(1 Numerical sure)

certain information in the image relatively Less importance than other this information is said to be psychovisually redundant

mapper: Reduce interpixel redundancy

quantizer: reduce psychovisual Redundancy

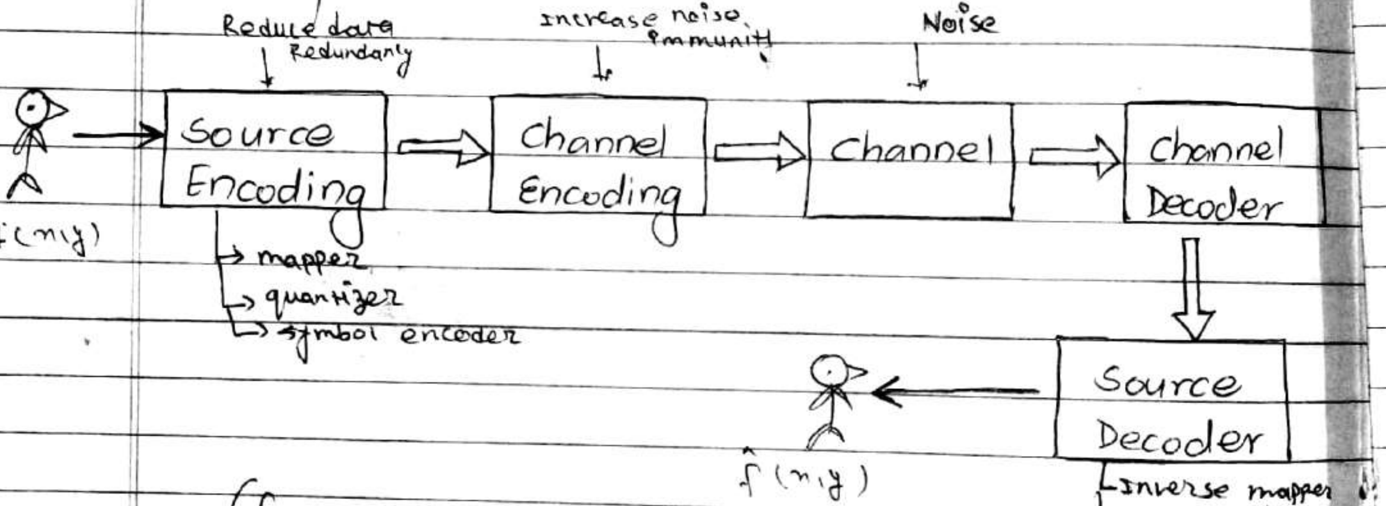
Symbol encoder: reduce coding redundancy

Chapter-5

Image Compression

[6 hours]

Introduction To Image Coding And Compression:



(fig: Block Diagram of Image Compression)

Image Compression addresses the problem of reducing the amount of data required to represent a digital image.

The underlined basis of the reduction process is removal of redundant data.

The compression image is decompressed to reconstruct the original image and approximation of it.

Thus, Image (data) compression is

the reduction of number of bits required to transmit the image probably without any remarkable loss of information.

The image encoding process through the channel is called image compression and image decoding process is known as image decompression.

Application of Data Compression:

(1) Broadcasting T.V
(2) Remote sensing through satellite
(3) Military communication
(4) Fascimile Transmission & Tele-communication

Data Redundancy:

It contains data that either provide no relevant information or simply restate that which is already known, this is called Data Redundancy.

Image (Data) Compression Technique:

The image (data) compression techniques are:

(1) Pixel coding Technique
(2) Predictive coding Technique
(3) Transform coding Technique
(4) And, other Methods.

(P.T.O)

DCPM: Differential

Compression Technique

- Pixel Coding
  - PCM / Quantization
  - Run length coding
  - Bit-plane coding
- Predictive coding
  - Delta Modulation
  - Line by Line DPCM
  - 2-D PCM
  - Interface Technique
  - Adaptive
- Transform coding
  - Threshold
  - Multi-dimension
  - Adaptive
- Other Methods
  - Hybrid
  - Color
  - vector

(fig: Different Image compression Technique)

(1) Pixel Coding Technique:

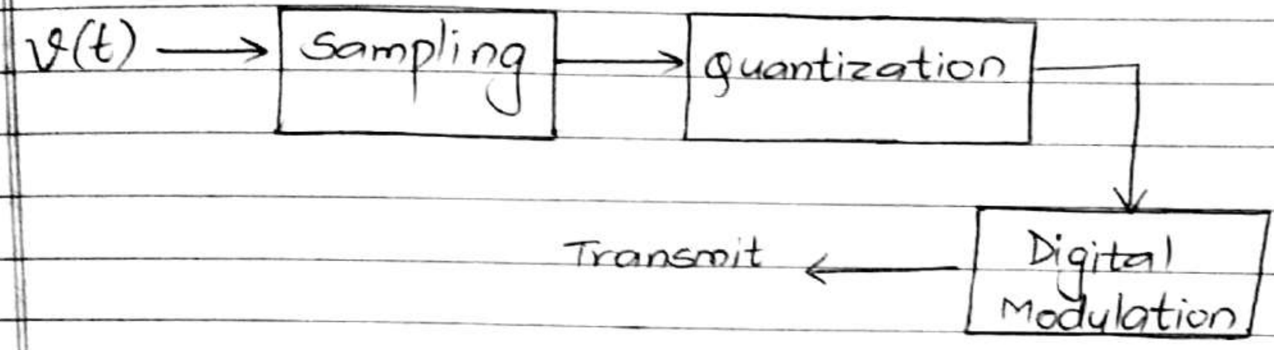
In this technique, each pixel is processed independently ~~ing~~ ignoring the inter-pixel dependencies. This technique firstly remove the inter-dependencies between the neighbourhood pixel and then removal of redundant data.

Pixel coding technique is further divided into following three types:

(a) PCM / Quantization
(b) Run-length coding
(c) Bit-Plane coding

sampling: Digitizing the spatial coordinates
quantization: Amplitude digitization (gray-level)

Imp (a) Pulse Code Modulation (PCM):



(fig: PCM / Quantization)

In PCM technique, the incoming video signal is sampled, quantized and coded by the suitable code word before inputing it to a digital modulator for transmission.

The quantizer output is coded by a fixed length binary code word having usually 8-bit.

The formula for achieving rate of compression of PCM is given by the rate distortion formula.

i.e

$$\boxed{R_{PCM} = \frac{1}{2} \log_2 \frac{\sigma v^2}{\sigma q^2}, \quad \sigma q^2 < \sigma v^2}$$

Where,

σv^2 = The variance of quantize i/p

σq^2 = The quantizer Mean square Deviation

(b) Run Length Coding:

It is a very simple form of the data compression which represents each subsequence of identical symbols by a peer like (L, a) where, L is the subsequence and a is the recurring symbol.

Run-length coding describes each row of the image by a sequence of length that describes successive runs.

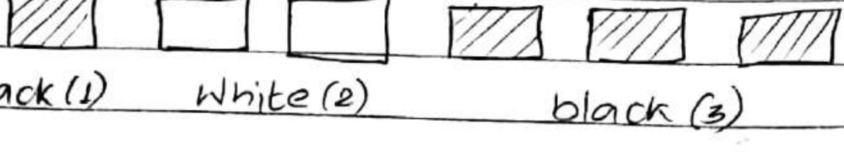
for example, aaabbbaaa is coded as 3a3b3a.

The most common methods are as follows:

(i) To specify the value of the first run of each row.

(ii) To assume that each row begins with a white ~~white~~ run.

for example :



From above figure, the run length code can be written as : 1 black 2 White 3 black

In this technique, the numeric run length coding can be also defined as:

000002222111 = 0!5 2!4 1!3

(Imp) **Entropy Coding:** (Replace input string by code word)

Entropy ~~codes~~ coding encodes the given set of symbols with minimum number of bits required to represent them.

The most important entropy coding is Huffman Coding. The theoritical minimum average of bits that are required to transmit a particular source string is known as entropy of source and it can be computed by using following formula:

$$\boxed{\text{Entropy } (H) = -\sum_{i=1}^{N} P_i \log_2 P_i}$$

It has two major parts. They are:

(i) Construction of Probability Tree
(ii) Assigning code to each node of the constructed tree.

It gives the variable length code words and highest probability assigns ~~sub-pa~~ short path and lowest probability assigns longest path i.e code length.

Huffman Coding Algorithm:

(1) Arrange the symbol probabilites 'P_i' in a decreasing order and consider them as leaf node of a tree.

(2) While there is more than one node;

(a) Manage the 2 nodes with smallest probability to form a new node.

(b) Assign '1' & '0' to each pair of branches merging into a node.

(3) Read sequentically from root node to leaf node.

(Imp) Q. Source generates the symbol S_1, S_2, S_3, S_4 and S_5 randomly with probability $P_1 = 0.4$, $P_2 = 0.2$, $P_3 = 0.2$, $P_4 = 0.1$ and $P_5 = 0.1$ respectively.

Generate the code word for each symbol using Huffman coding. Also, calculate the compression ratio and efficiency of the system.

Step(1):

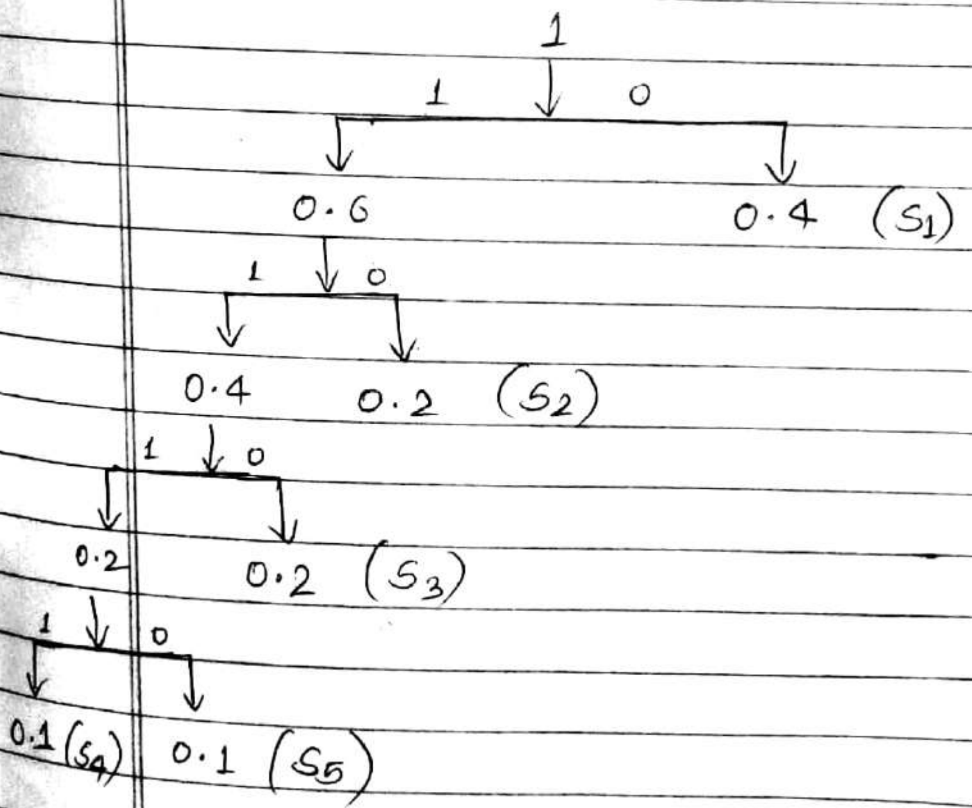
Soln:

Symbol	Original size Probability (P_i)	Step 1	Step 2	Step 3	Step 4
S_1	0.4	0.4	0.4	0.6	1
S_2	0.2	0.2	0.4	0.4	
S_3	0.2	0.2	0.2		
S_4	0.1	0.2			
S_5	0.1				

Step(2):

Construct a Huffman Tree:

Assume, Highest probability 0.6 = 1
Lowest probability 0.4 = 0



(P.T.O)

Step(3): Generate code Word length:

Symbol	Probability(P_i)	Code length(l_i)
S_1	0.4	0 = 1
S_2	0.2	10 = 2
S_3	0.2	110 = 3
S_4	0.1	1111 = 4
S_5	0.1	1110 = 4

Now,

$$\text{Compression Ratio (or } L_{avg}) = \sum_{i=1}^{N} P_i \cdot l_i$$

$$= 0.4 \times 1 + 0.2 \times 2 + 0.2 \times 3 + 0.1 \times 4 + 0.1 \times 4$$

$$= 2.2 \text{ bits / symbol}$$

$$\therefore \text{Entropy } (H) = -\sum_{i=1}^{N} P_i \log_2 P_i$$

$$= -\left[0.4 \times \log_2 0.4 + 0.2 \times \log_2 0.2 + 0.2 \times \log_2 0.2 + 0.1 \times \log_2 0.1 + 0.1 \times \log_2 0.1\right]$$

$$= 2.12$$

$$\text{Hence, Efficiency} = \frac{\text{Entropy}}{L_{avg}} \times 100$$

$$= \frac{2.12}{2.2} \times 100\%$$

$$= 96.36\%$$

Note: $\boxed{\log_2 P_i = \dfrac{\log_{10} P_i}{\log_{10} 2}}$

$\log_{\square} \square$

Q. Calculate the compression ratio and efficiency from below image:

Gray level (r)	0	1	2	3	4	5	6	7
No. of Pixel (n_K)	400	1350	659	2034	816	2560	250	1500

