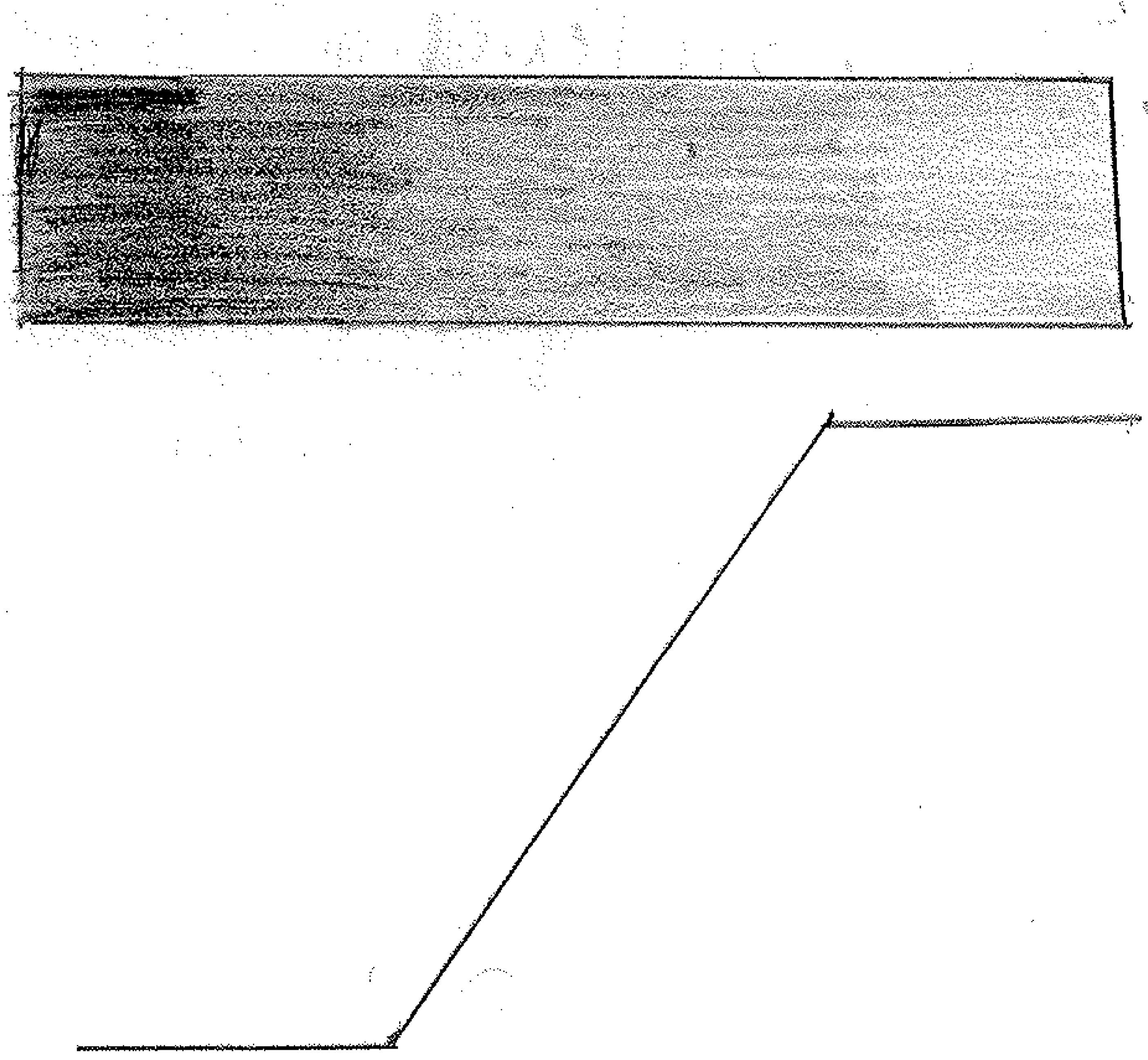
* These ideal edges occur over the distance of one pixel, provided that no additional processing is used to make them look real.

Model of an ideal digital edge.

→ Ramp Edge Model

* A ramp allows for a smoother transition between segments.

* A ramp edge is useful for modeling the blurred edges created from sampling a scene containing objects



* In digital images have edge that are blurred and noisy with the degree of blurring determined principally by limitations in the focusing mechanism.

Model of a ramp edge.

* The slope of the ramp is inversely proportional to the degree of blurring in the edge.

* The thickness of the edge is determined by the length of the ramp.

* The length is determined by the slope, which in turn is determined by the degree of blurring.
* Blurred edges will appear thick and sharp edge tends to be thin.

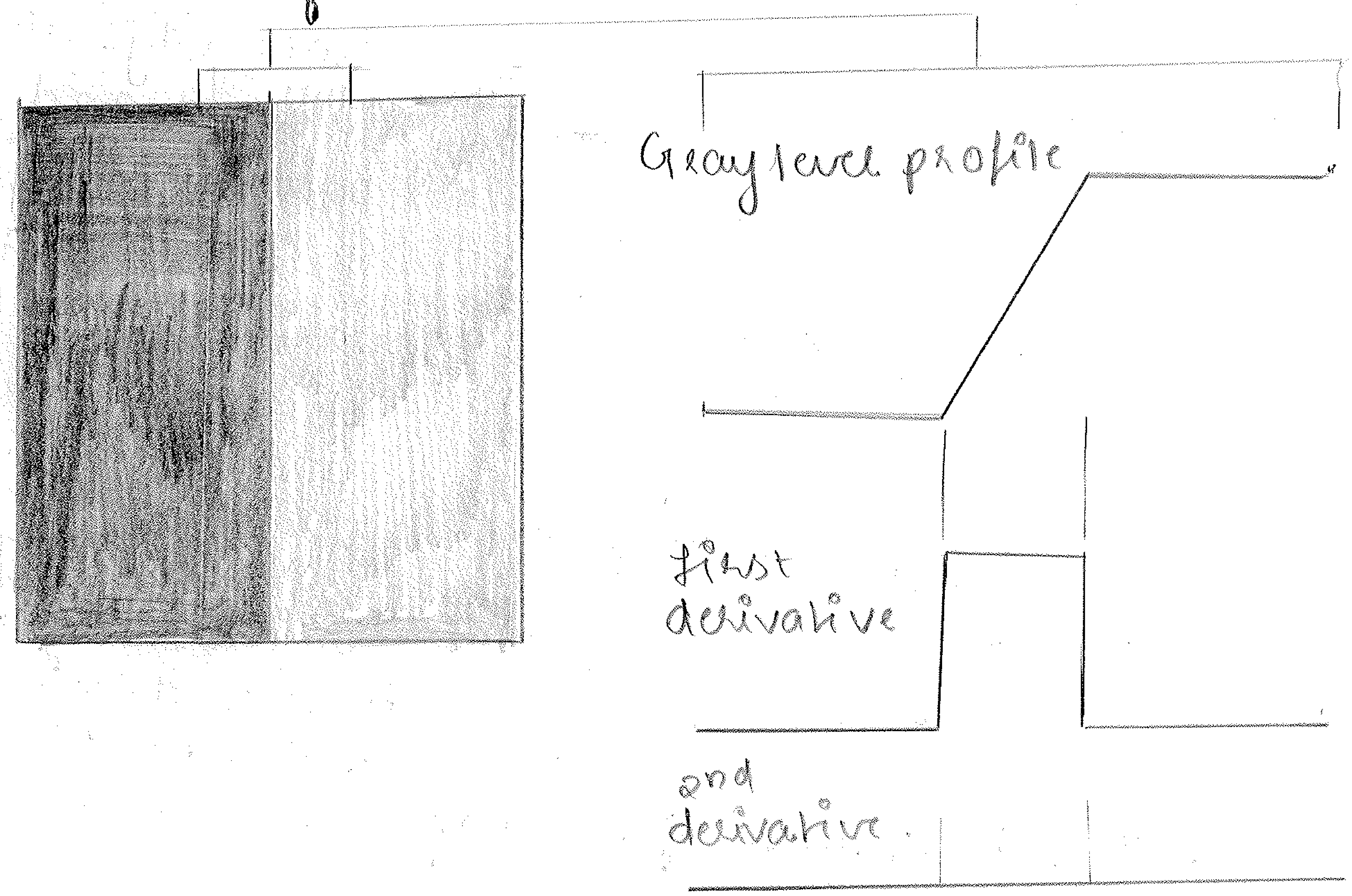
Use of the gradient operators

First order derivative is positive at the points of transition into and out of the ramp as move from left to right

* It is constant for points in the ramp.
* Zero in areas of constant gray level.

Second-order derivative

* It is positive at the transition associated with the dark side of the edge, and zero along the ramp and in areas of constant gray level.



Properties of 2nd order derivatives

* It produces two values for every edge in an image.
* An imaginary straight line joining the extreme positive and negative values of the 2nd derivative would cross zero near the midpoint of the edge.

This zero crossing property of 2nd derivative is quite useful for locating the centers of thick edges

Gradient Operators

* A gradient is a 2D vector that points to the direction in which the image intensity grows fastest.
* The tool of choice for finding edge strength and direction at location (x,y) of an image, f is the gradient, denoted by ∇f and defined as the vector.

$$\nabla f = \text{grad}(f) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Magnitude of vector ∇f

denoted as $M(x,y)$ defined $M(x,y)$

$$M(x,y) = \text{mag}(\nabla f) = \sqrt{G_x^2 + G_y^2}$$

Direction of the gradient.

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

i) Roberts cross gradient operators.

The Robert masks which are used for this calculations are

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

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

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

(a)
image.

-1	0
0	1

(b)

0	-1
1	0

(c).

Roberts.

ii) Prewitt operators

The simplest digital approximations to the partial derivatives using masks of size 3×3 is given by.

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

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

* The difference between third and first rows of the 3×3 region approximation of the derivative in the x-direction and the diffience between the third and first columns approximates the

derivative in the y direction. These masks are called the Prewitt operators.

iii) Sobel operator

A slight variation of the preceding two equations uses a weight of 2 in the center coefficients

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

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

The value 2 in the center location provides image smoothing. These masks are called the Sobel operators.

iv) Laplacian Operators

The Laplacian of a 2D function $f(x,y)$ is a second order derivative expressed as

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

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

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

Laplacian masks.

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

The digital approximation including the diagonal neighbours is given by the mask

$$\nabla^2 f = 8z_5 - [z_1 + z_2 + z_3 + z_4 + z_5 + z_6 + z_8 \; z_9]$$

Edge detection by Laplacian

The Laplacian cannot be used for edge detection in its original form due to:

* It is highly sensitive to noise
* Its magnitude produces double edges which makes the segmentation difficult.
* It is unable to detect edge direction.

The Laplacian can only be used for two purpose in segmentation.

* To find edge location using zero crossing property.
* To find whether a pixel is on the dark or light side of an edge.

Laplacian of a Gaussian (LoG)

To detect edges using a zero crossings, the Laplacian is convolved with the function $h(r)$. This blurs ie smoothens the image.

$$h(r) = -e^{-r^2/2\sigma^2}$$

$$r^2 = x^2 + y^2.$$

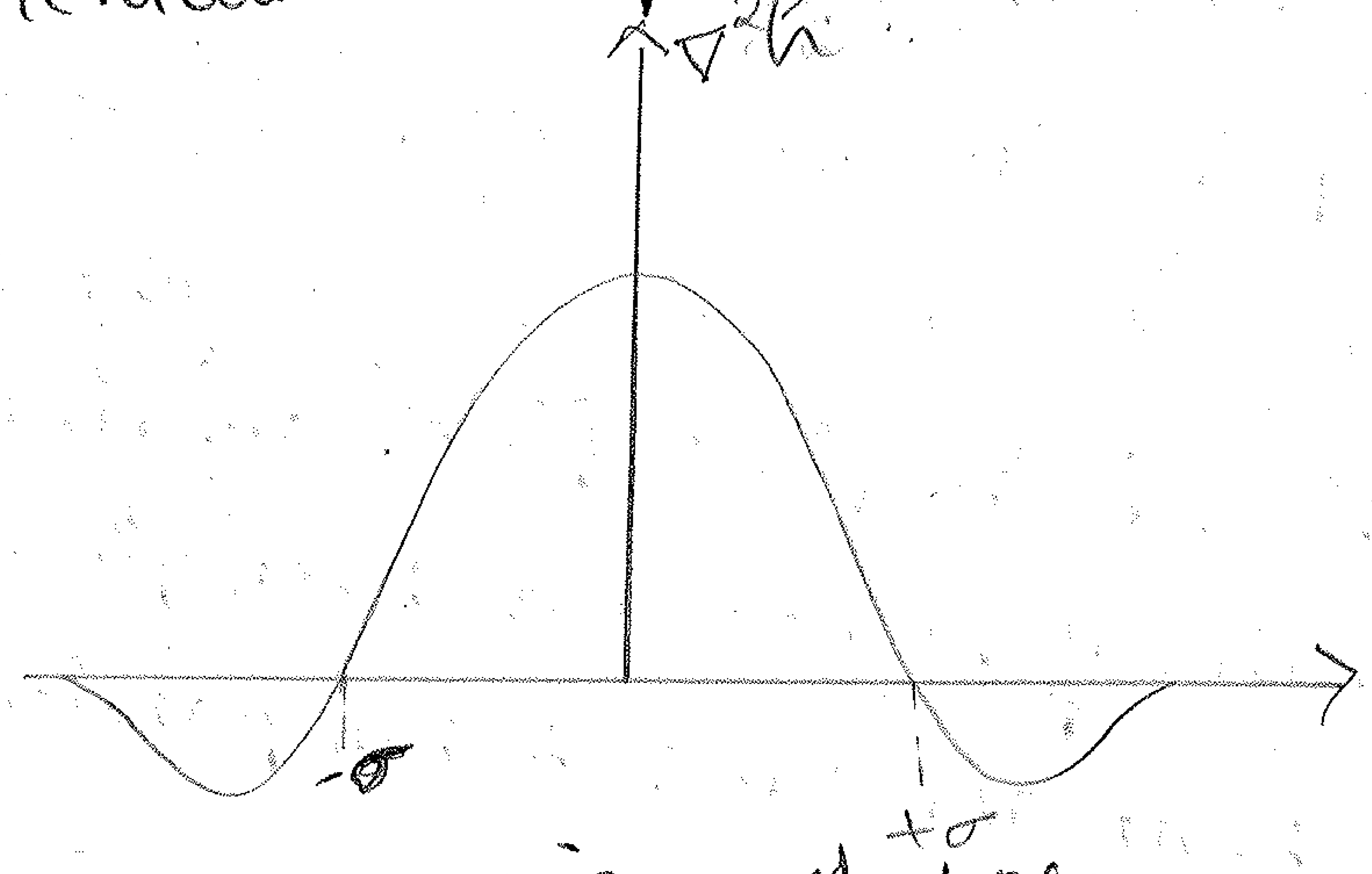
σ - Standard deviation which determines the degree of blurring.

$$\nabla^2 h(r) = -\left[\frac{r^2-\sigma^2}{\sigma^4}\right] e^{-r^2/2\sigma^2}$$

This equation forms a gaussian function, this function is referred as "Laplacian of a Gaussian" (LoG). This method has good noise reduction capacity.

