Chapter-1

Introduction To Digital Image Processing

(4 hours)

Introduction:

The pictorial representation of an object is called formation of an image

A **digital image** is a 2-D light intensity function $f(x, y)$ where x and y denotes spatial co-ordinates and value of function 'f' at any point (x, y) is proportional to the brightness (gray level) of an image at that point (x, y).

Digital Image Processing refers to the processing of digital image by using digital computer on the basis of digitization (sampling and quantization).

Digital Image Representation:

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

It is represented by 2-D integer array. The digitized brightness value is called gray level values.

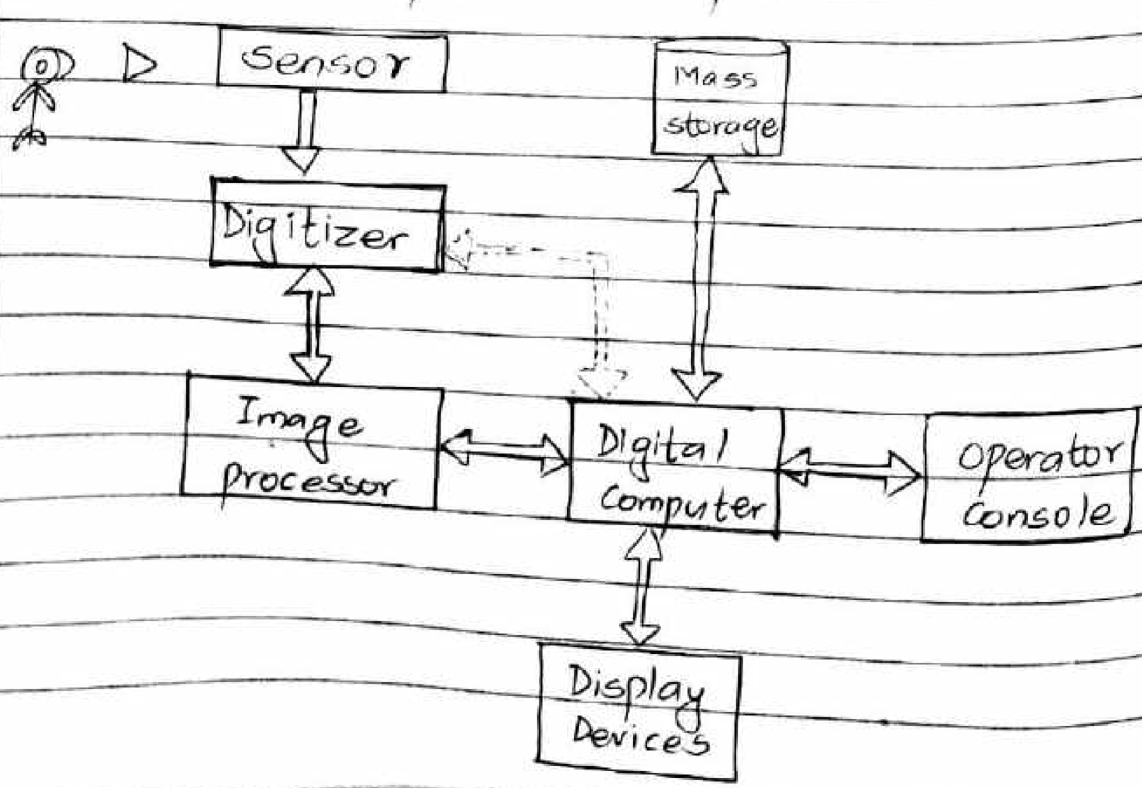
Digital image can be represented as a matrix as shown in figure below:

$$f(x,y) = \begin{bmatrix} f(0,0), & f(1,0) & \cdots\cdots \\ & & \\ \cdots\cdots & & f(x_{max}, y_{max}) \end{bmatrix}$$

(fig: Representation of Digital Image As Matrix)

Here,
the row and column identify the point in the image and the corresponding matrix elements values identify the gray level at that point.
The element in digital image is called Pixel or pels or Picture Points.

Elements (Components) of DIP:



classmate
Date
Page

1. Sensor:

It intercepts the energy propagation from the scene and transform it to produce an intensity image.

Camera and scanner are the examples of photo sensor.

2. Digitizer:

The digitizer produce an image by the digitization technique.

Normally, a digitizer converts an intensity image into numerical representation which is suitable for import into a digital computer.

3. Image Processor / Pre-Processor:

It interface with digital computer to provide easy of programming. For this purpose, We can use micro-computer or mainframe computer.

Specially, it does functions like image acquistion, low level processing, scaling and rotating.

4. Digital Computer:

It always digitize the required image by the help of image pre-processor.

5. Mass storage:

It is used to store digital image line by line. The storage devices may be magnetic tape, hard disk etc.

classmate
Date
Page

⑥ Display Devices:

It produce and shows a visual of numerical values stored in a computer.

The example of display devices are: CRT monitor, printer etc.

⑦ Operator Console:

It controls the full mechanism of image processing system.

Applications of Digital Image Processing:

① Office Automation:

BIO MECI

Ex: OCR (optical character Reader) & document processing etc.

② Industrial Automation:

Ex: Robotics, Automatic assembly etc.

③ Bio Medical Automation:

Ex: ECG, EEG, X-Ray etc.

④ Information Technology (IT):

Ex: Video phone, fax, video conferencing, etc.

classmate
Date
Page

(5) Entertainment:

Ex: Video game, visual Animations, etc.

(6) Military Application:

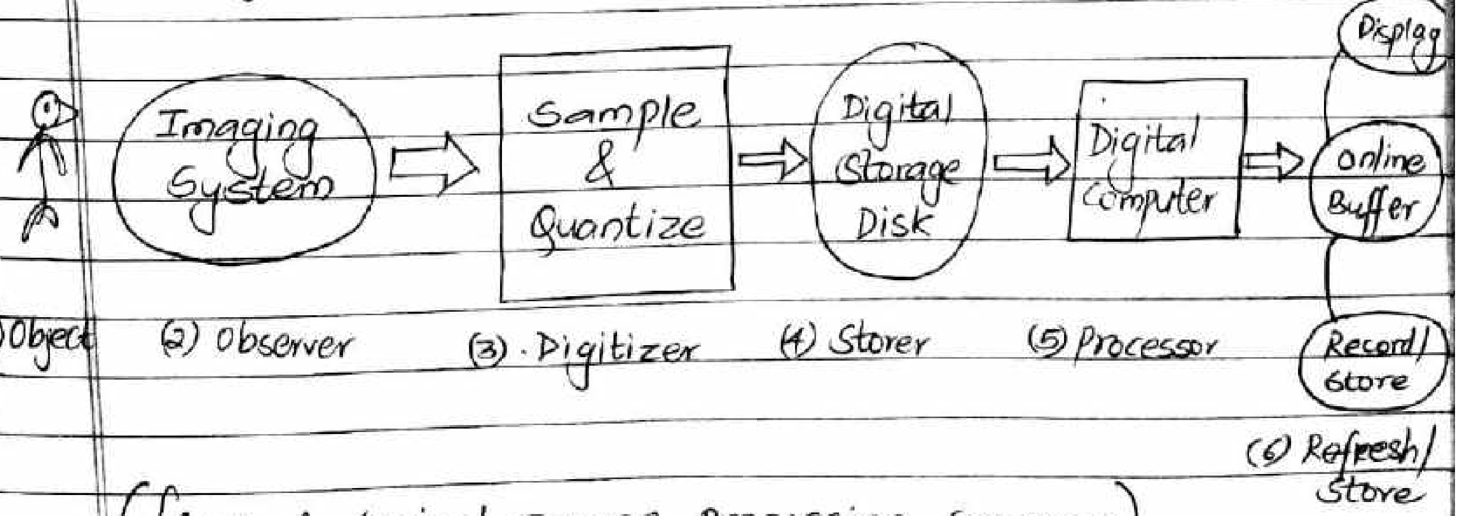
Ex. Target Identification and missile guidance and detection.

(7) Criminology:

Ex. Bio-metric detection such as finger print detection, face detection, retina detection, etc.

(2010, spring)

Imp # A typical Image Processing Sequence:



(fig: A typical Image processing sequence)

In typical image processing sequence involves various components.

The image of an object travels along the different blocks.

Firstly, the object is ~~cha~~ sensed by the Imaging system and image is formed. This image is digitized by the process of sampling and quantization.

Now, image is in a digital form. A digital image is stored in digital storage system. The digital computer is connected to digital storage disk where all the manipulation of digital image is done.

Here, also online buffer mechanism is available where the information are stored as memory.

The image is displayed through the display block which may be color monitor, color TV etc. Finally, the image is recorded on a recording block.

But sometimes some problem may occur when the image is processing sequentially into a digital computer.

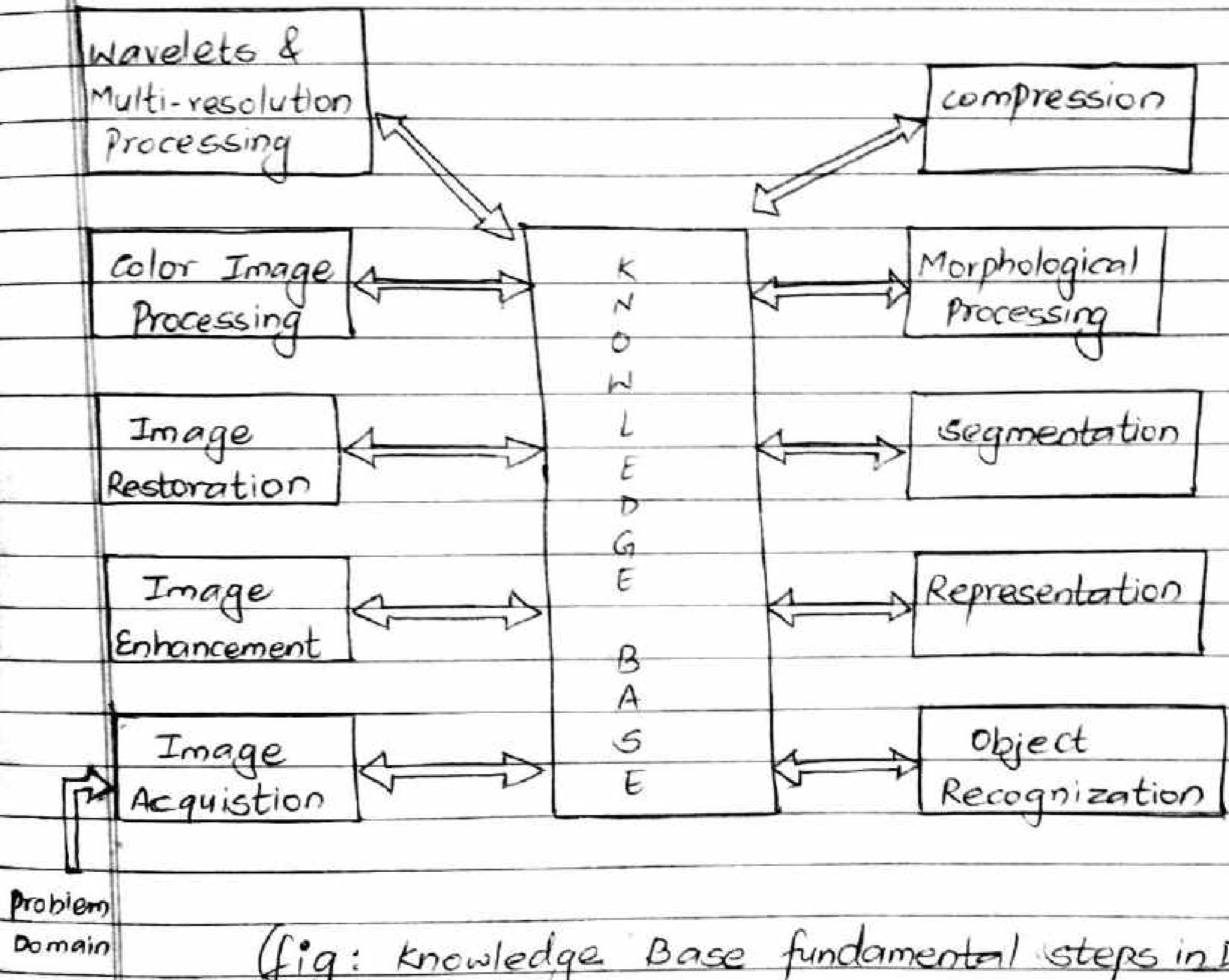
AER/ Cp WP/ 𝓃 / SRO

(2015 fall)

classmate
Date
Page

Imp

Fundamental steps in DIP:



(fig: knowledge Base fundamental steps in DIP)

① Image Acquistion:

It is the first process in which pre-processing activities like scaling and rotating are done.

② Image Enhancement:

In this stage, the original image is

processed in such a manner that the resultant image will be more suitable with compared to original image.

③ <u>Image Restoration:</u>

It is an area that also deals with improving apparance of an image. Basically,
image enhancement is subjective process but image restoration is objective process.

④ <u>Color Image Processing:</u>

With the help of color, object identification and abstraction from the scene is processed.

⑤ <u>Wavelets & Multi-Resolution Processing:</u>

Foundation for the representing image in various degree of resolution deals to the wavelets and multi-resolution.

Sometimes, it is also used for image compression.

classmate
Date
Page

6) Compression:

It deals with the technique for reducing the storage required to save an image to transmit it.

7) Morphological Processing:

It is the combination of structure. So, it deals with the tools for extraction of an image component that are useful in representation and description of shape.

8) Segmentation:

It deals with the partation of an image into different parts and objectives.

9) Representation:

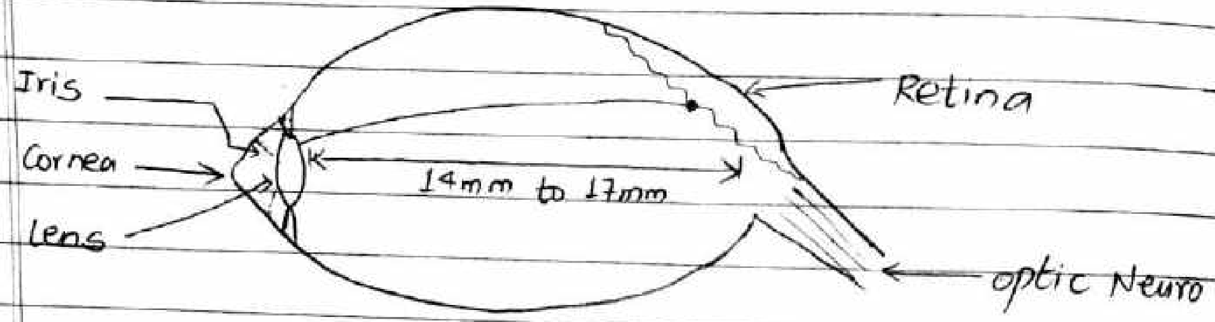
After an image has been segmented into sub-regions then the resultant segmented pixels is represent and described in a suitable form for further analysis.

10) Object Recognation:

It process that, the level to an object based on its descriptors.

(2011 fall)

Elements of Visual Perception:

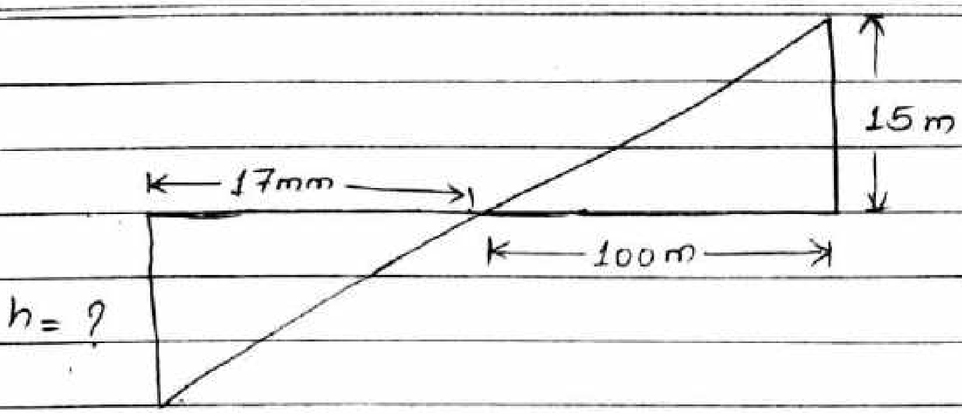


(fig: Cross Section of Human Eye)

Image analysis plays central role in the image processing. Therefore, the study of human visual perception is important which is represented as cross-section of human eye as shown in figure above.

The human eye nearly a sphere with an average diameter of 14mm to 17mm. The lens of the eye is flexible than the ordinary optical lens.

Let, the observer is looking at a tree 15m height at the distance of 100m. Find the height of retrival image.



Here,

$$\frac{h}{17\,mm} = \frac{15\,m}{100\,m}$$

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

$\therefore$ Height of Retrival Image (h) = 2.55 mm

V. Imp

Sampling & Quantization:

A physical image is changed into digital image by performing sampling and quantization. This process is also known as Digitization:

Sampling:

(It is the process where measurement are taken at regular space intervals.) Digitizing the spatial co-ordinates value (x, y) is called sampling.

classmate
Date
Page

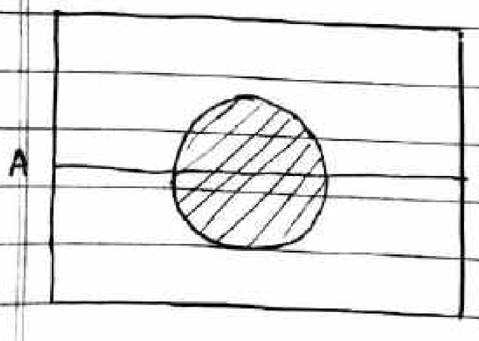
It is the process by which image is formed over a patch (domain) in a continous domain is mapped into a discrete point with integer co-ordinates.

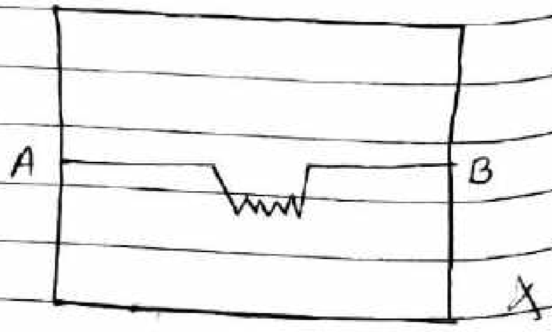
Hence, Sampling can be defined as a selection of discrete location in the continous 2-D space.

Quantization:

It is the process of mapping the measured intensity to one of the finite number of discrete levels.

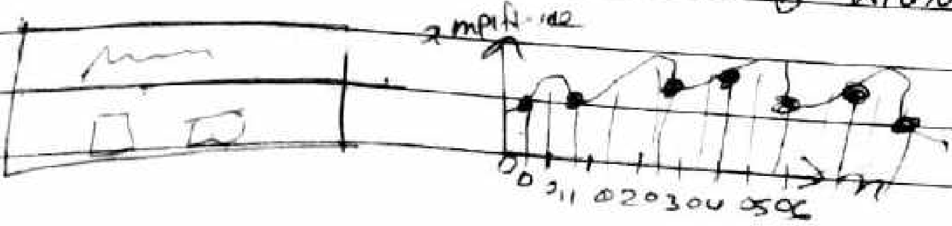
Hence, digitizing the amplitude values is called quantization or gray level quantization.



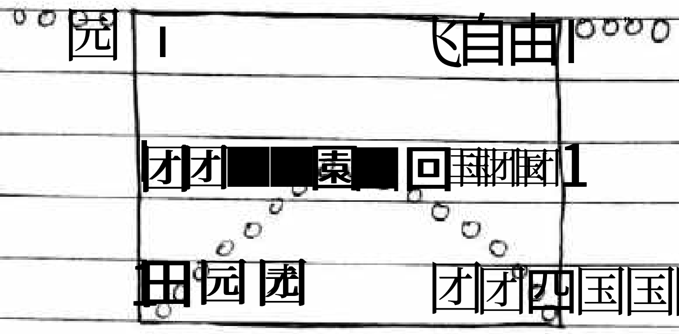


(1) Continuous Type

(2) Ideal Scan line A to A without noise



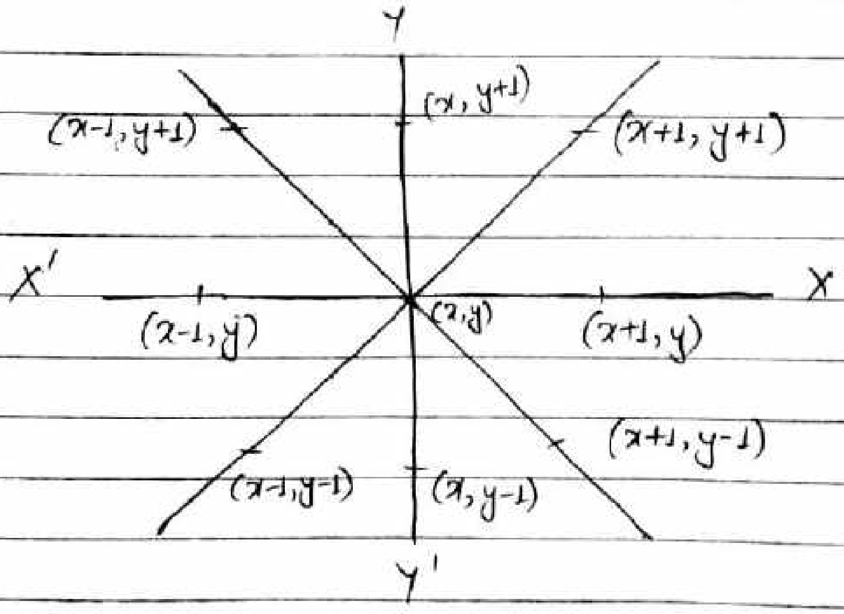
(a) Scan line A to B in image with noise



(b) Sampling and quantization

Neighbours of pixels:

(Arrangement of pixel in Image)



(fig: Neighbours of pixel)

A pixel at co-ordinate (x, y) has a four horizontal and four vertical neighbours whose co-ordinates are given by $(x-1, y+1)$, $(x, y+1)$, $(x+1, y+1)$, $(x+1, y)$, $(x+1, y-1)$, $(x, y-1)$, $(x-1, y-1)$, $(x-1, y)$.

These set of pixel is called four neighbours of horizontal and four neighbours of vertical.

Each pixel is at unit distance from (x,y) and called as $N_8(P)$.

To create a digital image, we need to convert the continous sensed data into digital form.

This involves two processes : Sampling & Quantization

An image may be continous with respect to x- and y-coordinates, and also in amplitude.

To convert it to digital form, we have to sample the function in both co-ordinates & in amplitude.

Digitizing the co-ordinate value is called Sampling.

And, Digitizing the amplitude value is called Quantization.

Neighborhood of Pixel

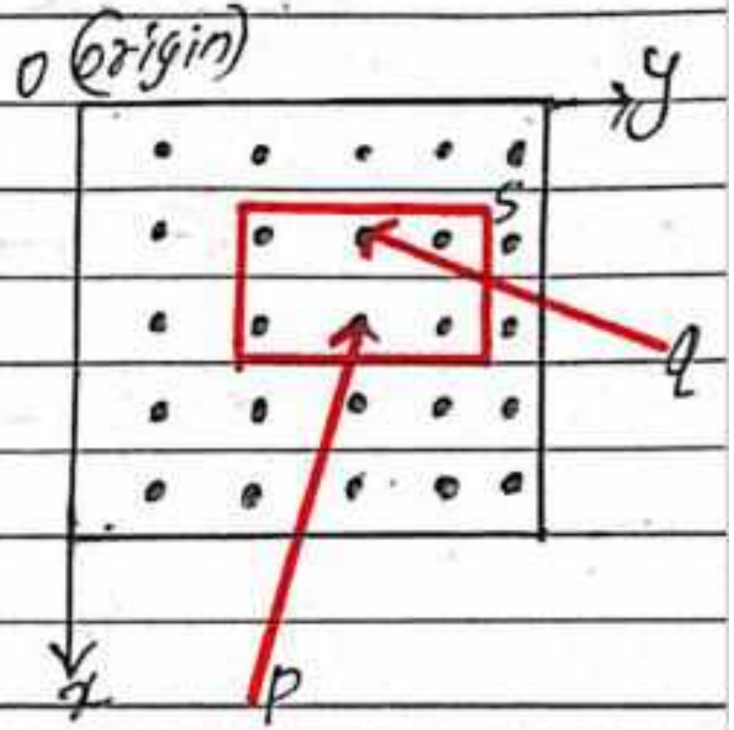
Basic Relationship between Pixel

Image → $f(x,y)$

S → subset of pixels

x represent no. of column

y represent " " rows



4-Neighbour's N4(P)

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

2 Horizontal

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

2 vertical

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

$$N4(P) = (x, y-1), (x, y+1), (x-1, y), (x+1, y)$$

Diagonal Neighbour $[N_D(P)] = (x+1, y-1), (x-1, y-1)$
$(x-1, y+1)\ (x+1, y+1)$

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

8 Neighbour $[N_8(p)]$

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

Connectivity / Adjacent

(1) 4- Adjacency
(2). 8- adjacency
(3) m- adjacency [mixed- adjacency]

Binary Image.