Soln: **Step(1):** Calculate the probability in each gray level

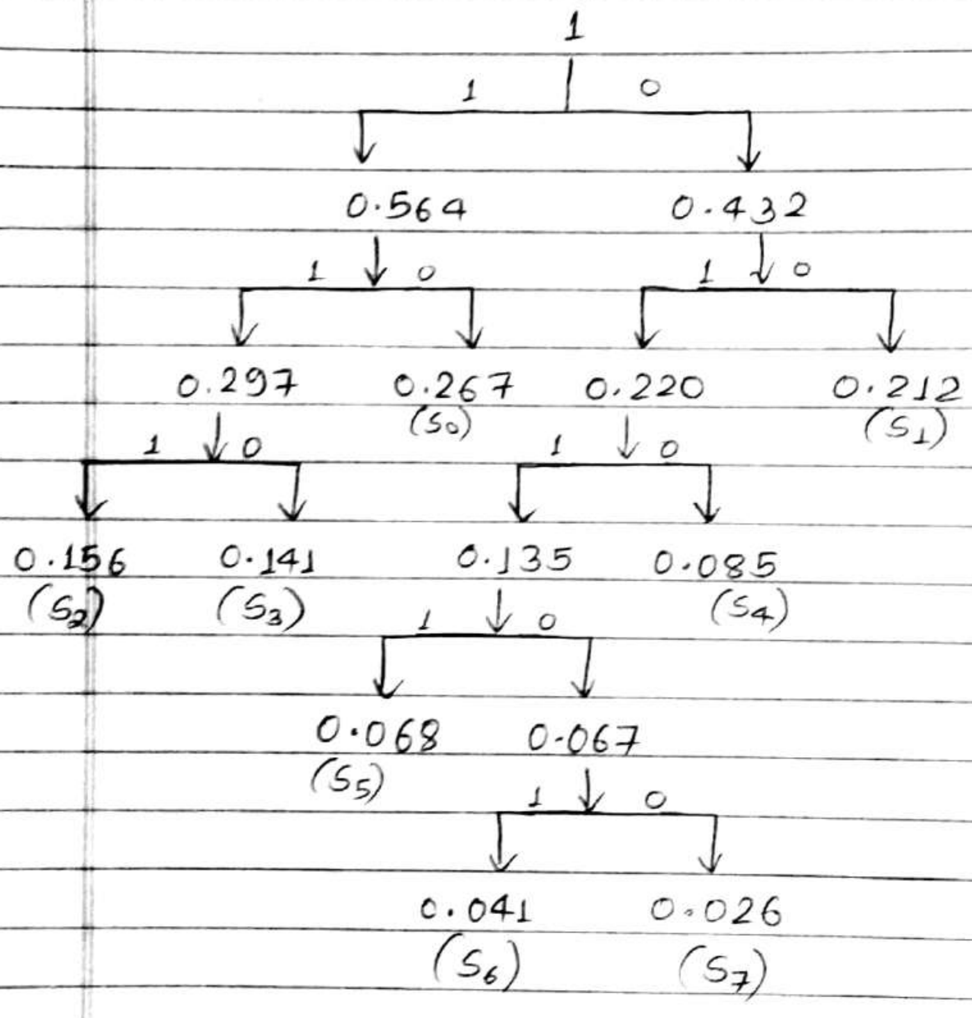
Gray level (r)	No. of Pixel (n_K)	Probability $(P_i) = \frac{n_K}{n}$
0	400	0.041
1	1350	0.141
2	659	0.068
3	2034	0.212
4	816	0.085
5	2560	0.267
6	250	0.026
7	1500	0.156
	$n = 9569$	

Step(2): Assume symbol for each probability in descending order.

Symbol	P_i	Step 1	2	3	4	5	6	7
S_0	0.267	0.267	0.267	0.267	0.297	0.432	0.564	1
S_1	0.212	0.212	0.212	0.220	0.267	0.297	0.432	
S_2	0.156	0.156	0.156	0.212	0.220	0.267		
S_3	0.141	0.141	0.141	0.156	0.212			
S_4	0.085	0.085	0.135	0.141				
S_5	0.068	0.068	0.085					
S_6	0.041	0.067						
S_7	0.026							

Step(3): Construct Huffman Tree

Assume, Highest probability 0.564 = 1
lowest probability 0.432 = 0



Step(4): Generate code word length

Symbol	Probability (P_i)	Code word length (l_i)
S_0	0.267	10 = 2
S_1	0.212	00 = 2
S_2	0.156	111 = 3
S_3	0.141	110 = 3
S_4	0.085	010 = 3
S_5	0.068	0111 = 4
S_6	0.041	01101 = 5
S_7	0.026	01100 = 5

Classmate

Date ______
Page ______

Now,

$\therefore$ Compression Ratio $(L_{avg}) = \sum_{i=1}^{N} P_i l_i$

$= (0.267 \times 2) + (0.212 \times 2) + (0.156 \times 3) + (0.141 \times 3) + (0.085 \times 3) + (0.068 \times 4) + (0.041 \times 5) + (0.026 \times 5)$

$= 2.711$

$A \log_{[2]} [A] \rightarrow$ CALC

$\therefore$ Entropy $(H) = -\sum_{i=1}^{N} P_i \log_2 P_i$

$= -[0.267 \log_2(0.267) + 0.212 \log_2(0.212) + 0.156 \log_2(0.156) + 0.141 \log_2(0.141) + 0.085 \log_2(0.085) + 0.068 \log_2(0.068) + 0.041 \log_2(0.041) + 0.026 \log_2(0.026)]$

$= -[-0.508 - 0.474 - 0.418 - 0.398 - 0.302 - 0.263 - 0.188 - 0.136]$

$= 2.687$

$\therefore$ Efficiency $(\eta) = \dfrac{\text{Entropy}(H)}{L_{avg}} \times 100$

$= \dfrac{2.687}{2.711} \times 100\%$

$\boxed{\eta = 99.11\%}$

→ Based on the concept of decomposing a multilevel (monochrome or color) image into a series of binary images and compressing each binary image via one of several well-known binary compression methods.

(c) Bit Plane Coding:

It breaks the image into bit planes and apply run length coding to each plane.

Let, 'I' be an image where every pixel value is 'n'-bit long. We can express every pixel in binary using n-bit from out of an image and n-binary matrix (called Bit plane) where, n^{th} matrix consist of ith bit of the pixel of 'I'.

It can be illustrated as:

$$I = \begin{bmatrix} 101 & 110 \\ 111 & 011 \end{bmatrix}$$, corresponding 3-bit plane and 2×2 matrix

Now,

$$I' = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Bit plane coding uses the run length coding but in case of m-bit gray level scale image at polynomial 2-base is defined as:

$$a_{m-1} \cdot 2^{m-1} + a_{m-2} \cdot 2^{m-2} + \dots + a_1 2^1 + a_0 2^0$$

The intensities of m-bit monochrome image can be represented in the form of base-2 polynomial as

→ Based on Eliminating the redundancies of closely spaced pixels - in space / time
- by extracting & coding only the new information in each pixel.

(2) Predictive Coding:

→ New Information of pixel is defined as the difference between the actual & predicted value of pixel.

Future value will be predicted on the basis of past values.

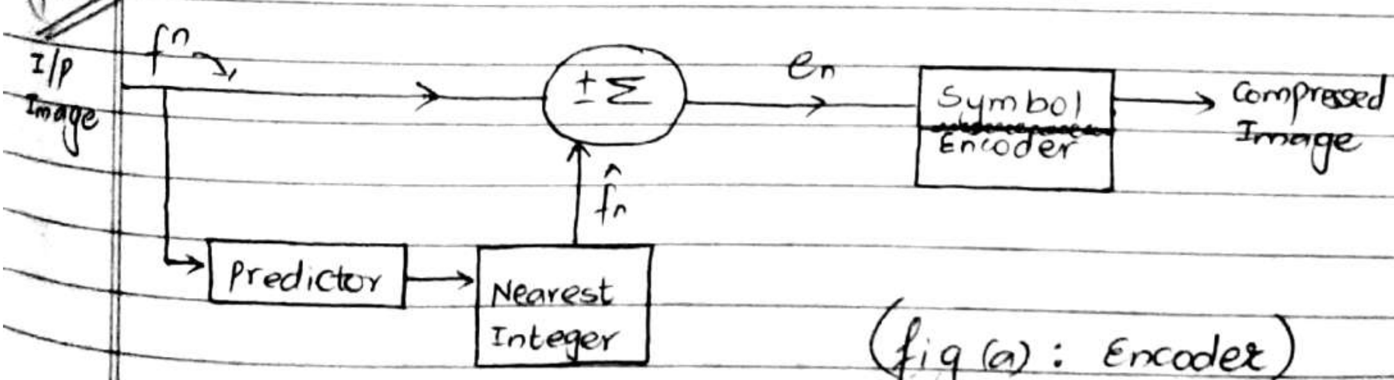
If we know ~~that~~ the past behaviour of a signal upto a certain point in time then, it is possible to make some prediction about its future values then such process is called Prediction.

It is of following two types:

(a) Lossless predictive coding
(b) Lossy Predictive coding

(a) Lossless Predictive Coding: (2015 fall)

(pre)



(fig (a): Encoder)

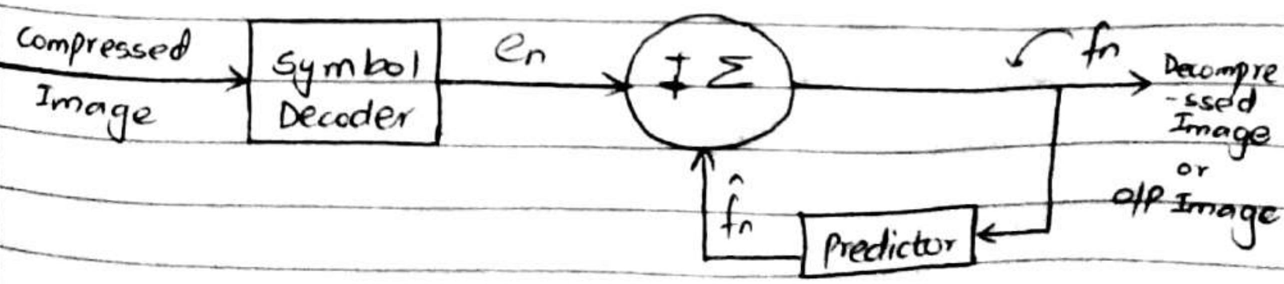


fig: Lossless Predictive coding]

(fig(b): Decoder)

There is no information loss and image can be reconstructed exactly the same as the original.

So, compression process is reversible.

The system consist of an Encoder and a Decoder each containing an identical predictor.
The $\boxed{\text{predictor}}$ generates anticipated value of that pixel Based on some number of past values. The output is then rounded to the nearest integer.

Then, the differences or prediction error is calculated as:

$$e_n = f_n + \hat{f}_n$$

Where,

e_n = Prediction Error
f_n = i/p Image symbol
$\hat{f}_n$ = prediction function

In case of encoding, $\boxed{\text{symbol encoder}}$ uses variable length coding technique to encode symbol.

In case of decoding, $\boxed{\text{symbol decoder}}$ uses decoding function as:

$$f_n = e_n + \hat{f}_n$$

Where,

f_n = Decompressed Image or o/p Image

In most cases, the predictor is formed by a linear combination of 'm' previous pixel.

ie

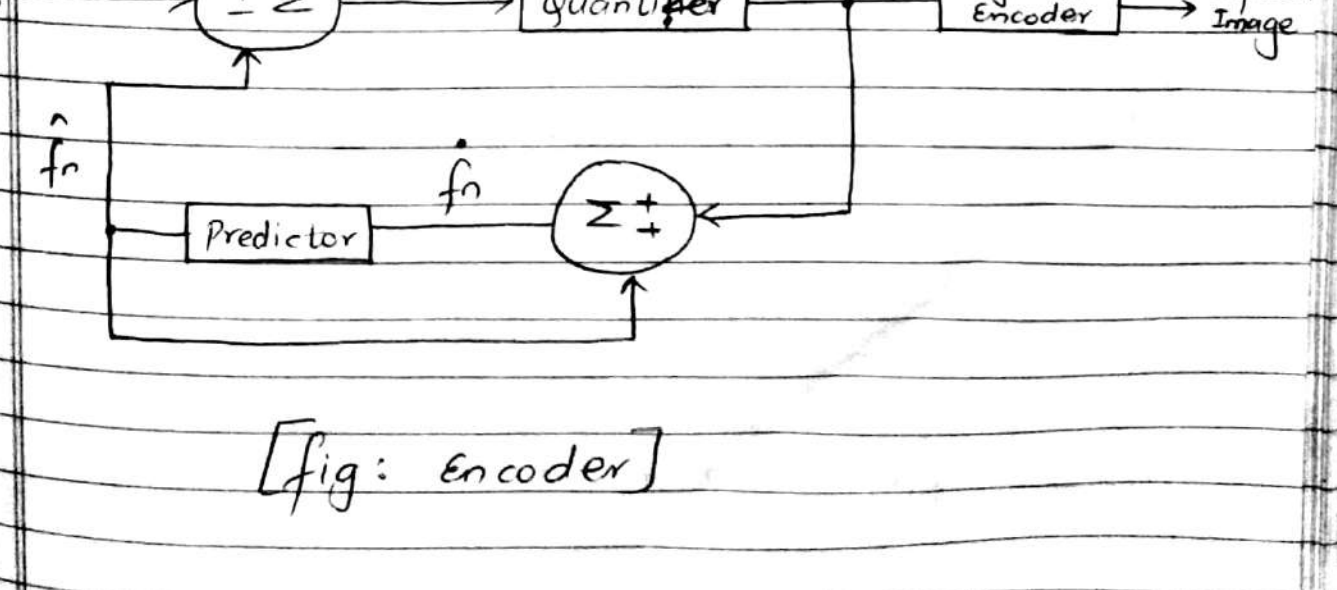
$$\boxed{\hat{f}_n = \text{round}\left[\sum_{i=1}^{m} \alpha_i f_{m-i}\right]}$$