Mexican Hat Function

The cross section of LoG, due to its shape of a hat, the "Laplacian of a Gaussian" (LoG) is also called "Mexican hat function".



cross section of LoG.

Edge Linking via Hough Transform

* Edge detectors do not produce continuous edge.
* The detected edges are not sharp and continuous due to the presence of noise and intensity variation.

The two principal ~~prop~~ properties used for establishing similarity of edge pixels.

* The strength of the response of the gradient operator used to produce the edge pixel.

* The direction of the gradient vector.

i) Global processing via Hough Transform

This method of edge linking determines whether or not a set of edges lie on a line.

* This method then links the edges by producing the line/curve.

* Given n points in an image, find the subset of these points that lies on straight lines.

* One solution is to first find all lines determined by every pair of points and then find all subset of points that are close to the particular lines

Drawback

* It involves finding $n(n-1)/2 \sim n^2$ lines

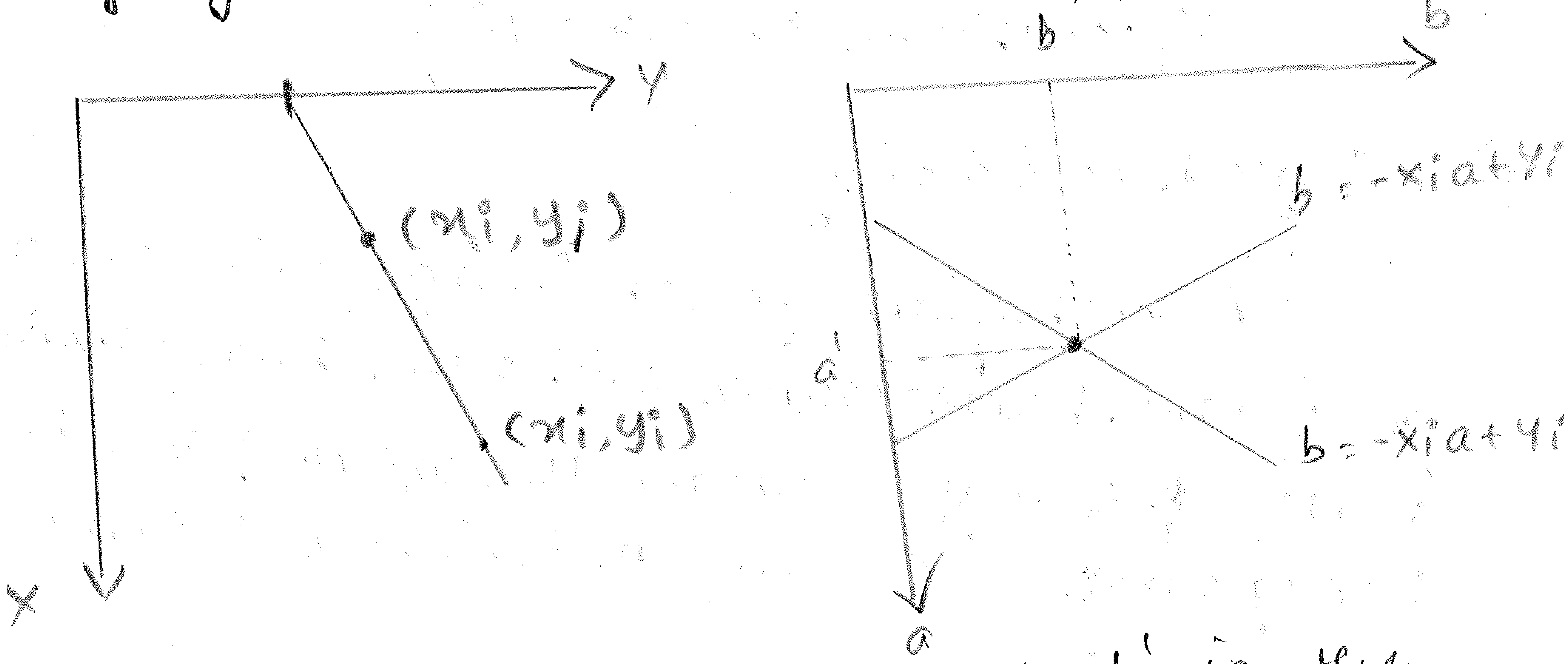
* Perform $n(n(n-1))/2 \sim n^3$ comparison of every points to all lines.

Hough Transform

* Hough proposed an alternative approach referred to as the Hough transform.
* Consider a point (x_i, y_i) and the general equation of a straight line in slope intercept form m,

$$y_i = ax_i + b.$$

* Many lines pass through (x_i, y_i) but they all satisfy the equation $y_i = ax_i + b$ for varying values of a and b.



where a' is the slope and b' is the intercept.

Hough transform algorithm

* Compute the gradient of an image and threshold it to obtain a binary image.

* Specify subdivision in the ρθ plane.
* Examine the counts of the accumulator cells for high pixel concentrations.
* Examine the relationship between pixels in a chosen cell.

Thresholding

It is used to seperate the objects present in an image from its background.
Depending on the objects present, thresholding is divided into.

i] Single thresholding
ii] Multilevel thresholding.

↱ Single thresholding

A histogram for an image with an object on a dark background will have two dominant modes. One way to extract the object from the background is to select a threshold T that seperate these modes.

↱ Any point (x,y) for which $f(x,y) > T$ is called object point.

↱ Any point (x,y) for which $f(x,y) < T$ is called background point.

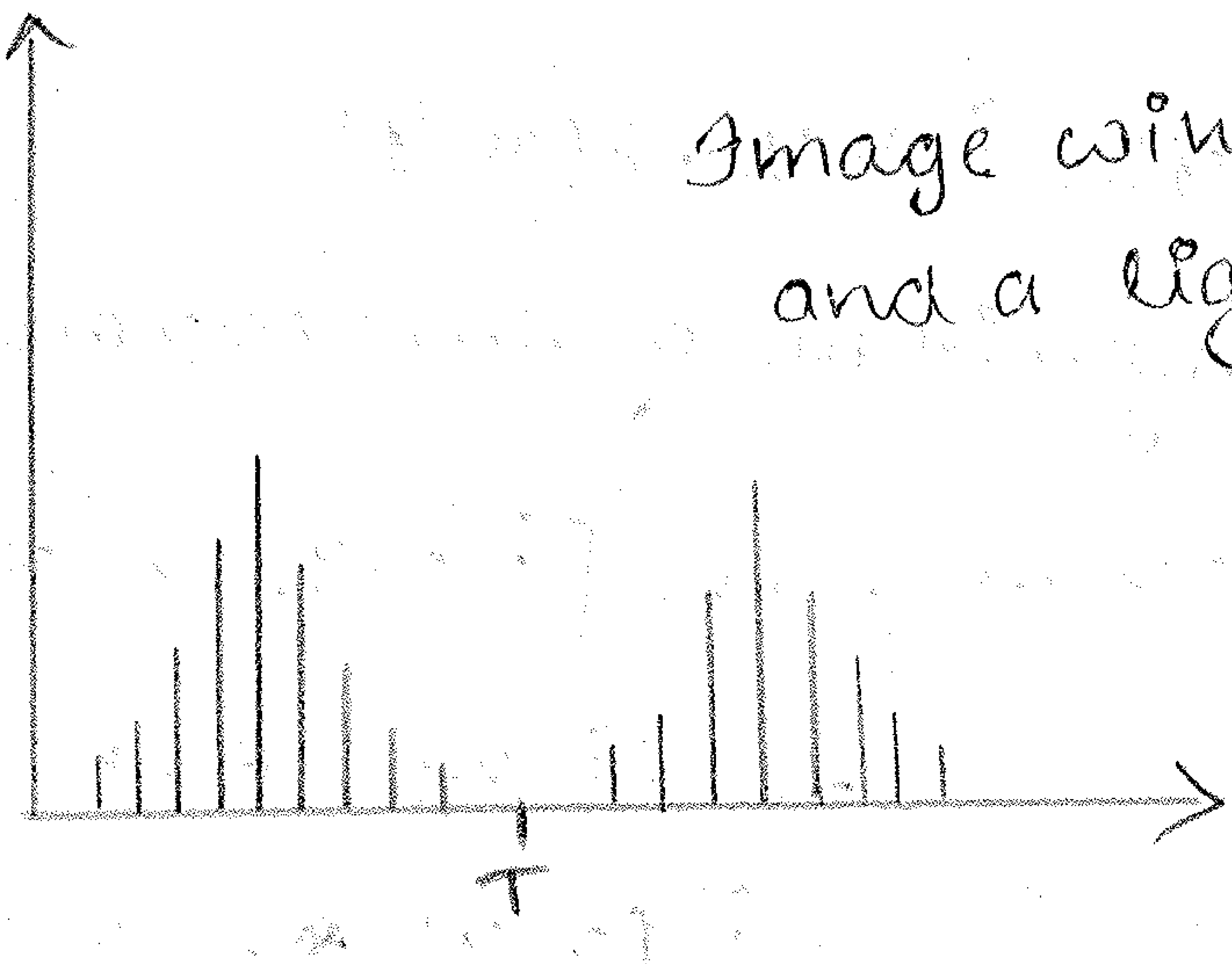


Image with dark background and a light image

→ Multilevel thresholding

* To an object class if $T_1 < f(x,y) \le T_2$
* To another object class if $f(x,y) > T_2$
* To background if $f(x,y) \le T_1$
* Two types of light object on a dark backgroun

The thresholded image $g(x,y)$ is defined as.

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

When T depends only on $f(x,y)$, the threshold is called global. If T depends on both $f(x,y)$ and $p(x,y)$, the threshold is called local. ~~This~~ The threshold called dynamic or adaptive, T depends on the spatial coordinates x and y.

Illumination

The image $f(x,y)$ is formed as the product of a reflectance component $r(x,y)$ and illumination component $i(x,y)$.

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

Taking logarithm of this equation

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

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

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

There are types of illumination

→ Uniform illumination

Here $i(x,y)$ is constant and thus $i'(x,y)$ is also a constant whose histogram will be a simple spike similar to an impulse.

→ Non-uniform illumination

- In this case, $i'(x,y)$ may have a broader histogram, which smear the basic shape of the histogram of $r'(x,y)$.
- The degree of distortion depends on the broadness of the histogram of $i'(x,y)$ which is proportional to the non-uniformity of the illumination function.

To compensate for the non-uniformity, the illumination pattern is to be projected on a constant, white reflective surface. This gives an image.

$$g(x,y) = K\, i(x,y)$$

↳ Global Thresholding

* It is the simplest and most widely used for all possible segmentation methods.
* When the intensity distribution of objects and background pixels are sufficiently.
* Segmentation is then accomplished by scanning the image pixel by pixel and labeling each pixel as object or background.

Algorithm:

↳ Select an initial estimate for the global threshold, T.

↳ Segment the image using T, forms two groups G_1 and G_2.

↳ Compute the average intensity values μ_1 and μ_2 for the pixels in G_1 and G_2.

↳ Compute a new threshold value.

$$T = \frac{1}{2}\left[\mu_1 + \mu_2\right]$$

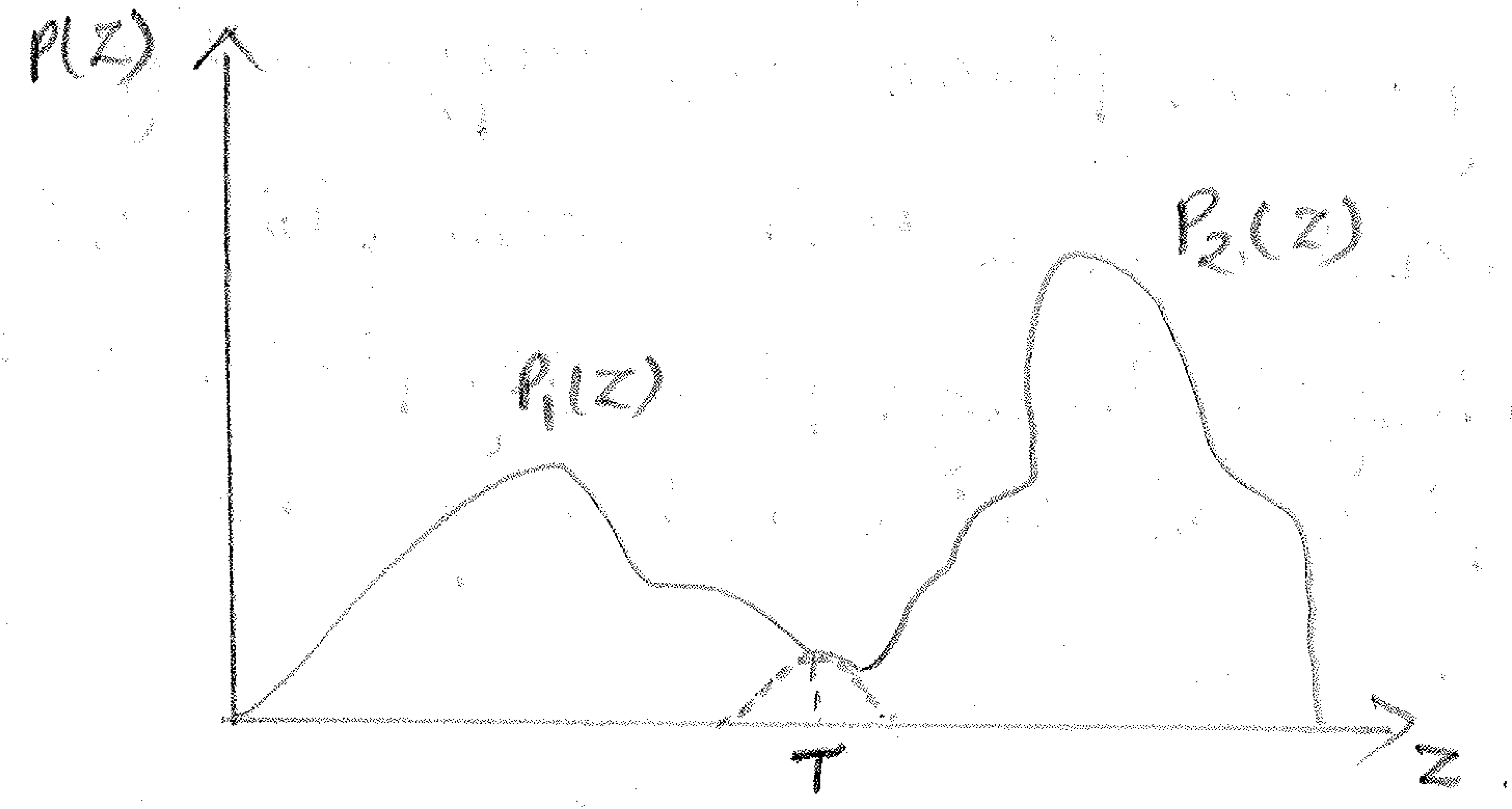
↳ Repeat 2 to 4 until the difference between values of T in successive iterations is smaller that a predefined parameter T_0.