$V = \{1\}$

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

Grayscale Image [0-255]

$V = \{0, 1, 2, \ldots 10\}$

54	10	100	[illegible]4
81	150	2	34
201	200	3	45
7	70	147	56

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

[mixed connectivity]

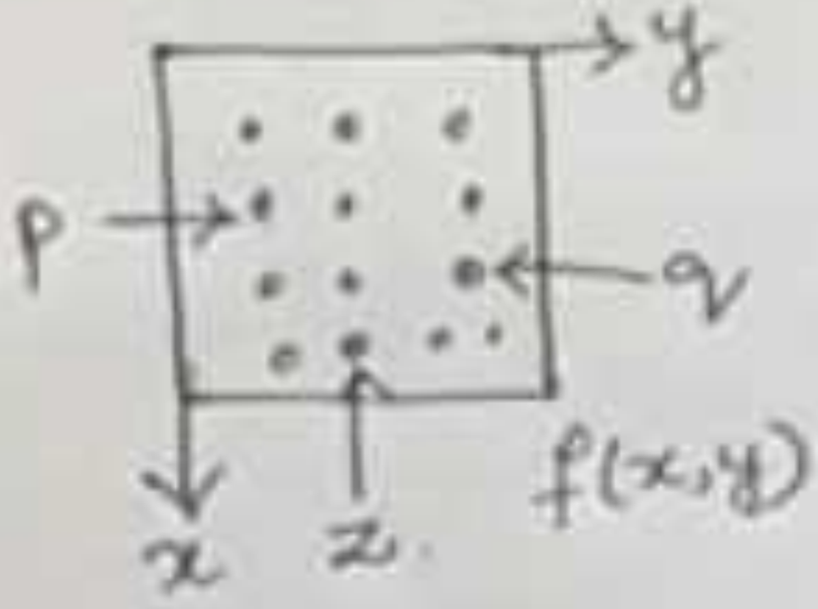
If you can able to connect in st. line. we should not connect in diagonal line

Distance Measures

DISTANCE MEASURE:

Image $\rightarrow f(x,y)$

$p, q, z \rightarrow$ particular pixels.



Distance function D

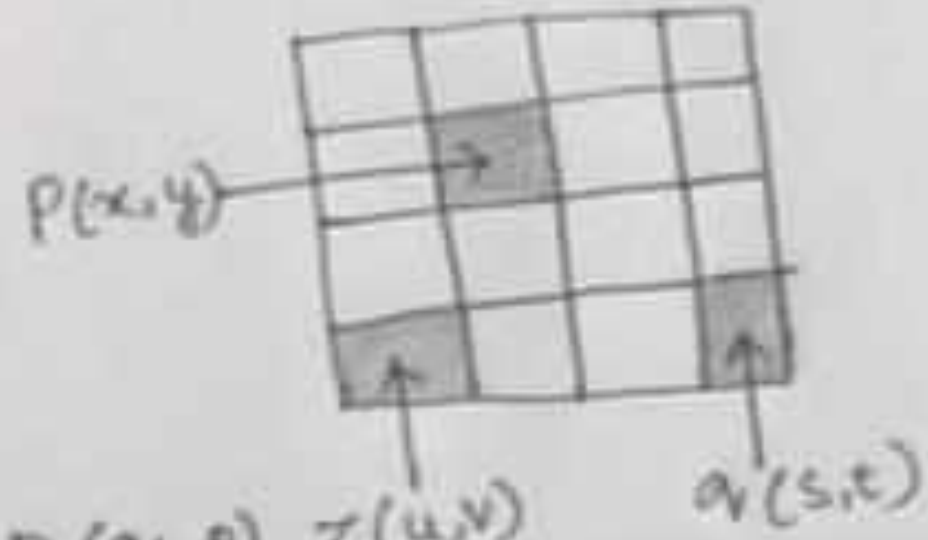
Properties of D

(i) $D(p, q) \geq 0$

(ii) $D(p, q) = 0$ if $p = q$

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

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



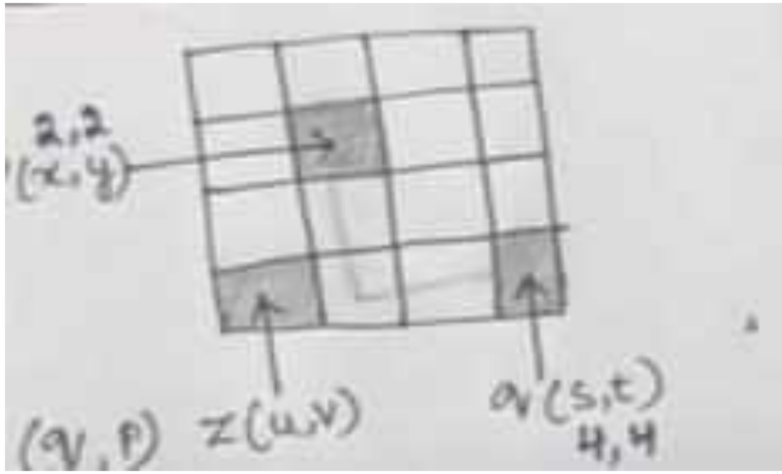
Distance measure

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

(ii) City Block : $D_4(p,q) = |x-s| + |y-t|$

(iii) Chess board : $D_8(p,q) = \max\{|x-s|, |y-t|\}$

Euclidean Distance



Ex:-

(a) $D_E(p,v) = [(2-4)^2 + (2-4)^2]^{1/2}$

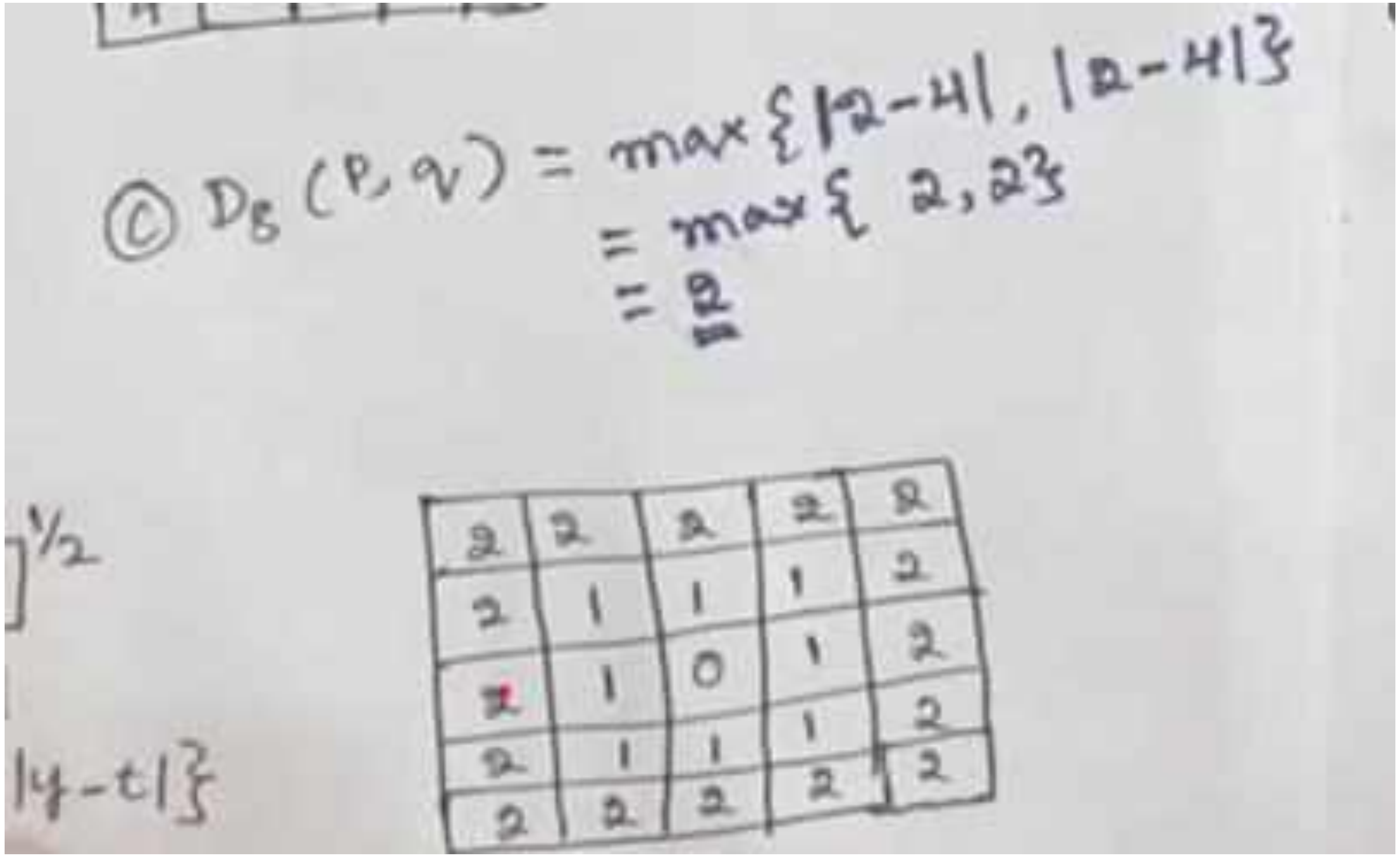
$= [8]^{1/2}$

City Block Distance

(b) $D_4(p,q) = |2-4| + |2-4| = 2+2$
$= \underline{4}$

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

Chess Board



classmate
Date ______
Page ______

Chapter - 2

Image Enhancement In Spatial Domain

[7 hours]

Introduction:

Image enhancement refers to the sharpening or to improve the quality of an image by enhancement of pixel or point in a spatial Domain.

Some basic gray level transfor-mation is needed to enhance the image in a spatial domain.

Some basic gray level transformation is listed as below:

(1) Point Operation
(2) Contrast Stretching
(3) Thresholding
(4) Digital Negative
(5) Intensity Level Slicing (Gray level slicing)
(6) Bit Plane slicing

[P.T.O]

(pre-university) (Defn)

Point Operation:

Point operation is defined as a function that are performed on each pixel of an image.

Point operation is a uninary point operation if single image is modified.

The operation can be binary which means that two images combined in the same manner.

The operation can be ternary if three images are used in same manner at operation.

Unary Operation can be defined as,

$$\boxed{T_{xy} = \text{function}\,(P(x,y))}$$

Where,

T_{xy} = Transformed Pixel Value

function = Numeric Transform

$P(x,y)$ = Input Pixel value

This operation can be used in many situation and for different purpose such as contrast stretching, noise clipping, window slicing and histogram modeling.

So, in point operation a given gray level $u \in [0, L-1]$ is mapped into a gray level $v \in [0, L-1]$ according to a transformation

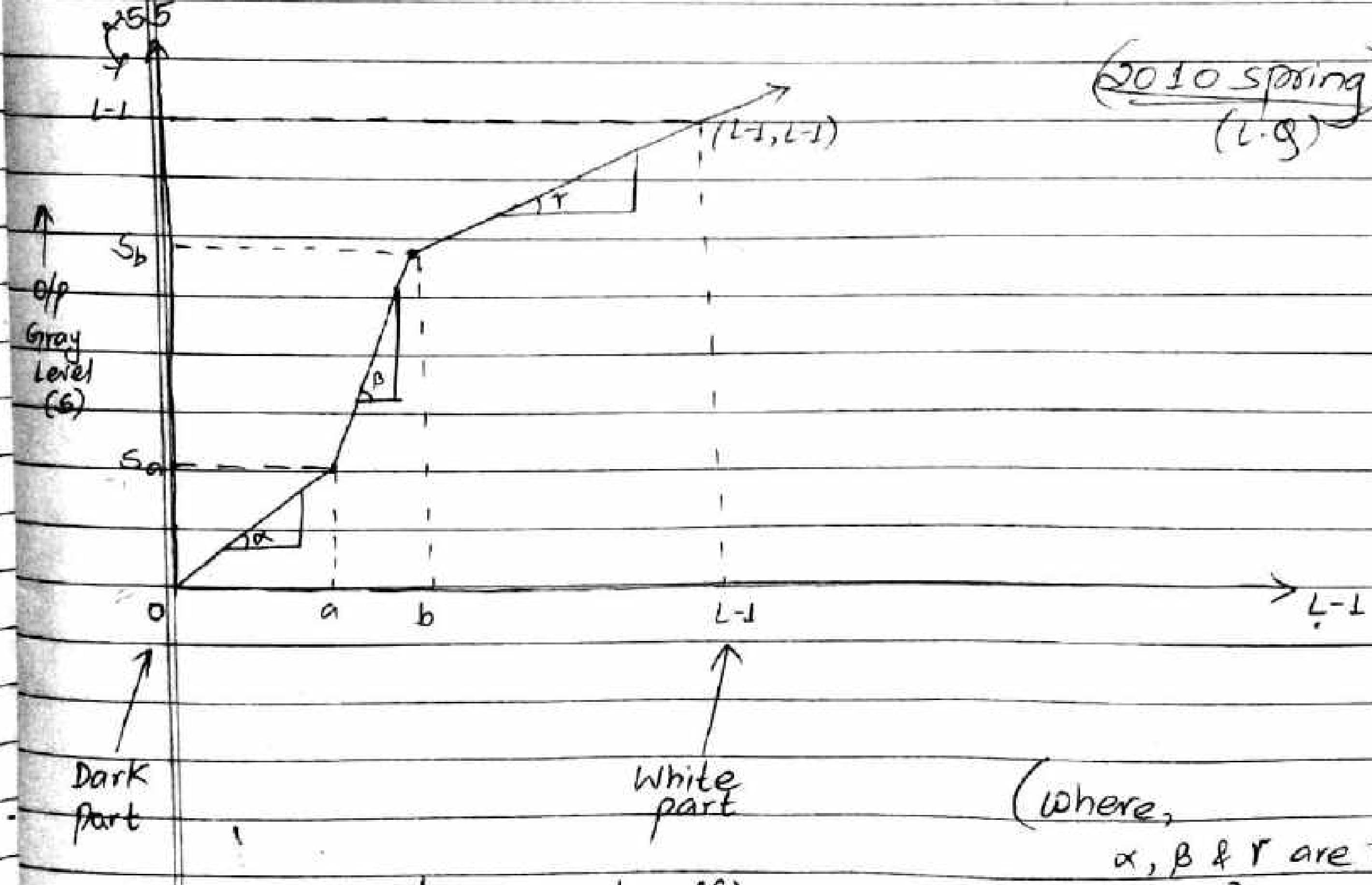
i.e $\boxed{v = f(u)}$ $\left(\begin{array}{l} L-1 = 255 \\ L \rightarrow \text{Max. Gray level} \end{array}\right)$

(Imp)

Contrast Stretching:

(Dec → 2012)

(2010 spring)
(1.9)



(where, α, β & γ are the slopes)

(fig: Contrast Stretching Transformation)

In many times, we obtain low contrast image due to poor illumination or due to the wrong setting of the lens aparture.

The idea behind the contrast stretching is to increase the dynamic range of gray level in the image.

The main idea is to increase the contrast of the image by making the dark portion darker and bright portion brighter.

To obtain the contrast stretching we should have some formulations as below:

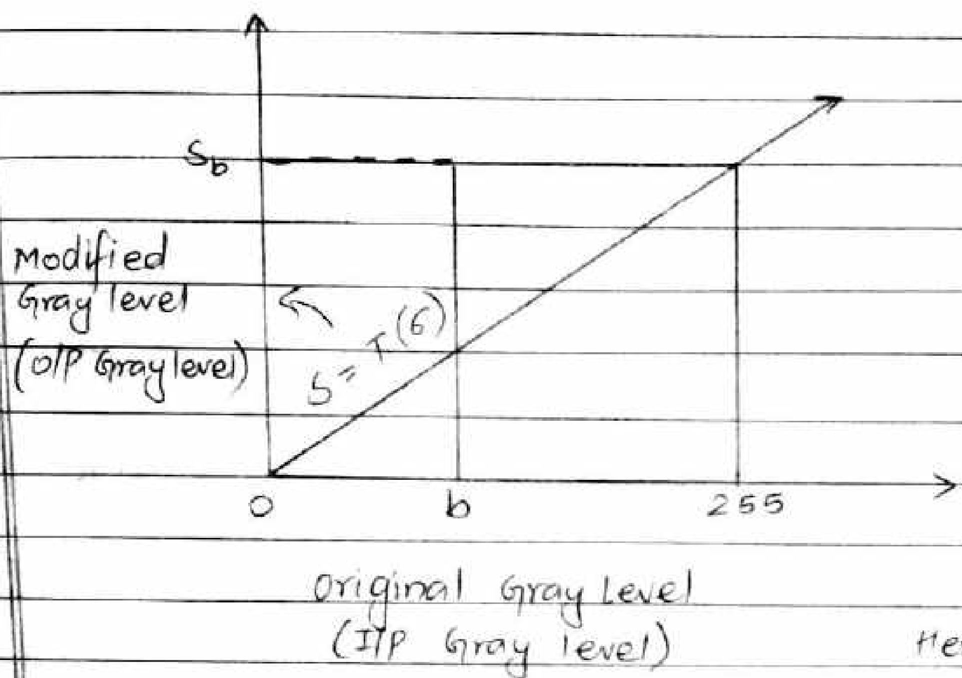
$$S = \begin{cases} \alpha r & , \quad 0 \le r < a \\ \beta(r-a) + S_a & , \quad a \le r < b \\ \gamma(r-b) + S_b & , \quad b \le r < k-1 \end{cases}$$

In the figure, 'r' represents the horizontal axis (i.e I/p gray level) and 's' represents the vertical axis (i.e o/p gray level.)

Here,

darker gray level is obtained by making a slope < 1 and brighter is obtained by making a slope > 1.

Thresholding (Binary Image)



Here,
$b = 100$ (Threshold value)

(fig: Thresholding)

Thresholding is the simplest method of the image processing. It can be used to create binary image.

If we observe the contrast stretching diagram, we notice that the first and last slopes are made to zero.

If $r_1 = r_2$ & $s_1 = 0$, $s_2 = L-1$

Now, The thresholding function is given below:

$s = 0$, if $r < b \Rightarrow 0$
$s = L-1$, if $r \geq b \Rightarrow 1$

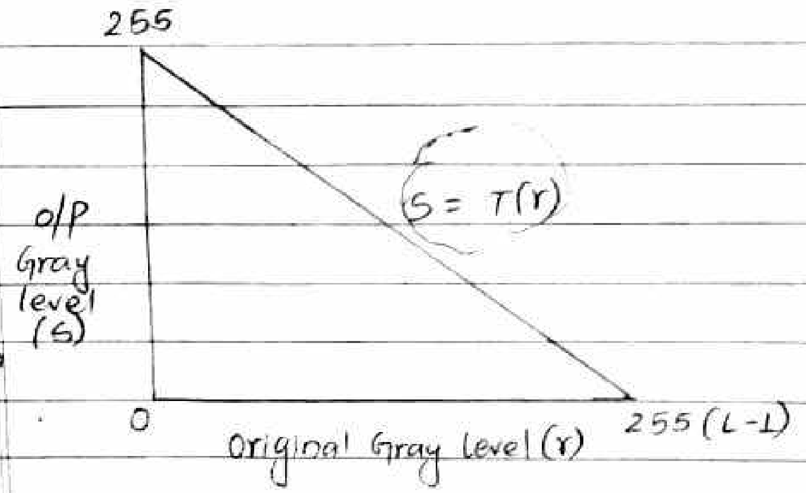
Where,

b = Thresholding value

L = Maximum Gray level

Important notice is that, threshold image has maximum contrast on it having black and white gray level only.

(Imp) # Digital Negative (x-ray): (2011 fall)



(fig: Digital Negative)

It is useful in a lots of application. The common example of digital negative is X-ray image.

As the name suggest, negative means simply inverting the gray level i.e black into white and vice-versa.

So, the negative of digital image

with the gray level in the range from 0 to L-1 is obtained by using negative transformation.

i.e
$$\boxed{s = T(r) = [(L-1) - r]}$$

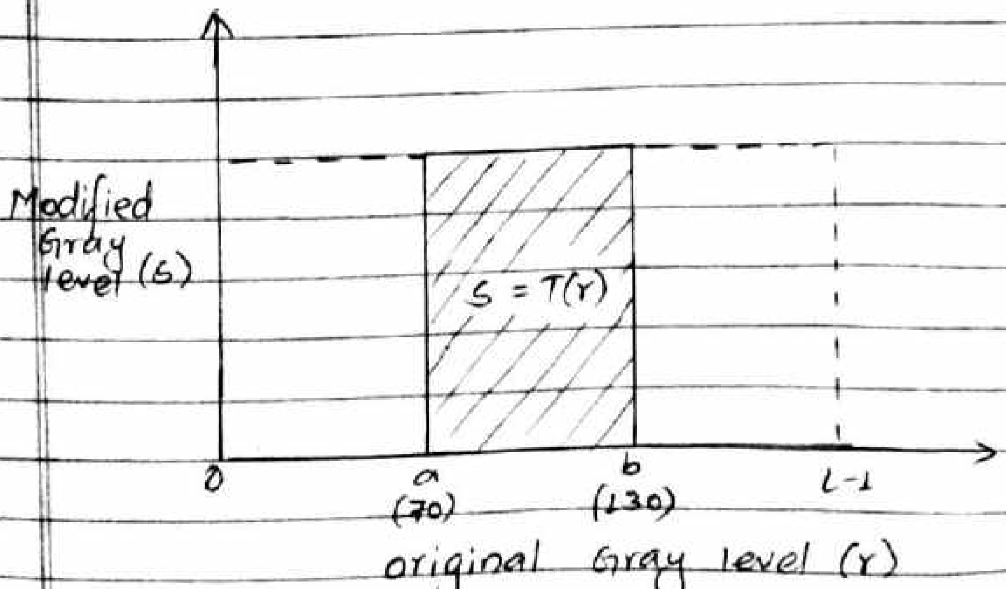
Where,

r = Input pixel value
L = No. of Gray level or Maximum Gray level
$T(r)$ = Transformation Function
s = output Gray level.

Hence,

$r = 0$, $s = 255$ (White)
$r = 255$, $s = 0$ (Black)

(IMP) # Intensity Level Slicing:
[Gray Level Slicing]



(fig: Slicing without Background)

5	2	6
7	8	3
9	1	2

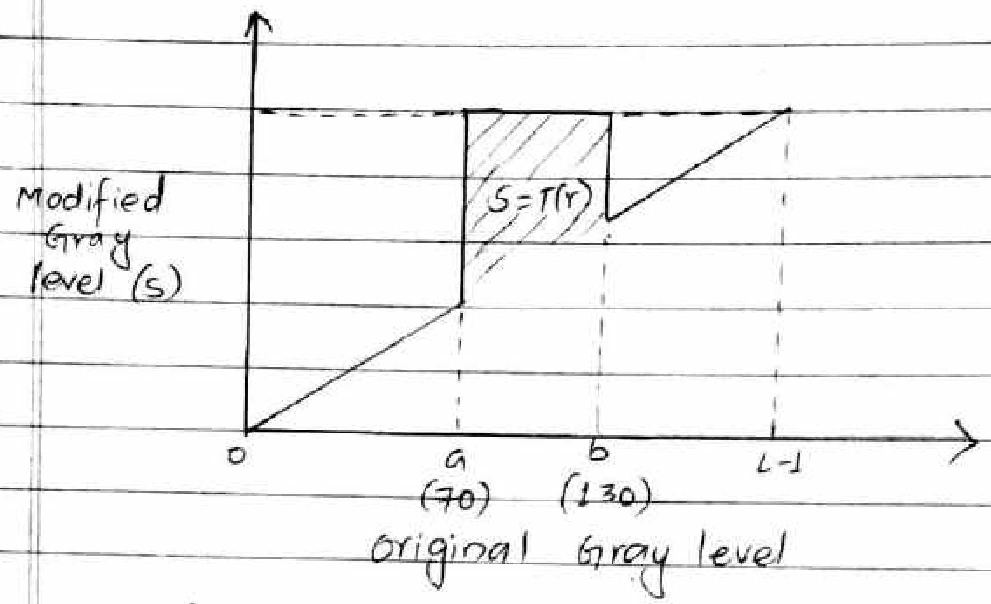
perform intensity level slicing for
$r_1 - r_2$ (3-5)
(i) with background
(ii) without background

9	2	1
7	8	9
9	1	2

with background

9	0	0
0	0	9
0	0	0

without background



(fig: Slicing With Background)

Thresholding splits the gray level into two parts. But many times we need highlight a specific range of gray level.