Where,

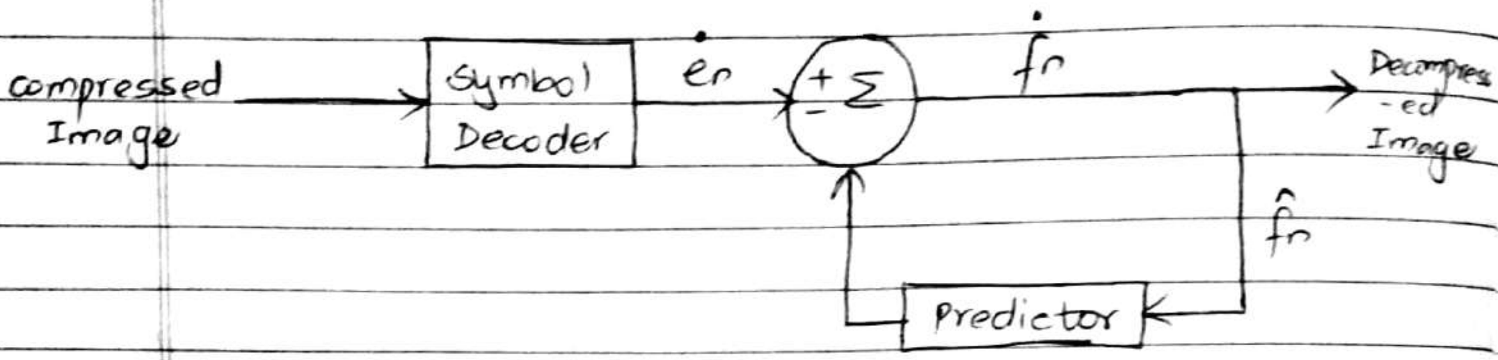
α = Prediction co-efficients

m = Order of linear prediction

(pre) b) <u>Lossy Predictive Coding:</u>



[fig: Encoder]



[fig : Decoder]

The major difference in this technique is that a quantizer is added between the symbol encoder and the point where the prediction error is calculated.

The quantizer absorbs the nearest value of a prediction error and maps it into a limited range of output denoted by $\dot{e}_n$ at ~~ene~~ encoding & $\dot{f}_n$ at decoding, which is defined as :

$\dot{e}_n$ is given by :

$$\dot{e}_n = \dot{f}_n - \hat{f}_n$$

And,

$\dot{f}_n$ is given by :

$$\dot{f}_n = \dot{e}_n + \hat{f}_n$$

Pro: Dr. V.K. Bairagi

→ compressing number of frame [redacted] together

→ compressing individual frame [redacted]

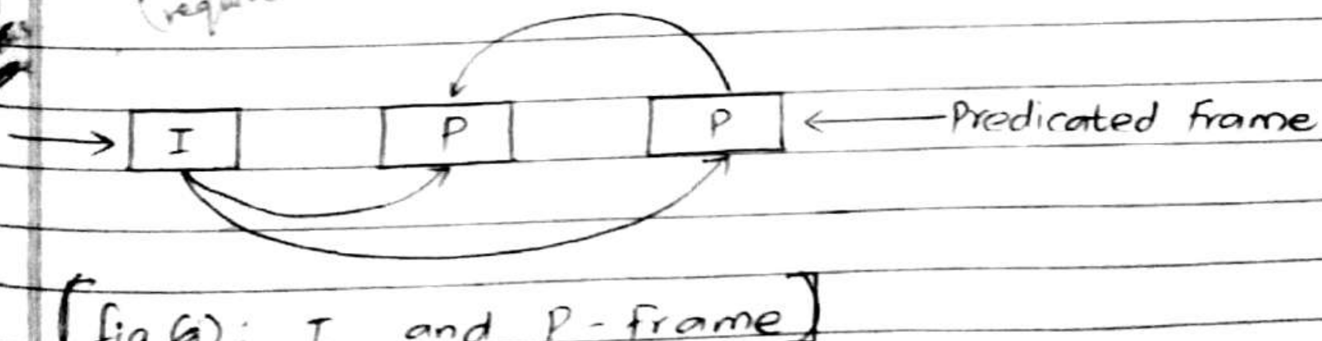
Interframe & Intraframe Coding:

It is used for video compression. The Intraframe coding is also known as I-frame. Intraframe compression uses only current frame for encoding.

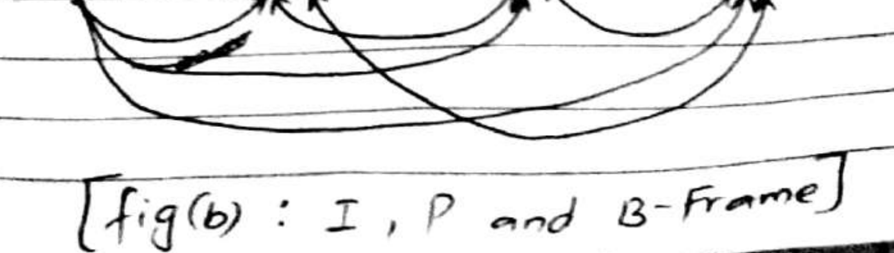
Interframe coding is encoding and compression technique between the video signal. (use number of frame)

Interframe & Intraframe coding is categorized into three types on the basis of predicted frame.

(1) Predictive Frame (P- Frame)
(2) Intraframe or Independent Frame (I - Frame)
(3) Bi-directional Frame (B- Frame)
(require both previous & [redacted] next frame for decoder.)



[fig(a): I and P - frame]



[fig(b) : I, P and B-Frame]

* morphology is a broad set of image processing operation that process image Based on Shape.
* A binary image is digital image that has only two possible values. (black & white)
* each pixel is stored as a single bit i.e 0 or 1

background pixel — foreground pixel

Chapter - 6

Introduction To Morphological Image Processing

[4 hours]

Logic Operation Involving Binary Images:

The language and theory of mathematical morphology often present a dual (but equivalent) view of binary images.
Thus,
We have considered a binary image to be a bitvalued function of spatial co-ordinates (x and y.

Morphological Theory views a binary image as a set of foreground (one valued) pixel and elements of which are in Z^2 (cartesian product) i.e pair of elements.

Set operation such as Union and Intersection can be applied directly to binary image sets.

For Example,
If 'A' & 'B' are binary images then,
$C = A \cup B$ is a binary image.
where, a pixel in c is a foreground pixel

* It is a type of Image processing in which structure of the object with in an image are modified.
* Dilation, erosion are fundamental morphological operation

Date
Page

if either or both the corresponding pixel in A and B are foreground pixel.

The function c is given by two view. They are :

(1) First View :

$$c(x,y) = \begin{cases} 1, & \text{if either } A(x,y) \text{ or } B(x,y) \text{ is } 1 \\ & \text{or} \\ & \text{if both are } 1 \\ 0, & \text{otherwise} \end{cases}$$

(2) Second View :

$$c = \{(x,y) : (x,y) \in A \text{ or } (x,y) \in B \text{ or } (x,y) \in (A \text{ and } B)\}$$

Where,

A & B are Binary Images where the elements of A and B are 1-valued.

Thus,

we see that the function point of view deals with the both foreground (1) and background (0) pixel simultaneously.

The set point of view deals only with foreground pixel and it is understood that all pixel are not foreground pixel

classmate

Date

Page

constitute the ~~diagram~~ background.

The set operation can be performed on binary images using following logical operations:

	Set operation	Logical operation for Binary Images	Name
(1)	$A \cap B$	A & B	AND
(2)	$A \cup B$	A | B	OR
(3)	A^c	~A	NOT
(4)	$A - B$	A & ~B	DIFFERENCE

(V.Imp) # Dilation & Erosion:

Dilation:

(pre) (V.N)

Dilation is an operation that grows or thickens objects in an image. The specific manner and extent of this thickening is controlled by a set referred to as a structuring element.

Graphically, structuring elements can be represented either by a matrix of 0's and 1's or as a set of foreground (1-valued) pixels.

So, the origin of the structuring element must be clearly identified by using the both representation.

Dilation operation can be done by using following steps:

① Representation of original Image with Rectangular object.

let,

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

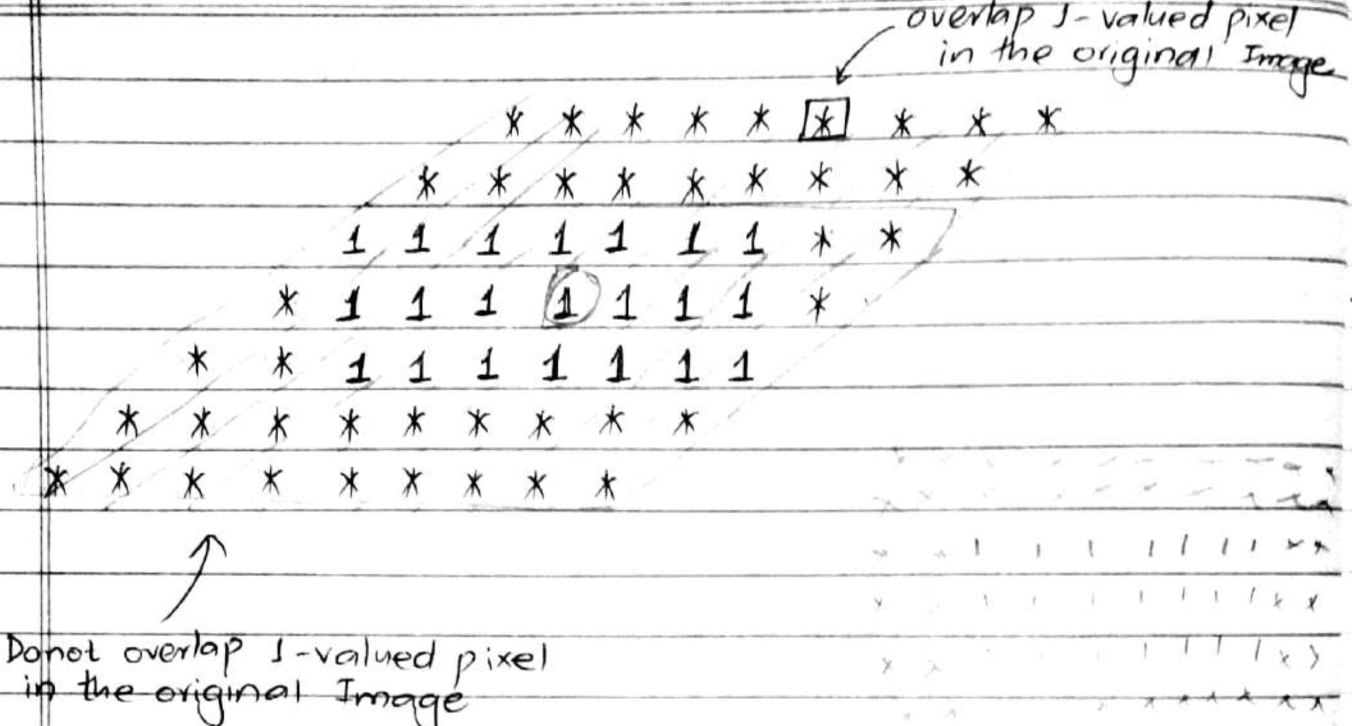
② Structuring Element with 5-pixels arranged in a diagonal line.

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

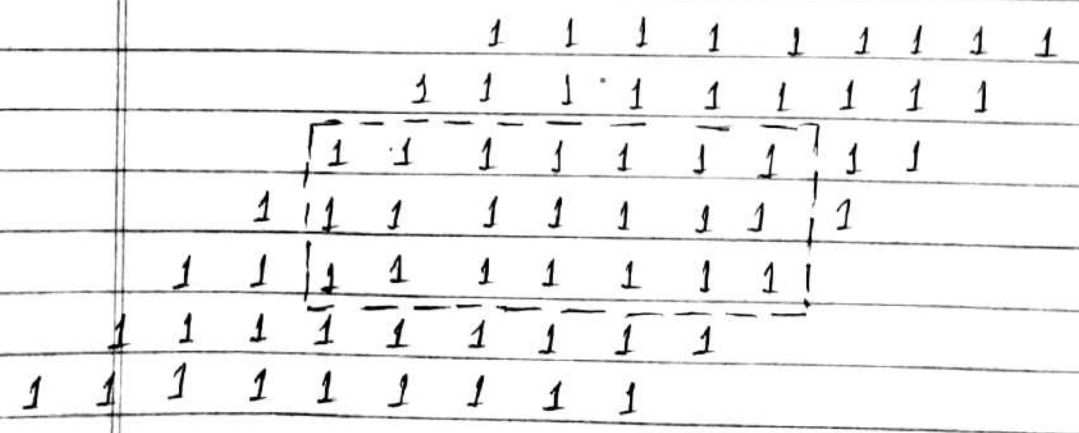
← origin or center of the structuring element as shown with a dark border.