↳ Adaptive Thresholding

* It is also known as dynamic thresholding
* In adaptive thresholding, the image is divided into many ~~of~~ overlapping subimages.

* utilize a different threshold to segment each subimages.

Since the threshold used for each pixel depends on the location of the pixel in terms of the subimages, this type of thresholding is adaptive.



$$p(z) = P_1 \cdot p_1(z) + P_2 \cdot p_2(z) + \ldots$$

$P_1 \rightarrow$ Probability that a random pixel is an object pixel.

$P_2 \rightarrow$ Probability that a random pixel is a background pixel.

$$P_1 + P_2 = 1$$

The probability of erroneously deciding a background point as a object point is

$$E_1(T) = \int_{-\infty}^{T} p_2(z)\,dz.$$

Similarly, the probability of erroneously deciding an object point as background is

$$E_2(T) = \int_{T}^{\infty} p_1(z)\,dz.$$

The overall probability of error is

$$E(T) = P_2 E_1(T) + P_1 E_2(T).$$

→ If object points will never occur, $P_1 = 0$

→ If background points will never occur, $P_2 = 0$

→ If both object and background points are equally likely to occur $P_1 = P_2 = 0.5$.

Gaussian density,

$$P(z) = \frac{P_1}{\sqrt{2\pi}\sigma_1} e^{(z-\mu_1)^2/2\sigma_1^2} + \frac{P_2}{\sqrt{2\pi}\sigma_2} e^{-(z-\mu_1)^2/2\sigma_2^2}$$

Where μ_1 and μ_2 are mean, σ_1 and σ_2 are the variance of the gaussian density.

Using this gaussian density, the solution for T is obtained as

$$AT^2 + BT + C = 0$$

where, $A = \sigma_1^2 - \sigma_2^2$

$B = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)$

$C = \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 4\sigma_1^2\mu_2^2 \ln\left(\frac{\sigma_2 P_1}{\sigma_1 P_2}\right)$

→ If the variance are equal, single optimal threshold is obtained as

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln\left(\frac{P_2}{P_1}\right)$$

→ If $P_1 = P_2$ or $\sigma = 0$, the optimal threshold is the average of the means given by

$$T = \frac{\mu_1 + \mu_2}{2}.$$

Minimum mean square error approach

$$e_{ms} = \frac{1}{n} \sum_{i=1}^{n} [p(z_i) - h(z_i)]^2$$

where $p(z_i)$ - mixture density [continuous]

$h(z_i)$ - n point image histogram [discrete].

Boundary characteristics for histogram improvement

* If the histogram peaks are tall, narrow, symmetric and seperated by deep valleys a good threshold can be selected which leads to an efficient segmentation.

* Therefore, the shape of the histogram should be enhanced, which is possible by the use of boundary characteristics of an image.

Ways for improvement

i] The shape of the histogram can be improved by considering only the pixels that lie on ~~[illegible]~~ near the edges between objects and background ii] All the peaks of the histogram will have almost equal height if only the pixels on or near the edges are used. Thus, it will improve the symmetry of the histogram peaks. 21

→ Local thresholding

It is a method used to obtain detailed information of the image. It uses the gradient and Laplacian operators.

Gradient, ∇f

The gradient of a pixel is used to find whether that pixel is on an edge or not.

$$\nabla f = \sqrt{G_x^2 + G_y^2} = |G_x| + |G_y|$$

Laplacian, $\nabla^2 f$.

The Laplacian provides information, whether a pixel lies on the dark or light side of an edge can be found.

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

$$= 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

By calculating the gradient and laplacian at every point (x,y), a three level image can be formed as,

$$s(x,y) = \begin{cases} 0, & \nabla f < T \\ +, & \nabla f \geq T \text{ and } \nabla^2 f \leq 0 \\ -, & \nabla f \geq T \text{ and } \nabla^2 f < 0 \end{cases}$$

where $T \rightarrow$ Threshold

$0, +, - \rightarrow$ different gray levels.

* All pixels that are not on an edge are labeled as '0'.
* All pixels on the dark side of an edge are labeled '+'.
* All pixels on the light side of an edge are labeled '-'.

→ Multispectral thresholding

In previous techniques, only gray levels are used for thresholding. The thresholding technique which uses more than one variable to characterize each pixel is called multispectral thresholding.

Consider a colour image in which each pixel is characterized by the RGB values. The procedure for multispectral thresholding is

i) Construct a 3D histogram of the image with each pixel characterized by RGB values.

ii) Thresholding finds cluster ie group of points in 3D space.

iii) Segment image by assigning one arbitrary value to pixels whose RGB components are closer to one cluster and arbitrary value to other pixels in the image.

Advantages

This method can be easily extended to more components ie variables and clusters.

Disadvantages

As the number of variables increases, cluster searching becomes more complex.

Region based segmentation

It is entirely different from other segmentation techniques.

* Detection of discontinuities method finds boundaries between regions based on discontinuities in intensity levels.

* Thresholding methods segments an image using thresholds based on the pixel properties like gray level values or colour.

Condition

a) $\bigcup_{i=1}^{n} R_i = R$

b) R_i is a connected set.

c) $R_i \cap R_j \neq \Theta$ for i and j where Θ is the null set

d) $P(R_i) =$ True for $i = 1, 2, \ldots n$

e) $P(R_i \cup R_j) =$ False for $i \neq j$.

Region Growing

One of the major disadvantages of the thresholding algorithms is that they will produce isolated regions. Thus it is necessary to produce coherent regions. the segmented image to produce coherent region

* Region growing is a procedure that group pixels or subregions into larger regions.

* The simplest of these approaches is pixel aggregation, which starts with a set of 'seed' points.

Similarity criteria

* Intensity values
* Texture
* Color
* Size.
* Shape of the region being grown

There are 3 main issues during region based segmentation

* Selection of seed points.

This selection depends on the type of problem given. When no proper information is available, set of same properties are computed for every pixel.

* Selection of similarity criteria

It depends upon both given problem as well as the type of image data available. Here the pixels are grouped without considering connectivity properties.

* Formulation of a stopping rule.

The region growing should be stopped when there is no more pixels which satisfy the similarity criteria for that particular region.

Region Splitting and Merging

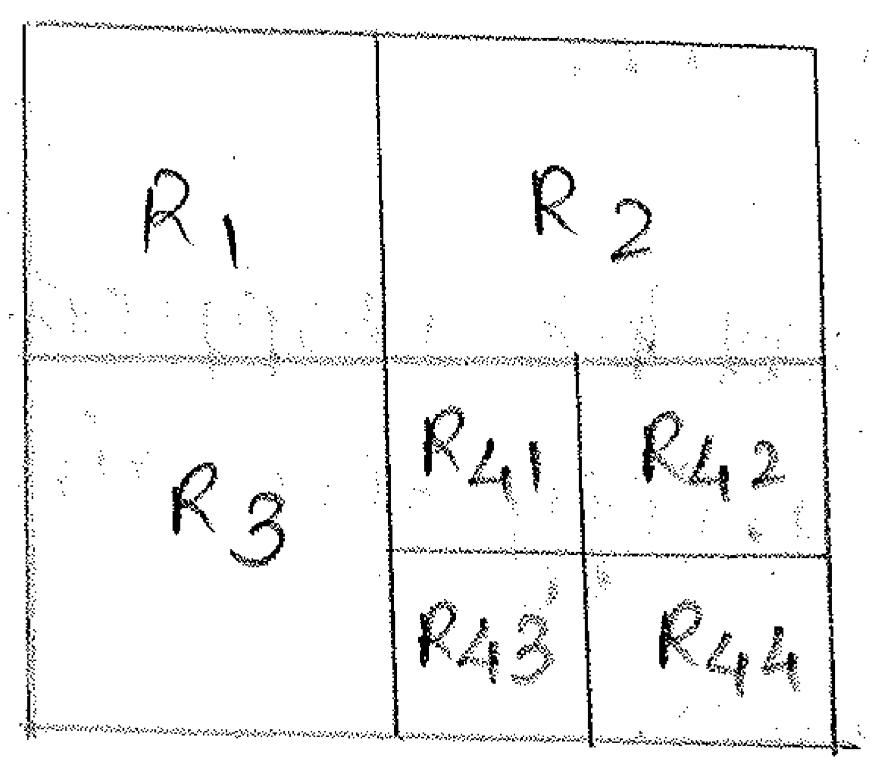
Region splitting and merging is a segmentation process in which the image is subdivided.

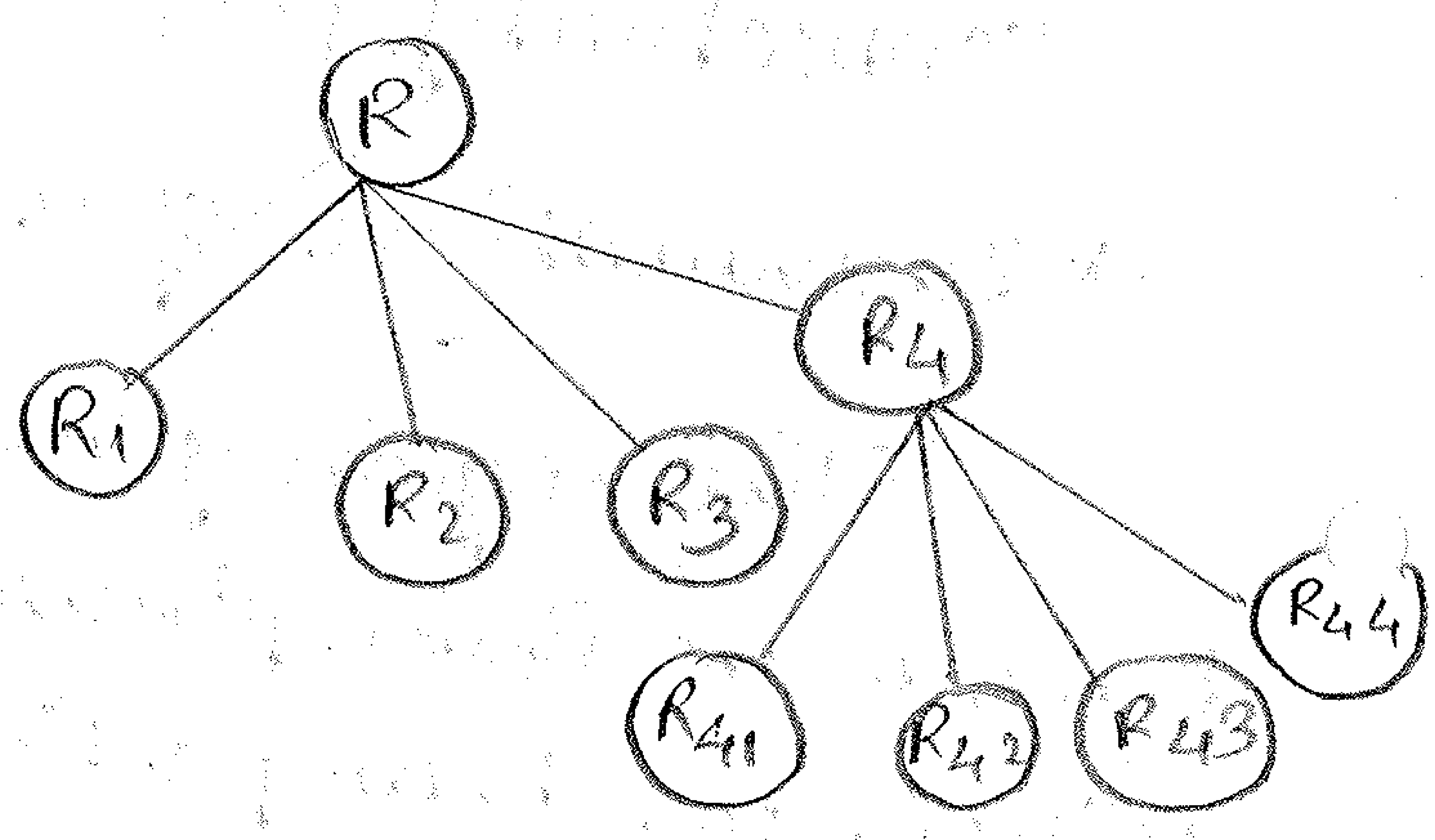
Region Splitting

Let R represent the entire image region and select a predicate P. One approach for segmenting R is to subdivide it successively into smaller and smaller quadrant regions.

If $P(R) = \text{False}$, divide the image into quadrants.

If P is false for any quadrant, subdivide into subquadrants and so on. This particular splitting technique has a convenient representation in the form of a 'quadtree'.