In such a situation, use a transformation function on a gray level called as Gray Level Slicing.

There are two types of slicing:

(1) Slicing without background:

Here,
We have completely lost the background and this can be implemented by using the following function:

$s = L-1$, if $a \le r \le b$
$s = 0$, otherwise

classmate

Date

Page

(2) Slicing with Background:

Here,
in some applications we donot only need of enhancement to the band of gray level but also need to retain the background.
This can be implemented by using the following formation:

$$s = L-1 \;,\; \text{if } a \leq r \leq b$$
$$s = r \;,\; \text{otherwise}$$

(2011 fall)

Bit Plane Slicing (clipping):

* heighlighting • interesting aspect

It is used to enhance features with an image. We must provide a threshold ~~value~~ level to determine how the clipping occurs.

* is any particular bit

The threshold level or value is a reference value where other values set changed.

In general, it is performed for maximum and minimum operators. An expression that contain an image with threshold level.

For example:
If we have image variable and want to slice it to > or = to 125, then the

resulting image stated as:

Clipped Image = (Image (original)) > 125

Here,
it means if the element value is less than 125 then it is set as equal to 125. Otherwise, it is unchanged.

(V. Imp) # Introduction to Histogram:

Histogram of an image provides a global description of the apparance of an image. The information obtained from histogram is innermost and hence histogram modeling is important in digital image processing (DIP).

By the defination,
the histogram of an image represents the relative frequency of occurance of various gray level in an image.

Histogram of an image can be plotted as:

(i) The X-axis has the gray level and Y-axis has the number of pixel in each gray level.

(ii) The x-axis represent the gray level and Y-axis represents the probability of occurance of that gray level.

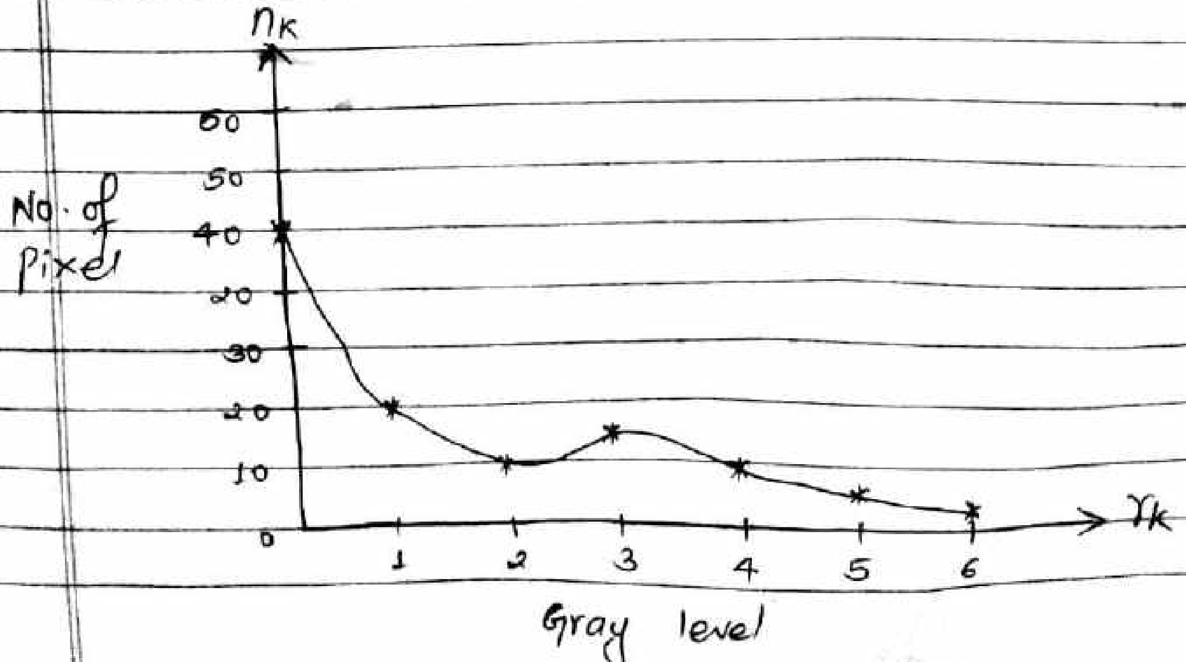
The method of histogram is shown as below:

Step(1): Original Image

Gray level (r_K)	No. of pixel (n_K)
0	40
1	20
2	10
3	15
4	10
5	3
6	2

Step(2): Histogram Modeling

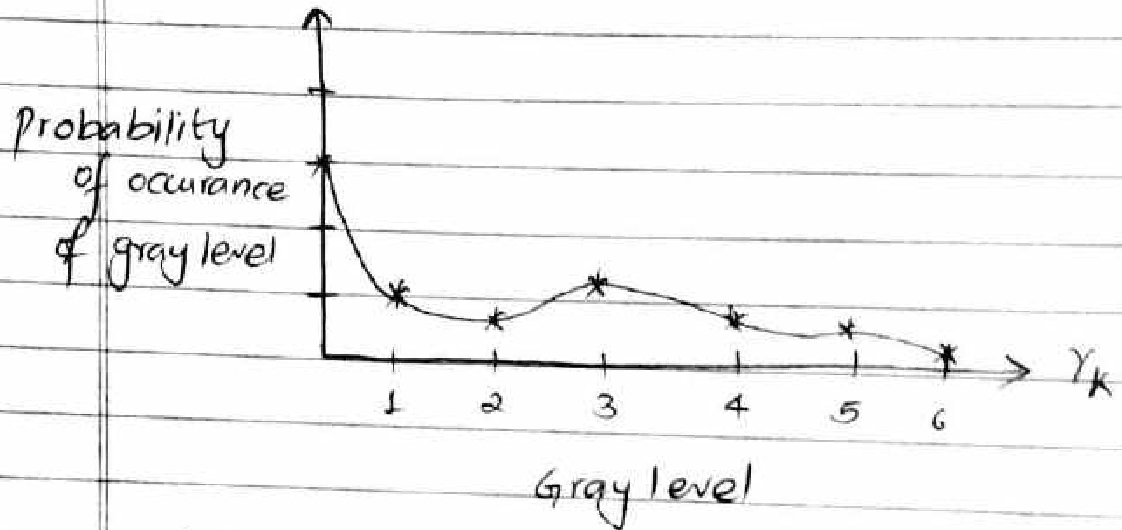
Method (1):



classmate
Date
Page

Method (2):

$P(r_K) = n_K/n$



(fig: Histogram Representation by 2-methods)

Here,

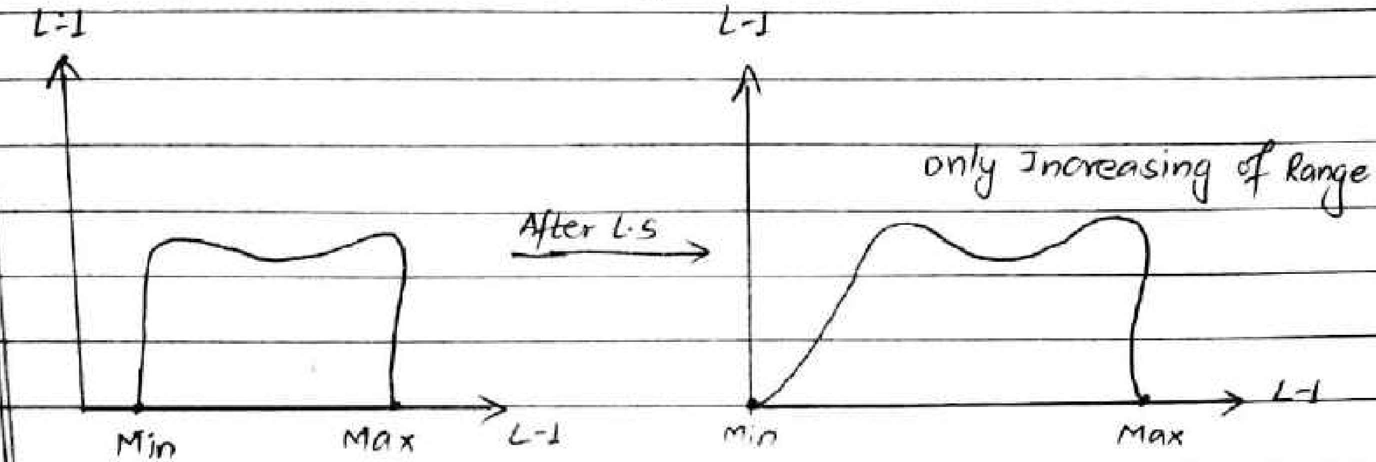
r_K = K^{th} Gray level
n_K = No. of pixel in K^{th} Gray level
n = Total no. of pixel

Histogram Techniques:

(1) Histogram stretching / Linear stretching
(2) Histogram Equalization / Histogram linearization
(3) Histogram specification / Histogram Matching

(non-university)

① Histogram stretching :



(fig: Linear strectching)

To increase the dynamic range by using a technique is known as Histogram streching.

In this method, we donot change basic shape of the histogram but spread it over the entire dynamic range.

For this, we use this technique on the basis of slope. It can be given as:

$$\text{slope}(m) = \frac{S_{max} - S_{min}}{r_{max} - r_{min}} \quad \text{——— (i)}$$

For transformation of histogram stretching then, transformation can be defined as:

$$S = T(r) = \frac{S_{max} - S_{min}}{r_{max} - r_{min}} (r - r_{min}) + S_{min} \quad \text{——— (ii)}$$

$$\boxed{s = m\,(r - r_{min}) + s_{min}}$$

Where,

s_{max} = Max^m gray level in o/p image
s_{min} = Min^m gray level in o/p image
r_{max} = Max^m gray level in i/p image
r_{min} = Min^m gray level in i/p image

This transformation function stretches or increase the dynamic range of given image.

Example: stretch the contrast of histogram over the entire range.

[2011 fall]

(r) Gray Level	0	1	2	3	4	5	6	7
No. of pixel	0	0	50	60	50	20	10	0

Now, perform histogram stretching. show that the new image has a dynamic range of entire range ie [0, 7]

Soln: Given,

$r_{min} = 2$ $\qquad$ $s_{min} = 0$
$r_{max} = 6$ $\qquad$ $s_{max} = 7$

Now, We perform transformation function on histogram stretching as:

$$s = T(r) = \frac{s_{max} - s_{min}}{r_{max} - r_{min}}(r - r_{min}) + s_{min}$$

(S)

Note:

Now,

When $r=2$: $s = \frac{7-0}{6-2}(2-2) + 0 = 0$

$\therefore \boxed{r=2,\ s=0}$

When $r=3$: $s = \frac{7-0}{6-2}(3-2) + 0 = 1.75 \approx 2$

$\therefore \boxed{r=3,\ s=2}$

When $r=4$: $s = \frac{7-0}{6-2}(4-2) + 0 = 3.5 \approx 4$

$\therefore \boxed{r=4,\ s=4}$

When $r=5$: $s = \frac{7-0}{6-2}(5-2) + 0 = 5.2 \approx 5$

$\therefore \boxed{r=5,\ s=5}$

When $r=6$: $s = \frac{7-0}{6-2}(6-2) + 0 = 7$

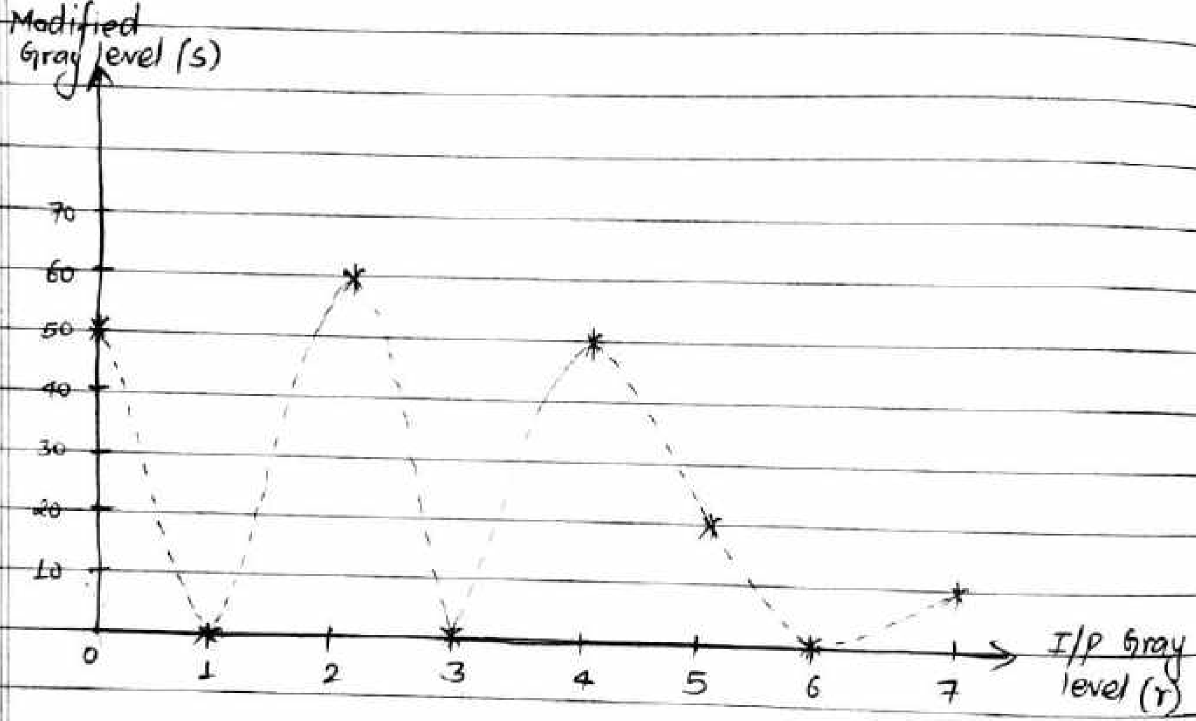
$\therefore \boxed{r=6,\ s=7}$

Hence,
Modified histogram or stretched range can be illustrated as:

(S) → Gray Level	0	1	2	3	4	5	6	7
No. of pixel	50	0	60	0	50	20	0	10

Note: $\boxed{s=0}$ ⇒ old Table मा $\boxed{r=2}$ को value हेर्ने
त्यो value New Table $s=0$ मा राख्ने and so on

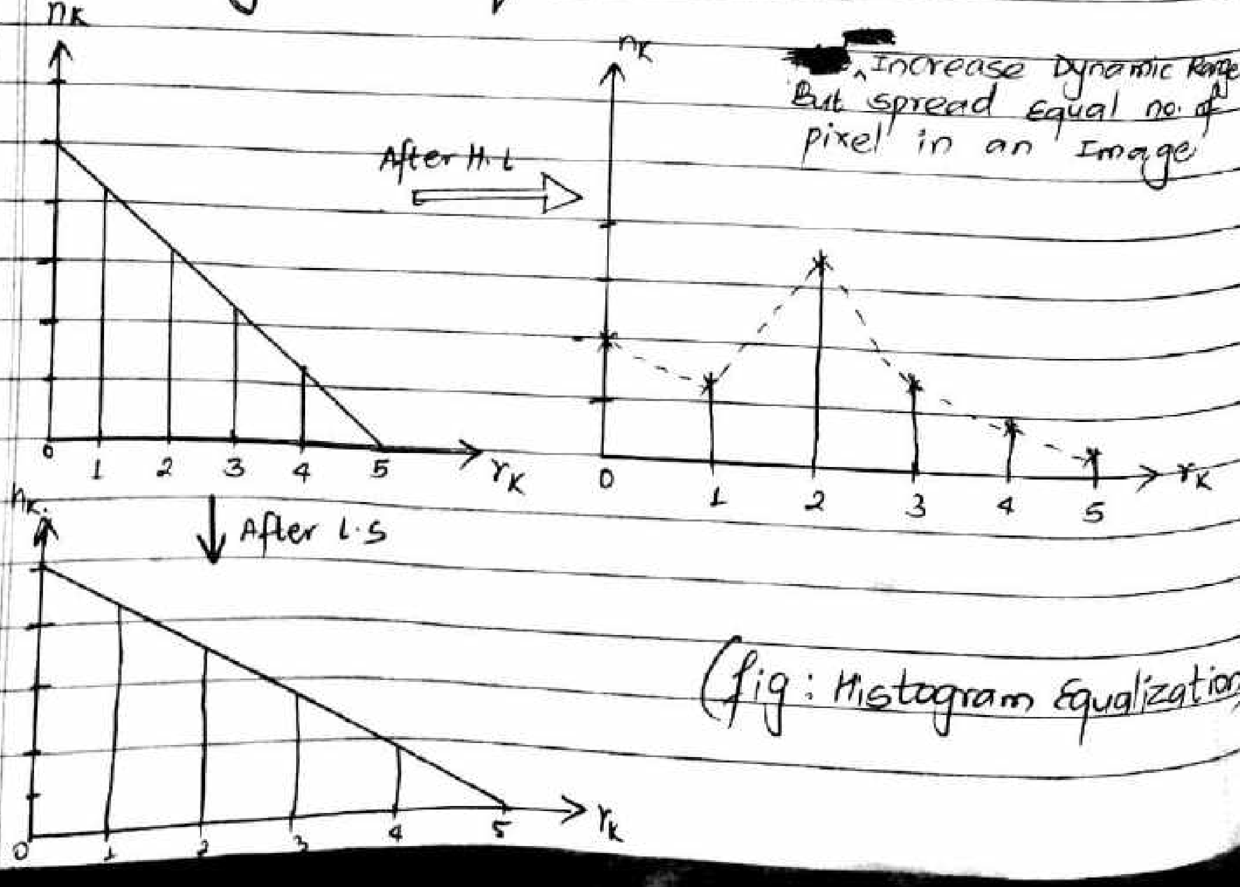
Now, The output histogram image is shown as below:



(fig: stretched / Modified / o/p Histogram Image)

(Imp) ② Histogram Equalization:

(2012 fall)



(fig: Histogram Equalization)

classmate

(3) The gray level for continous variable can be characterized by proability density function (Pdf) ie $P_r(r)$ and $P_s(s)$

We know that, we need to find the transformation which could be the flat histogram.

(1) linear stretching is a good technique but shape remain same. There are many applications where we need a flat histogram so ~~that~~ this cannot be achieved by histogram stretching.

(2) Hence, a technique that can be used to obtain a uniform histogram is known as Histogram Equalization or Histogram linearization.

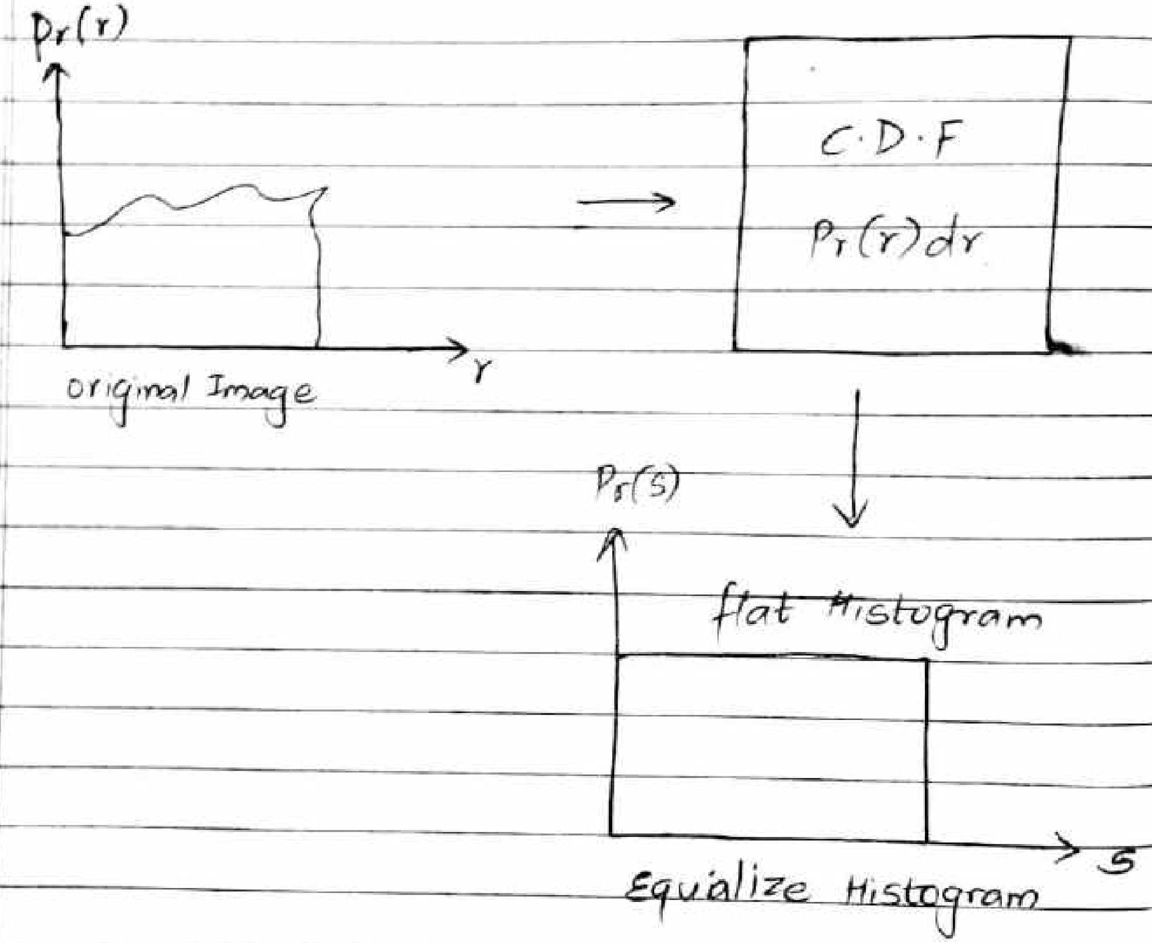
A perfect image is one which has equal number of pixels in all gray level. Therefore, In this technique,
→ image is not only spread over the dynamic range but also to have equal number of pixel in all gray level.