③ Structuring Element translated to several location in the image.

(P.T.O)



④ Output Image i.e shaded region will be replaced by the 1's in the original Image.



Now,

the dilation of A by B denoted as $A \oplus B$ is defined as the set operation as :

reflected & translated & overlaps

At least 1 element of A.

$$A \oplus B = \{ z \mid (\hat{B})_z \cap A \neq \phi \}$$

Where,

ϕ = Empty set

B = Structuring Element

The dilation of A by B is the set consisting of all the structuring element origin location where the reflected and translated B overlaps at least one element of A.

It is a convention in image processing to the first operand of $A \oplus B$ be the image and second operand be the structuring elements which usually is much smaller than the image.

Dilation is of following two types:

→ Dilation is associative

$$A \oplus (B \oplus C) = (A \oplus B) \oplus C$$

→ Dilation is commutative

$$A \oplus B = B \oplus A$$

(pre)

Erosion :

(S.N)

Erosion is an operation that thins or srinks objects in a binary image.

As in dilation, the manner and extent of erosion is controlled by a structuring element.

The erosion process can be done by using following steps:

1. Representation of Original Image with Rectangular Object.

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

2. Structuring element with 3-pixel arranged in vertical line.

$$\begin{matrix} 1 \\ \boxed{1} \\ 1 \end{matrix} \longleftarrow \text{Origin of the structuring element is shown with dark border.}$$

3. Structuring element translated into several location in the image.

1 1 1 1 1 1 $\boxed{1}$

1 1 1 $\boxed{1}$ 1 1 1

1 1 1 1 1 1 1

The result is 0 at this location in the o/p image because all or part of the structuring element overlaps the background.

The result is 1 at this location in the o/p Image because the structuring element fits entirely within the foreground.

(4) Output Image. The shaded region shows the location of 1's in the original image.

0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0

0 0 1 1 1 1 1 1 1 0 0

0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0

Now,

the erosion of A by B denoted as: $A \ominus B$ is defined as:

$$A \ominus B = \{ z \mid (B)_z \subseteq A \}$$

Where,

$(B)_z \subseteq A$ means that $(B)_z$ is a subset of A.

This equation says that, the erosion

of A by B is the set of all points z such that B translated by z is contained in A.

Thus, B is contained in A, is equivalent to B not sharing any elements with the background of A.

The equivalent expression as the definition of erosion is defined as:

$$A \ominus B = \{ z \mid (B)_z \cap A^c = \phi \}$$

properties: $A \ominus B \neq B \ominus A$,

Imp
L.Q

Opening & Closing:

Opening:

The morphological opening of A by B is denoted as $A \circ B$ and defined as the erosion of A by B and followed by a dilation of the result by B.

i.e

$$A \circ B = (A \ominus B) \oplus B$$

Now,

an equivalent formulation of opening is:

$$A \circ B = \cup \{ (B)_z \mid (B)_z \subseteq A \}$$

Where,

$\cup \{ \}$ denotes the union of all sets inside the braces.

$A \circ B$ is the union of all transactions of B that fits entirely within A.

closing:

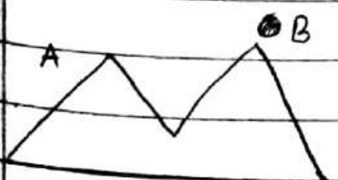
The morphological closing of A by B is denoted as $A \bullet B$ and defined as the dilation of A by B and followed by a erosion of result by B.

i.e

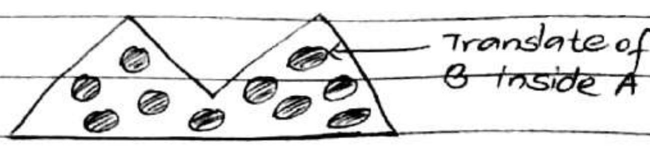
$$A \bullet B = (A \oplus B) \ominus B$$

Geometrically, $A \bullet B$ is the complement of the union of all the translation of B that donot overlap A.

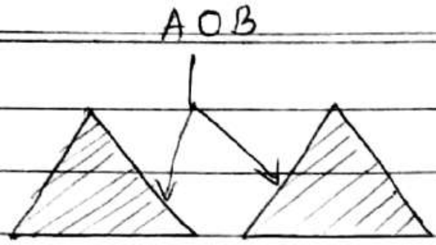
The opening & closing can be illustrated as:



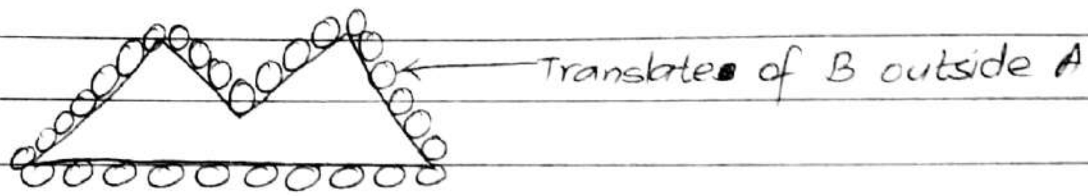
fig(a): set A and structing element B



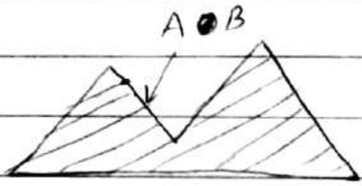
fig(b): Translation of B that fit entirely within set A.



(fig (c) : The complete opening)



(fig (d) : Translate of B outside the border of A)



(fig (e) : The complete closing)

Fig : Opening & closing as Unions and translated structuring element.

classmate
Date
Page

Chapter- 5
Image Compression

[Remaining Topic]

v.v.imp

Redundancy:

In digital image processing, three basic data redundancy can be identified and ~~expor~~ exploited as follows:

(1) Coding Redundancy
(2) Interpixel Redundancy
(3) Psychovisual Redundancy

CIP

(1) Coding Redundancy:

We know that, average number of bits required to represent each pixel is given by:

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

$$L_{avg} = \sum_{i=0}^{N} P_i l_i$$

Hence, the number of bits required to represent the image is $n \times L_{avg}$, where n is the total number of pixel in the given image.

Maximum compression ratio is

achieved when Lavg is minimized.
But,
Coding the gray level in such a way that the Lavg is not minimized then resulting image is said containing coding redundancy.

(2) Interpixel Redundancy:

It is related to interpixel corelation within an image. Usually, the value of certain pixel in the image can be reasonively predicted from the values of its neighbours in the image.

Thus, the values of the individual pixel carries relatively small amount of information and much more information about pixel value that can be inferred on the basis of its neighbours concluded value.

These dependencies between pixel value in the image is called interpixel redundancy.

(P.T.O)

classmate

Date
Page

(3) Psychovisual Redundancy:

The eye doesnot respond with equal sensitivity to all visual information. Human perception searches for important features (edges, texture) and doesnot perform quantitive analysis of every pixel in the image.

So, Psychovisual Redundancy takes into perception of the human visual system.

classmate
Date
Page

Chapter-7
Image Segmentation
[7 hours]

Introduction:

Image segmentation is a process of partitioning a digital image into multiple regions (set of pixels).

The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze.

It is typically used to locate objects and boundaries (line and curves) in an image.

The result of an image segmentation is a set of regions that collectively cover the entire image.

Each of the pixel in a region are similar with respect to the some characteristics such as colour intensity or texture.

Segmentation accuracy determines the success or failure of computerized analysis procedure.

Some practical application of image segmentation are as follows:

(1) Medical Imaging:

- Measurement of tissue volumes
- Diagonosis
- Treatment Planning
- study of anatomical structure

(2) locate Object in satellite Image

(3) Face Recognition

(4) Finger print Recognition

(5) Machine Vision

(6) Iris Recognition

Image segmentation algorithm are generally based on two properties:

(a) Discontinuity
(b) Similarity (Continuity)

Discontinuity: (pre-university)

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

There are three basic types of gray level discontinuities in an image i.e points or spots, lines and edges.

Continuity:

(pre-university)

This approach is based on partitioning image into regions that are similar according to pre-defined criteria such as thresholding, ~~raising~~ region growing, region merging and splitting.

[V. Imp]

Types Of Image Segmentation On the property of Similarity: [Segmentation Technique]

(1) Region Based Segmentation
(2) Segmentation Using Region Growing
(3) Using Region splitting & Merging
(4) Segmentation By Thresholding
(5) Basic Global Thresholding

(1) Region Based Segmentation:

The objective of image segmentation is to partitioned an image into multiple regions.

This technique is based on finding the regions directly. To apply this technique following basic formulation condition must be applied:

let, R represent the entire image region, segmentation partition R into n-subregions as $R_1, R_2, R_3, R_4 \ldots R_n$ such that:

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

It indicates that every pixel must be in a region.

(b) $R_i \cap R_j = \phi$

It indicates that the multiple regions must be disjoint.

(c) $P(R_i) = \text{TRUE}$

It indicates that properties must be satisfied by all the pixels in a segmented regions.

(d) $P(R_i \cup R_j) = FALSE$

It indicates that ~~resu~~ properties of region R_i and R_j are different.

(2) Segmentation Using Region Growing:

Region growing is a technique that groups pixels or sub-regions into larger regions based on pre-defined criteria.

let us, pick up an arbitary pixel (r, c) from the domain of an image to be segmented. This pixel is called seed Pixel.

The basic approach is to start with a set of seed points. Then, examine the nearest neighbour (i.e 4 or 8 neighbour) of (r, c) one by one and neighbourhood pixel accepted belongs to the same region as (r, c).

Once, a new pixel is accepted due to the homogeneity property of a region as a member of the current region.

~~Now~~ Then, the nearest neighbour of this new pixel are examined. This

(ii) If the formula or the condition is satisfied then merge the all sub-regions

This technique starts somewhere at the middle level. Suppose we start with rectangular region of size $m * n$ pixels.

To each region homogenity property is tested. If the test fails, the region is split into four quadrants each of size $m/2 * n/2$.

Now, If the region satisfies the homogenity property then merging process is followed to form a size $2m * 2n$.

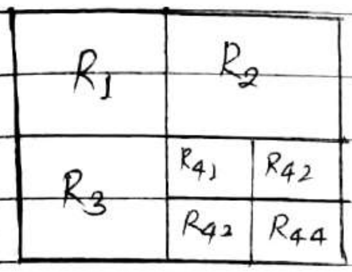
So, the procedure is summarized as:

(i) Split the region R_i into 4 disjoint quadrant any region R_i for which $P(R_i)$ = FALSE

(ii) Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$ = TRUE

(iii) Stop when no further splitting or merging is possible.

Example:



fig(a): Partitioned Image

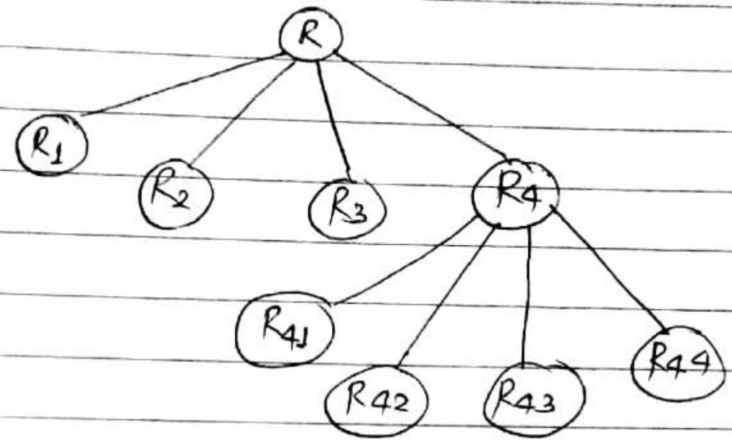


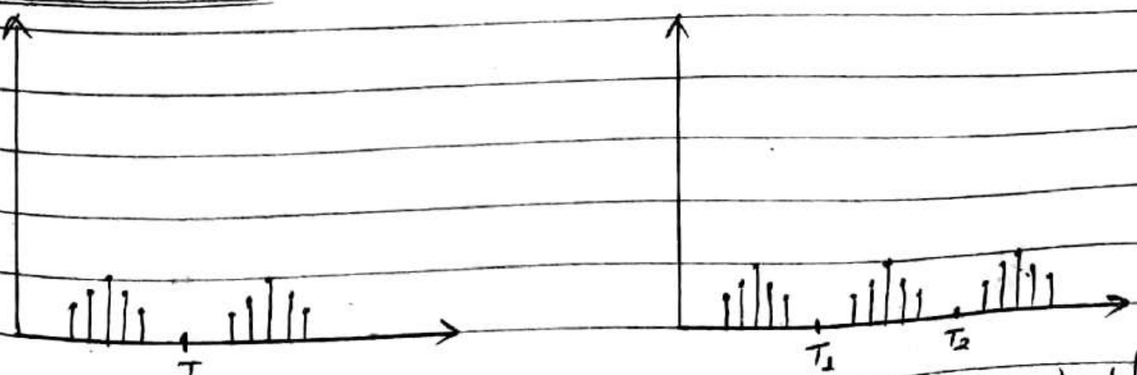
fig (b) : Corresponding quadtree

(IMP) (4) Segmentation By Thresholding : (2012 fall) (2015 fall)

An image contains various regions corresponding to different object. The pixel comparing a region and receive information from point of corresponding object.