Region Merging

The combined pixels of two adjacent regions satisfy the predicate P, they are merged. Consider two adjacent regions R_j and R_K are merged

$$P(R_j \cup R_K) = \text{True}.$$

Algorithm

* Split into 4 disjoint quadrants any region R_i where $P(R_i) = False$.
* Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k) = True$.
* Stop when no further merging or splitting is possible.

Advantages

* Both of them uses same ~~quadtree~~ quadtree for splitting as well as merging.

Morphological image processing

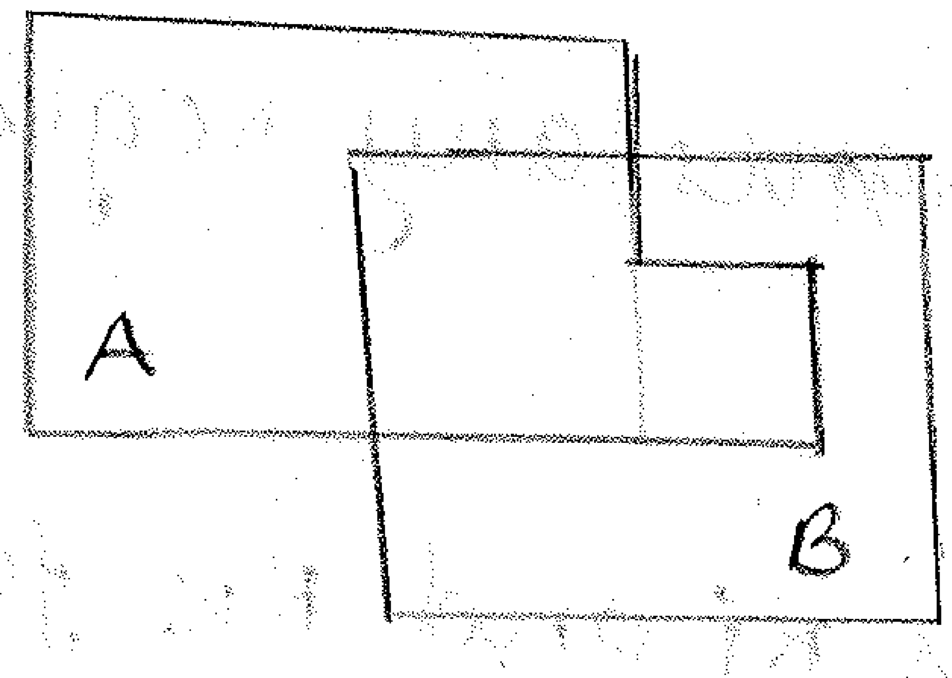
Morphology: Denotes a branch of biology that deals with the form and structure of animals and plants.

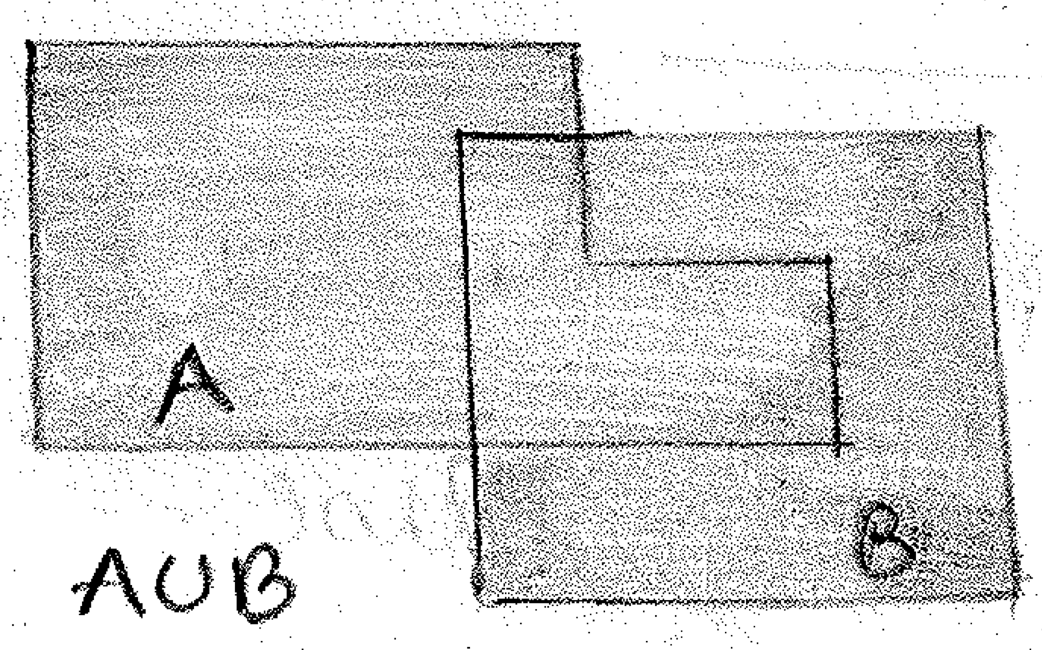
Mathematical morphology represent objects in an image:

i) Binary image → '0' for black and '1' for white.

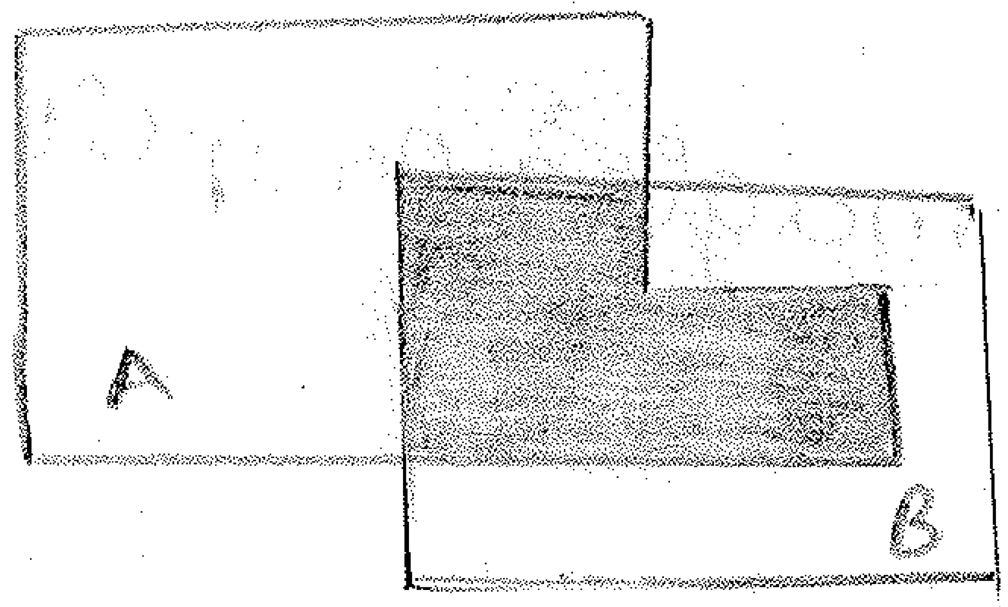
ii) Gray scale image

iii) Several operations such as union, intersection etc.

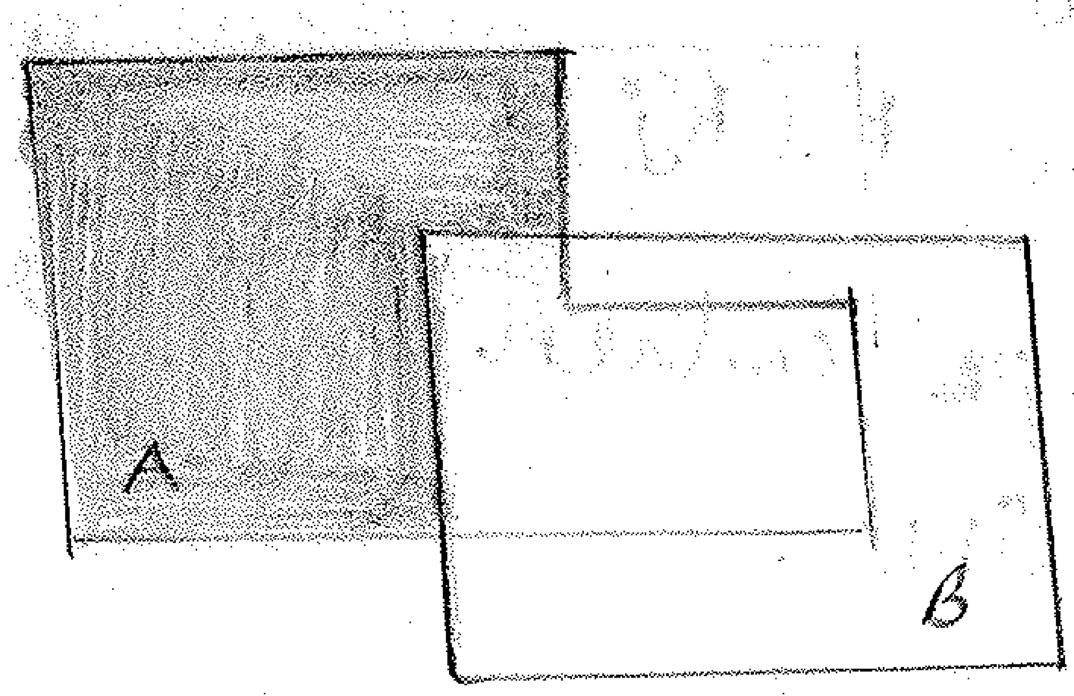




AUB



A∩B



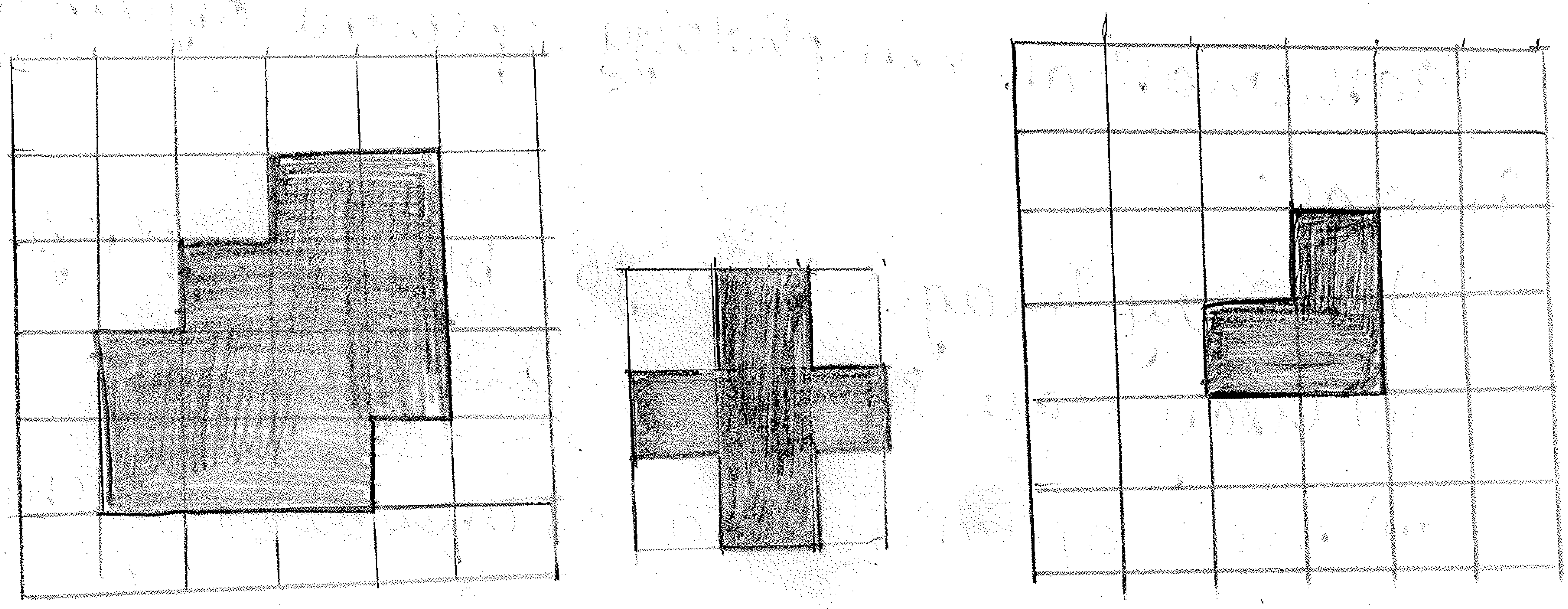
A-B

EROSION AND DILATION

Erosion

* It is used for shrinking of element A by using element B.
* Erosion for set A and B in Z^2, is defined by

$$A \ominus B = \{ z \mid (B)_z \subseteq A \}.$$



Segmentation using morphological watersheds

* Watershed are defined as the lines seperating catchment lines which belongs to different minima.

* The region that the watershed seperates are called catchment basins.

* Based on visualizing an image in 3 dimensions.
  a) Points belong to the regional minimum.
  b) Points at which a drop of water, if placed at the location of any of those points, would fall with certainity to a single minimum.
  c) Points at which water would be equally likely to fall to more than one such minimum.

Catchment Basin / Watershed

* For a particular regional minimum, the set of points satisfying condition 'b' is called the catchment basin or watershed of that minimum.

* The points satisfying condition 'c' form crest lines on the topographic surface are termed divide lines or watershed lines.

If a hole is punched in each regional minimum and that the entire topography is flooded from below by letting water rise through the holes at a uniform rate.

* When the rising water is distinct catchment basins is about to merge, a dam is built to prevent the merging.
* The water is still allowed to rise and it finally reaches a stage at which only the top of the dams are visible above the water.
* These dam boundaries correspond to the divide lines of the watershed.