We know that,

$$s = T(r) = \int_0^r P_r(r)\, dr, \quad 0 \leq r \leq 1$$

classmate
Date
Page

Above equation shows the transformation function of cumulative pdf (c.d.f) and it can be shown as:



Example: (V. Imp)

Equalize the given histogram:

Gray level	0	1	2	3	4	5	6	7
No. of pixel	790	1023	850	656	329	245	122	81

Soln: Here,

Step(1):

classmate
Date
Page

Gray level	n_k	P.D.F $P_r(r_k) = n_k/n$	C.D.F $S_k = \Sigma P_r(r_k)$	$7 * S_k$	Rounding off
0	790	0.19	0.19	1.33	1
1	1023	0.25	0.44	3.08	3
2	850	0.21	0.65	4.55	5
3	656	0.16	0.81	5.67	6
4	329	0.08	0.89	6.23	6
5	245	0.06	0.95	6.65	7
6	122	0.03	0.98	6.86	7
7	81	0.02	1	7	7
	n = 4096				

Step (2): Representing New Gray Level

Old Gray Level	No. of pixel	New Gray Level
0	790	1
1	1023	3
2	850	5
3	656	6
4	329	6
5	245	7
6	122	7
7	81	7

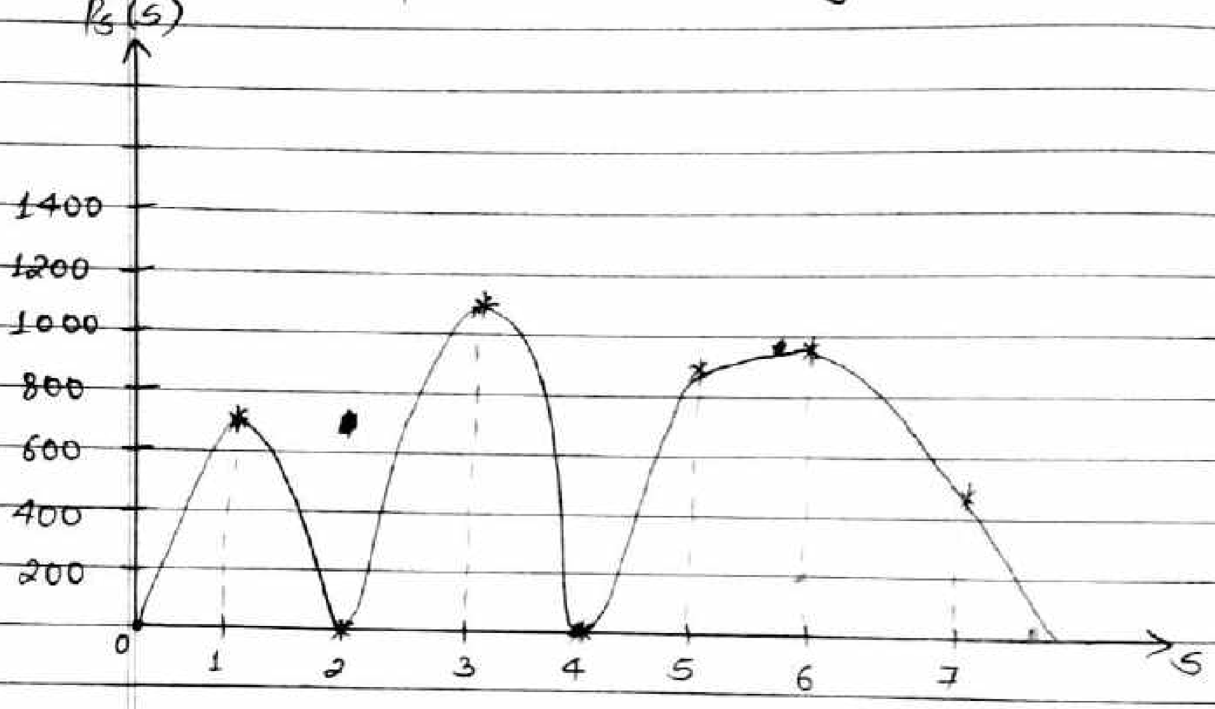
Step (3): Modified Histogram

Gray level	0	1	2	3	4	5	6	7
No. of pixel	0	790	0	1023	0	850	985	448

(656 + 329)

(245 + 122 + 81)

Hence, Equalized Histogram:



(fig: Equalized Histogram)

③ Histogram specification:

(Th + Nm)

Histogram specification is not interactive because it always gives one result i.e an approximation to an uniform histogram.

It is at time desirable to have an interactive methods in which certain gray levels are highlighted.

let, us suppose that $P_r(r)$ is the original pdf (Probability density function) and $P_z(z)$ is the ~~original~~ desired pdf.

Suppose that, histogram equalization is first applied on the original image.

i.e

$$S = T(r) = \int_0^r P_r(r)\, dr$$

Histogram equalization of desired image 'z' is that:

$$v = G(z) = \int_0^z P_z(r)\, dz$$

Now,

The inverse process $z = G^{-1}(s)$ which have the desired pdf that provide histogr-am specification.

i.e $z = G^{-1}(s) = G^{-1}(T(r)) = G^{-1}\int_0^r P_r(r)\, dr$

Example:

Matching the given histogram:

Histogram for image (a):

Gray level	0	1	2	3	4	5	6	7
No. of pixel	790	1023	850	656	329	245	122	81

Histogram for image (b):

Gray level	0	1	2	3	4	5	6	7
No. of pixel	0	0	0	614	819	1230	819	614

Soln: Here,
Step(1):
Equalized Histogram of Image (a):

Gray level	0	1	2	3	4	5	6	7
No. of pixel	0	790	0	1023	0	850	985	448

Now,
Equalized Histogram for Image (b) is:

Gray level	n_k	PDF $P_r(r_k) = n_k/n$	C.D.F $S_k = \Sigma P_r(r_k)$	$S_k * 7$	Rounding off
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	614	0.149	0.149	1.05	1 · 790
4	819	0.20	0.35	2.45	2
5	1230	0.30	0.65	4.55	5
6	819	0.2	0.85	5.77	6
7	614	0.15	1	7	7
	n = 4096				

Equalized Histogram of Image (b) is:

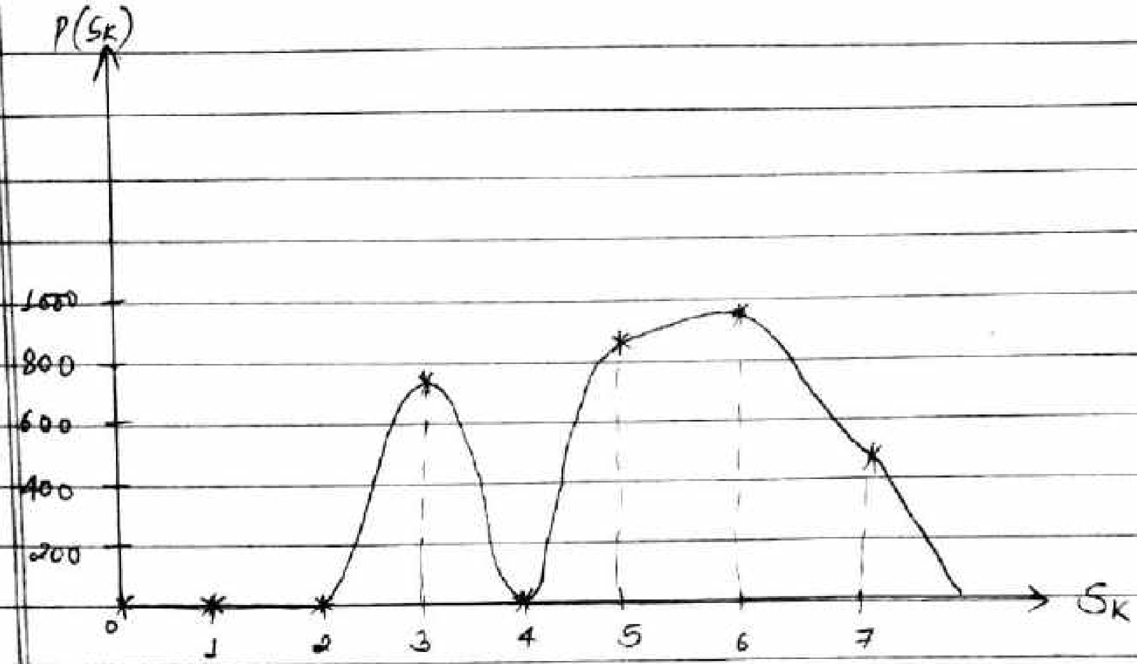
Gray level	0	1	2	3	4	5	6	7
No. of pixel	0	614	819	0	0	1230	819	614

Step(2):
To obtain histogram specification we apply inverse transformation comparing both equalized histogram.

Image (b) → Gray level → Rounding off no.
Then Image (a) → Same gray level
Then, No. of pixel of Image (a)

Gray level	0	1	2	3	4	5	6	7
No. of pixel	0	0	0	790	0	850	985	448

Hence,
The matched or resultant histogram is shown as below:



(fig: Specified Image)

(S.N) # Spatial Operations:

Many image enhancement techniques are based in spatial operations. Most of the spatial operations are performed on the basis of local neighbourhood of the input pixel.

Some of the example of spatial operation for image enhancement are as below:

classmate
Date
Page

(1) Neighbourhood processing
(2) Spatial Averaging
(3) Smoothing spatial Filter
(4) Zooming
(5) Low Pass, High Pass and Band Pass

(1) Neighbourhood Processing:

l

K	-1	0	1
1	$g(x-1, y+1)$	$g(x, y+1)$	$g(x+1, y+1)$
0	$g(x-1, y)$	$g(x, y)$	$g(x+1, y)$
-1	$g(x-1, y-1)$	$g(x, y-1)$	$g(x+1, y-1)$

(fig: 3 x 3 neighbourhood pixel)

We change the value of pixels based on the values of its 8 neighbourhood as shown in figure above.

Instead of 3×3 neighbourhood, we can also use 5×5, 7×7 etc.

Neighbourhood processing deals to the point operation. The above neighbourhood pixel can be masked into window as shown in the figure below:

classmate
Date
Page

$$
\begin{array}{c|ccc}
 & & L & \\
K & -1 & 0 & 1 \\
\hline
-1 & W_1 & W_2 & W_3 \\
0 & W_4 & W_5 & W_6 \\
1 & W_7 & W_8 & W_9
\end{array}
$$

(fig : 3×3 window)

The above figure is a one type of tablet which is also called as mask or window.

To achieve the neighbourhood processing, we place this 3×3 mask on the image corresponding to each of the component and get some information at the centre.

If g is the original image and f is the modified image then, f is derived as :

$$
\begin{aligned}
\therefore f(x,y) = {} & g(x-1, y+1) * W_1 + g(x, y+1) * W_2 + \\
& g(x+1, y+1) * W_3 + g(x-1, y) * W_4 + \\
& g(x, y) * W_5 + g(x+1, y) * W_6 + \\
& g(x-1, y-1) * W_7 + g(x, y-1) * W_8 + \\
& g(x+1, y-1) * W_9
\end{aligned}
$$

classmate

Hence,
the sum of products of the mask coefficients with corresponding pixel takes place.

(2) Spatial Averaging:
[Averaging Filtering]

(2010 spring)

In spatial averaging, each pixel is replaced by the weighted average of its neighbourhood pixel i.e

$$\boxed{v(m,n) = \sum\sum_{(k,l)\in W} a(k,l)\, y(m-k,n-l)}$$

where,

$y(m-k, n-l)$ = i/p image
$v(m,n)$ = o/p image
w = Choosen window
$a(k,l)$ = Filter Weight

A common class of spatial average filter has all equal weight i.e

$$\boxed{v(m,n) = \frac{1}{N_w} \sum\sum_{(k,l)\in w} y(m-k,n-l)}$$

where,

$a(k,l) = \frac{1}{N_w}$; N_w is the no. of pixel in window.

Now,
Each pixel is replaced by its

average with the average of its nearest pixel value.

Spatial average is used for noise smoothing, low pass filtering and sub-sampling of the image.

Example of spatial averaging can be represented as :

$K \backslash L$	-1	0	1
-1	1/9	1/9	1/9
0	1/9	1/9	1/9
1	1/9	1/9	1/9

(fig: 3 x 3 window)

$k = 2$, $L = 2$

1/4	1/4
1/4	1/4

(fig: 2x2 window)

Smoothining filters are used for blurring and for noise reduction

(3) Smoothing Spatial Filter :
[Median & Mean Filter]

Median Filter:

The median filter is also a spatial filter but it replaces the central value in the window with median of all the pixel values in the window.

The kernel or template or mask is usually square but can be initaited in different window.

An example of median filtering is

classmate

Date ______

Page ______

i.e

unfiltered values in order like

0, 1, 2, 3, 4, 6, 10, 15, 19, 97

Now,

$$\text{o/p value OR Median value} = \left(\frac{n+1}{2}\right)^{th}$$

$$= \left(\frac{9+1}{2}\right)^{th}$$

$$= 5^{th}$$

$$= 6$$

Hence,

6	2	0
3	97	4
19	15	10

Before filtering

$\xrightarrow[\text{filtering}]{\text{Median}}$

6	2	0
3	6	4
19	15	10

After filtering

The center value 97 is replaced by the median value of all nine values i.e 6.

Formula for median filtering is :

$$\boxed{\vartheta(m,n) = \text{Median}\{y(m-k, n-l)\ (k,l) \in W\}}$$

Mean Filter:

Mean filter is a spatial filter that replaces the center value in the window with the average of all the pixel values in the window.

Here, the window can be of any shape but can be initaited as a square masked. Therefore, average of smoothing filter shows the properties of lowpass filter. So, it reduces the noise from an image.

An example of mean filter is:
Unfiltered values in order like:
5, 3, 6, 2, 1, 9, 8, 4, 7

Now,

o/P value OR Mean value $= \frac{\Sigma x}{N} = 5$

Hence,

5	3	6
2	1	9
8	4	7

Before filtering

$\xrightarrow[\text{filtering}]{\text{Mean}}$

5	3	6
2	5	9
8	4	7

After filtering

The center value '1' is replaced by the mean value '5'

Spatial Filtering:
• Many image enhancement techniques are based on spatial operation which are performed on local neighborhood of the input pixel.
• Spatial masks aka window, filter kernel, template.
• They are used and convolved over the entire image for local enhancement.
• Spatial filters are characterized by the following two concept:
o Concept of neighborhood
o Convolution operation over neighborhood.
• It can be further classified on the basis of nature of response:
1) Linear Spatial Filter (Output is linear combination of input filters, mean, median)
2) Non-Linear Spatial Filter (output is not linear combination of input filters, min max)

(P.T.O)

classmate
Date
Page

V. Imp

(4) Zooming:
[Magnification & Interpolation]

(2012)

Another important application is image enhancement where spatial domain neighbourhood operation is used.

Such type of image enhancement technique is known as zooming.

Actually, it is the process of making an image larger. It involves two steps:

(1) Creating new pixel location.
(2) Assigning gray level to those new location.

Different types of zooming technique can be used to zoom. They are:
(i) Replication
(ii) linear Interpolation

(i) Replication:

In replication, we simply replicate each pixel and then replicate row.

let, us consider an image as:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 8 & 6 & 7 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

As we start from the first row, first we replicate each pixel and then replicate each row.

Now,
The first row looks as:

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

We will now replicate this row then it looks as:

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

Now,
Performing replication operation on an entire image then we obtain output image as:

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \\ 5 & 5 & 6 & 6 & 7 & 7 & 8 & 8 \\ 5 & 5 & 6 & 6 & 7 & 7 & 8 & 8 \\ 9 & 9 & 8 & 8 & 6 & 6 & 7 & 7 \\ 9 & 9 & 8 & 8 & 6 & 6 & 7 & 7 \\ 0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\ 0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \end{bmatrix}$$

Here,
4×4 image is zoomed to 8×8 image.
This method can be repeated to get larger image but no new data is added.

classmate
Date
Page

Zooming By zero - Interlace Technique:

step(1): Adding zero ~~to every~~ after every pixel of the first row then we obtain as:

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

step(2): Adding zero at every column.
Now,
Insert row of zero and then this is known as zero-Interlacing.

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

step(3): Now,
Performing zero-Interlace technique on an entire image then we obtain output image as:

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 3 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 6 & 0 & 7 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 0 & 8 & 0 & 6 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(fig: 8 x 8 final Pseudo Image)

Hence, this image is known as zero-Interlace Image.

(ii) Linear Interpolation:

In this method, instead of replicating each pixel, average of the two adjacent pixel along the row is taken and placed between the two pixels.

Accordingly, same operation is then performed along the column.

Linear Interpolation is also derived from the zero-interlace technique.

step(1): Zero-Interlace

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 3 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 6 & 0 & 7 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 0 & 8 & 0 & 6 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

step(2): Interpolate row

$$\begin{bmatrix} 1 & 1.5 & 2 & 2.5 & 3 & 3.5 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 5.5 & 6 & 6.5 & 7 & 7.5 & 8 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 8.5 & 8 & 7 & 6 & 6.5 & 7 & 3.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 1 & 1.5 & 2 & 2.5 & 3 & 3.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Step(3): Interpolate column

$$\begin{bmatrix} 1 & 1.5 & 2 & 2.5 & 3 & 3.5 & 4 & 2 \\ 3 & 3.5 & 4 & 4.5 & 5 & 5.5 & 6 & 3 \\ 5 & 5.3 & 6 & 6.5 & 7 & 7.5 & 8 & 4 \\ 7 & 7 & 7 & 6.25 & 6.5 & 7 & 7.5 & 3.25 \\ 9 & 8.5 & 8 & 7 & 6 & 6.5 & 7 & 3.5 \\ 4.5 & 4.5 & 4.5 & 4.5 & 4 & 4.5 & 5 & 2.5 \\ 0 & 0.5 & 1 & 1.5 & 2 & 2.5 & 3 & 1.5 \\ 0 & 0.25 & 0.5 & 0.75 & 1 & 1.25 & 1.3 & 0.7 \end{bmatrix}$$

(5) Filtering:

(i) Low Pass Filtering:

It is employed to remove high frequency (unnecessary noise) from a digital image.

This type of filtering is usually used to attenuate the image noise.

Here, an image is $\boxed{\text{smoothed}}$ by decreasing the disparity between the pixel value by average value.

Example of mask or kernel or window of low pass filter can be shown as below:

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

(fig: 3 × 3 mask)

classmate
Date
Page

1/16	1/16	1/16	1/16
1/16	1/16	1/16	1/16
1/16	1/16	1/16	1/16
1/16	1/16	1/16	1/16

(fig: 4 x 4 mask)

(ii) High Pass Filtering: (Sharpening filters)

A high pass filtering is the basis for most sharpening methods. sharpening is the opposite of smoothing.

The sharpening can make it possible to highlight the boarders between homogenous region.

This is used to return high frequency information.

The kernel of the high pass filter is designed to increase the brightness of the center pixel.

Example of mask or kernel or window of high pass filter is as shown below:

~~1/9~~	~~1/9~~	~~1/9~~
~~1/9~~	~~-1/9~~	~~1/9~~
~~1/9~~	~~1/9~~	~~1/9~~

$$\frac{1}{9}\begin{array}{|c|c|c|}\hline -1 & -1 & -1 \\ \hline -1 & 8 & -1 \\ \hline -1 & -1 & -1 \\ \hline\end{array}$$

(fig: 3x3 mask high pass filter)

classmate
Date
Page

(iii) Band Pass Filtering :

This filter is used to remove selected frequency region between low and high frequency.

It is also used for image restoration. It passes a limited range of frequency.

Calculating spatial averaging for band pass filtering is :

$$\boxed{v(m, n:\theta) = \frac{1}{N_\theta} \sum\sum_{(k,l)\in W_\theta} y(m-k,\, n-l)}$$

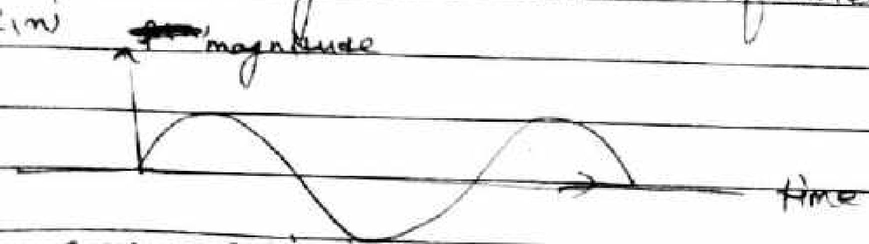
Where,

W_θ = Neighbourhood mask along direction θ

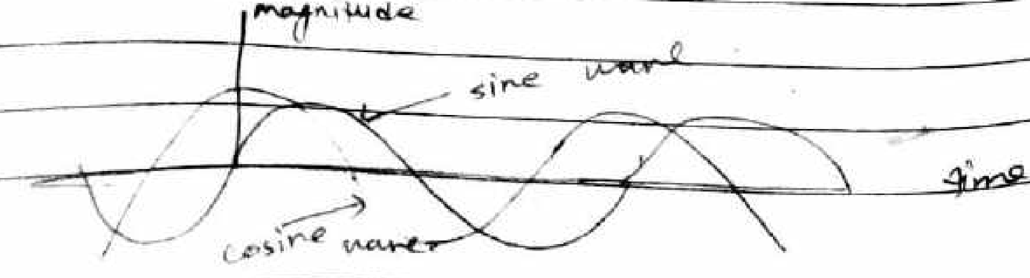
N_θ = No. of pixel within this neighbourhood.

fourier series

* A series of cosine and sine term that represent the periodic signal is called fourier series.



$f(n) \rightarrow$ periodic signal



Assignment:

1. Equalize the given histogram.

Gray Level (I)	0	1	2	3	4	5	6	7
Frequency n(I)	32	160	128	512	192	256	704	64

2. Compute the histogram equalization from the given data

   Input image: 4*4 image

   Gray Scale =[0,9]

2	3	3	2
4	2	4	3
3	2	3	5
2	4	2	4

3. Computer the histogram equalization from the given data

Gray Level (I)	0	1	2	3	4	5	6	7
Frequency n(I)	5320	1000	500	525	1236	956	128	856

4. Computer the histogram equalization from the given data

Gray Level (I)	0	1	2	3	4	5	6	7
Frequency n(I)	750	1020	850	660	340	240	123	80

5. What are the uses of histogram processing? Perform the histogram equalization on the following 3-bit gray level image

3	2	4	5
1	2	1	2
7	3	1	2
7	6	4	7

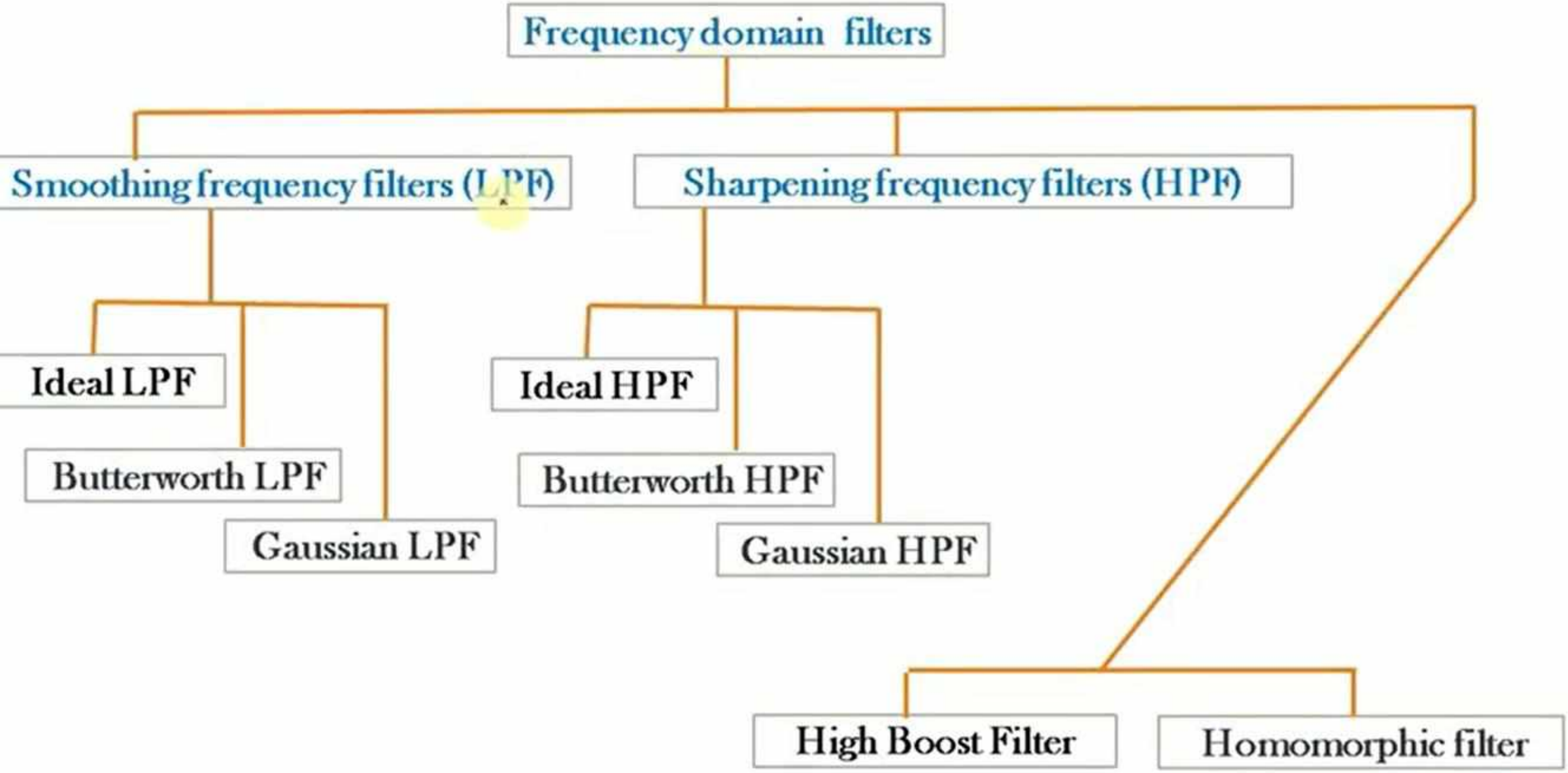
1. Perform the histogram stretching on the following 3-bit gray level image

3	2	4	5
1	2	1	2
7	3	1	2
7	6	4	7

Image Enhancement in Frequency Domain

- Image Enhancement
  - Spatial domain
    - Point processing
      - E.g. Negative Image, contrast stretching, thresholding etc.
    - Neighbourhood processing
      - E.g. Averaging filter, median filtering etc.
  - Frequency domain
    - E.g. Image sharpening using Gaussian high pass filters, unsharp masking, highboost filtering etc.