Hence, Thresholding is the Simplest method for image segmentation. It plays central ~~position~~ role in application of an image segmentation.

Foundation:



(a) A Single Threshold

(b) Multiple Threshold

fig: Gray level Histogram that can be Partitioned

Suppose that, the gray level histogram of an image $f(x,y)$ composed of light objects on a dark background as shown in figure above.

The figure (a), it shows to extract the object from the background by selecting threshold T that seperate the object from background. Then any point (x,y) for which $f(x,y) > T$ is called an object point. Otherwise, background point.

In figure(b), it shows a slightly more general case of this approach (multiple thresholding). There are two types of light objects on the dark background.

Here, multi-level thresholding classifies a point (x,y) as belonging to one object class. It can be expressed as:

If,
$T_1 < f(x,y) \leq T_2$, to the other object.

Class if,
$f(x,y) > T_2$, to the background.

If, $f(x,y) \leq T$, to the background.

Threshold can be viewed as by operation that involves test aganist a function T of the form,

$$\therefore T = T\left[(x,y),\ P(x,y),\ f(x,y)\right]$$

Where,

$P(x,y)$ = Some local property of the point

$f(x,y)$ = Gray level of point (x,y)

A threshold image can be defined as,

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

Thus, if T depends on both $f(x,y)$ & $P(x,y)$, the threshold is called as local and if T depends on spatial co-ordinates, then, it is called Dynamic or Adaptive.

(5) Basic Global Thresholding:

automatic threshold value (2015 fall)

In this approach, first partition the image histogram by using a single global threshold 'T'.

Segmentation is then accomplished by scanning the image pixel by pixel and labeling each pixel as object or background depending whether the gray level of the pixel is greater

① initial value of threshold = T_1 (select)
② make two group of pixel G_1 & G_2 ← may have similar im..
③ mean μ_1 for group G_1, μ_2 for G_2
Repeat 4. $T = \mu_1 + \mu_2$
5. $|T_i - T_{i+1}|^2 < T_i$

or less than the value of T.

The pixel are assigned and labeled are also assigned. This technique is specially used in the industrial areas where control of illumination is possible.

How to obtain T using global thresholding as follows:

(i) Select an initial estimate for T.

(ii) Segment the image using T. This will provide two group of pixels i.e $G_1 \leq T$ and $G_2 > T$

(iii) Compute the average gray level values μ_1 and μ_2 for the pixel in the regions G_1 and G_2.

The average gray level threshold value is given by:

$$\boxed{T = \frac{1}{2}(\mu_1 + \mu_2)}$$

(iv) Repeat the step (ii) through step (iii) until the difference in T in successive iteration is smaller than predefined parameter.

$$\text{i.e } |T_i - T_{i+1}| < T'$$

← predefined threshold

Threshold value in ith iteration

T_{i+1} iteration

- create a small sub image (template) of a object to be found
- do pixel by pixel matching of the template with the image by placing centre of the template at every possible pixel of main image

with correlation

(IMP)

Template Matching : (2012 fall) (LQ)

Template Matching is a technique in DIP for finding small parts of an image that match a template image.

It can be used in manufacturing as a part of quality control as a way to detect edges in an image.

Here,
We have a template $g(i, j)$ and we wish to detect its instances in an image.

The incoming is compared directly to copies (templates) stored in the memory. So, it has various different applications such as: Face Recognition, Finger Print Detection & Medical Image System.

The logic behind this technique is,
In this technique, the behind logic is that comparing between pixels or pattern. This method is also called as **Linear spatial Filtering**. This is done only thrugh the use of correlation Theorem.

(P.T.O)

* using co-relation find the matching bet^n template pixel & original image pixel.

uses of object recognization

Correlation Theorem:

It is the measure of degree to which two variables agree or not in a actual value.

Here,

two variables are the corresponding pixel values in two images i.e template and source image.

So, it is used to detect location of a certain object inside an image.

Therefore, Correlation coefficient is defined as:

$$\text{correlation} = \frac{\sum_{i=0}^{N-1} (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=0}^{N} (x_i - \bar{X})^2 \sum_{i=0}^{N} (y_i - \bar{Y})^2}}$$

Where,

x_i = Template Gray level Image

y_i = Source Gray level Image

$\bar{X}$ = Average Gray level Template Image

$\bar{Y}$ = Average Gray level Source Image

classmate

Date

Page

Edge Linking & boundary ~~detection~~

Edge or Boundary linking & Detection:

① Using Template Matching:

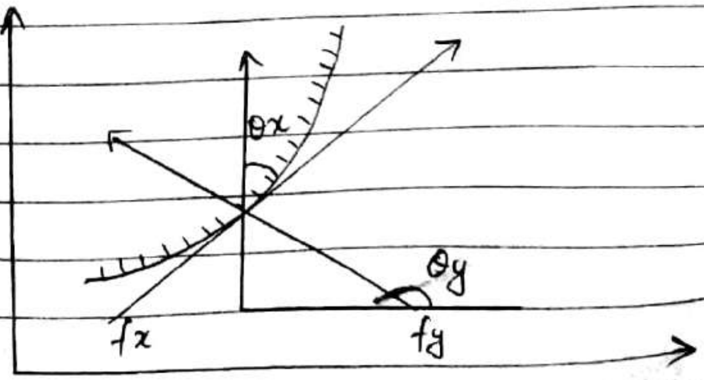
Edge detection extract edges of object from an image. Edge characterize object boundaries and therefore it is useful for segmentation, registration and identification of an object in the scenes.

The variation of an image features usually brightness to edges. So, the edges are the representation of discontinuties of an image intensity function.

So, edge detection algorithm is essentialy a process of detection of dis continuties in an image.

Hence, the change in brightness level indicates edges.

⇒ Edge Detection by using template matching can be derived as follows:



(fig: Gradient of $f(x,y)$ along x-direction using Template matching) gradient: directional change in image intensity.

classmate
Date
Page

For the continous image $f(x,y)$, its derivatives assumes a local maximum in the direction of edge.
Therefore, measuring the gradient of $f(x,y)$ can be used for edge detection.

It can derived as:

$\therefore$ Gradient of Horizontal direction:

$$\Delta f = \left[\frac{df}{dx}, 0\right]$$

$\therefore$ Gradient of vertical direction:

$$\Delta f = \left[0, \frac{df}{dy}\right]$$

$\therefore$ Gradient in θ(any) direction:

$$\Delta f = \left[\frac{df}{dx}, \frac{df}{dy}\right]$$

(I)

Now,

$$\frac{df}{dr} = \left[\frac{df}{dx} \cdot \frac{dx}{dr} + \frac{df}{dy} \cdot \frac{dy}{dr}\right]$$

$$\therefore \frac{df}{dr} = f_x \cdot \cos\theta + f_y \cdot \sin\theta$$

Then,

The Max^m value of $\frac{df}{dr}$ is obtained When,

$$\frac{d}{d\theta}\left(\frac{df}{dr}\right) = 0$$

Hence,

$$f_x \sin\theta_g + f_y \cos\theta_g = 0$$

$$\boxed{\theta_g = \tan^{-1}\left(\frac{f_y}{f_x}\right)}$$

$$\boxed{\left(\frac{df}{dr}\right)_{max} = \sqrt{f_x^2 + f_y^2}}$$

Where,

θ_g is the direction of edge &

$\left(\frac{df}{dr}\right)_{max}$ is the gradient value that indicates the edges.

IMP (2) Using Gradient Model (operator):

(A) First Derivative:

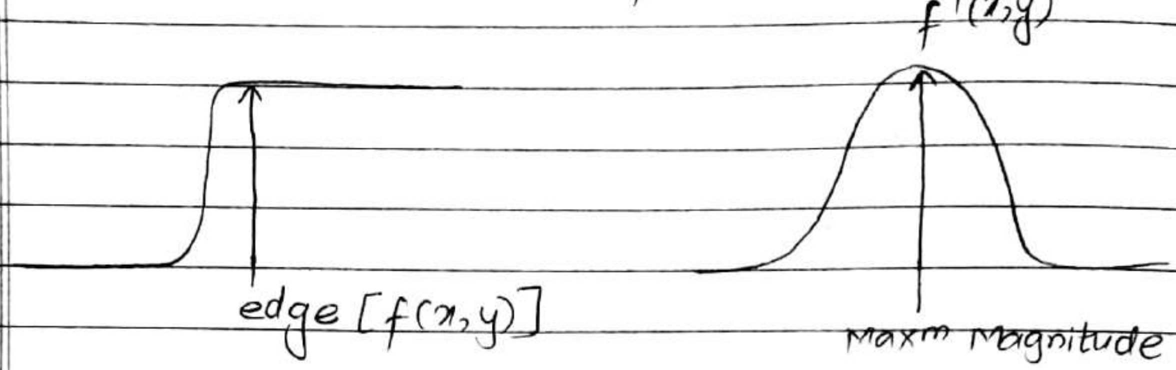
(i) Roberts operator
(ii) Prewitt operator
(iii) Slobel operator

(B) Second Order Derivative

(C) ~~Compass Operator~~

(A) First Derivative:

[ie Difference operator]



fig(a) : cross section of an image edge

fig(b) : cross section of an image edge using first derivatives

The derivatives which yeild high values at places where gray level changes rapidly are used to find gradient of an image as shown in the figure above.

Above figure shows that, first derivative produces thicker edge because the first derivative is positive at the point of transition into and out of the ramp as we move from left to right.

If $\left(\frac{df}{dx}\right)$ and $\left(\frac{df}{dy}\right)$ are the rates of changes of 2D- function $f(x,y)$ along x and y axis then the direction in which rate of change has the greatest magnitude is defined as:

$$\theta = \tan^{-1}\left[\frac{(df/dy)}{(df/dx)}\right]$$

And,

$$\text{Magnitude} = \sqrt{\left(\frac{df}{dx}\right)^2 + \left(\frac{df}{dy}\right)^2}$$

The vector having this magnitude and direction is called the gradient of $f(x,y)$ and is denoted by $f'(x,y)$ for digital image.

First derivative is also called as Difference Operator which is used to find the differences like:

$$f(r,c) - f(r-1, c) = d_1$$

$$f(r,c) - f(r, c-1) = d_2$$

$\therefore$ The magnitude $f'(r,c) = \sqrt{d_1^2 + d_2^2}$

And,

The direction of the greatest step as,

$$\theta(r,c) = \tan^{-1}\left(\frac{d_2}{d_1}\right)$$

There are various gradient operator only using first derivate are as follows:

(i) Roberts Operators
(ii) Prewitt operators
(iii) Slobel operators

RPS

Let the 3×3 area shown in the figure below represents the gray level in a neighbourhood of an image.

$$\begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix}$$

(fig: 3×3 Image)

(i) Roberts Operators:

It is defined as:

$$d_1 = G_x = \frac{df}{dx} = z_9 - z_5$$

$$d_2 = G_y = \frac{df}{dy} = z_8 - z_6$$

The derivative can be implemented for the entire image by using masks which is given as:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & -1 & 0 \\ z_7 & 0 & 0 \end{bmatrix}$$

d_1 $\qquad$ d_2

Here,
the mask of 2×2 are upward because they donot have clear center.

(ii) Prewitt Operators :

It is defined as :

$$d_1 = G_x = [(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)]$$

$$d_2 = G_y = [(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)]$$

Then, The prewitt operator mask will be :

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

for d_1 $\qquad$ for d_2

Here,
the difference between the first & third row of an 3×3 image region approximates the derivatives in the x-direction and the difference between the 3^{rd} & 1^{st} column approximates the derivates in y-direction.

(iii) Slobel Operators :

In slobel operators higher weights are assigned to the pixels, close to the candidate pixel.

So, it is defined as :

$$d_1 = G_x = \left[(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \right]$$

$$d_2 = G_y = \left[(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \right]$$

Then,
the corresponding 3x3 mask for slobel operator will be :

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

for d_1 for d_2

(B) Second Order Derivative : (pre-university)

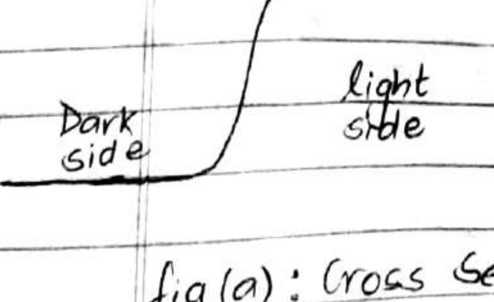


fig (a) : Cross Section of ramp edge

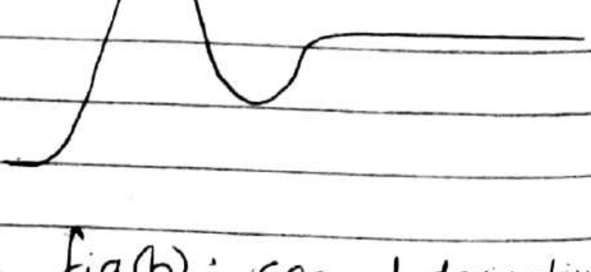


fig (b) : second derivative of ramp edge

The second derivative is $\boxed{\text{positive}}$ at the transition intoramp i.e associated with the dark side of the edge and $\boxed{\text{negative}}$ at the transition associated with the light side of the edge i.e out of the ramp. And, $\boxed{\text{zero}}$ along the ramp in area of constant gray level.