Application

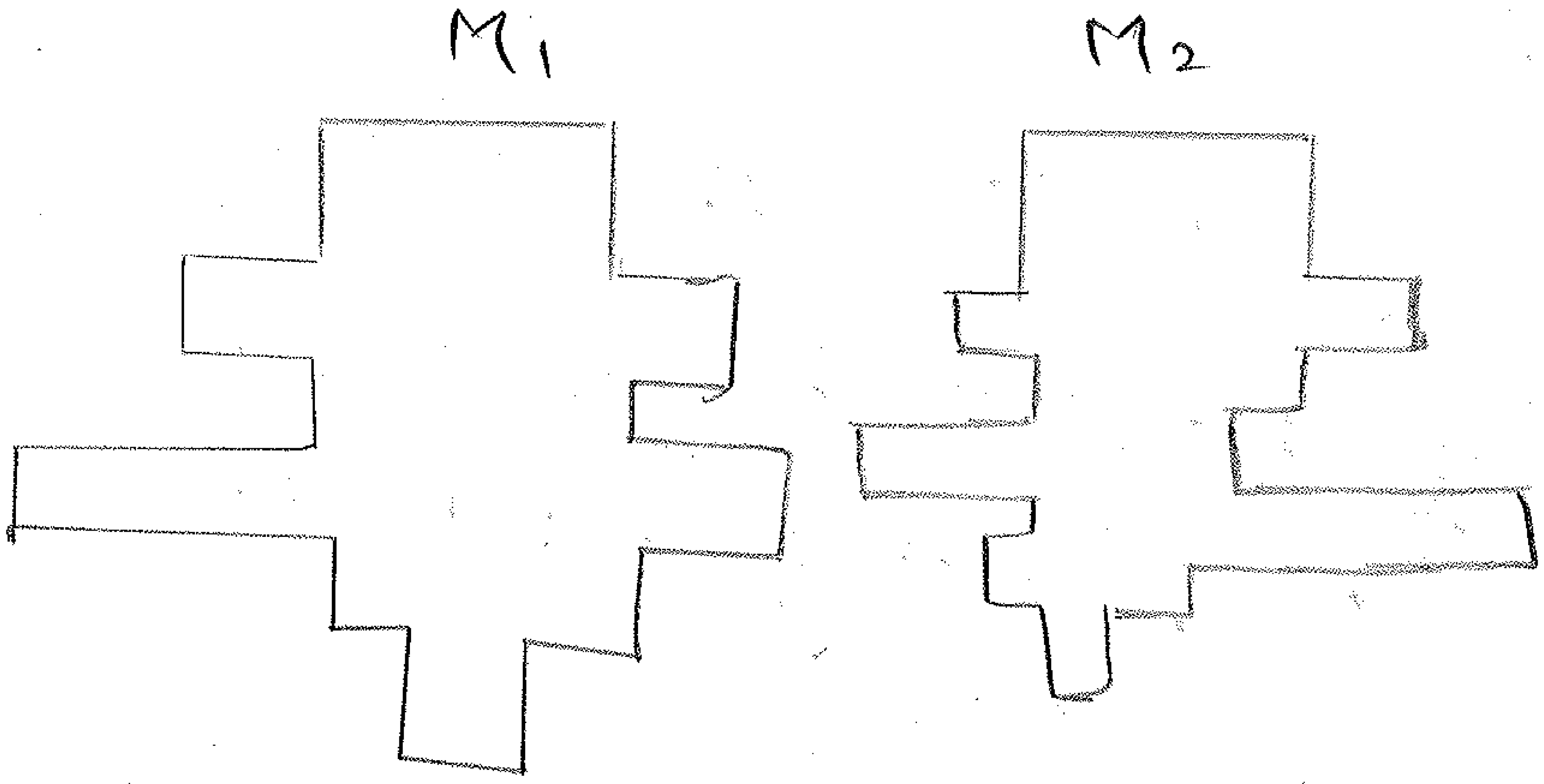
* Extraction of nearly uniform objects from the background.

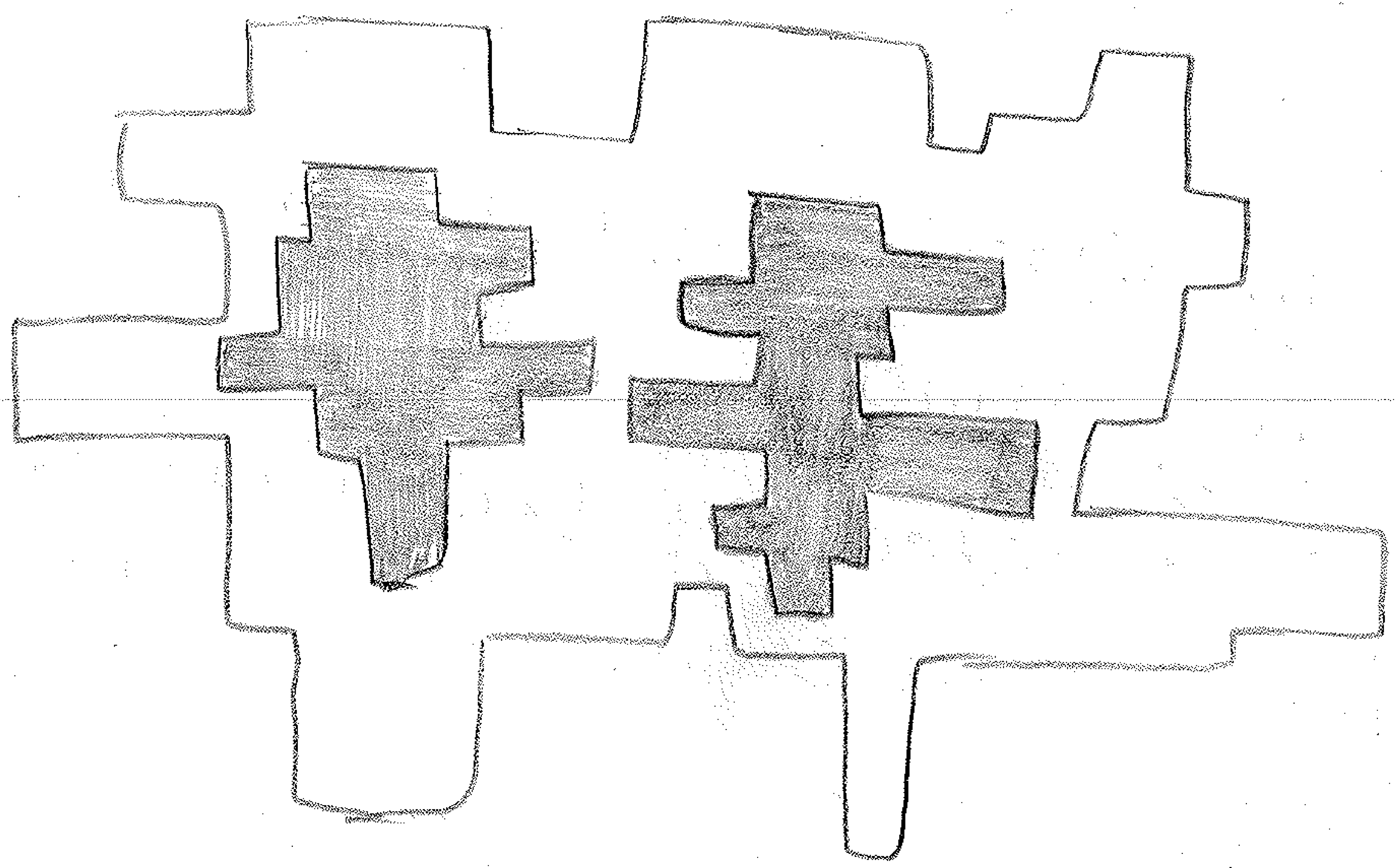
DAM CONSTRUCTION

* It is very important part of watershed segmentation technique
* Based on binary morphological dilation.

→ Dam construction is based on binary images which are members of 2D integer space Z^2.

→ The simplest way to construct dams separating sets of binary points is to use morphological dilation.





At each step of the algorithm, the binary image is obtained in the following manner.

→ Initially, the set of pixels with minimum gray level are 1, others 0.

→ In each subsequent step, we flood the 3D topography from below and the pixels covered by the rising water are 1s and others 0s.

→ M_1 and M_2 represent the sets of coordinates of points in the two regional minima.

→ $C_{n-1}(M_1)$ and $C_{n-1}(M_2)$ represents the sets of coordinates of points in the catchment basin

→ $C[n-1]$

union of $C_{n-1}(M_1)$ and $C_{n-1}(M_2)$

* At flooding step n-1, there are two connected components. At flooding step n, there is only one connected component.

* use 'q' to denote the single connected component.

* Repeatedly dilate $C_{n-1}(M_1)$, $C_{n-1}(M_2)$ by 3×3 structuring element.

* Dam is constructed by the points on which the dilation would cause the sets being dilated to the image.

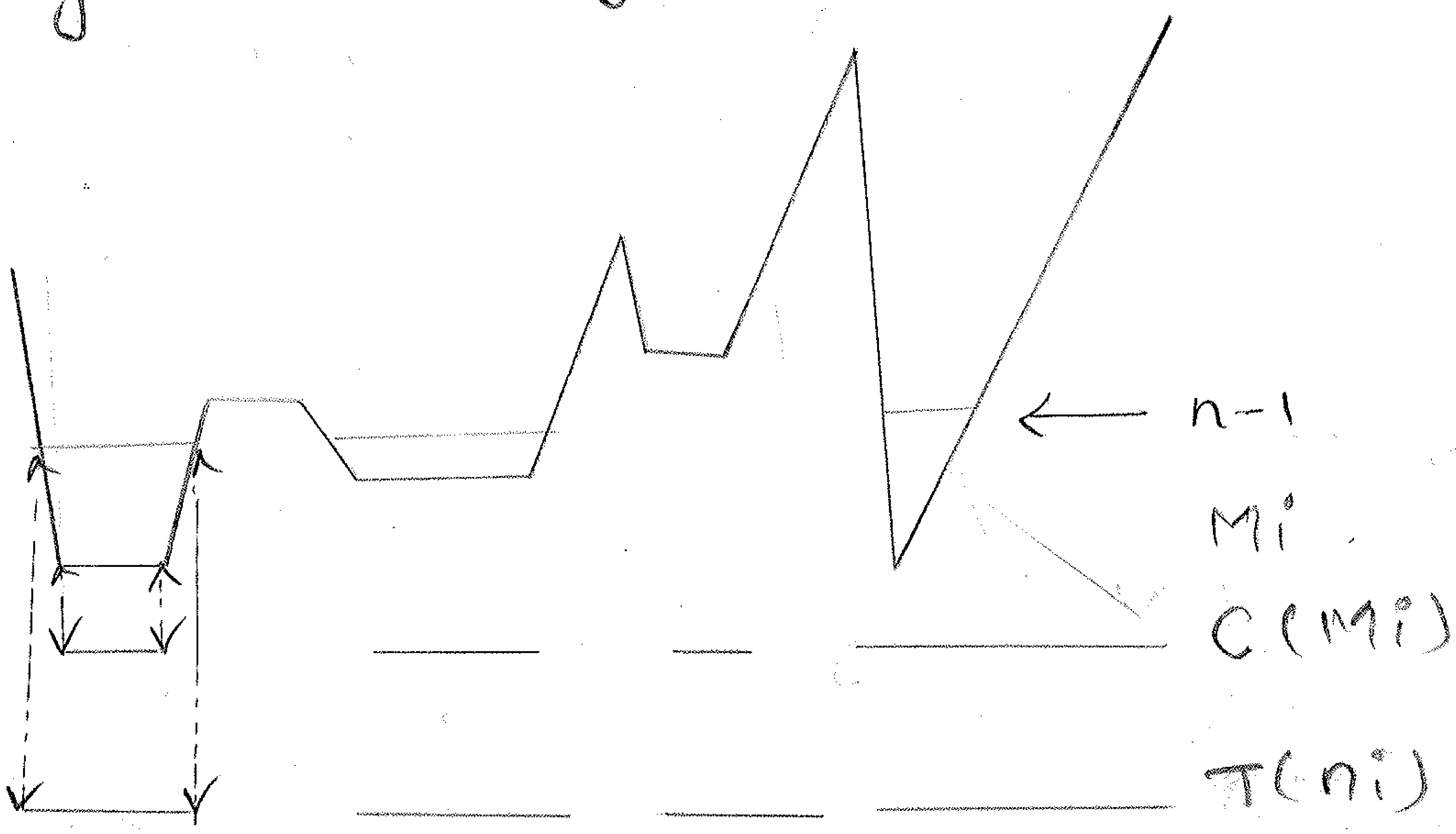
* Setting the gray level at which each point in the resultant path to a value greater than the maximum gray value of the image.

Watershed Segmentation Algorithm

* Powerful tool for image segmentation.

* Here, the image is considered as a topographic surface

* Gray level of the image represents the altitudes
* The region with constant gray level constitutes the flat zone of an image.
* Region edges corresponds to high watersheds and low gradient region interiors correspond to catchment basins.
* catchment basins represents the region of the segmented image
*



Watershed transform.

Algorithm

* Let $M_1, M_2, \ldots M_k$ be set of denoting the coordinates of the points in the regional minima of the image $g(x,y)$.
* $C(M_i) \rightarrow$ coordinates of the points in the catchment basin associated with regional minima

* $T[n] = \{(s,t) / g(s,t) < n\}$.

* $C_n(M_i)$ denote the set of coordinates of points in the catchment basins associated with minima M_i.

$$C_n(M_i) = C(M_i) \cap T[n].$$

* union of all catchment basins.

$$C(max+1) = \bigcup_{i=1}^{R} C_n[M_i].$$

It is known that $C[n-1]$ is a subset of $C[n]$, becau the flooded portion of the catchment basin will be more at stage n than the previous stage $(n-1)$.

Markers

* An approach used to control oversegmentation is based on the concept of markers.
* It is a connected component belonging to an image.

There are two types of markers available.

⇒ Internal markers.

These are related with the objects of interest in an image.

⇒ External markers.

These are related with the background of an image.

Marker selection

It consist of two steps.

i) Preprocessing

ii) Definition of a set of criteria that the markers must satisfy.

→ Because of their size, many of these minima really are irrelevant detail.

→ An effective method for minizing the effect of small spatial detail is to filter the image with a smoothing filter.

Watershed based image segmentation

* use internal markers to obtain watershed lines of the gradient of the image to be segmented.
* use the obtained watershed lines as external markers.
* Each region is defined by the external markers and contains a single internal marker and part of the background.

Advantages

* It produces more stable segmentation results.
* This method eliminates the problems of broken segmentation lines.

* This approach provides a simple framework for incorporating knowledge based constraints in the segmentation process.