- The frequency content of an image refers to the rate at which the gray levels change in time. Rapidly changing brightness values correspond to high frequency terms, slowly changing brightness values correspond to low frequency terms. The Fourier transform is a mathematical tool that analyses a signal (e.g. images) into its spectral components depending on its frequency content.


Frequency domain filters
Smoothing frequency filters (LPF)
Sharpening frequency filters (HPF)
Ideal LPF
Butterworth LPF
Gaussian LPF
Ideal HPF
Butterworth HPF
Gaussian HPF
High Boost Filter
Homomorphic filter

Frequency domain vs. Spatial domain

	Frequency domain	Spatial domain
1.	is resulted from Fourier transform	is resulted from sampling and quantization
2.	refers to the space defined by values of the Fourier transform and its frequency variables (u,v).	refers to the image plane itself, i.e. the total number of pixels composing an image, each has spatial coordinates (x,y)
3.	has complex quantities	has integer quantities

$F(u,v) \longrightarrow \boxed{H(u,v)} \longrightarrow G(u,v)$

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

multiplication

$g(x,y) = F^{-1}(G(u,v))$

$f(x,y) \rightarrow \boxed{h(x,y)} \rightarrow g(x,y)$

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

❖ Advantages of filtering in frequency domain:

- The frequency domain filtering is advantageous because it is computationally faster to perform two 2D Fourier transforms and a filter multiply than to perform a convolution in the spatial domain.

- Frequency domain gives you control over the whole images, where we can enhance(eg. Edges) and suppress(eg. Smooth shadow) different characteristics of the image very easily.

- Frequency domain has a established suit of processes and tools that can be borrowed directly from signal processing in other domain.

- Image enhancement in the frequency domain is straightforward. The idea of blurring an image by reducing its high frequency components, or sharpening an image by increasing the magnitude of its high frequency components is intuitively easy to understand.

Chapter-3

Image Enhancement In Frequency Domain

[6 hours]

Fourier Transform :

Fourier Transform is a mathematical tool to analyze and design linear system. It is used to reduce the number of calculation to a fraction.

It also helps to quantify the effects of digitizing system.

The application of Fourier transform is:

- Image Enhancement
- X-Ray
- Enhancement of TV transmission
- and video signal.

(A) 1-D Fourier Transform:

Let, $f(x)$ be a continous real variable of x.

Then,

Fourier Transform $= F[f(x)] = F(u)$

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

P

classmate
Date
Page

And,

Inverse Fourier Transform is :

$$F^{-1}[F(u)] = f(x)$$

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

(B) 2-D Fourier Transform :

A 2-D function f(x) has 2-D transform and,

$$\boxed{F[f(x,y)] = F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \cdot e^{-j2\pi(ux+vy)} \, dxdy}$$

Now,

Inverse of fourier transform is :

$$\boxed{F^{-1}[F(u,v)] = f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v) \cdot e^{j2\pi(ux+vy)} \, dxdy}$$

ASS CREP

Properties of Fourier Transform:

(1) Addition Theorem:

$$F[f(x,y)+g(x,y)] = F(u,v) + G(u,v)$$

(2) Shift Theorem:

$$F[f(x-a, y-b)] = F(u,v)\, e^{-j2\pi(U_a, V_b)}$$

(3) Similarity Theorem:

$$F[f(ax, by)] = \frac{1}{|ab|} F(U/a, V/b)$$

(4) Convolution Theorem:

$$F[f(x,y) * g(x,y)] = F(U,V) * G(U,V)$$

(5) Rayleigh's Theorem:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x,y)|^2\, dx.dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |F(U,V)|^2\, dU.dV$$

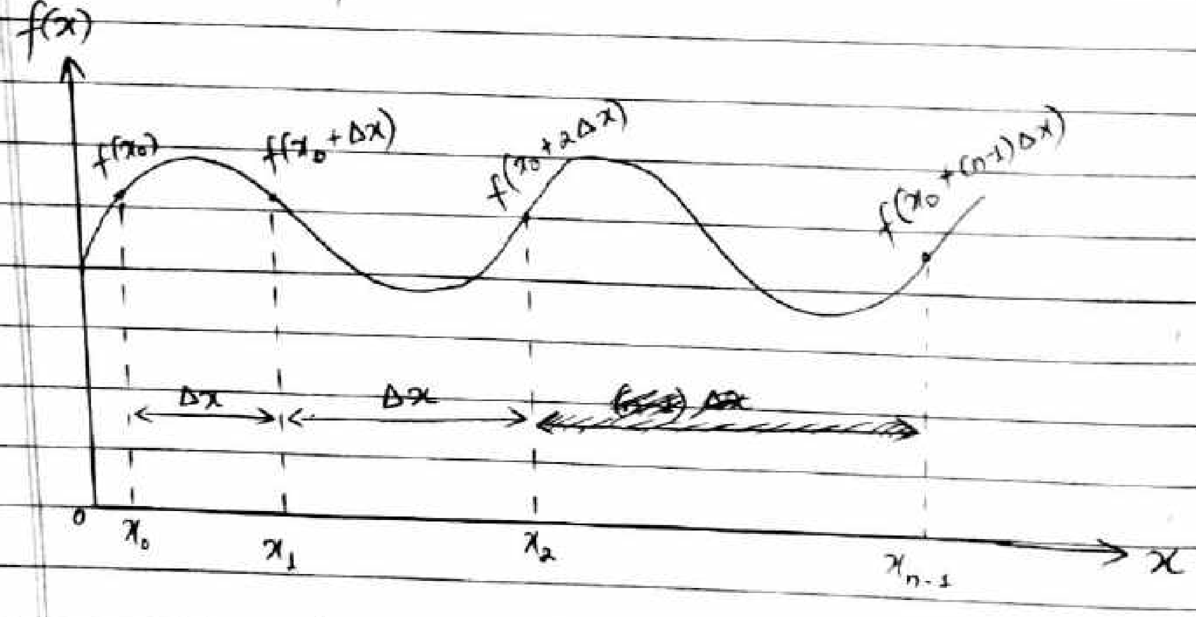
(6) Evenness & oddness:

$$fe(x) \Rightarrow f(x) = f(-x) \text{ and}$$
$$fo(x) \Rightarrow f(-x) = -f(x)$$

(7) Power Theorem:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) * g^*(x,y)\, dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v) * G^*(u,v)\, du\, dv$$

Discrete Fourier Transform (DFT) and It's Importance:



(fig: DFT)

Here,

$f(x)$ is a continous function to be discretized by taking 'n' sample i.e

$$\{f(x_0), f(x_0+\Delta x), f(x_0+2\Delta x), \ldots f(x_0+(n-1)\Delta x)\}$$

It is defined as:

$$\boxed{F(u) = \frac{1}{n}\sum_{x=0}^{n-1} f(x)\cdot e^{-j2\pi\frac{ux}{n}}}$$

Inverse Discrete Fourier Transform is defined as :

$$\boxed{f(x) = \sum_{u=0}^{N-1} F(u) \cdot e^{j2\pi \frac{ux}{n}}}$$

These two given equation are discrete fourier transform and it is used to calaulate the following :

(1) Fourier spectrum:

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

(2) Power spectrum:

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

(3) Phase Angle (spectrum):

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

Where,
'U' is the transform frequency variable.
'x' is sapital or Image Variable.

Discrete Fourier Transform is most widely used orthogonal transform in the field of image processing and it can diagonilized circuilant matrix.

However, in this transform, it takes n^2 multiplication and $n(n-1)$ addition to calculate 1-D discrete fourier transform of 'n' data points.

Discrete Cosine Transform And It's properties : (DCT)

(2010 spring)

Discrete cosine transform is a Fourier related transform similar to DFT but it is only concerned with real number only.

It can be defined as:

$$\text{DCT} \Rightarrow \boxed{C(U) = \alpha(U) \sum_{x=0}^{N-1} f(x) \cdot \cos\left[\frac{(2x+1)U\pi}{2N}\right]}$$

Where,

$$U = 0, 1, \dots N-1$$

$$\alpha(U) = \sqrt{\frac{1}{N}} = \frac{1}{\sqrt{N}} \text{ for } U = 0$$

$$= \sqrt{\frac{2}{N}} \text{ for } U = 1, \dots N-1$$

classmate
Date
Page

∴ Inverse DCT is defined as:

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

Where,

$$u = 0, 1, \dots N-1$$

<u>For 2-D</u>:

$$\text{DCT} \Rightarrow c(u,v) = \alpha(u,v) \cdot \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)u\pi}{2N}\right]$$

Where,

$$\alpha(u,v) = \frac{1}{\sqrt{N}} \quad \text{for } u, v = 0$$

$$= \sqrt{\frac{2}{N}} \quad \text{for } u, v = 1, \dots N-1$$

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

where,

$$x = y = 0, \dots N-1$$

(P.T.O)

Arthogonal : $A^T A = I$
symmetric : $A^T = A$

classmate
Date
Page

Properties of DCT:

(1) The casine Tranform is real and orthogonal.

i.e

$$C = C^* \Rightarrow C^{-1} = C^T$$

Where,

C^* = conjugate of C
C^T = Transform of C

(2) The Casine Transform is not real part of Unitary DFT.

i.e

$$T^{-1} = T^{*T}$$

(3) It is a fast transform. The cosine transform of N elements can be calculated in $O(N \log_2 N)$ operation through N point FFT (fast fourier Transform).

(4) It has excellent energy compaction for highly co-related data.

Uses of DCT:

DCT is used in image processing and signal processing especially for loose data compression. ~~dense~~

Since it has strong anergy compaction features.

It is also used for JPEG image compression, MPEG video compression and MJPEG video compression.

(Imp) # Hadamard Transform:

(2015 fall) (2012 fall)

Hadamard Transform is a fast transform which is real, symmetric and orthogonal.

1-D Hadamard Transform can be implemented in (Big oh) $O(N \log_2 N)$ addition and subtraction. It has only binary values i.e 1 or -1 in its kernel matrix.

It is defined as:

$$H(U) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x) \cdot \underbrace{(-1)^{\sum_{i=0}^{n-1} b_i(x) \cdot b_i(U)}}_{\text{Kernel}}$$

For $N=2$:

$$N = 2^n \Rightarrow 2 = 2^n \Rightarrow n = 1$$

Where, $U = 0, \ldots, N-1$

Now,

Hadamard Matrix can be defined as:

$$H_n = \frac{1}{\sqrt{N}} \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -(H_{n-1}) \end{bmatrix}$$

Hadamard Transformation

→ It is 1D non-sinusodial signal
→ It is a transform for a complete set of orthogonal square function.
→ It has all its cofficient real & their values will be either +1 or -1 only
→ This has ~~be~~ a very low computational complexity
→ It is similar to walse Transformation

The 1D Hadamard Transform is represented by

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

where

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

Here, $b_i(x)$ & $b_i(u)$ is binary representation of x, & u.

Scanned with CamScanner

2 D Hadamard Transform is

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

2D Function is

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) (-1)^{\sum_{i=0}^{p-1} b_i(x) b_i(u) + b_i(y) b_i(v)}$$

Where
N = order of function
x = time index
u = frequency index
p = no. of bit required to represent n

$p = \log_2 N$

Scanned with CamScanner

Kernel Design of Hadamard Transform order $N=2$

Algorithm

Step 1: To determine the Kernel for Hadamard Transform, let us first consider the equation as

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

Step 2: , Here $N=2$

Step 3: Since $N=2$, the values time & frequency indexes of x & u are 0, 1

$\therefore$ size of matrix is 2×2.

	$x=0$	$x=1$
$u=0$		
$u=1$		

Step 4:- No of bits req^d to represent order

$$p = \log_2 N = \log_2 2 = 1$$

Step 5:- For Hadamard Transfer function, When $x=0$ & $u=0$ Then

$$g(x,u) = \frac{1}{N} = \frac{1}{2} \quad \text{i.e } g(0,0) = \frac{1}{2}$$

Scanned with CamScanner

PNC

step 6: Determine the value of $b_i(x)$, as the value of $N=2$, $x=0,1$

(Representation in binary form)

n	$b_0(n)$
0 →	0
1 →	1

(one bit मात्रै ले 0 ले represent गर्छ)

step 7: Determine the coefficient for every ~~part~~ point of transfer function $g(x,y)$, by expanding Hadamard function & substitute the binary parameter values obtained in step ⑥

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

$$= \frac{1}{2}(-1)^{\sum_{i=0}^{1-1} b_i(x)\, b_i(u)}$$

$$= \frac{1}{2}(-1)^{\sum_{i=0}^{0} b_i(x)\, b_i(u)}$$

$$= \frac{1}{2}(-1)^{b_0(x) \cdot b_0(u)}$$

Scanned with CamScanner

8. Determine the varies values of $g(x,u)$ by substituting the values of x & u then we get

$$g(0,0) = \frac{1}{2}(-1)^{b_0(0).b_0(1)}$$

$$= \frac{1}{2}(-1)^0$$

$$= \frac{1}{2}$$

$$\cancel{g(0,1) = \frac{(-1)}{2} b_0(0).b_0(1)}$$

$$\cancel{= \frac{1}{2}}$$

$$g(0,1) = \frac{1}{2}(-1)^{b_0(0).b_0(1)}$$

$$= \frac{1}{2}(-1)^0$$

$$= \frac{1}{2}$$

$$g(1,0) = \frac{1}{2}(-1)^{b_0(1).b_0(0)}$$

$$= \frac{1}{2}(-1)^{1.0}$$

$$= \frac{1}{2}(-1)^0$$

$$= \frac{1}{2}$$

$$g(1,1) = \frac{1}{2}(-1)^{b_0(1).b_0(1)}$$

$$= \frac{1}{2}(-1)^1$$

$$= -\frac{1}{2}$$

$$\begin{array}{c|cc} & x=0 & x=1 \\ \hline u=0 & \frac{1}{2} & \frac{1}{2} \\ u=1 & \frac{1}{2} & -\frac{1}{2} \end{array} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}_{2\times2}$$

Scanned with CamScanner

Hadamard Transform Kernel for Order $N=4$

1. Let us consider the equation

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

2. Here size of kernel $N=4$

3. Since $N=4$, x & u varies from 0 to $N-1$ takes vales from [0, 1, 2, 3]

Size of maxtrix is 4×4

4. No. of bits reqd to represent 4 i.e (order)

$$p = \log_2 N = \log_2 4 = 2$$

5. For Hadamard Transfer function, when $x=0$ & $u=0$ gives

$$g(x,u) = \frac{1}{N}$$

$$g(0,0) = \frac{1}{4}$$

Scanned with CamScanner

6. Representing all $x = 0, 1, 2, 3$ values in binary

n	$b_1(n)$	$[b_0(n)]$
0	0	0
1	0	1
2	1	0
3	1	1

7. Determining the coeff for every point of transfer function ~~g(x,u)~~ $g(x,u)$ by expanding Hadamard function & substituting the binary parameter values obtained in 6, we have

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

$$= \frac{1}{4}(-1)^{\sum_{i=0}^{2-1} b_i(x) \cdot b_i(u)}$$

$$= \frac{1}{4}(-1)^{b_0(x)b_0(u) + b_1(x) \cdot b_1(u)}$$

	0	1	2	3
0				
1				
2				
3				

Scanned with CamScanner

8. Determining values ~~from~~ varies of $g(x,u)$ by substituting values of x & u

$$g(0,0) = \frac{1}{4}(-1)^{b_0(0)\cdot b_0(0) + b_1(0)\cdot b_1(0)}$$

$$= \frac{1}{4}(-1)^{0\cdot 0 + 0\cdot 0}$$

$$= \frac{1}{4}(-1)^{0} = \frac{1}{4}$$

$$g(0,1) = \frac{1}{4}(-1)^{b_0(0)\cdot b_0(1) + b_1(0)\cdot b_1(1)}$$

$$= \frac{1}{4}(-1)^{0\cdot 1 + 0\cdot 0} = \frac{1}{4}(-1)^{0} = \frac{1}{4}$$

$$g(0,2) = \frac{1}{4}(-1)^{b_0(0)\cdot b_0(2) + b_1(0)\cdot b_1(2)}$$

$$= \frac{1}{4}(-1)^{0\cdot 0 + 0\cdot 1} = \frac{1}{4}(-1)^{0} = \frac{1}{4}$$

$$g(0,3) = \frac{1}{4}$$

Scanned with CamScanner

$g(1,0) = \frac{1}{4}$

$g(1,1) = -\frac{1}{4}$

$g(1,2) = \frac{1}{4}$

$g(1,3) = -\frac{1}{4}$

$g(2,0) = \frac{1}{4}$

$g(2,1) = \frac{1}{4}$

$g(2,2) = -\frac{1}{4}$

$g(2,3) = -\frac{1}{4}$

$g(3,0) = \frac{1}{4}$ $u=0$

$g(3,1) = -\frac{1}{4}$ $u=1$

$g(3,2) = -\frac{1}{4}$ $u=2$

$g(3,3) = \frac{1}{4}$ $u=3$

$x=0$	$x=1$	$x=2$	$x=3$
$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$
$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$

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

Scanned with CamScanner

classmate
Date
Page

where,

$\frac{1}{N}$ = sample value

$$\begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -(H_{n-1}) \end{bmatrix} = \text{Indicate kernel}$$

$$H_0 = 1$$

<u>Hadamard Matrix at 1</u> : (H_1) i.e $n=1$

$$H_1 = \frac{1}{2} \begin{bmatrix} H_{1-1} & H_{1-1} \\ H_{1-1} & -(H_{1-1}) \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} H_0 & H_0 \\ H_0 & -H_0 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Also,

<u>For N = 4</u>, $N = 2^n \Rightarrow 4 = 2^n \Rightarrow n = 2$

$$\therefore H_2 = \frac{1}{4} \begin{bmatrix} H_{2-1} & H_{2-1} \\ H_{2-1} & -(H_{2-1}) \end{bmatrix}$$

$$= \frac{1}{4} \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

(2012 Fall) (pre-university)

Properties of Hadamard Transform:

(1) It is real, symetric and orthogonal i.e

$$H = H^* \Rightarrow H^{-1} = H^T$$

(2) It is a fast transform and 1-D Hadamard Transform can be implemented in $O(N\log_2 N)$ addition and subtraction.

(3) It has only binary values i.e 1 or -1 in its kernel matrix. No multiplications are required in the transform.

(4) It is used for digital image processing and digital signal processing.

(5) It has good energy compaction for highly co-related image.

(6) It is also used in digital hardware implementation of image processing algorithm.

(S.N) 2010 spring 2012 fall

(S.N) # Haar Transform:

Haar Transform is the orthogonal transform which is derived from the Haar Matrix.

It is also a fast transform and can be implemented in $O(N\log_2 N)$ operation.

The main advantage of this transform is that it samples the input data sequence into fine resolution and takes the difference between the adjacent pairs.

The resolution increases by a power of 2 and as a result, in the transform domain, differential energy is more and more localized.

It only exist for $N = 2^n$, where n is an integer.

Haar Transform is defined as:

$$\boxed{H_T = H_K(x)}$$

Where, $k = 0, ---- N-1$

k is given as:

$$\boxed{K = 2^p + q - 1}$$

And,

$0 \leq p \leq (n-1)$
$1 \leq q \leq 2^p$, for $p \neq 0$
$q = 0, 1$, for $p = 0$

Date
Page

Haar Transform

→ It is transform which is based on the class of orthogonal, non-sinusoidal, non-square functions whose elements are multiples of −1, +1, 0 by the power of $\sqrt{2}$.

→ It is a computational efficient transfer function of N part whose vectors require $2(N-1)$ additions & ~~m~~ N multiplications.

Steps:-

1. Determine the value of 'N', which represent the order of the system or the size of the kernel.

2. Determine the total no. of Bits required based ~~on~~ upon the order of the function using expression $n = \log_2 N$

3. Determine the values of 2 specific terms 'p' & 'q':

(i) $p \in [0, n-1]$ ie $0 \le p \le n-1$

(ii) If $p = 0$, then the value of q is either 0 or 1.

(iii) if $p \neq 0$, then $q \in [1, 2^p]$

i.e $1 \leq q \leq 2^p$

Step 4:- Determine the value of K which tells us that total no. of rows in the kernel by eq^n

$$K = 2^p + q + 1$$

Step 5:- Determine the value of Z where

$$Z \in \left[0, \frac{1}{N}, \frac{2}{N}, \dots \frac{N-1}{N}\right]$$

Step 6:- If $K=0$, then (for First row) k-represent - row, z " - column

$$H_0(Z) = \frac{1}{\sqrt{2}} \ \forall Z$$

else

$K \neq 0$ (for remaining row)

$$H_k(Z) = H_{pq}(Z) = \frac{1}{\sqrt{N}} \begin{cases} +2^{P/2} : \frac{q-1}{2^p} \leq Z < \frac{q-1/2}{2^p} \\ -2^{P/2} : \frac{q-0.5}{2^p} \leq Z \leq \frac{q}{2^p} \\ 0 \quad \text{otherwise} \end{cases}$$

Scanned with CamScanner

Design kernel for Haar Transform of order = 2

step 1:- Here order = 2, so, $N = 2$, which represent order of the system or size of kernel

2:- total no. of bits req^d based upon order of function is

$$n = \log_2 N = \log_2 2 = 1.$$

step 3:- Determine p & q

(i): p ranges from 0 to n-1 i.e

$$0 \le p \le 1-1$$

$\therefore p = 0$

(ii) if $p = 0$, q can be '0' or '1'.