The sign of second derivatives can be used to determine whether an edge pixel lies on the dark or light side of the edge.

So, there are additional properties of the second derivatives around the edge.

For second order derivatives following two point is mostly important:

→ It produces two values of every edge in an image.

→ An imaginary straight lines joining the extreme positive and negative value of the second derivative would cost zero near the mid-point of the edge.

Second order derivatives is derived by the Laplacian filter or operator.

The laplacian of a 2D-function (x, y) is a second order derivative which is defined as:

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

i.e

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

$$\begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix}$$

For a 3X3 region, one of the two features encountered most frequently as:

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

The laplacian mask is:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

(ii) A digital approximation including the diagonal neighbour is given by:

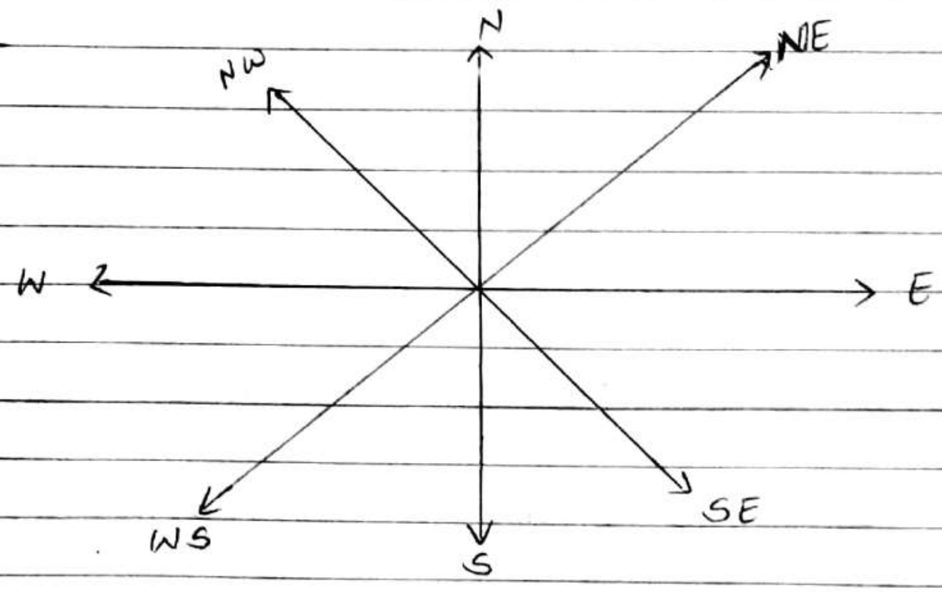
$$\Delta^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

Then,

The laplacian mask will be:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

③ Compass Operator:



(fig: compass operator)

Compass Operators measures the gradient in a selected number of direction.

The ~~metrics~~ matrix representation for different direction is:

(i) N ↑ $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$

N → +ve
S → −ve

(ii) $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$ ← W

W → +ve
E → −ve

(iii) ↓ S $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$ → E

NW → +ve

(v) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$ NW ↑ (vi) $\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ → NE

(vii) $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$ ← SW (viii) $\begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ → SE

(fig : Direction wise Matrix Representation of compass operator)

Line & Spot Detection :

① Line Detection / Curve Detection:

It is an important step in image processing and analysis. Lines are features in any scene taken by sensing devices.

We can detect the line by using following compass Gradient i.e 4 line detection kernel are as below :

(i) $\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$

→ Horizontal → +45°

-45°

(iii)
$$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

→ vertical

(iv)
$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

→ -45°

② Spot Detection :

These are most easily detected by comparing the value of a pixel with an average or median of the neighbourhood pixel.

The point detection mask or spot is given by :

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

(fig : Spot Detection kernel)

The point (spot) has been detected at the location on which the mask is centered if $|R| \geq T$.

where, T = Non-Negative Threshold
R = Response of mask at any point in image

The response of mask R is given by :

$$\boxed{R = \sum_{i=1}^{n} w_i z_i}$$

Where,

W_i = Window or mask coefficient
z_i = Gray level of the i^{th} pixel.

The idea is that, a spot will be detected easily from its surrounding by this method.

Chapter-8

Representation & Description

[3 hours]

Introduction to Representation:

The segmentation technique is a decomposition of an image generally referred to as raw data in the form of pixels along a boundary contained in a region.

Although these data sometimes are used directly to obtain descriptors that compact the data into representation that are considerably more useful in the computation of descriptors.

So, Representation refers to the dealing of naming for segmented image dedicated to the object point.

Representation can be classified into following two types:

(1) Chain Codes

(2) Signatures

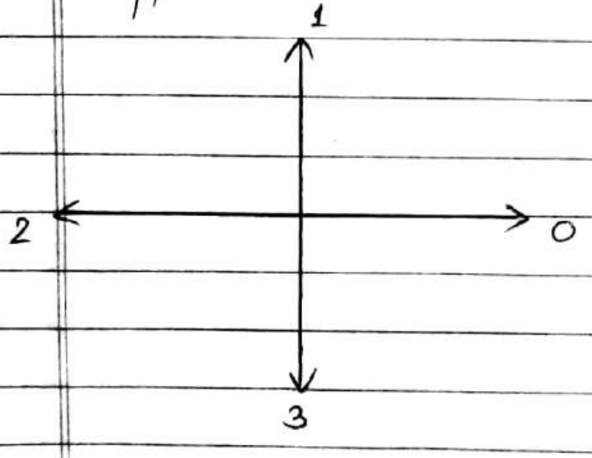
(1) Chain Codes :

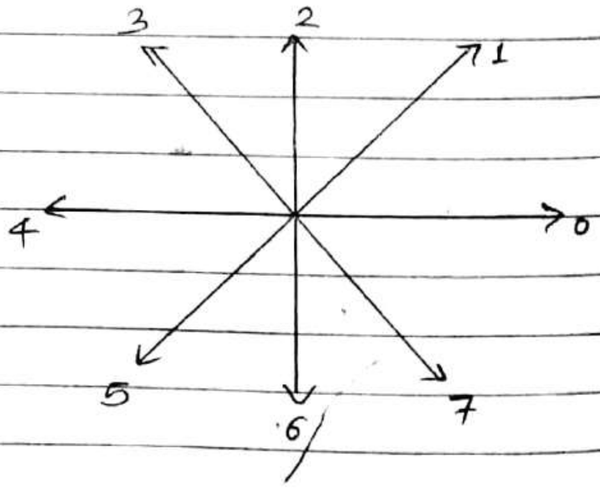
(2015 fall)

chain codes are used to represent a boundary by a connected sequence of straight lines segments of specified length and direction.

Typically, this representation is based on 4 or 8 connectivity of the segments. The direction of each segment is coded by using a number.

Chain code on this scheme are reffered as:





fig(a): Direction Number for 4-directional chain code

fig(b): Direction Number for 8-directional chain code

(The chain code of a boundary depends on the starting point.) However, the code can be normalized with respect to the starting point by treating it as a circular sequence of direction numbers and redefining the starting point so that the resulting sequence of numbers is a integer magnitude.

We can normalize by using rotation i.e 90° or 45° for the codes using the difference of the chain code.

Above figure shows the best example of 4 and 8 - directional chain codes.

② Signatures : (pre-university)

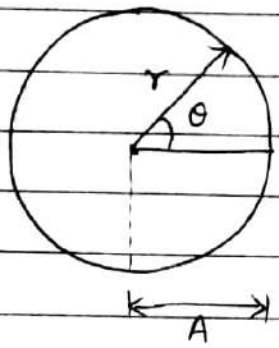
Signature is a 1-Dimensional functional representation of a boundary and may be generated in various ways.

One of the simplest method is plot the distance from a centroid point to the boundary as a function of angle.

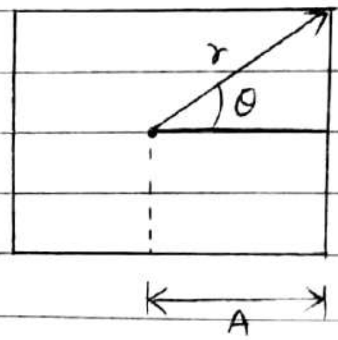
Signature generated by the this approach just described ~~are~~ is invariant to translation but they depends on rotation and scaling.

Normalization with respect to the rotation can be achieved by finding a way to select the same starting point to generate the signature.

The signature can be defined by using the following object:

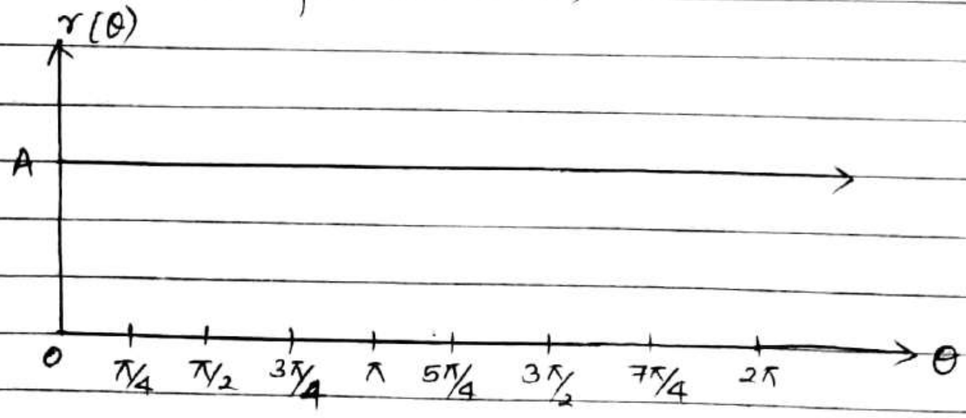


fig(a): Circular object

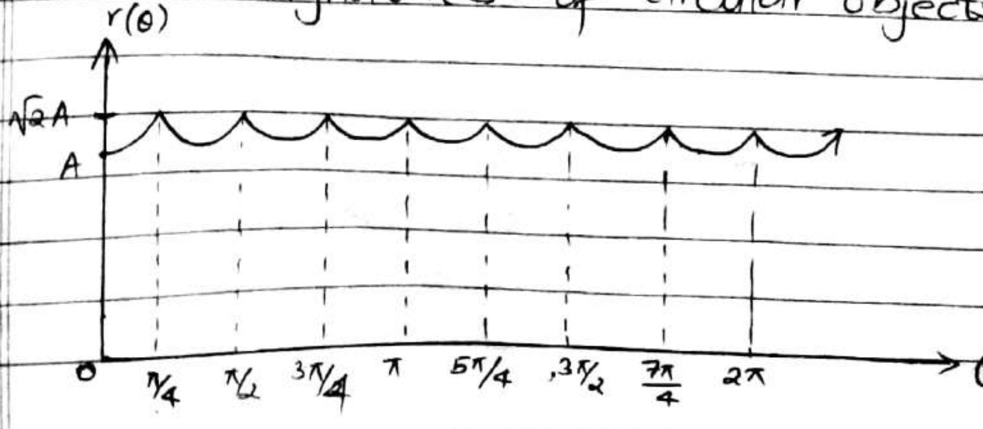


fig(b): square object

The corresponding angle signature can be defined as,



fig(c): Corresponding distance Vs Angle Signatures of circular objects



fig(d): Corresponding distance Vs Angle signature of square object

(A similar set of commands yeilded the plot.) Simply counting the number of prominent peaks in the signature is sufficient to differentiate between the fundamental shape of the boundaries as shown in the figure below:

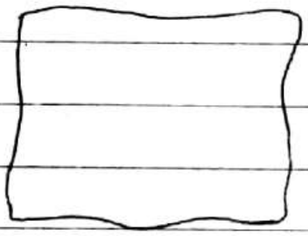


fig (a): Boundaries of an irregulare square

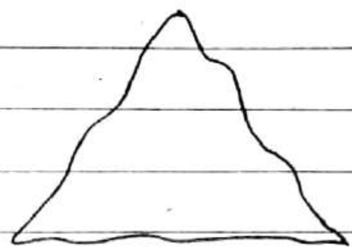


fig (b): Boundaries of irregular triangle

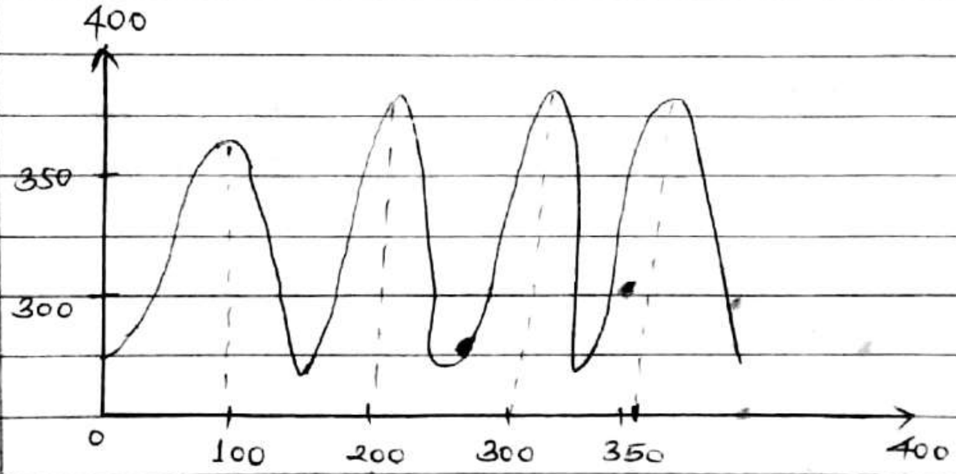


fig (c): corresponding signature of an irregular square

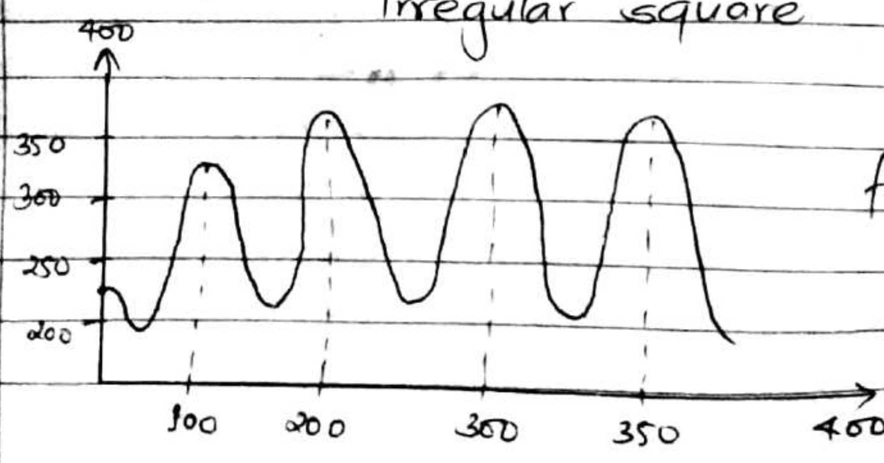


fig (d): corresponding signature of irregular triangle

Introduction To descriptors:

Descriptors are also called features of an image that are useful when working with region boundaries.

Many of the descriptors are applicable to region and grouping of descriptors in the tool box. The length of a boundary is one of its simplest descriptors.

The length of a four connected boundary is defined as the number of pixel - 1.

Descriptor is usually categorized into two types based on the boundary:

(1) Shape Number
(2) Fourier Descriptor

(1) Shape Number:

2015 fall

The shape number of a boundary generally based on four directional chain codes, is defined as the first difference of smallest magnitude.

The order n of a shape number is defined as the number of digits in its representation.

eover, 'n' is even for a closed boundary and its value limits the number of possible different shapes.

Thus, the shape number of a boundary is given by the parameter and order of shape number (length).

The four directional chain codes can be made insensitive to the starting point by using the integer of minimum magnitude and made rotation that are multiples of 90° by using the first difference of the code.

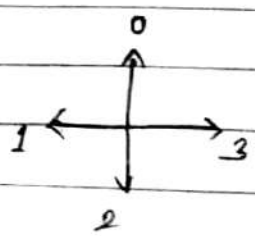
Thus, shape numbers are insensitive to the starting point and rotation that are multiple of 90°.

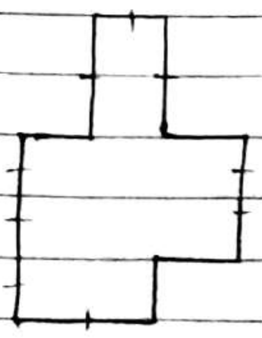
Shape number is illustrated step by step as:

①

②

③





④

$\therefore$ Chain code = 0000300 3223 2221211

Difference = $\overset{3}{0}$00031311311 00310310

Shape no. = 000310330130031303

(fig: steps in the generation of shape number)

(2) Fourier Descriptor:

(pre university)

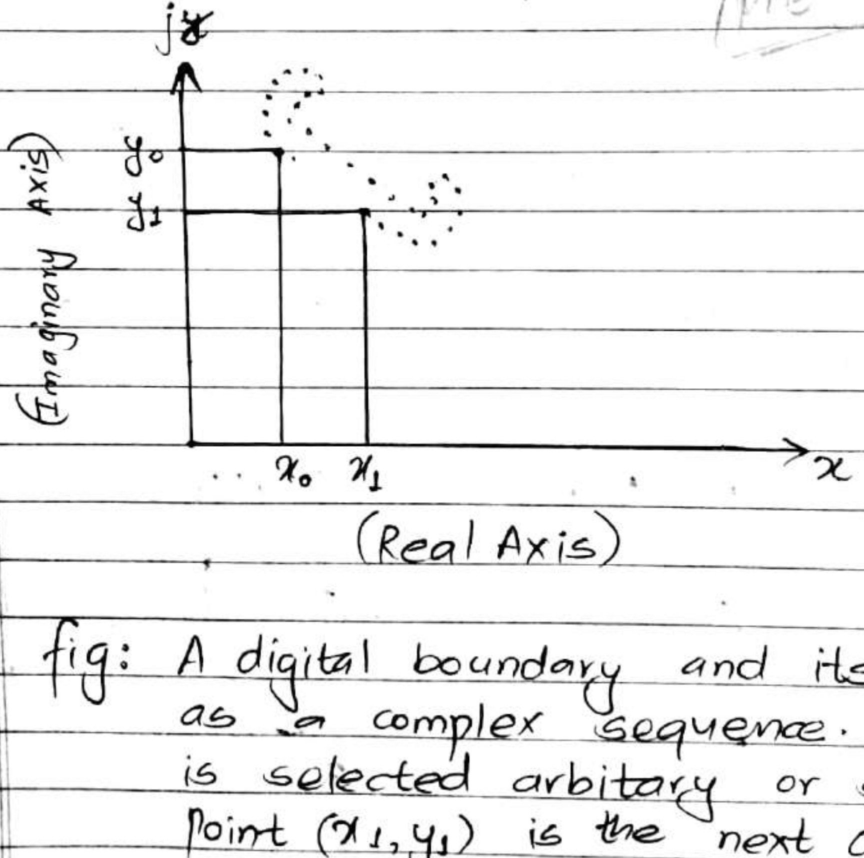


fig: A digital boundary and its representation as a complex sequence. point (x_0, y_0) is selected arbitary or starting point. Point (x_1, y_1) is the next counter clockwise point in a sequence.

Fourier descriptors can be explained by using the following steps:

(i) Define the arbitary point (x_0, y_0) and the co-ordinate point or pairs $(x_0, y_0), (x_1, y_1), (x_2, y_2) \ldots (x_{K-1}, y_{K-1})$ are encountered in traversing the boundary in counter-clockwise direction.

(ii) These co-ordinates can be expressed in the form $x(k) = x_k$ and $y(k) = y_k$. The boundary can be represented as the sequence of co-ordinates as:

$$S(k) = [x(k) \cdot y(k)] \quad \text{for } k = 0, 1, 2, \ldots K-1$$

The each co-ordinate pair can be treated as a complex number so that

$$S(k) = x(k) + j\,y(k).$$

(iii) The ~~descrip~~ discrete fourier transform of 1D sequence, $S(k)$ can be written as:

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} S(k) \cdot e^{-j2\pi uk/K}$$

Where,

$$u = 0, 1, 2, \ldots K-1$$

(iv) The complex coefficient $a(u)$ are called the "Fourier Descriptors". of the boundary.

The inverse fourier transform of these coefficient is:

$$s(k) = \sum_{u=0}^{K-1} a(u) \cdot e^{j2\pi uk/K}$$

(V) Therefore,
for boundary descriptors, the Fourier Descriptors can be explained as Fourier Coefficient which is defined as:

$$\boxed{\hat{s}(k) = \sum_{u=0}^{P-1} a(u) \cdot e^{j2\pi uk/P}}$$

Where,

P = Preceeding equation for $a(u)$
Where $p = 0, 1, \ldots P-1$

$\hat{s}(k)$ = Approximation to $s(k)$

Chapter - 9
Object Recognition [3 hours]

Classification or learning of an Image or Decision Theoretic Methods:

Object Recognition & classification

- (1) Unsupervised classification (No target or teacher or sample signal)
- (2) supervised classification (Presence of target or teacher or sample signal)
  - (1) statistical supervised classification (Bayesian classifier)
  - (2) Distribution free (Neighbourhood classification)

(fig: Decision Theoretic Methods)

① Unsupervised classification:

A very common task concerning in image processing is classification which is done in order to use the image for mapping or further analysis.

In unsupervised classification learning, there is no teacher or target or sample signal in the same way, it is also computerized method without direction of analysis.

It means that the output are based on software without providing user signal or sample classes.

The computer uses the technique to determine which pixel is detected and what classes belongs together.

In this technique, the user or human doesnot interface to the unsupervised classification.

Cluster is one of the unpervised learning technique or classification

For unsupervised classification, the figure below shows the technique of object recognition:

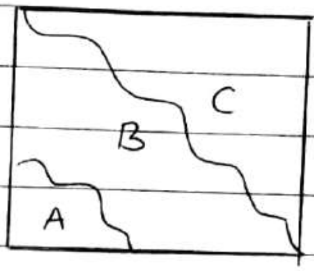
101	105	105	105
101	101	105	105
103	101	101	101
103	103	101	101

i/p image

$\Longrightarrow$

B	C	C	C
B	B	C	C
A	B	B	B
A	A	B	B

$\Longrightarrow$



class classification

fig: Example of unsupervised learning or classification using clustering technique

where,

A = Water
B = Land
C = Rock

② Supervised classification:

In supervised classification, it doesnot use the computer software to create the classes.

Here, ~~th~~ the analysis identify the several areas in an image which represents different futures. These known areas are referred to as Training site. The computer only do the assignment of pixel to the classes.

It is based on the idea that, a user can select sample pixel or target signal in an image i.e representative of specific classification.

It means that, the provided user signal or teacher signal can help to make the classification without using computer tool.

Supervised classification is categorized into following two types:

(i) Statistical Supervised classification
(ii) Distribution free

(i) Statistical Supervised classification:

It is a type of supervised classification on which the hypothetical or probability of that classification is determined using presence of user signal or sample signal or teacher signal or target signal.

This still can be extended to classification of features or pattern where similar type of pattern are grouped into one class that same as the characteristics of user signal.

The term similar ~~is~~ in the containing objects is closeness to the features between sample signal and segmented region.

~~We~~ To justify this classification, we use Bayesian classifier.

ॐ

- one conditional probability
- two unconditional probability

(Imp)

$P(A/B) = P(A \cap B) / P(B)$
$P(B/A) = P(A \cap B) / P(A)$
$\Rightarrow P(B/A) \cdot P(A) = P(A/B) \cdot P(B)$

classmate Date Page

Bayesian classifier:

A Bayesian classifier is $\boxed{\text{simple probabilistics classifier}}$ based on the Bayesian Theorem.

The Bayesian classifier classified the features to contribute the probability of hypothetical evidence.

i.e

$$P\left(H_i/E\right) = \frac{P(E/H_i) \cdot P(H_i)}{\sum_{n=1}^{K} P(E/H_n) \cdot P(H_n)}$$

Where,

$P(E/H_i)$ = The probability that we will observe evidence "E" given that Hypothesis H_i is true.

$P(H_i)$ = A prior probability that the Hypothesis H_i is true.

$P(H_i/E)$ = The probability that we will observe hypothesis H_i given that evidence "E" is true.

K = The no. of possible Hypothesis.

(ii) Distribution Free:

It is a type of supervised classification and donot require knowledge of any density function on p.d.f based on the heuristic information or algorithm.

The example of distribution free is nearest neighbourhood classification so that it is also called as Neighbourhood classification.

A very commonly used classification method is called Nearest Neighbourhood classification.

Therefore, it works as follows:

(1) We create a database of any object for which we already known that the correct classification should be known.

When the system is given a query i.e new object classify the system. It finds the nearest neighbourhood of the query in the database.

(2) Then, the system classify the query as belongs to the same classes as its nearest neighbours i.e database images.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 5 & 0 & 2 \end{bmatrix} \xrightarrow{\text{I/p query '4'}} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 5 & 0 & 2 \end{bmatrix}$$

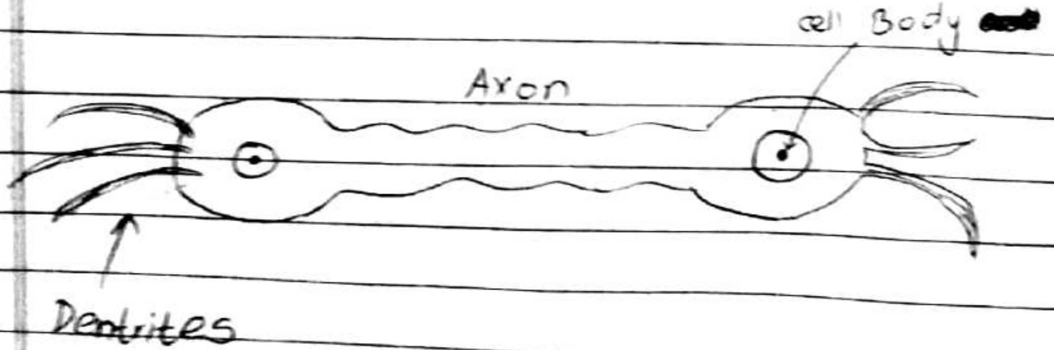
(fig: ~~Recording~~ Recognizing digit using Neighbourhood Classifier)

(3) Then, the system classify the query as an image of 4.