UNIT V

IMAGE COMPRESSION AND RECOGNITION

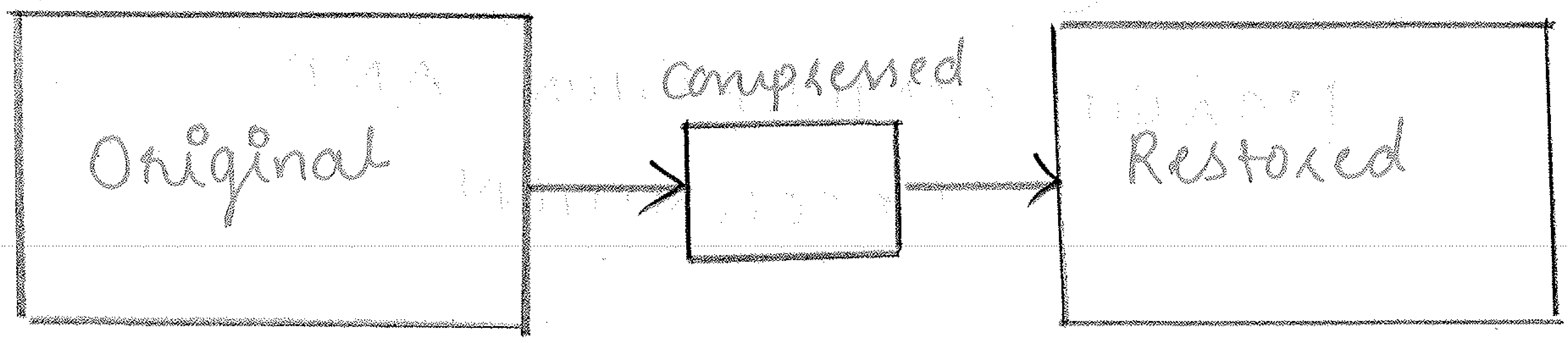
Image compression

It is defined as the process of reducing the amount of data needed to represent a digital image. This is done by reducing/removing the redundant data.

Data compression can be classified into two types
i) Lossless Data compression.
ii) Lossy Data compression.

Lossless data compression

* Compress data/file without losing an original files quality and data.
* Here, the file size is reduced but the quality of data remains the same.
* Able to restore the original data in its original form after the decompression.
* Mainly used in sensitive documents, confidential information and PNG, BMP file formats.

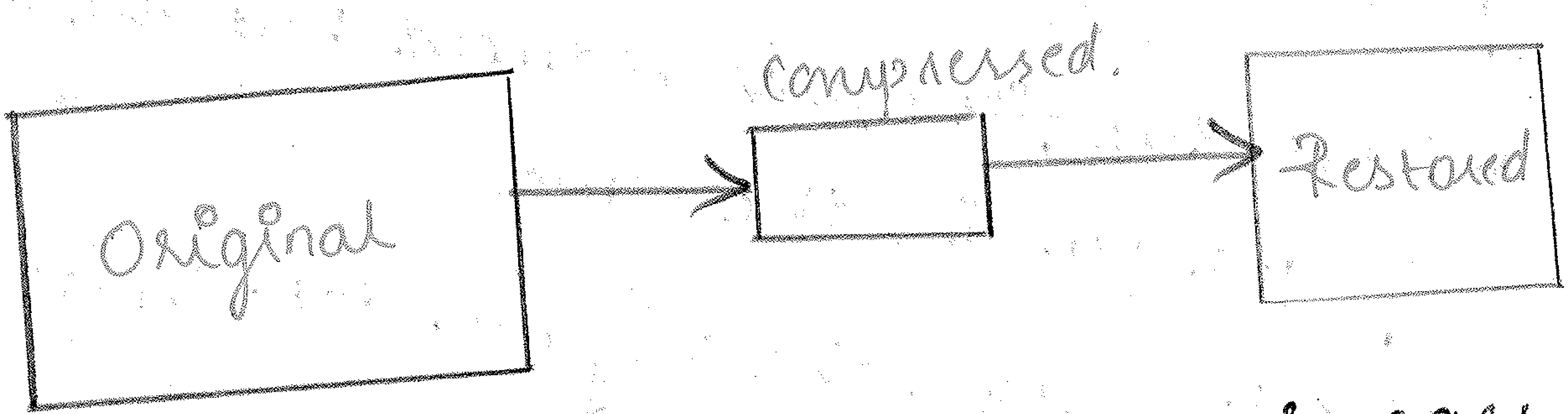


Lossless data compression techniques are

* RunLength Encoding [RLE]
* LZW [Lempel Ziv-Welch]
* Huffman coding
* Arithmetic coding

Lossy data compression

* It is used to compress larger file into smaller files.
* Here, some specific amount of data and quality are removed from the original file.
* It takes less memory space from the original file due to the loss of original data and quality



* It is most widely used in JPEG images, MPEG video and MP3 audio formats.

Lossy data compression techniques are

* Transform coding
* DCT [Discrete cosine Transform]
* DWT [Discrete wavelet Transform].

Applications

* Tele video
* Remote sensing
* Facsmile Transmission [FAX].
* Document and medical imaging.

$$\boxed{\text{Data} = \text{Information} + \text{Redundant data}}$$

Need for data compression

* To speed up file transfer
* To save storage capacity
* To decrease costs for storage hardware and network bandwidth.

For eg:, if we want to store a 1600 X 1200 colour image then the space required to store the image is

$1200 \times 1600 \times 8 \times 3 = 46,080,000$ bits

$= 5,760,000$ bytes

$= \underline{5.76 \text{ Mbytes}}$

The maximum space available in one floppy disk is 1.44Mb.

For 3 floppies $= 3 \times 1.44$

$= \underline{5.76\text{Mb}}$

Thus minimum 4 floppies are required to store an RGB image of size 1600x1200.

Redundancy

Data redundancy occurs when the same piece of data exist in multiple places.

The relative data redundancy

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

where C_R is ~~the~~ common called compression ration

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

Where n_1 and n_2 are numbers of information carrying units in two datasets.

Case 1

$n_1 = n_2$, $C_R = 1$ and $R_D = 0$

Case 2

$n_2 << n_1$, $C_R = \infty$ and $R_D = 1$

Case 3

$n_2 >> n_1$, $C_R = 0$ and $R_D \to -\infty$.

$$\text{Compression Ratio} = \frac{\text{Message size before compression}}{\text{Code size after compression}}$$

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

Question 1

If the original size is 256x256 pixels 8 bits per pixel gray scale. This file is 65,536 in size. After compression the image file is 6554 bytes. The compression ratio is ______.

$$C_R = \frac{651536}{6554} = 9.999$$

$$\simeq 10$$

HUFFMAN CODING

* It is a type of variable length coding.
* The most popular technique for removing coding redundancy.
* optimal code.

Algorithm.

- List the symbols and sort them.
- Pick two symbols having the least probabilities.
- Create new code. Add new probabilities of the symbols. Selected in step and label the new code with it.
- Repeat step 2 and 3 until a reduced source with two symbols.
- Start assigning code 0 for the smallest symbol and code 1 for the other symbol.
- Working back to the original source.

Question :

0.1 0.4 0.06 0.1 0.04

Alphabets with probability {0.4, 0.2, 0.2, 0.1, 0.

for symbols $\{s_1, s_2, s_3, s_4, s_5\}$. Find

Step 1 & 2

Source reduction + code assignment.

Original source — Source reduction

Symbol	probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6 0
a_6	0.3	0.3	0.3	0.3 00	0.4 1
a_1	0.1	0.1	0.2 010	0.3 01	
a_4	0.1	0.1 0100	0.1 011		
a_3	0.06 01010	0.1 0101			
a_5	0.04 01011				

Average length of this code

$= 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 3 + 0.1 \times 4 + 0.06 \times 5 + 0.04 \times 5$

$= \underline{2.2}$ bits/symbol.

Advantages

* It provides the smallest number of code symbols per source symbols.
* It creates an optimal code for a set of symbols and probability.
* coding/decoding process can be done in a simple lookup table manner.

Disadvantages

* Only one symbols can be coded at a time.

* For N source symbols, it requires N-2 source reductions and N-2 code assignments, therefore the computational complexity is more.

RUN LENGTH CODING

Inorder to reduce interpixel redundancies, a binary image or bit plane can be compressed using three coding techniques, known as

* coding Area coding (CAC)
* One dimensional Run Length coding
* Two dimensional Run Length coding.

One dimensional RLC

* Effective when long sequences of the same symbol occurs.
* RLC exploits the spatial redundancy by coding the number of symbols in a run.
* RLC can be classified into two
  - 1D RLC
  - 2D RLC.

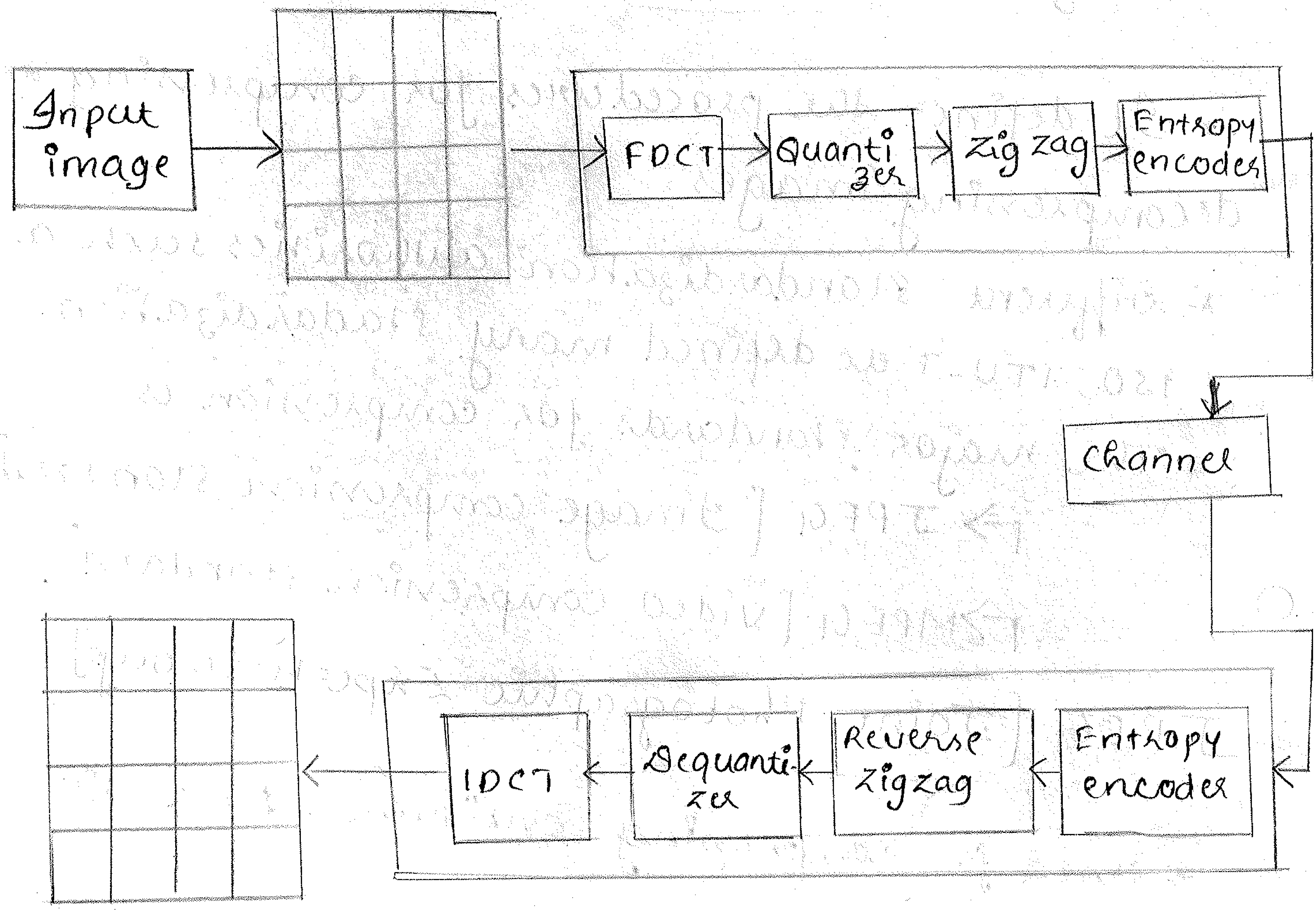
Image Compression Standards

* It defines the procedures for compressing and decompressing images.
* Different standardization authorities such as ISO, ITU-T etc defined many stadardization.
* The major standards for compression is
  - JPEG [Image compression standard]
  - MPEG [video compression standard].

JPEG [Joint Photographic Experts Group].

* used for compressing continuous tone still images.
* Introduced in 1992 for compression techniques and reconstruction of still photographic image
* Normal JPEG is a lossy baseline coding system.
* uses quantized DCT coefficients on 8 x 8 image blocks.
* Coding is done using Huffman coding / Run Length coding.
* Most widely used over internet.

Procedure



DCT computation

* FDCT stands for Forward Discrete Cosine Transform and IDCT stands for Inverse Discrete Cosine Transform
* The input image is partitioned into an 8x8 sub block.
* FDCT is computed on each of the 8x8 blocks of pixels.
* The coefficient with zero frequency in both dimensions is called the 'DC coefficient' and the remaining are called the AC coefficients.

* The role of DCT is to decompose the original signal into its dc and ac components and its role of inverse DCT is to reconstruct the signal.