Step 4:- Determine value of K which is total no. of rows in the kernel

$K = 2^p + q - 1$ for $p = 0, q = 0, K = 2^0 + 0 - 1$
$= 0$

for $p = 0, q = 1, K = 2^0 + 1 - 1$
$= 1$

$\therefore K = 0, 1$

Scanned with CamScanner

PNC

Step 5 :- value of $Z \in \left[0, \frac{1}{N}, ----, \frac{N-1}{N}\right]$

i.e 0 to $\frac{N-1}{N}$ i.e 0 to $\frac{2-1}{2} = \frac{1}{2}$

$$Z \in \left[0, \frac{1}{2}\right]$$

Step 6 :- (I) if $k=0$, then $H(Z) = \frac{1}{\sqrt{2}}$ for all Z

i.e

$$H(Z) = \frac{1}{\sqrt{2}} \quad \forall Z \in [0, 1/2]$$

When

(II) $K = 1$,

$$H_K(Z) = H_{pq}(Z) = \frac{1}{\sqrt{2}} \begin{cases} +2^{p/2} : \frac{q-1}{2^p} \leq Z < \frac{q-0.5}{2^p} \\ -2^{p/2} : \frac{q-0.5}{2^p} \leq Z < \frac{q}{2^p} \\ 0 : \text{otherwise} \end{cases}$$

Here for $k=1$, $p=0$, $q=1$

Then,

$$\frac{q-1}{2^p} \leq Z < \frac{q-0.5}{2^p} \Rightarrow \frac{1-1}{2^0} \leq Z < \frac{1-0.5}{2^0}$$

$$= 0 \leq Z < 0.5$$

Scanned with CamScanner

Date ____ Page ____

Also,

$$\frac{q-0.5}{2^p} \le z < \frac{q}{2^p} \Rightarrow \frac{1-0.5}{2^0} \le z < \frac{1}{2^0}$$

$$= 0.5 \le z < 1$$

So, we get,

$$H_k(z) = H_{pq}(z) = H_{01}(z) = \frac{1}{\sqrt{2}} \begin{cases} 2^{p/2} & : 0 \le z < 0.5 \\ -2^{p/2} & : 0.5 \le z < 1 \\ 0 & \text{otherwise} \end{cases}$$

~~for z the~~

for $z \in [0, \frac{1}{2}]$

$$H_1(0) = H_{01}(0) = \frac{2^0}{\sqrt{2}} = \frac{1}{\sqrt{2}}$$

$$H_1(\tfrac{1}{2}) = H_{01}(\tfrac{1}{2}) = \frac{-2^0}{\sqrt{2}} = -\frac{1}{\sqrt{2}}$$

We have kernel for the ~~hard~~ haar transform is

$$\begin{array}{c|c} & \begin{array}{cc} z=0 & z=\frac{1}{2} \end{array} \\ \begin{array}{c} k=0 \\ k=1 \end{array} & \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \end{array} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Q.NO. 14

Date ______
Page ______

Haar Transform of $N=4$

$$H = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$H = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{bmatrix}$$

$$H(Z) = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}_{8\times 8}$$

Q. Compute 2D Discrete Hadamard Transform of the given image.

$$f = \begin{bmatrix} 5 & 6 & 8 & 10 \\ 6 & 6 & 5 & 7 \\ 4 & 5 & 3 & 6 \\ 7 & 8 & 3 & 5 \end{bmatrix}$$

Apply Hadamard Transform

$$T = K \cdot f \cdot K^T, \text{ where } K = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 6 & 8 & 10 \\ 6 & 6 & 5 & 7 \\ 4 & 5 & 3 & 6 \\ 7 & 8 & 3 & 5 \end{bmatrix} \cdot K^T$$

$$= \frac{1}{4} * \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 6 & 8 & 10 \\ 6 & 6 & 5 & 7 \\ 4 & 5 & 3 & 6 \\ 7 & 8 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \frac{1}{16} \begin{bmatrix} 94 & -12 & 0 & 6 \\ 0 & -2 & -14 & 0 \\ 12 & 2 & 14 & 0 \\ 10 & 0 & 0 & -2 \end{bmatrix}$$

Scanned with CamScanner

Q.NO. 14

P.N.

Date ____
Page ____

Q. compute Haar Transform T from given matrix F. And reconstruct the original image F by performing Inverse Haar Transform on T.

$$F = \begin{array}{|c|c|c|c|} \hline 1 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 1 & 1 \\ \hline \end{array}$$

Apply Haar Transform to obtain T.

$$T = H \cdot F \cdot H^T$$

$$= \frac{1}{\sqrt{4}} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{vmatrix} H^T$$

$$= \frac{1}{\sqrt{4}} * \frac{1}{\sqrt{4}} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{vmatrix}$$

$$= \begin{bmatrix} 2.5 & 0 & \frac{3}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ 0 & \frac{1}{2} & -\frac{1}{2\sqrt{2}} & -\frac{3}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2} & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & \frac{1}{2} \end{bmatrix}$$

Scanned with CamScanner

(Pre-University)

classmate
Date
Page

Properties Of Haar Transform:

(1) It is symmetric, seperable unitary transform that uses haar function for its basis.

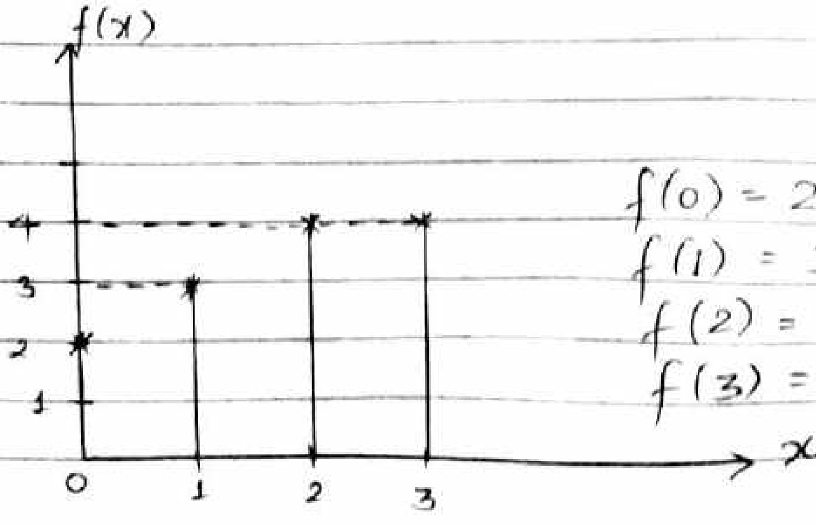
(2) It is orthogonal and real.
ie
$$T^{-1} = T^{T} \Rightarrow H_r = H_r^{*}$$

(3) It is a fast transform and can be implemented in $O(N)$ operation. Where, N is a number of samples.

(4) It exist for $N = 2^n$, where 'n' is an integer.

(5) It has poor energy compaction property.

Q. Find the Fourier spectrum of the function at point 0 and 1 as shown in figure below:



$f(0) = 2$
$f(1) = 3$
$f(2) = 4$
$f(3) = 4$

classmate

Date ______

Page ______

Soln: Here,

Given that:

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

We know that:

DFT $\Rightarrow$ $$\boxed{F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \cdot e^{\frac{-j2\pi ux}{N}}}$$

Where,

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

Now,

for $u=0$:

$$F(0) = \frac{1}{4} \sum_{x=0}^{3} f(x) \cdot e^{-j2\pi \frac{0 \cdot x}{4}}$$

$$= \frac{1}{4} \sum_{x=0}^{3} f(x)$$

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

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

$$= 13/4$$

classmate
Date
Page

$\therefore$ Fourier spectrum at point 0 is:

$$|F(0)| = \sqrt{R^2(0) + I^2(0)}$$

$$= \sqrt{\left(\frac{13}{4}\right)^2 + 0}$$

$$= \frac{13}{4} \quad \text{Ans}$$

Note:

Power spectrum at point 0 is:

$$|P(0)| = |F(0)|^2 = R^2(u) + I^2(u)$$

$$= \left(\frac{13}{4}\right)^2$$

$$= \frac{169}{16}$$

Now,

for $u=1$:

$$F(1) = \frac{1}{4} \sum_{x=0}^{3} f(x) \cdot e^{-j2\pi \frac{1 \cdot x}{4}}$$

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

$$= \frac{1}{4}\left[f(0) \cdot e^{\frac{-j\pi \cdot 0}{2}} + f(1) \cdot e^{\frac{-j\pi \cdot 1}{2}} + f(2) \cdot e^{\frac{-j\pi \cdot 2}{2}} + f(3) \cdot e^{\frac{-j\pi \cdot 3}{2}} \right]$$

$$= \frac{1}{4}\left[f(0)\cdot e^{0} + f(1)\cdot e^{-\frac{j\pi}{2}} + f(2)\cdot e^{-j\pi} + f(3)\cdot e^{\frac{-j\pi 3}{2}}\right]$$

$$= \frac{1}{4}\left[2\times 1 + 3\times\left(\cos\frac{\pi}{2} - j\sin\frac{\pi}{2}\right) + 4\times\left(\cos\pi - j\sin\pi\right) + 4\times\left(\cos\frac{3\pi}{2} - j\sin\frac{3\pi}{2}\right)\right]$$

$$= \frac{1}{4}\left[2 + 3\times(0 - j\cdot 1) + 4\times((-1) - j0) + 4\times(0 - j(-1))\right]$$

$$= \frac{1}{4}\left[2 - 3j - 4 + 4j\right]$$

$$= \frac{1}{4}\left[j - 2\right]$$

$$= \frac{1}{4}j - \frac{1}{2}$$

$$\therefore\ F(1) = \frac{j}{4} - \frac{1}{2}$$

Now,

Fourier spectrum at point 1 is:

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

$$= \sqrt{\left(\frac{-1}{2}\right)^2 + \left(\frac{1}{4}\right)^2}$$

$$= \sqrt{\frac{1}{4} + \frac{1}{16}}$$

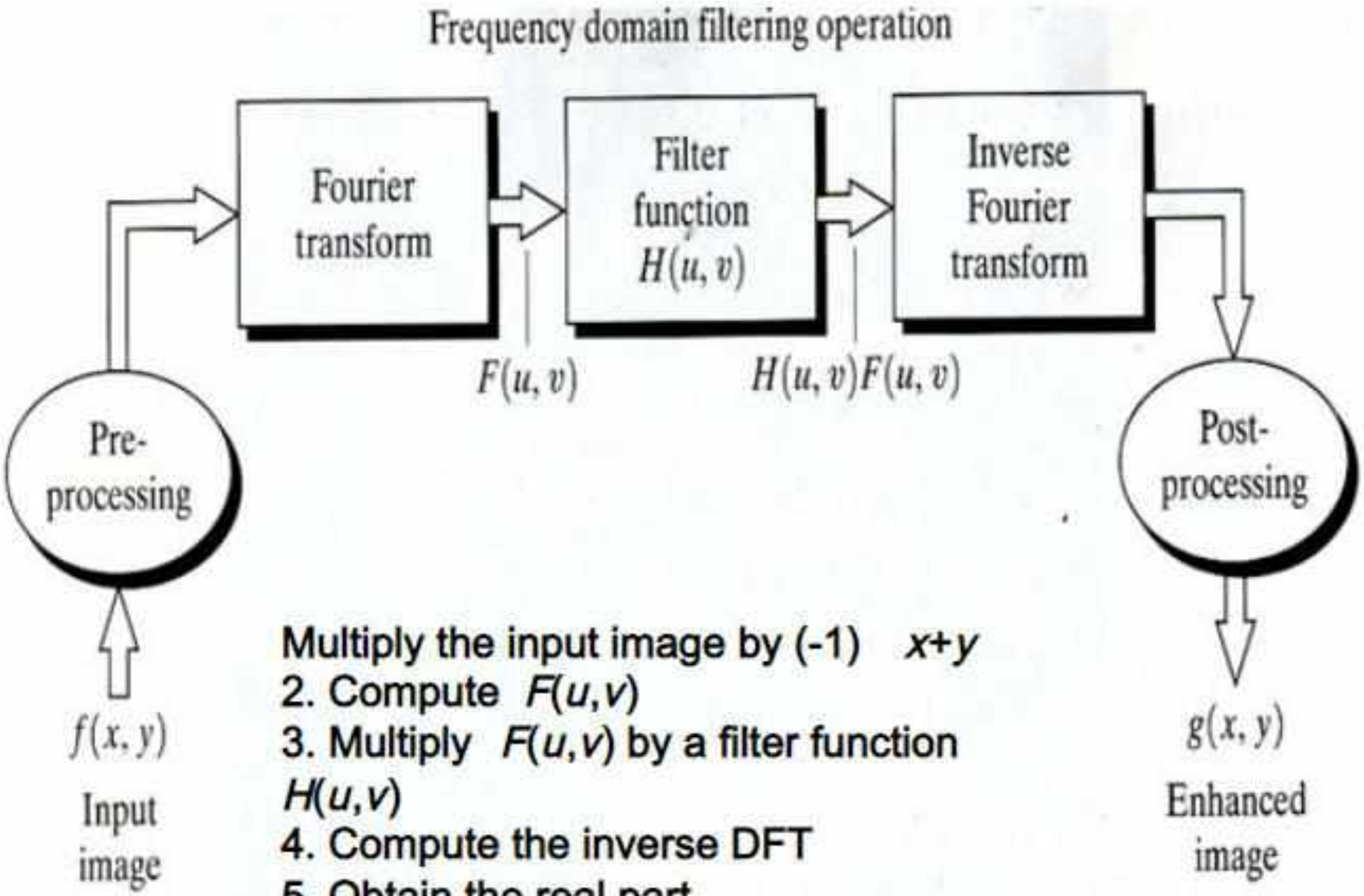
$$= \sqrt{\frac{5}{16}}$$

$$= \frac{\sqrt{5}}{4}$$

$$\therefore \; |F(1)| = \sqrt{5}/4 \quad \underline{\underline{\text{Ans}}}$$

FFT (Fast Fourier Transform)

Steps of frequency domain filter



Multiply the input image by $(-1)^{x+y}$
2. Compute $F(u,v)$
3. Multiply $F(u,v)$ by a filter function $H(u,v)$
4. Compute the inverse DFT
5. Obtain the real part
6. Multiply the real part by $(-1)^{x+y}$

Steps of frequency domain filter

1. Given an input image $f(x,y)$ of size MxN, select the padding parameters $P=2M$ and $Q=2N$.
2. Form a padded image of size PxQ by appending the necessary zeros to $f(x,y)$.
3. Multiply padded image by $(-1)^{(x+y)}$ to center its transform.
4. Compute the DFT $F(u,v)$ of the image from step 3.
5. Generate a real, symmetric filter function $H(u,v)$ of size PxQ. Form the product $G(u,v)=H(u,v)F(u,v)$.
6. Find the processed image by computing the real part of the IDFT of $G(u,v)$

7. Multiply above by by $(-1)^{(x+y)}$
8. get $g(x,y)$ from above step by extracting an MxN region from top left quadrant.

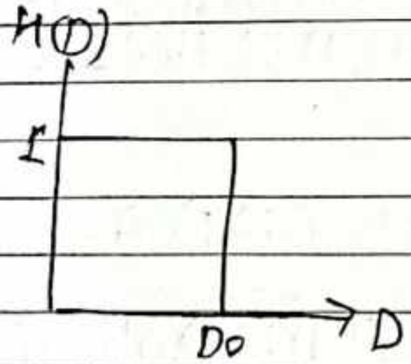
Image smoothing in Fequenly Domain

1. IDeal Low Pass filter
2. Butterworth LPF
3. Gaussia LPF

1. 1D ILPF

→ allows all freq. upto certain cutoff freq. D_0 and removes all the freq beyond that.

$$H(D) = \begin{cases} 1 & D \leq D_0 \\ 0 & D \geq D_0 \end{cases}$$



→ gives smooth effect to image

* 2D ILPF

→ H(D) first apply on row & get intermediate image

→ H(D) again apply on column of the intermediate image to get 2D mask.

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

(D_0 = cutoff)

Scanned with CamScanner

PNC

2. Butterworth LPF

→ effective low PF
→ removes ringing effect

→ Butterworth LPF of order n with cutoff D_0 is

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

where,
H is filter mask magnitude range from 0 & 1

→ if n increases (↑), filter becomes sharper but ringing effect increases

→ If n = 1, no ringing effect exist

→ if n = 2, ~~ringing~~ ringing is present but imperceptible.

Scanned with CamScanner

③ Gaussian LPF

→ used to avoid ringing effect

→ G-LPF in 2D

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

→ $H(u,v)$ is magnitude of 2D gaussian mask. It range from 0 to 1

→ Let $\sigma = D_0$ Then

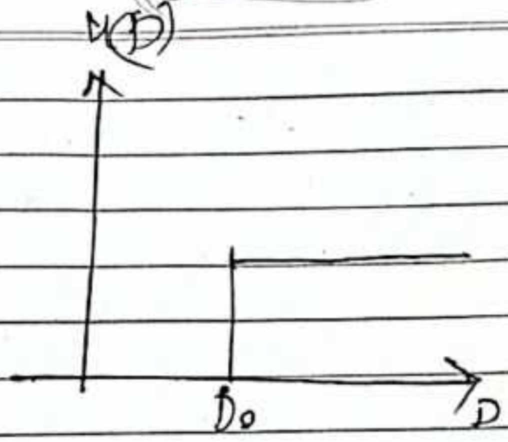
$$H(u,v) = e^{\frac{-D^2(u,v)}{2D_0^2}}$$

Scanned with CamScanner

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

Image sharpening Filters:

① Ideal H.P.F



1 D.

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

2D

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

② Butterworth H.P.F

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

n → determines Sharpness if cutoff value & amount of ri[illegible]

③ Gaussian H.P.F

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

→ smother than I.H.P.F & B.H.F.P.

Scanned with CamScanner

Chapter 3

Image Enhancement In Frequency Domain

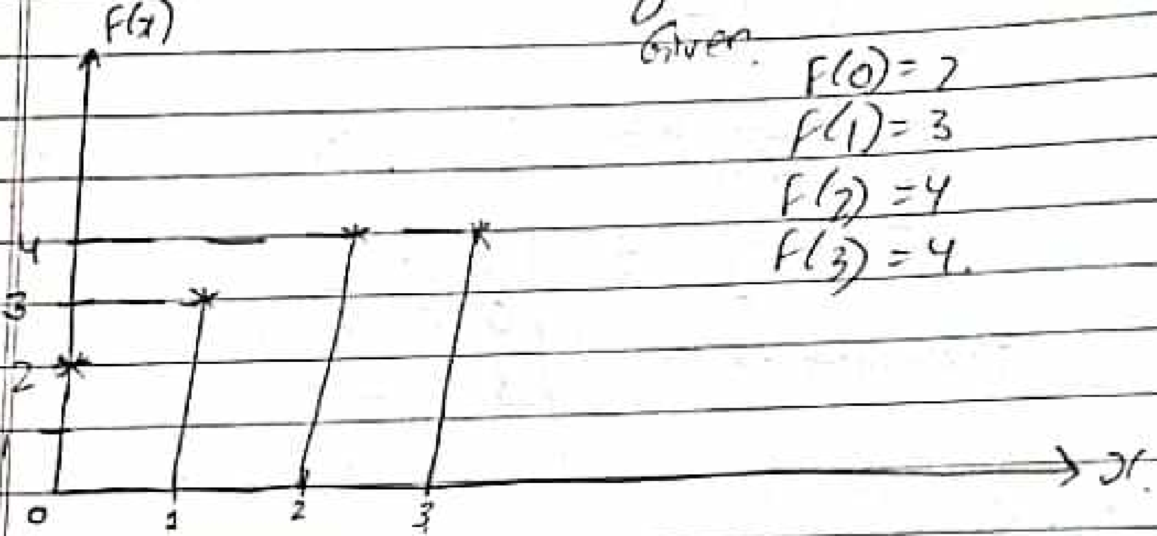
pg. 52

Inverse Fourier Transform. $F(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du.$

correction

FT

\# Find the fourier spectrum of the function at point 0 & 1 as shown in figure below:



Given.

$F(0) = 2$

$F(1) = 3$

$F(2) = 4$

$F(3) = 4.$

~~Given,~~ f we know that.

$$DFT \Rightarrow F(u) = \frac{1}{N} \sum_{x=0}^{N-1} F(x) \cdot e^{-\frac{j2\pi ux}{N}}$$

where, $N = 4$ $(0, 1, 2, 3)$

Now,

For $u = 0$;

$$F(0) = \frac{1}{4} \sum_{0}^{3} f(x)\, e^{-0}.$$

$$= \frac{1}{4} \sum_{x=0}^{3} f(x).$$

$e^{-j\pi} = \cos\pi - j\sin\pi$

$e^{-jn\pi} = \cos n\pi - j\sin n\pi$

$= \frac{1}{4}[2+3+4+4] \Rightarrow 13/4$

$\therefore$ Fourier spectrum at point 0 is:

$|F(0)| = \sqrt{R^2(0)+I^2(0)} \Rightarrow \sqrt{(\frac{13}{4})^2+0} \Rightarrow 13/4$

$\Rightarrow$ ~~For u = 1~~ Power spectrum at point 0 is:

$|P(0)| = |F(0)|^2$ $\quad (\because |P(u)| = |F(u)|^2 = R^2(u)+I^2(u))$

$= \frac{169}{16}$

phase $\measuredangle = \tan^{-1}\left(\frac{0}{(13/4)}\right) \Rightarrow 0$

For u = 1

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

$= \frac{1}{4}\sum_{x=0}^{3} f(x)\, e^{\frac{-j2\pi \cdot 1 \cdot x}{2\times 2}}$

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

$= \frac{1}{4}\left[f(0)e^{0} + f(1)e^{\frac{-j\pi}{2}} + f(2)e^{-j\pi} + f(3) \cdot e^{-j\frac{3\pi}{2}}\right]$

$= \frac{1}{4}\left[2\times 1 + 3\times\left(\cos\frac{\pi}{2} - j\sin\frac{\pi}{2}\right) + 4\times(\cos\pi - j\sin\pi) + 4\times\left(\cos\frac{3\pi}{2} - j\sin\frac{3\pi}{2}\right)\right]$

$= \frac{1}{4}[2 + 3\times(0 - j\,1) + 4\times((-1)-0) + 4\times(0 - j(-1))]$

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

$\Rightarrow \frac{1}{4}j - \frac{1}{2}$

Scanned with CamScanner

$e^{-j\pi x} = \cos\pi x - j\sin\pi x$

Now, Fourier spectrum at point 1 is:

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

$$= \sqrt{\left(\frac{-1}{2}\right)^2+\left(\frac{1}{4}\right)^2} \Rightarrow \sqrt{\frac{1}{4}+\frac{1}{16}} \Rightarrow \sqrt{\frac{5}{16}} \Rightarrow \frac{\sqrt{5}}{4}.$$

$$\therefore |F(1)| = \sqrt{5}/4.$$

Power spectrum at point 1 is:

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

$$= \frac{1}{4} + \frac{1}{16}$$

$$= 5/16.$$

$$\text{Phase } \phi = \tan^{-1}\left(\frac{1/4}{-1/2}\right)$$

$$= -0.4636^c$$

$$= -26.565^\circ$$

For $u = 2$.

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

$$= \frac{1}{4}\sum_{x=0}^{3} f(x)\, e^{\frac{-j2\pi \cancel{2}x}{\cancel{4}}}$$

$$= \frac{1}{4}\sum_{x=0}^{3} f(x)\, e^{-j\pi x}.$$

$$= \frac{1}{4}\left[2\times e^{0} + 3\times e^{-j\pi} + 4e^{-j2\pi} + 4e^{-j3\pi}\right]$$

$$= \frac{1}{4}\left[2 + 3(\cos\pi - j\sin\pi) + 4(\cos 2\pi - j\sin 2\pi) + 4(\cos 3\pi - j\sin 3\pi)\right]$$

-1 0 1 0 -1 0

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

$= \cancel{5/4} \; -\frac{1}{4}.$ $\quad |F(2)| = \sqrt{\left(\frac{-1}{4}\right)^2 + 0} \Rightarrow \frac{1}{4}$

For $u = 3.$

$$F(3) = \frac{1}{4}\sum_{x=0}^{3} f(x)\, e^{-j2\pi 3x/4}$$

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

$$= \frac{1}{4}\left[2 + 3e^{-j\frac{3\pi}{2}} + 4e^{-j\frac{3\pi \cdot 2}{2}} + 4e^{-j\frac{3\pi 3}{2}}\right]$$

$$= \frac{1}{4}\left[2 + 3\left(\overset{0}{\cos\frac{3\pi}{2}} - j\overset{-1}{\sin\frac{3\pi}{2}}\right) + 4\left(\overset{-1}{\cos 3\pi} - j\overset{0}{\sin 3\pi}\right) + 4\left(\overset{0}{\cos\frac{9\pi}{2}} - j\overset{1}{\sin\frac{9\pi}{2}}\right)\right]$$

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

$$= \frac{1}{4}\cancel{[-6 + 3j]} \; [-2 - j]$$

$$= -\frac{1}{2} - \frac{1}{4}j.$$

$$|F(3)| = \sqrt{\left(\frac{-1}{2}\right)^2 + \left(\frac{-1}{4}\right)^2}$$

$$= \frac{\sqrt{5}}{4}.$$

$$F(u) = \begin{bmatrix} 3/4 \\ -1/2 + j/4 \\ -1/4 \\ -1/2 - j/4 \end{bmatrix}$$

Scanned with CamScanner

* HADAMARD TRANSFORM

IMP.