(4) Now, the image is identified and classify by the distribution free supervised classification.

Here, query 4 is the target signal by the user.

Neuron, Neural Network (NN) & Artificial Neural Network (ANN) :



(fig: Neural Components)

Neuron:

A neuron is a cell in the brain whose principle function is the collection, processing and dissemination of electrical signal.

The cell body is a part of the cell containing the nucleus and maintaining protein synthesis.

A neuron may have many dentrites which branch out tree like structure and receive signals from other neurons.

The axon conducts electric signal generated at the axon hillock.

Neural Network (NN) & Aritificial Neural Network (ANN):

Neural network is the branch of the field known as AI. A neural network can be considered as black-box i.e able to prediction of output pattern when it recognizes the given input pattern.

Artificial Neural Network (ANN) is the type of Neural Netwook which is inspired as artifically to show the human behaviour.

ANN shows the brain process information system.

Why use NN ?

NN is only used for following:

(1) Self organization
(2) Real time operation
(3) Information is distributed over the entire network.
(4) Adaptive learning

Applications of Neural Network (NN):

(i) Automotive, Automobile and Automatic Guidance system

(ii) Aerospace and Aircraft component simulation.

(iii) Banking cheque reader

(iv) ~~Defe~~ Different facial recognition

(v) Manufacturating product design (CAD)

(vi) Telecommunication automated information

(vii) Robotics control system

Disadvantages of Neural Network (NN):

(i) NN needs training to operate.

(ii) The architecture of NN is different from the architecture of micro-processor.

(iii) Requires high processing time for large neural network.

(iv) Hard to setup programming.

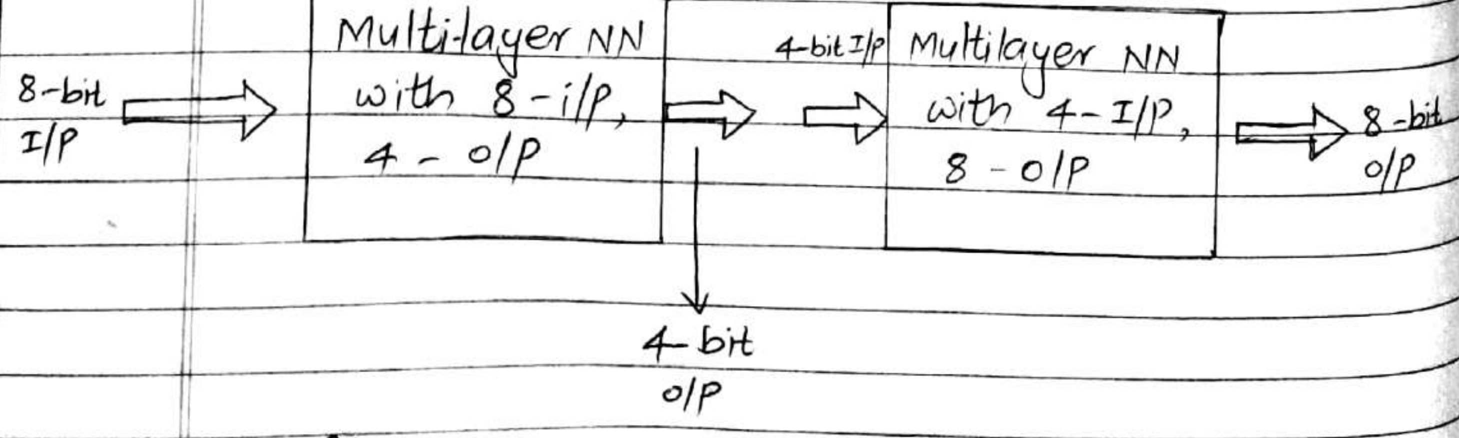
(V) lack of well manpower.

(V.V Imp)

Applications Of NN In Image Processing:

(1) NN for Image Compression
(2) NN for Pattern Recognition
(3) NN for Perceptron (Perception)

(1) NN for Image Compression:



(fig: Block Diagram of Image compression)

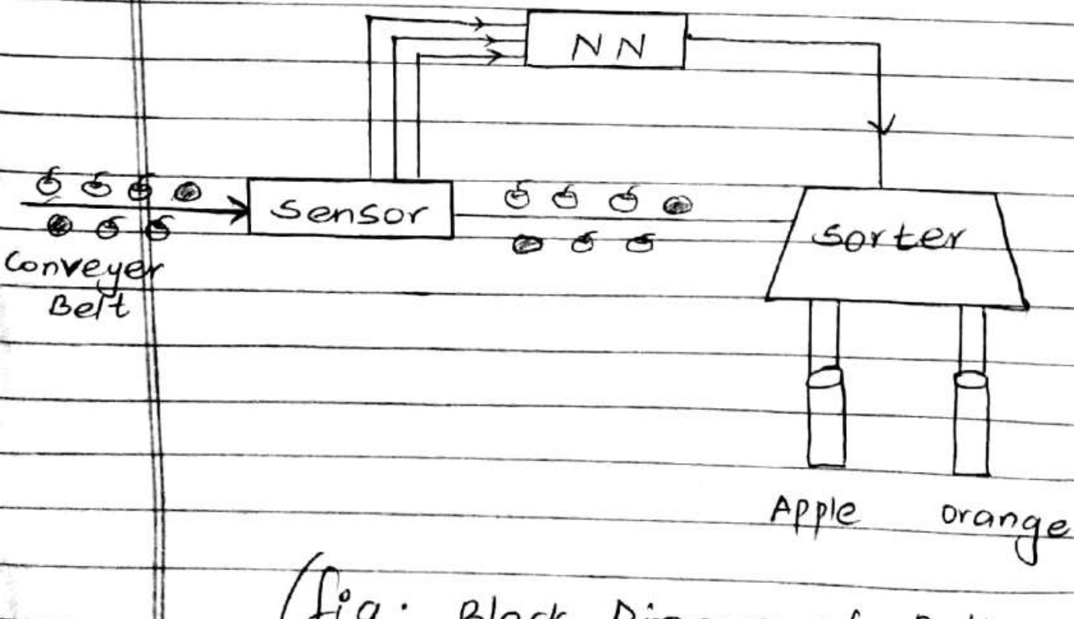
NN model have received much attention in many fields where high compression ratio is required.

Many NN approaches for image compression gives superior performance than the discrete traditional approach.

Digital Image generated with 8-bit may be reduced in size by feeding it to 8 x 4 multi-layer NN.

Neural network (NN) on the other side then process the compressed image by 4 x 8 multilayer NN for reconstruction.

(Imp) ② NN for Pattern Recognition: (2012 fall)



(fig: Block Diagram of Pattern Recognition (PR))

The NN is ideal tool for pattern recognition. Any recognition system needs to be trained to recognize different patterns.

NN is also simplification of human neuron network system. It is more likely to adapt the human way of solving the recognition problem than other techniques.

The design of NN system for PR starts from collecting data on each of objects that is to be recognized by the system.

for Example: A dealer has a warehouse that stores variety of fruits and vegetables that are mixed together.

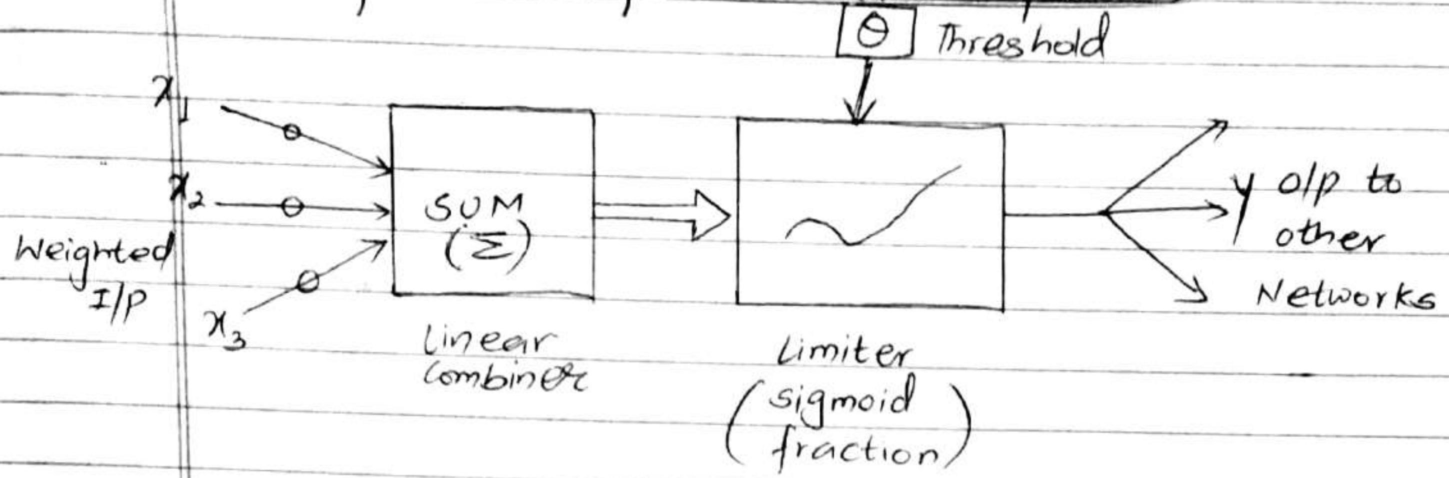
The dealer wants a machine that will sort the fruits according to their type.

The fruit is loaded on the conveyer belt and the fruit passes through the sensor which measures the shape, texture, color and weight properties.

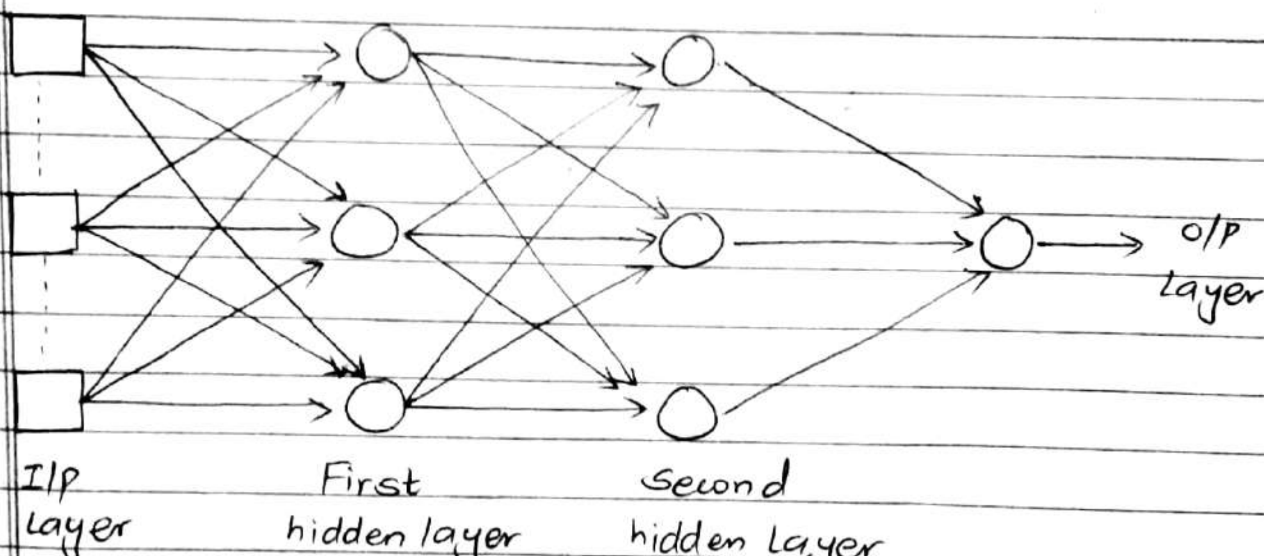
The output will be input to the purpose of NN ~~to~~ decided what kind of fruit is on the conveyer belt. So, fruit can be directed to the correct storer.

(2015 fall) L.Q

③ NN for Perceptron (Perception):



(fig: Block Diagram of Perceptron)



(fig: Formation of o/p by Perceptron)

It is one of the earliest neural network model.

Here, i/p^s from one or more previous neurons are individually weighted and then summed.

The result is non-linearilly

scaled between 0 and +1. And then, the result is passed to the neurons in the next layer for output result.

Several perceptions can be combined to form multi-layer perception (MLP). So MLP is a development from the simple perception in which extra hidden layers are added.

Here, more than one layer can be used.

Generally, connections are allowed from input layer, hidden layer and so on. At first i/p are feeded into the i/p layer and get multiplied by inter-connection weights as they are passed from the i/p layer to the first hidden layer.

After first layer then go to the second layer and finally data is multiplied by inter-connection weight and so on.

Back propagation never take the perception methods but at the time of sensing accuracy it is used.

(T/P)

(Assignment - I)

FFT (Fast Fourier Transform on the perspective of Image Enhancement)

pramod pararuli@gmail.com

SSM | D.C | OOSE | E.S | NET Tech |
A | A- | A | A- | A | A

6th sem

Perceptron

- simplest kind of feed forward neural n/w
- The Model consist of linear combiner followed by an activation function
- The weighted sum of i/ps is applied to the activation function, which produces an o/p equal to +1, if its i/p is +ve & -1 if its i/p is -ve.