The 2D-DCT of the input image $f(m,n)$ is given by $F(K,l)$ as.

$$F(K,l) = \alpha(K)\alpha(l) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m,n) \cos\left[\frac{(2m+1)\pi K}{2N}\right] \cos\left[\frac{(2n+1)\pi l}{2N}\right].$$

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

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

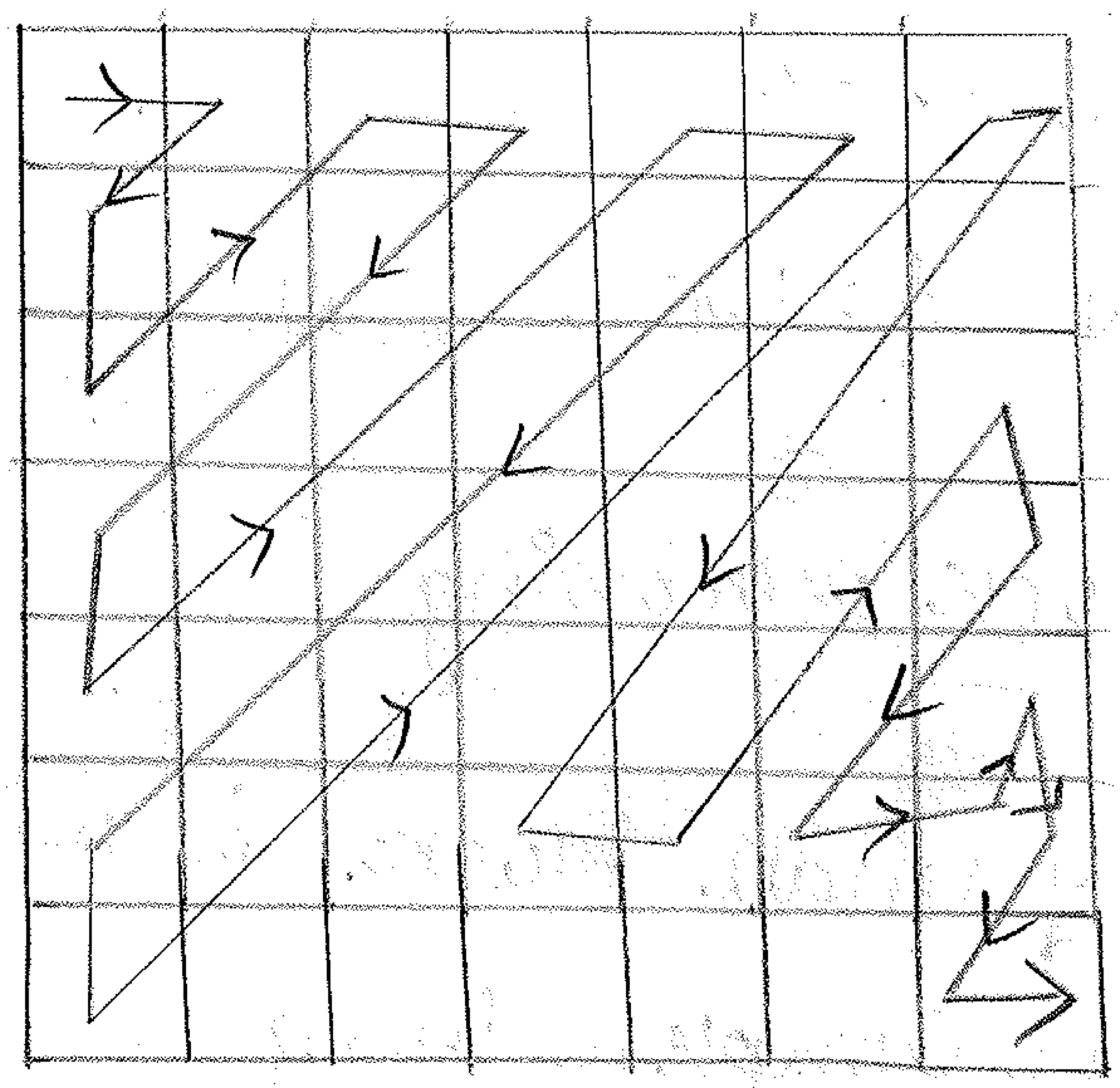
Advantages of block processing

* The transform of small blocks is much easier to compute than the complete image.
* The pixel correlation will not exceed usually 16 or 32 pixels.

Drawback of block processing

* DCT based transform coding is the introduction of block artifacts.

* The FDCT processing step lays the foundation for achieving data compression by concentrating most of the signal in the lower spatial frequencies.
* The 64 DCT coefficients are scalar quantised using uniform quantisation tables based upon psychovisual experiment.
* After the DCT coefficients are quantised, the coefficients are ordered according to the zigzag scan.
* The purpose of zigzag scanning is based upon the observation that most of the high frequency coefficients are zero after quantisation.



* Taking DCT for an entire image require for taking DCT in block manner requires large memory hence it is not preferable.
* An N-point DCT consist of N real basis vectors with cosine functions.

* Blocking artifacts is caused by the discontinuities that result from the rectangular windowing of the image data.

Minimisation of blocking artifact

The blocking artifact can be minimised by the use ~~by~~ of

* Overlapping blocks.
* Low pass filtering of boundary pixels and
* Lapped Orthogonal Transform (LOT)

The use of overlapping blocks results in increased bitrate, and hence higher compression ratio cannot be achieved.

Zonal coding

* It is based on the fact that the transform coefficients of maximum variance carry the ~~least~~ most picture information.
* The locations of the coefficients, with the K largest variances are indicated by means of zonal mask.
* To design a zonal mask, variances of each coefficients can be calculated based on a global image model such as Gauss Markov model.

Threshold coding

* Each transform coefficient is compared with a threshold.
* If it is smaller than the threshold then it is set to zero.
* If it is larger than the threshold, it will be retained for quantisation and encoding.
* The thresholding method is an adaptive method where only those coefficients whose magnitudes are above a threshold are retained within each block.

Quantiser

* The purpose is to remove the components of the transformed data that are unimportant to the visual appearance of the image and to retain the visually important components.
* The process of assigning a particular sample to a particular level is called quantisation.

Advantages

* The reordering of quantizes coefficients may result in long runs of zeroes. This makes the computational error.
* Instead of default coding tables and quantisation arrays, the user is allowed to construct custom table.

MPEG [Motion Picture Experts Group].

* MPEG-1

* used for video compression.
↱ Supported by ISO/IEC
↱ Standard for CD ROM applications with non interlaced video upto 1.5 Mbps.
↱ Frame predictions are based on previous frames next frame or an interpolation of both.
↱ Supported by almost all computers and DVD layers.

MPEG-2

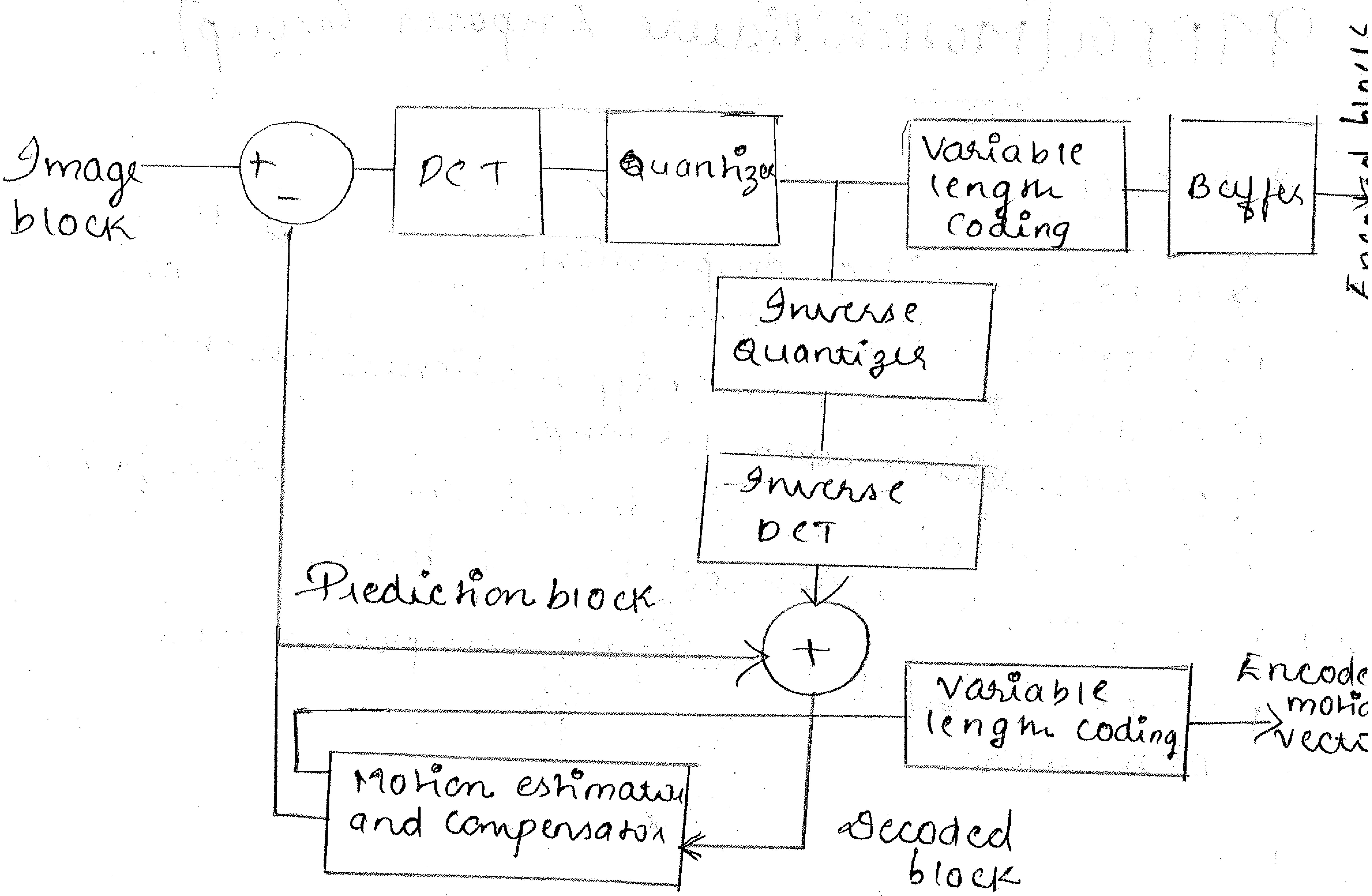
↱ Supported by ISO/IEC.
↱ Extension of MPEG-1 for DVDs with transfer rate 15 Mbps.
↱ Supports interlaced video and HDTV.
↱ Most useful video standard.

MPEG-4

↱ Supported by ISO/IEC.
↱ Extension of MPEG-2 supports variable block sizes and differential prediction within frames.

There are 3 basic types of encoded output frames.

* Intra frame or independent frame (I-frame)
* Predictive frame (P-frame)
* Bidirectional frame (B-frame).



Intraframe / Independent frame (I-frame).

* An I frame is compressed independently of all previous and video frames.

* I-frames provide the highest degree random access, ease of editing and greatest resistance to the propagation of transmission error.

Predictive frame (P-frame)

* P-frame is the compressed difference between the current frame and a prediction of it based on the previous I- or P-frame.

* The computed motion vector is variable length coded and transmitted as an integral of the encoded data stream, motion estimation etc.

Bidirectional frame (B-frame)

* B frame is the compressed difference between the current frame and a prediction of it based on the previous I or P-frame and next P-frame.

* The encoded frames are therefore reordered before transmission, the decoder reconstructs and displays them in the proper sequence.

Boundary Representation

* The binary images in which objects and background points are labeled as 1 and 0.

* The images are padded with the border of 0's to eliminate the possibility of an object merging with the image border.

Step 1

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

	c_0	b_0	1	1	1	
	1			1		
		1		1		
	1			1		
	1	1	1	1		

The starting point b_0 be the uppermost, leftmost point in the image that is labeled 1. Denote by c_0 the west neighbor of b_0.

Step 2

Let $b = b_1$ and $c = c_1$.

Step 3

Let the 8 neighbors of b starting at c and proceeding in a clockwise direction denoted by $n_1, n_2, \ldots n_8$ find the first n_k labeled 1

Step 4

Let $b = n_k$ and $c = n_{k-1}$

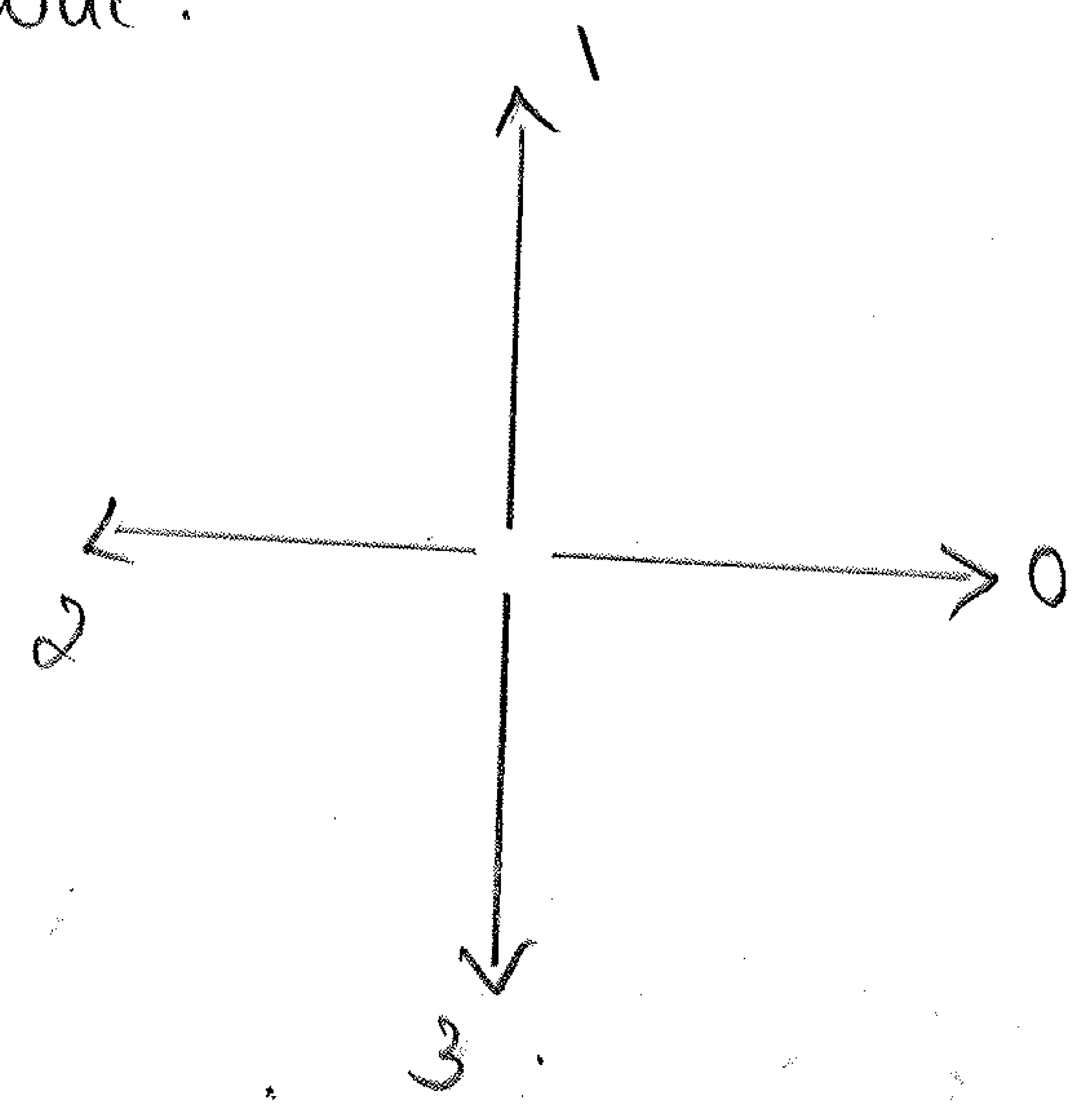
Step 5

Repeat step 3 and 4 until $b = b_0$ and the boundary point found is b_1.