Order $N = 2$, $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Order $2N$ generated by kronecker product operation

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

$N = 2$

$$H_{2N} = H_{2\times 2} \Rightarrow H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \leftarrow \text{Kernel} \ldots (u,x)$$

For $N = 4$,

$$H_{2\times 4} = H_8 = \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix}$$

$$= \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{array}\right]$$

Aba matrix multiplication

Kroneker Product

$F = H \times f(x)$

$F = H \times f(x, y) \times \ldots$

• good compression for highly correlated data

high • DIP & DSP

Q Find the 1D Hadamard Transform for the image matrix as:

$$f(x) = \{1, 2, 0, 3\}_{1\times 4}$$

we have,

$$F = H * f(x)$$

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

$$= \begin{bmatrix} 1+2+0+3 \\ 1-2+0-3 \\ 1+2-0-3 \\ 1-2-0+3 \end{bmatrix} \Rightarrow \begin{bmatrix} 6 \\ -4 \\ 0 \\ 2 \end{bmatrix}$$

Q Find the Hadamard Transform for the given image matrix

Yo pani w pata

$$f(m,n) = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 1 & 2 & 1 \end{bmatrix}$$

we have, $F = H * f(m,n) * H^T$

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

$$\Rightarrow \begin{bmatrix} 1+2+1+2 \to 6 & 2+1+3+1 \to 7 & 2+1+2+2 \to 7 & 1+2+1+1 \to 5 \\ 1-2+1-2 \to -2 & 2-1+3-1 \to 3 & 2-1+2-2 \to 1 & 1-2+1-1 \to -1 \\ 1+2-1-2 \to 0 & 2+1-3-1 \to -1 & 2+1-2-2 \to -1 & 1+2-1-1 \to 1 \\ 1-2-1+2 \to 0 & 2-1-3+1 \to -1 & 2-1-2+2 \to 1 & 1-2-1+1 \to -1 \end{bmatrix} * \text{"}$$

Scanned with CamScanner

$$\begin{bmatrix} 6 & 7 & [illegible] & [illegible] \\ -2 & 3 & 1 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix} * \begin{bmatrix} [illegible] & & & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6+7+7+5 & 6-7+7-5 & 6+7-7-5 & 6-7-7+5 \\ -2+3+1-1 & -2-3+1+1 & -2+3-1-1 & -2-3-1-1 \\ 0-1-1+1 & 0+1-1-1 & 0-1+1-1 & 0+1+1+1 \\ 0-1+1-1 & 0+1+1+1 & 0-1-1+1 & 0+1-1-1 \end{bmatrix} \rightarrow \begin{bmatrix} 25 & 1 & 1 & -3 \\ 1 & -3 & 1 & [illegible] \\ -1 & -1 & -1 & [illegible] \\ -1 & 3 & -2 & [illegible] \end{bmatrix}$$

* Walse Transform

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

i) sign change (R to L [illegible])

0
3
1
2

ii) Ascending [illegible]

0
1
2
3

iii)

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

we have, $F = W * f(m,n) * W^T$

Scanned with CamScanner

$$F = \begin{bmatrix} 25 & 1 & -3 & 1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \\ 1 & 1 & -7 & -3 \end{bmatrix}$$

$F(x) = \{1, 2, 0, 3\}_{1\times 4}$
we have, $F = w * f(x)$

$F = [6 \quad 0 \quad 2 \quad -4]$

* kernel of 4 point DCT

$$\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.6532 & 0.2706 & -0.2706 & -0.653 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.2706 & -0.6532 & 0.6532 & -0.2706 \end{bmatrix}$$

+ + + +
+ + − −
+ − − +
+ − + −

1) DCT = Kernel * F(x)
2) DCT = F(K, l) = Kernel * F(x, y) * $Kernel^T$

Scanned with CamScanner

Q Find 2D DCT of $f(x,y)$ matrix

$$f(x,y) = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 1 & 2 & 1 \end{bmatrix}_{4\times4}$$

Q Find the

$$= \begin{bmatrix} 3 & 3.5 & 3.5 & 2.5 \\ -0.3702 & 0.1152 & -0.2642 & 0.2738 \\ 0 & -0.5 & 0.5 & -0.5 \\ -0.9235 & 1.577 & 0.6532 & -0.6532 \end{bmatrix}$$

Scanned with CamScanner

Q. Find the DCT of $F(x) = [1, 2, 4, 4]$

1) $DCT = [1, 2, 4, 4]$ $[5.5 \quad -2.438 \quad -0.5 \quad 0.4946]$

☆ HAAR Transform

Kernel for 2*2 matrix

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}_{2*2}$$

$$H_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}_{4*4}$$

main advantage
• samples the i/p data into fine resolution & takes the diff. btw adjacent pair
• resol^n increase by power 2 so in transform domain energy is more → more local^n

Q. Find the Haar Transform of the image.

$$F(x,y) = \begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix}_{2*2}$$

⇒ The 2D HAAR Transform of image $F(x,y)$ is given by $F(K,L)$ where,

$$F(K,L) = H_2 * f(x,y) * H_2^T$$

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$F(K,L) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Scanned with CamScanner

$$F(K,L) = \begin{bmatrix} 4 & 2 \\ -1 & 3 \end{bmatrix}$$

Assignment:

Q) Steps for filtering in frequency domain
i) Fast Fourier transform
ii)

•) Frequency Domain Filters are used for smoothing & sharpening of image by removal of high or low frequency components. Sometimes it is possible of removal of very high & very low frequency. Frequency domain filters are different from spatial domain filters as it basically focuses on the frequency of the images. It is basically done for two basic operation i.e smoothing & sharpening.
These are of 3 types:

Scanned with CamScanner

attenuate: reduce the force, effect, or value of

$g(x,y) = F^{-1}[G(u,v)]\,(-1)^{x+y}$

$F(x,y) \rightarrow g(x,y)$

Page No. 31

Frequency Domain Filters

Low Pass Filters | High Pass Filters | Band Pass Filters

i) Low Pass Filter

- It removes the high frequency components that means it keeps low freq. components
- It is used for smoothing the image
- It is used to smoothen the image by attenuating high frequency component & preserving low freq. component

Mechanism of low pass filtering in freq domain is given by:

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

where, $F(u,v)$ is the Fourier Transform of original image

$H(u,v)$ is the F. T of filtering mask

ii) High Pass Filter

- It removes the low frequency components that means it keeps high frequency components
- It is used for sharpening the image
- It is used to sharpen the image by attenuating low frequency components & preserving high freq components ~~& preserving high freq. component~~

Mechanism:

$$H(u,v) = 1 - H'(u,v)$$

where, $H(u,v)$ is the fourier transform of high pass filtering & $H'(u,v)$ is the F T of low pass filtering

Scanned with CamScanner

iii) Band Pass Filter:
- It removes the very low freq. & very high freq components that means it keeps the moderate range band of frequencies
- Band pass filtering is used to enhance edges while reducing the noise at the same time.

Fast Fourier Transform

Good to know side

- In layman's terms, the FT is a mathematical operation that changes the domain (x-axis) of a signal from time to frequency
- FT can speed up the training process of CNN.
  ~~In~~: a convolutional layer overlays a kernel on a section of an image & performs bit-wise multiplication with all of the values at that location. The kernel is then shifted to another section of the image & the process is repeated until it has traversed the entire image.
- F.T can speed up convolutions by taking advantage of following property.

$$x[n]*y[n] = \sum_{k=-\infty}^{\infty} x[k]\,y[n-k] \xleftrightarrow{FT} X(e^{j\omega})Y(e^{j\omega})$$

The above eqn states that the convolution of 2 signals is equivalent to the multiplication of their F.T. Therefore, by transforming the i/p into freq space, a convolution becomes a single element-wise multiplication. In other words, the i/p to a convolutional layer & kernel can be converted into frequencies using the F.T, multiplied once & then converted back using the inverse F.T

Discrete Fourier Transform

DTF

$$x[k] = \sum_{n=0}^{N-1} x[n]\, e^{\frac{-j2\pi kn}{N}}$$

discrete signal

N = size of its domain

To calculate the DFT of a signal (using computer), it would need to perform N (multiplications) x N (additions) = $O(N^2)$ operations
∵ we multiply each of $x[n]$ value by e raised to some function of n & then sum the results obtained for a given n.

- As the name implies, the FFT is an algorithm that determines DTF of an i/p significantly faster than computing it directly. In computer science lingo, the FFT reduces the no. of computations needed for a problem of size N from $O[N^2]$ to $O(N \log N)$

N	1000	10^6	10^9
N^2	10^6	10^{12}	10^{18} ns ⇒ 31.2 yrs
$N \log_{10} N$	10^4	20×10^6	30×10^9 ns ⇒ 30 sec

[1 operation takes 1 ns] say.

Scanned with CamScanner

Assignment - 9

FFT algorithm.

A FFT is an algorithm that calculates the discrete fourier transform (DFT) of some sequence. The discrete fourier transform is a tool to convert specific types of sequences of function into other types of representations

Another way to explain DFT is that transformed the structure of the cycle of a waveform into sine components

The FFT is a fast algorithm for computing the DFT. If we take the 2 points DFT & 4 point DFT & generalize them to 8-point, 16-point ... 2^n point we get FFT algorithm. (i.e divided into odd & even parts)

The FFT algorithm computes the DFT using $O(N \log N)$ multiplies & adds

—x—————x—————x—

STEPS for filtering in frequency domain

1) Multiply the i/p image by $(-1)^{x+y}$ to centre the transform.

2) Compute the DFT $F(u,v)$ of the resulting image

3) Multiply $F(u,v)$ by a filter $H(u,v)$

4) Compute the Inverse DFT transform $h^{-1}(x,y)$

Scanned with CamScanner

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

5th.

5) obtain the real part $h(x,y)$ of g
6) multiply the result by $(-1)^{x+y}$.

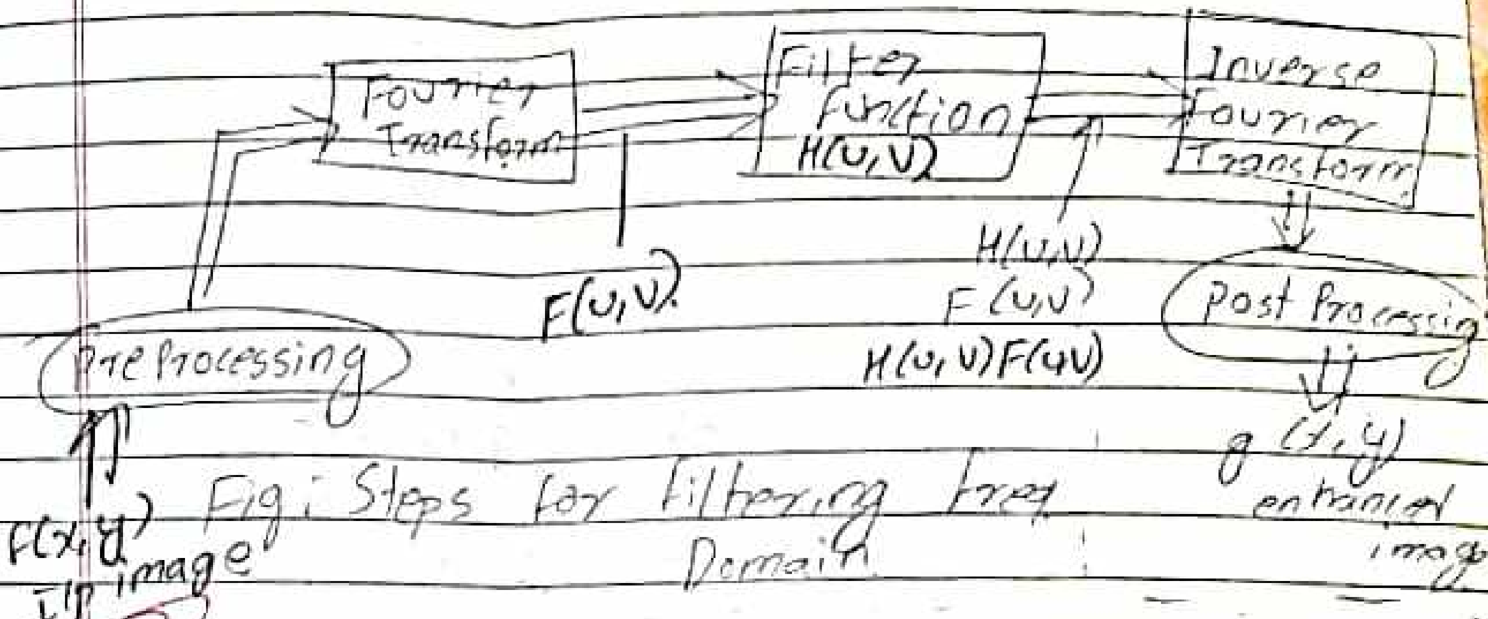


Fig: Steps for filtering Freq Domain

Imp

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

Ans

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

Sol,

Step 1:

Symbol	Original size Prob. (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				

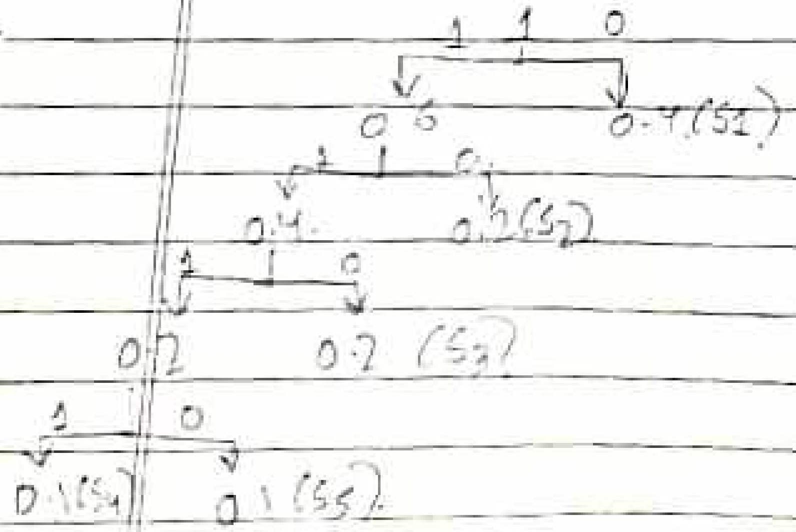
(pair) sano sano jodna prob. ani move forward.

Step 2:

Construct a Huffman Tree

Assume, Highest prob. 0.6 = 1

lowest prob 0.4 = 0



Scanned with CamScanner

Step③ : Generate code word length:

Symbol	Probability (P_i)	Code length (l_i)
S_1	0.4	0 = 4
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,

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

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

$= 2.2$ bits/symbol

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

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

$= 2.12$

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

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

$= 96.36\%$

Q. Calculate the compression ratio & 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

Step 1:

Calculate the probability in each gray level

Gray level (r_k)	No. of Pixel (n_k)	Prob. (P_i) = 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 of each probability in descending order

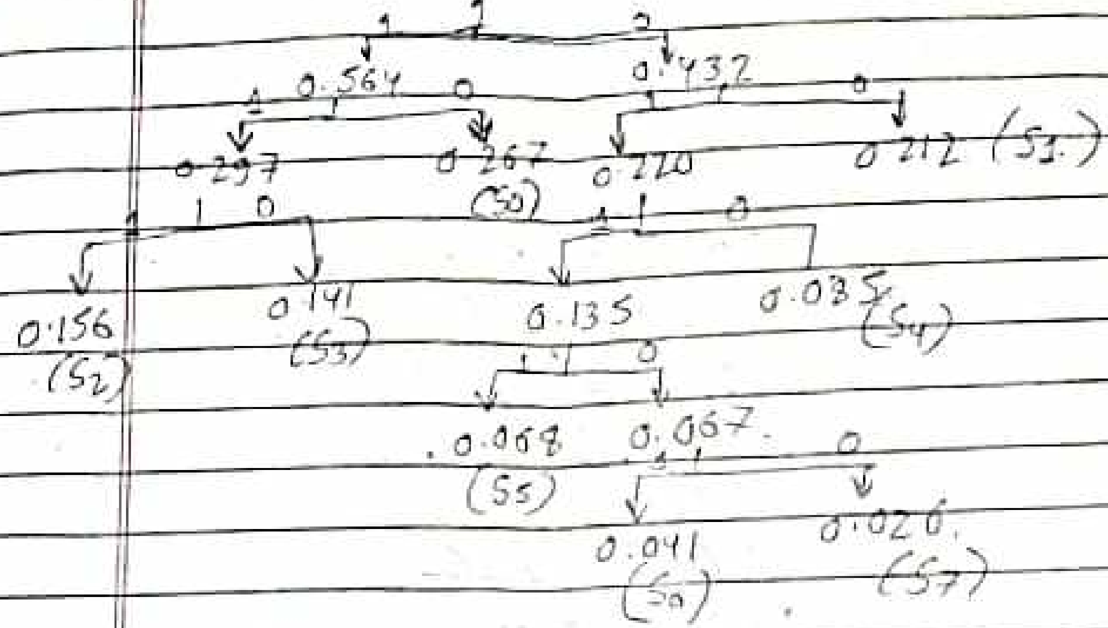
Symbol	P_i	1	Step 2	3	4	5	6	7
S_0	0.267	0.267	0.267	0.267	0.297	0.432	0.564	0.996 (1)
S_1	0.212	0.212	0.212	0.220	0.267	0.297	0.432	
S_2	0.156	0.156	0.156	0.212	0.220	0.267		
S_3	0.141	0.141	0.141	0.156	0.212			
S_4	0.085	0.085	0.135	0.141				
S_5	0.068	0.068	0.085					
S_6	0.041	0.067						
S_7	0.026							

Step 3: Construct Huffman Tree

Assume, Highest prob. 0.564 = 1

Lowest prob. 0.432 = 0

Scanned with CamScanner



Step 4: Generate Code word length

Symbol	Probability (P_i)	Code word	length
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	0.045	010	3
S_5	0.068	0111	4
S_6	0.041	01101	5
S_7	0.026	01100	5

Now, compression Ratio or (L_{avg}) $= \sum_{i=1}^{N} P_i l_i$

$= 0.267 \times 2 + 0.212 \times 2 + 0.156 \times 3 + 0.141 \times 3 + 0.045 \times 3 + 0.068 \times 4 + 0.041 \times 5 + 0.026 \times 5$

=

Vector addition

A⊙B = ...

$\begin{bmatrix} a \\ c \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} a+b \\ c+d \end{bmatrix}$

A = {(1,0), (1,1), (1,2), (2,2), (0,3), (0,4)}
B = {(0,0), (1,0)}

Vector addition
$\overline{A + B} =$
Then show A =

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

B =

	0
0	1
1	1

Now, A⊕B =

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

{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4), (2,0), (2,1), (2,2), (3,2), (1,3), (1,4)}

A⊖B = {(1,0), (1,1), (1,2), (2,2), (0,3), (0,4), (0,0), (0,1), (0,2), (1,2), (-1,3), (-1,4)}

	0	1	2	3	4
0					
1			1		
2					
3					

Here, the common is (1,2)

Opening & closing operation

→ Dilation & erosion operation are most fundamental to Image Morphology

→ Opening:

It is usually Erosion followed by a Dilution, using same structuring element. If A is the image & B is the " element then opening of A by B is given by

OPEN(A, B) = D(E(A)) D → Dilution, E → Erosion

It can be written as:

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

→ Closing:

It's dilution followed by Erosion, using same structuring element.

Close(A, B) = E(D(A))

It can be written as:

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

Q

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

$$B = \begin{array}{c|c} & 0 \\ \hline 0 & 1 \\ 1 & 1 \end{array}$$

Scanned with CamScanner

D = Add max
E = Sub min lina
centerpos

(1,1) (0,0)

$A \oplus B$	0	1	2	3	4	5	6	7	8	9
0		1	1	1			1	1	1	
1		1	1	1	0	0	1	1	1	
2		1	1	1	1	1	1	1	1	
3		1	1	1	1	1	1	1	1	
4		1	1	1			1	1	1	
5		1	1	1			1	1	1	
6		1	1	1			1	1	1	
7										
8										
9										

A = { (2,1), (2,2), (2,3), (3,1), (3,2), (3,3),
(4,1), (4,2), (4,3), (5,1), (5,2), (5,3),
(6,1), (6,2), (6,3), (4,4), (4,5), (2,6)
(2,7), (2,8), (3,6), (3,7), (3,8), (4,6), (4,7)
(4,8), (5,6), (5,7), (5,8), (6,6), (6,7), (6,8) }

B = { (0,0), (0,1,0) }

$A \ominus B$	0	1	2	3	4	5	6	7	8	9
0										
1										
2										
3		1	1	1		0	1	1	1	
4		1	1	1			1	1	1	
5		1	1	1			1	1	1	
6		1	1	1			1	1	1	
7										
8										
9										

1 → check if matching

1 tala 1 terma

if ya vana select else leave it

a wtai pos mila
.. tyo pos ma 1

center lai target pos. ma halna
erosion: saltai match vaya 1 natra 0

OPEN (A, B) =
$(A \oplus B) \ominus B$ =

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

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

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

satisfy vako ugna matra jana oni ie. lie. — 4. connectivity ma matra hunna

Page No.: 49
Date: / /

Q1 Apply region growing on the following image w initial point (2,2) & threshold value as 2. Use connectivity.

	0	1	2	3
0	0	1	2	0
1	2	5	6	1
2	1	4	7	3
3	0	2	5	1

T=2

→ seed point.

→ condition: absolute difference <= 2 4 way connectivity

7-6 = 1 ✓2

7-5 = 2 ✓2

7-4 = 3 ×4

7-3 = 4 × 3

6-1 = 5 × 1

6-5 = 1 ✓2

6-2 = 4 × 2

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

⇒

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

→ segmented image

2. Apply region growing on the following image with seed point & threshold value as 3.

5^a 6^a 6^a 7^a 6^a 7^a 6^a 6^a
6^a 7^a 6^a 7^a 5^a 4^a 4^a 7^a
6^a 6^a 4^a 4^a 3 2 5^a 6^a
5^a 4^a 5^a 4 2 3 4^a 6^a
0 3 2 3 3 2 4^a 7^a
6 0 0 0 0 2 2 5^a 6^a
1 1 0 1 0 3 4^a 4^a
1 0 1 0 2 3 5^a 4

~~Yesma~~ sardena
~~6 to wari~~
pari matra
herna

seed point = 6
Threshold (T) = 3 (a)
Absolute difference < = 3
8 way connectivity

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

segmented image

Assignment.
Edge or Boundary linking & Detection. — Uday
* Using Gradient Model
(A) 1st Derivative
(B) Second "

Scanned with CamScanner

(19)

Chapter-4
Image Restoration
[4 hours]

Introduction:

The objective of image restoration is to improve a given image in some pre-defined sense.

Image Enhancement is a subjective process whereas Image Restoration is an objective process.

Image Restoration is to re-construct or recover an image that has been degraded by using a prior knowledge of the degradation phenomena.

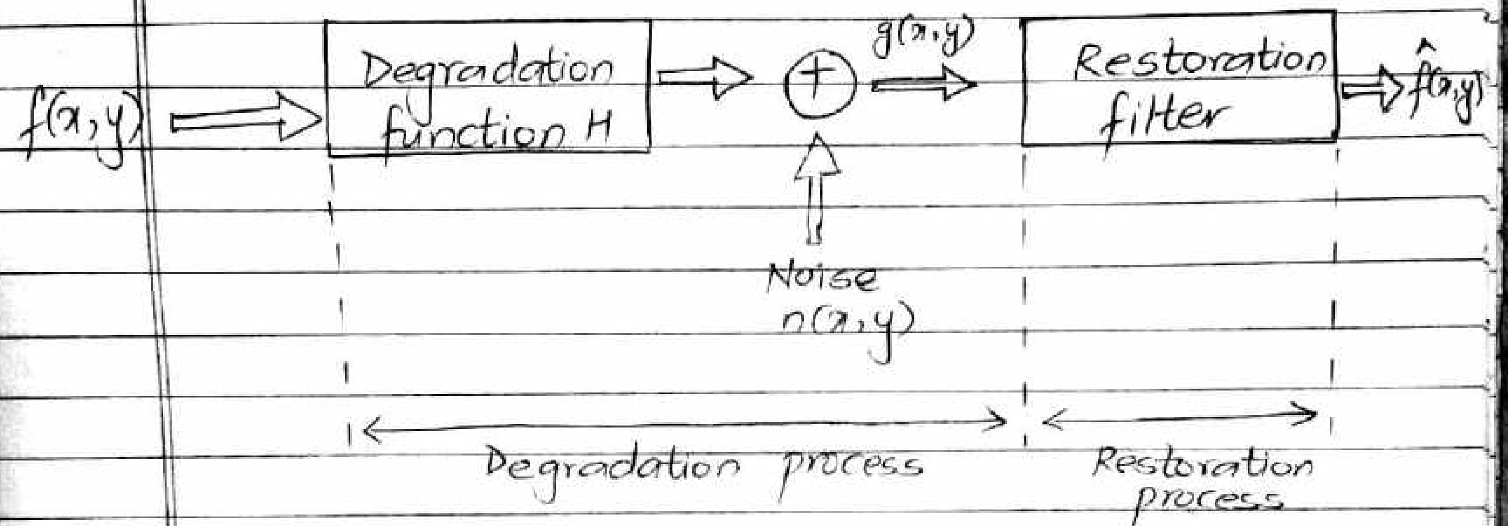
So, this technique is a process of degradation and applying the inverse process in order to recover the original image.