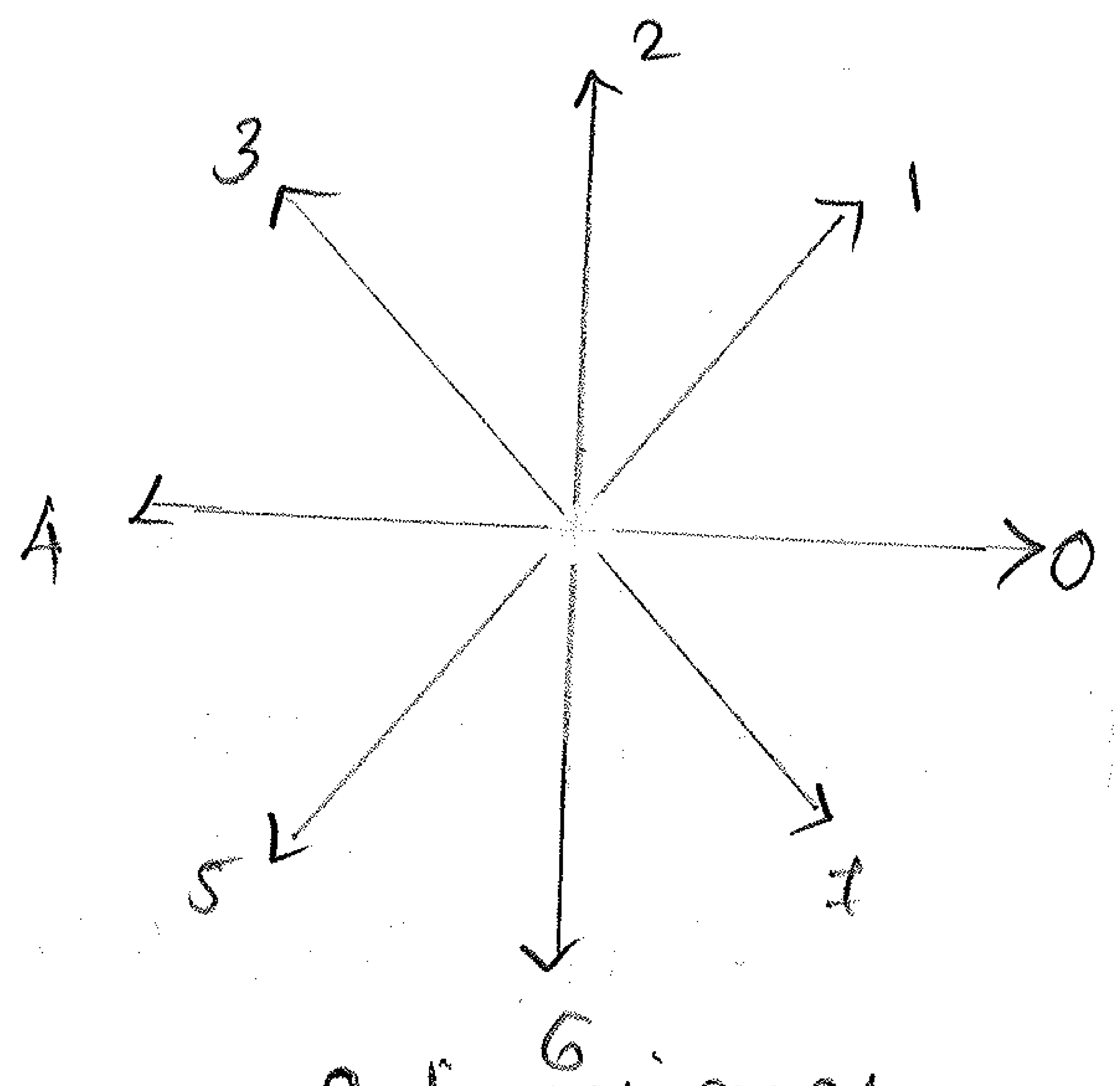
Chain codes

The method of chain code was introduced in 1961 by Freeman.

* Chain codes are used to represent a boundary by a connected sequence of straight line segments of specified length and direction.
* This representation is based on 4 or 8 connectivity of the segments.
* The direction of each segment is coded by using a numbering scheme.
* A boundary code formed as a sequence of such directional numbers is referred to as freeman chain code.



4 directional



8 directional

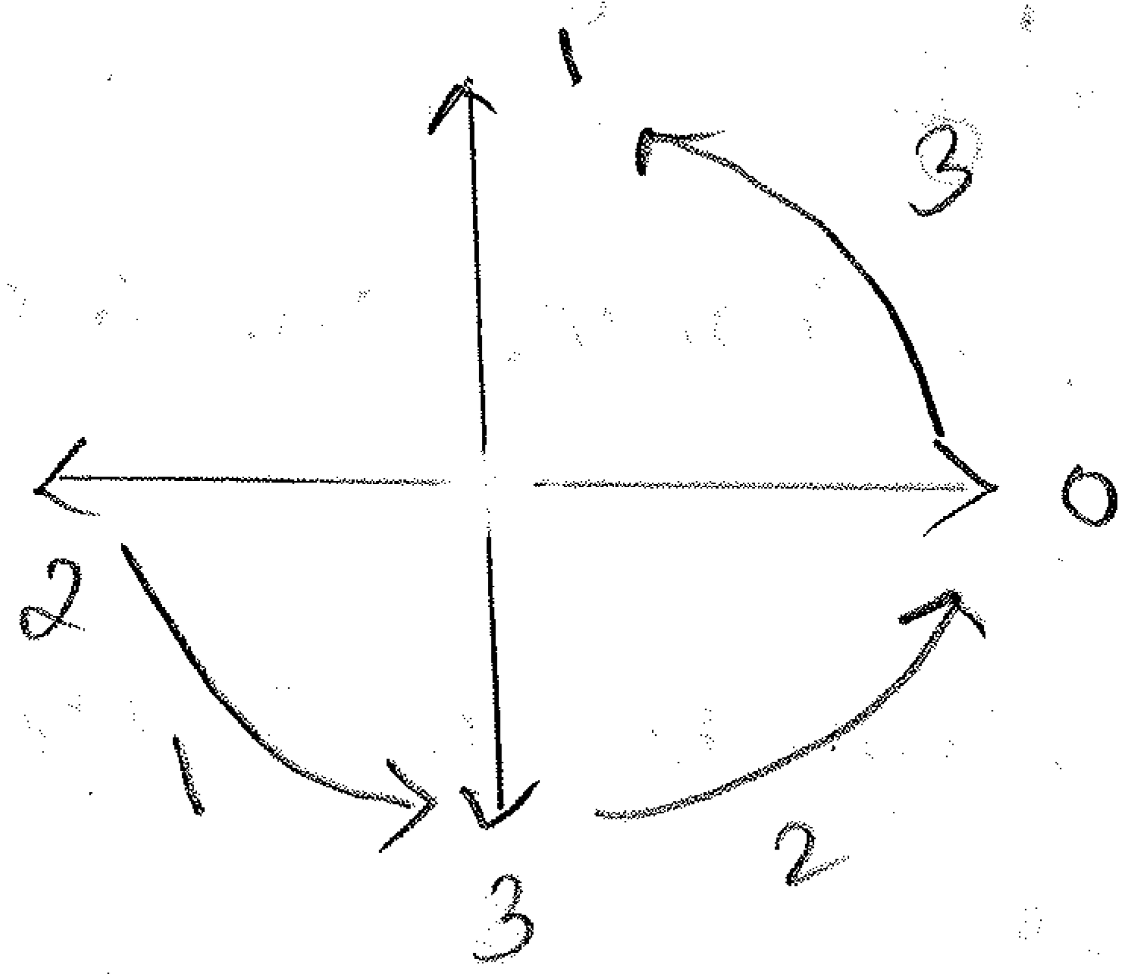
* The chain of code is too long.
* Any small disturbance along the boundary due to noise or imperfect segmentation causes changes in the code that may not be related to the shape of the boundary.

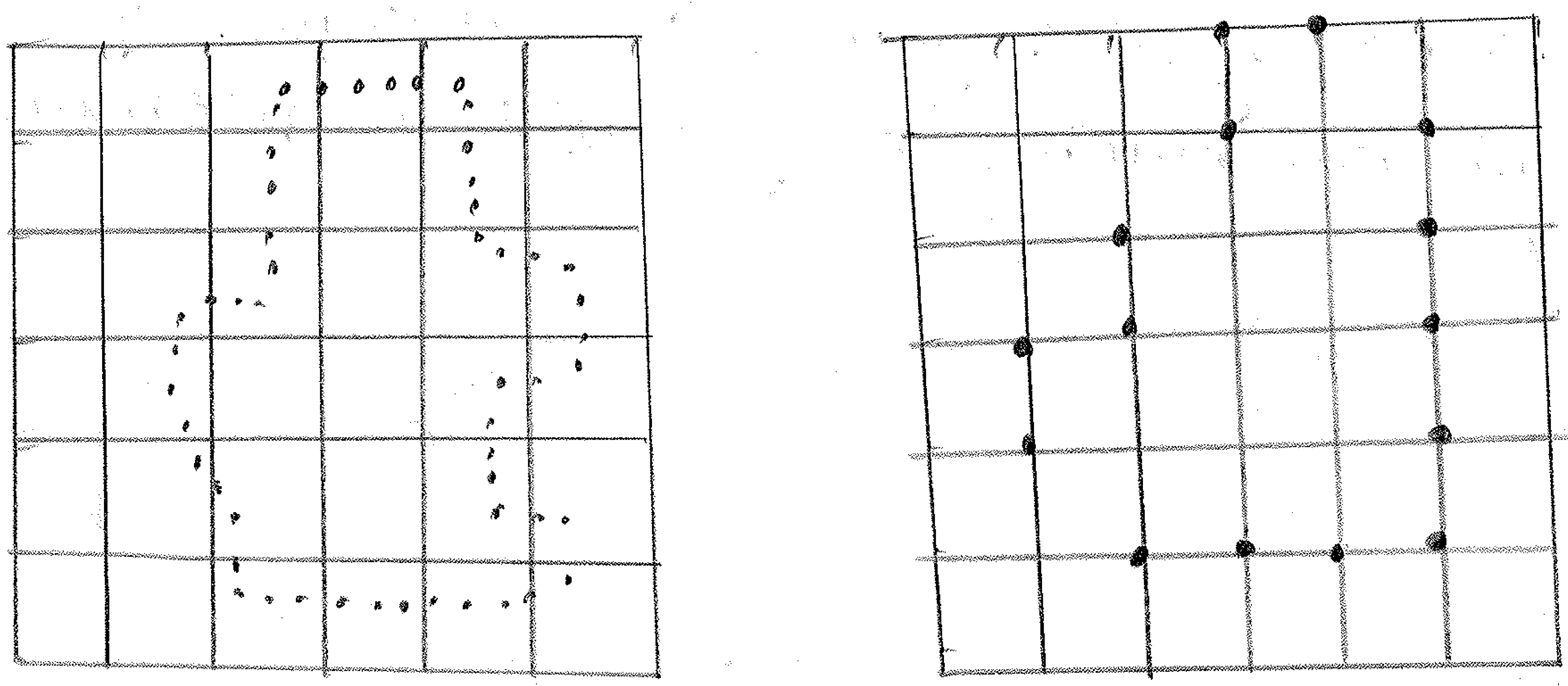
consider chain code 0 3 3 2 1

chain code : 0 3 3 2 1

first difference : 3 3 0 3 3.

Here 1st difference between 1 and 0 is 3,





Boundary Descriptors

It describes the boundary of a region using the features of the boundary.

i) Length.

The length of a boundary is one of the simplest descriptors. It gives the no. of pixels along a boundary

Length = no. of vertical and horizontal components

× $\sqrt{2}$ Number of diagonal components.

ii) Diameter.

The diameter of a boundary B is defined as

$$Diam(B) = \max_{i,j} [D(P_i, P_j)]$$

where

D = A distance measure.

P_i, P_j = Points on the boundary.

iii) Major Axis

The major axis of a boundary is defined as the line segments connecting the two extreme points of its diameter.

iv) Minor Axis

The minor axis of a boundary is defined as the line perpendicular to the major axis.

v) Eccentricity

The ratio of major axis to the minor axis is called the eccentricity of a boundary.

$$\text{Eccentricity} = \frac{\text{Major axis}}{\text{Minor axis}}$$

vi) Curvature

Curvature of a boundary is defined as the rate of change of its slope.

Regional Descriptors

i) Area

The area of a region is defined as the number of pixels in the region.

ii) Perimeter

The perimeter of a region is defined as the length of its boundary.

iii) Compactness

$$\text{compactness} = \frac{(\text{Perimeter})^2}{\text{Area}}$$

iv) Circularity Ratio

Ratio of the area of a region to the area of a circle, having same perimeter.

The area of the circle with perimeter length

P is $\frac{P^2}{4\pi}$.

$$R_c = \frac{4\pi A}{P^2}$$

where A is the area of the region and P is the length of its perimeter.

Texture

Texture content is an important quantity used to describe a region.

Texture of region provides measure of properties such as smoothness, coarseness and regularity.

Pattern and Pattern classes

A pattern is an arrangement of descriptors

* The name features is used often in the pattern recogniti literatures to denote a descriptor.
* A pattern class is a family of patterns that share some common properties.