If there is a presence of noise and image bluring we use the concept of image restoration.

$\boxed{\text{Degraded Image} = \text{Original Image} + \text{Noise}}$

classmate
Date
Page

A model of the Image Degradation/Restoration Process:



(fig: A model of Image Degradation/ Restoration process)

The Degraded image $g(x,y)$ can be produced as:

$$g(x,y) = H[f(x,y)] + n(x,y) \quad \text{------(i)}$$

Where,

$f(x,y)$ = Input Image
$n(x,y)$ = Additive Noise term
H = Degradation function

Now,

The degraded image can be defined in spatial Domain as:

$$g(x,y) = h(x,y) * f(x,y) + n(x,y) \quad \text{---(ii)}$$

in function analysis convolution is a mathematical operation of two function (f & g) that produce the 3rd function

classmate
Date
Page

Where,

$h(x,y)$ = spatial Representation of the degradation function

Then,

the degraded image in spatial domain is equivalent to the convoluted frequency domain represented as:

$$G(u,v) = H(u,v) * F(u,v) + N(u,v) \text{ ...... (iii)}$$

$\therefore$ The degradation process is also referred to as convolving the image with a point spread ~~from~~ function (PSF).

Similarly,

Restoration process is also referred to as De-convolution.

Hence,

Restoration Process can be obtained as,

$$\hat{f}(x,y) = R[g(x,y)]$$ in Linear spatial Invariant Form

$$\hat{f}(u,v) = R[G(u,v)]$$ in Deconvolution form.

(2015 fall)

(in) ~~Image~~ spatial Domain)

"Noise Models Restoration" In the Presence of Noise only Spatial Filtering:

The method of choice for reducing the noise is spatial filtering. After filtering the noise, there should have a restoration process to have or obtain original image.

For noise reduction, following summarize and implement two types of spatial filters:

(1) Spatial Noise Filters
(2) Adaptive spatial filters

(1) Spatial Noise Filters:

In spatial noise filters, image is independently off how image characteristics vary from one location to another location.

In this condition, the spatial noise filters can be used and explained as:

(i) Arithmetic Mean,

$$\boxed{A = \frac{1}{n} * \sum_{i=1}^{n} x_i}$$

value of individual item

number of terms

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

(ii) Geometric Mean,

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

(iii) Harmonic Mean,

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

(iv) Median,

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

(v) Maximum,

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

(vi) Minimum,

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

Where,

m & n = Variables denotes image rows & columns spanned by filter respectively

S_{xy} = Denotes an $m*n$ subImage of the input noisy image g.

$\hat{f}(x,y)$ = filter response at (x,y) co-ordinate

(x,y) = co-ordinates of sub-image.

(2) Adaptive Spatial Filter:

In some application, results can be ~~implement~~ improved by using filters (capable of adapting behaviour) based on the characteristics of the image in the region being filtered.

The adaptive median filtering algorithm uses two processing levels for adaptive spatial filter that can be denoted as:

Level A: $\left(\begin{array}{l} A_1 = Z_{med} - Z_{min} \\ A_2 = Z_{med} - Z_{max} \end{array} \right)$

Algorithm → If $A1 > 0$ AND $A2 < 0$. go to level B

* If $Z_{min} < Z_{med} < Z_{max}$ go to level B
* Else Increase the window size

→ If window size $\leq S_{max}$, repeat level A

→ Else output Z_{med}

$$\begin{pmatrix} B_1 = Z_{xy} - Z_{min} \\ B_2 = Z_{xy} - Z_{max} \end{pmatrix}$$

classmate
Date
Page

algorithm **Level B:** (If $B_1 > 0$ AND $B_2 < 0$, output Z_{xy})

If $Z_{min} < Z_{xy} < Z_{max}$, output Z_{xy}

Else output Z_{med}

Where,

Z_{min} = Min^m intensity (gray level) value in S_{xy}

Z_{max} = Max^m intensity value in S_{xy}

Z_{med} = Median of the intensity value in S_{xy}

Z_{xy} = Intensity value at co-ordinates (x,y)

S_{xy} = Maxm allowed size of the adaptive filter window

(Frequency Domain)

Periodic Noise Reduction Using Frequency Domain Filtering:

Periodic Noise produces impluse like burst that often are visible in the Fourier Spectrum.

The principle approach for filtering these components is to use Notch reject filtering.

The general expression for Notch Reject Filtering having "Q" notch pair can be defined as:

$$H_{NR}(U,V) = \prod_{K=1}^{Q} H_K(U,V)\ H_{-K}(U,V)$$

$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

where,

$H_K(U, V)$ & $H_{-K}(U, V)$ = high pass filter with centers at (U_K, V_K) & $(-U_K, -V_K)$ respectively.

These translated centers are specified with respect to the center of the frequency rectangle $(M/2, N/2)$. Therefore, the distance computation for the filters are given by following expression:

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

&

$$D_{-K}(U, V) = \left[(U - M/2 + U_K)^2 + (V - N/2 + V_K)^2 \right]^{1/2}$$

→ In the video stream, periodic noise is typically caused by the presence of electrical or electromagnetic interference during video acquisition or transmission.

→ periodic noise can be reduced with frequency domain filtering.

→ frequency domain filtering isolate the frequency occupied by the noise and suppresses them using ~~band~~ Notch - Reject filter.

Image Restoration

Chapter-4

Image Restoration

Enhancement

- Subjective process
- Cosmetic changes in the brightness and contrast
- These algorithms are heuristic in nature
- Procedure in simple.

Restoration

- Objective Process
- It requires degradation of modeling
- Algorithms are well defined
- Procedure in complex

What is Image Restoration?

- Image restoration is to restore a degraded image back to the original image
- Image enhancement is to manipulate the image so that it is suitable for a specific application.

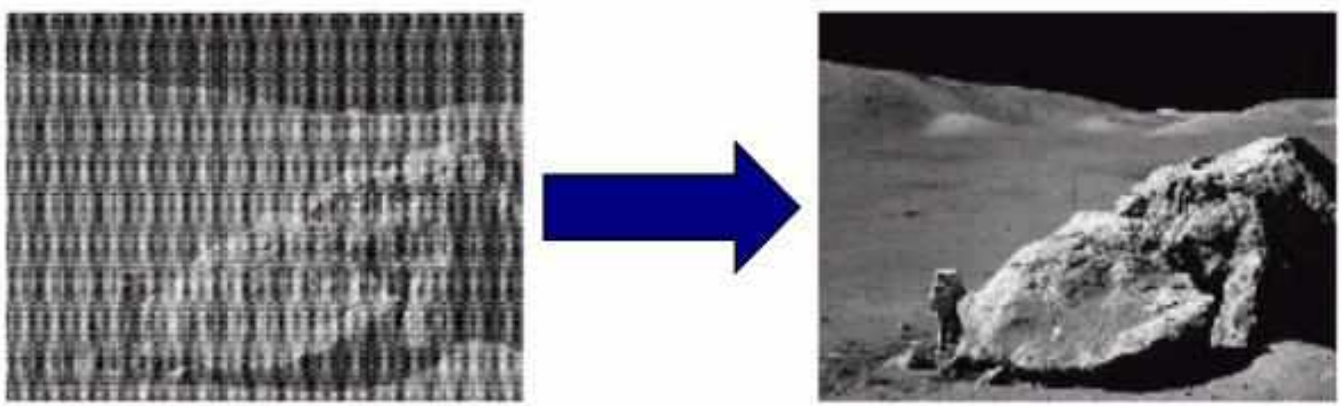
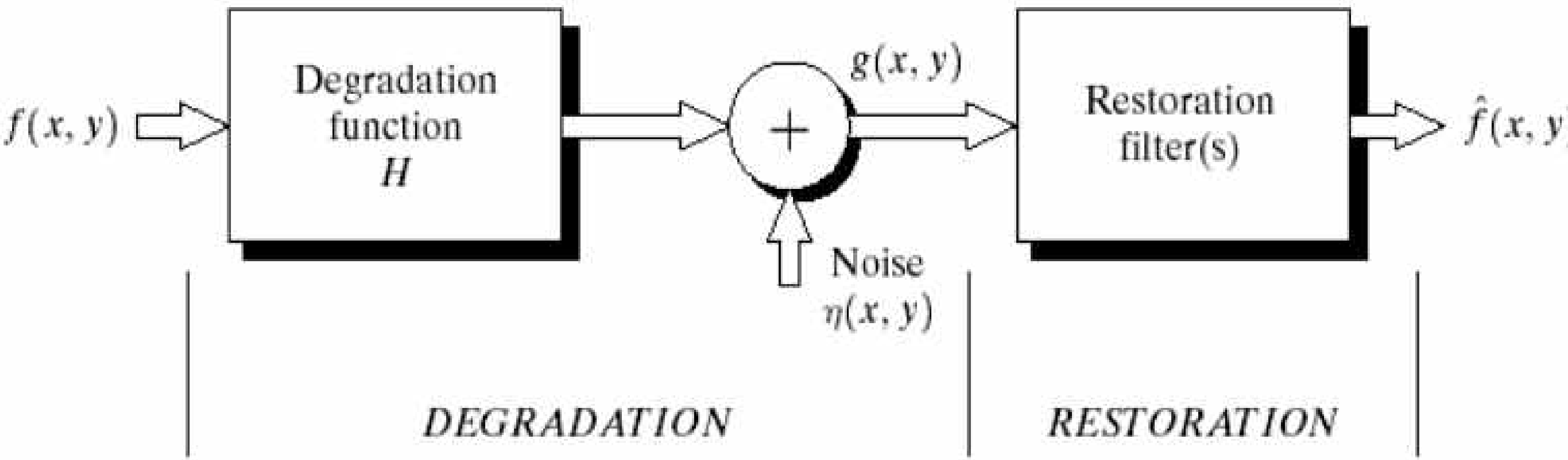


Image Restoration

Image restoration attempts to restore images that have been degraded

- Identify the degradation process and attempt to reverse it
- Similar to image enhancement, but more objective

A model of the image degradation/restoration process

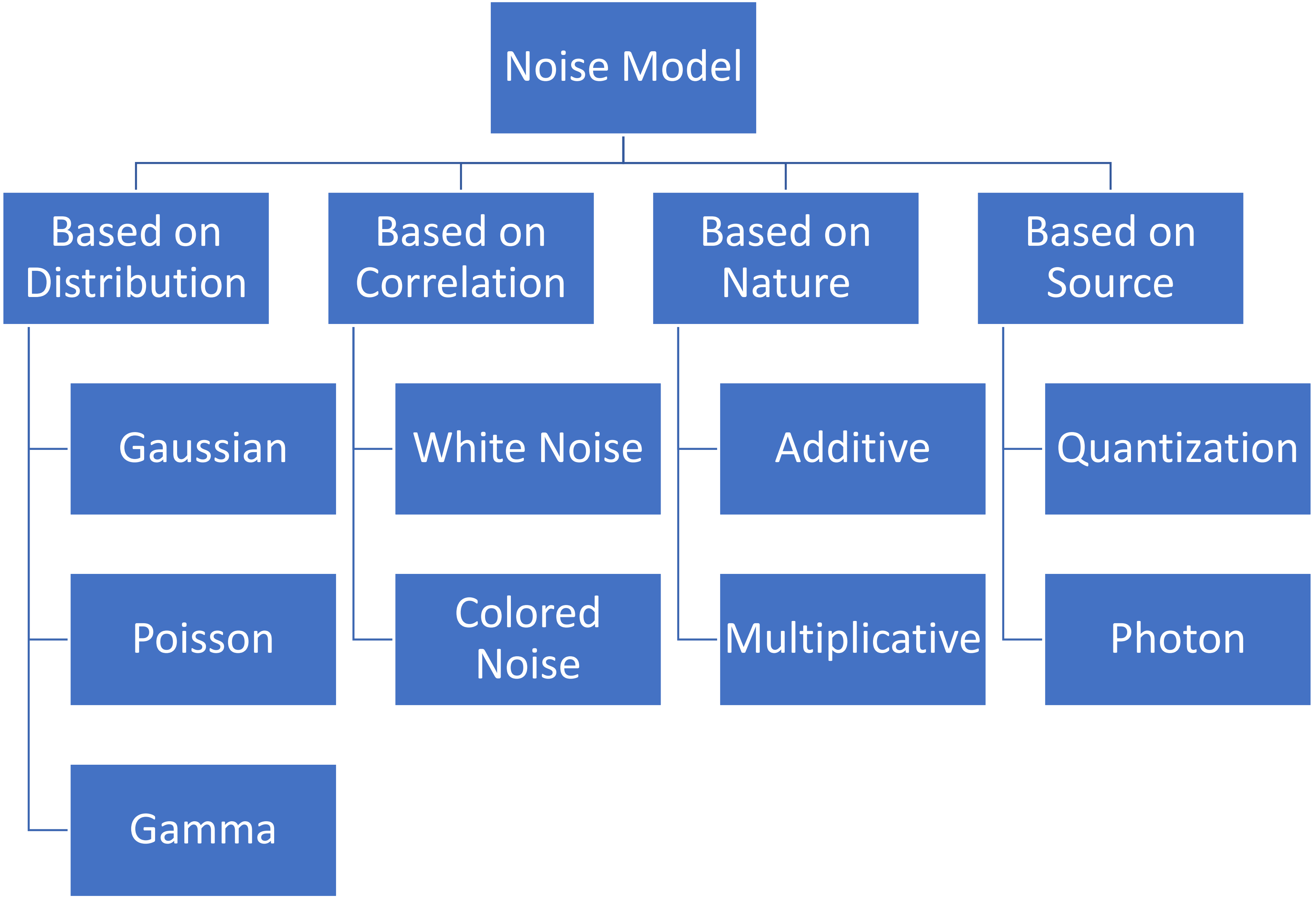


$$
\begin{cases}
g(x,y)=f(x,y)*h(x,y)+\eta(x,y) & \text{– Spatial domain} \\
G(u,v)=F(u,v)H(u,v)+N(u,v) & \text{– Frequency doma}
\end{cases}
$$

A model of the image degradation/ restoration process

- Where,
  - $f(x,y)$ - input image
  - $f^(x,y)$ - estimated original image
  - $g(x,y)$ - degraded image
  - $h(x,y)$ - degradation function
  - $\eta(x,y)$ - additive noise term

Noise Modelling (4 Category)



Types of Noise Based on Distribution

- Noise is a fluctuation in pixel values and it is characterized by Random Variable.
- A random variable probability distribution is an equation that links the values of statistical result with its probability of occurrence.

1. Gaussian Noise

- Random Noise that can be modelled as Gaussian or Normal Distribution.
- It affects both dark and light areas of image.

Gaussian noise

- Mathematical tractability in spatial and frequency domains
- Used frequently in practice
- Electronic circuit noise and sensor noise

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

Intensity mean variance

Note: $$\int_{-\infty}^{\infty} p(z)\,dz = 1$$

2. Impulse Noise

- Also known as salt and pepper noise and Binary Noise.
- Occurs mostly because of sensor and memory problem because of which pixel are assigned incorrect maximum values.
- PDF of I is

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

Z is pixel value
a and b is gray level

If either P_a or P_b is zero, it is called *unipolar*.
Otherwise, it is called *bipolar*.

If b>a, it means level a appears as lighter spot.

3. Poisson Noise

- This type of noise manifests as a random structure or texture in images.
- Very common in x-ray images.
- PDF of Poisson noise is
- $P(z) = \frac{(np)^2}{z!} e^{-np}$
- Where n= total number of pixels
- P=ratio of noise pixels to the total number of pixels

4. Exponential Noise

- It occurs mostly due to illumination problems
- Present in laser imaging.
- PDF is given by

$$p(z) = \begin{cases} ae^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases}$$

The mean and variance of this density are given by

$$\mu = 1/a \text{ and } \sigma^2 = \frac{1}{a^2}$$

5. Gamma (Erlang) Noise

- Also occurs due to illumination problem.
- PDF is given by:

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

  - The mean and variance of this density are given by $\mu = b / a$ and $\sigma^2 = \dfrac{b}{a^2}$
  - a and b can be obtained through mean and variance

6. Rayleigh Noise

- Mostly present in range noise. (used in remote sending application.PDF is given by:

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

−The mean and variance of this density are given by

$$\mu = a + \sqrt{\pi b/4} \text{ and } \sigma^2 = \frac{b(4-\pi)}{4}$$

−a and b can be obtained through mean and variance

7. Uniform Noise

- Occurs because of quantization and usually not visible

$$p(z) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le z \le b \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Mean:} \quad \mu = \frac{a+b}{2}$$

$$\text{Variance:} \quad \sigma^2 = \frac{(b-a)^2}{12}$$

Objective:- Simply Additive Noise

+ Image Restoration in Spatial Filter
[only in presence of Noise].

- Mean Filter
  - Arithmetic M.F
  - Geometric M.F
  - Harmonic MF
  - Contra-Harmonic MF
- order static Filter
  - Median Filter
  - Max "
  - Min "
  - Median "
  - Midpoint Filter
  - Alpha

① Mean Filter

Ⓐ Arithmetic mean Filter

→ removes local variation within image
→ Similar to low pass filter.
→ removes Gaussian noise & uniform noise

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

Scanned with CamScanner

2. Geometric M.F

- eliminate Gaussian Noise
- achives smoothing than A.MF

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

3. Harmonic M.F

→ works well for salt noise but fails for pepper noise
→ eliminate Gaussian Noise

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

Scanned with CamScanner

4. Contra-Harmonic M.F

- For salt & pepper noise

→ $Q>0$ for pepper noise

$Q<0$ " salt "

→ Q is the order of filter

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

Scanned with CamScanner

Order statistic Filter

→ also termed as rank, rank order or order filters.

→ are not based on convolution

1. Median Filters

→ is an example of non-linear filter

→ simply sorts list and find median

→ center pixel is replaced by median

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

2. Maximum Filter

→ select largest value in sorted list

→ used for removing pepper noise

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

Scanned with CamScanner

3. Minimum Filter

→ select smallest value
→ remove salt noise

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

4. Midpoint filter
→ select the mid-point
→ is the average of maximum & minimum values.

→ eliminate Gaussian & uniform Noise

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

Scanned with CamScanner

5. Alpha-Trimmed MF

- based on the concept of computation of the average of the pixels that falls with in window

→ removes Gaussian & salt pepper noise

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

- we delete the $d/2$ lowest & $d/2$ highest intensity values of $g(s,t)$ in the neighbourhood of S_{xy}

→ so g_r (s,t) represents the remaining $(mn-d)$ pixels

d ranges from 0 to $mn-1$

when $d=0$, it reduces to mean filter

" $d=mn-1$, it becomes median filter

Scanned with CamScanner

RNC

Date ______
Page ______

* Periodic Noise

→ arises due to electrical & electro magnetic ~~intere~~ inferente

→ give rise to regular noise pattern in an image

→ can be removed by band pass, Band-reject and notch filter.

① Band- Pass Filter

- allow frequency within particular range to pass through and attenuate all other frequency.

→ allow frequency if they fall in the range $D_l - D_h$ ($D_{low} - D_{high}$)

$$H(D) = \begin{cases} 1 & \text{for } D_l \leq D \leq D_h \\ 0 & \text{for } D > D_0 \text{ (cut off frequency)} \end{cases}$$

Scanned with CamScanner

Date ______
Page ______

2D band pass filter

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

Butterworth Band pass filter

$$H_{bp}(u,v) = \frac{\left[\dfrac{D(u,v)+W}{D^2(u,v) - D_0^2}\right]^{2n}}{1 + \left[\dfrac{D(u,v)*W}{D^2(u,v) - D_0^2}\right]^{2n}}$$

Gaussian band pass filter

$$H(u,v) = e^{-\frac{1}{2}\left[\frac{D^2(u,v) - D_0^2}{D(u,v)\,W}\right]^2}$$

Scanned with CamScanner

PNC

2. Band-Reject Filter

→ attenuate limited range of frequency & leaving all other freq.

→ It is complement of Band Pass Filter

$$H_{br}(u,v) = 1 - H_{bp}(u,v)$$

$$H(D) = \begin{cases} 0 & \text{for } D_l \le D \le D_h \\ 1 & \text{for } D > D_0 \end{cases}$$

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

* Gaussian-B-R-Filter

$$H(u,v) = 1 - e^{-\frac{1}{2}\left[\frac{D^2(u,v) - D_0^2}{D(u,v) * W}\right]^2}$$

3. Butterworth B-R Filter

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

Notch Filter (specifies within given range)

$$H(u,v) = \begin{cases} 0 & \text{if } D_1(u,v) < D_0 < D_2(u,v) \\ 1 & \text{else} \end{cases}$$

$$D_1(u,v) = \left[\left\{\left(u - \frac{M}{2}\right) - u_0\right\}^2 + \left\{\left(v - \frac{N}{2}\right) + v_0\right\}^2\right]^{\frac{1}{2}}$$

[$\because$ Distance Formula]

$$D_2(u,v) = \left[\left\{\left(u - \frac{M}{2}\right) - u_0\right\}^2 + \left\{\left(v - \frac{N}{2}\right) + v_0\right\}^2\right]^{\frac{1}{2}}$$

Scanned with CamScanner

Butterworth N·F

$$H(u,v) = \frac{1}{1 + \left[\frac{D_0^2}{D_1(u,v) \cdot D_2(u,v)}\right]^n}$$

Gaussian N·F

$$H(u,v) = 1 - e^{-\frac{1}{2}\left[\frac{D_1(u,v) \cdot D_2(u,v)}{D_0^2}\right]}$$

Scanned with CamScanner

Adaptive Median Filter:

→ It has 3 main purposes
- (a) To remove salt-and-pepper noise (Impulse noise)
- (b) To provide smoothing of other noise that may not be impulse
- (c) To reduce distortion

Let

Z_{min} = minimum intensity value in S_{xy}
Z_{max} = maximum " " " S_{xy}
Z_{mid} = median " " " "
Z_{xy} = intensity value at coordinate (x,y)
S_{max} = maximum allowed size of S_{xy}

→ Adaptive median Filter works in 2 stages

stage A: $A1 = Z_{med} - Z_{min}$
$A2 = Z_{med} - Z_{max}$
If $A_1 > 0$ ($Z_{min} < Z_{med}$) AND
$A2 < 0$ ($Z_{max} > Z_{med}$), goto step B.

else
increase the window size

If window size $\leq S_{max}$ repeat stage A
else o/p Z_{med}.

Scanned with CamScanner

Stage B:

$B1 = Z_{xy} - Z_{min}$

$B2 = Z_{xy} - Z_{max}$

If $B1 > 0$ ($Z_{xy} > Z_{min}$) AND
$B2 < 0$ ($Z_{xy} < Z_{max}$)
o/p Z_{xy}
else o/p Z_{med}.

→ Purpose of stage A is to determine Z_{med} is an impulse or not

If $Z_{min} < Z_{med} < Z_{max}$ hold Z_{med} can't be an impulse
In this case, we goto stage B and test if Z_{xy} itself an impulse

→ If condition $B1 > 0$ & $B2 < 0$ is true, then $Z_{min} < Z_{xy} < Z_{max} \Rightarrow Z_{xy}$ cannot be impulse
→ o/p unchanged pixel value Z_{xy}

→ If $B1 > 0$ & $B2 < 0$ is false, then either $Z_{xy} = Z_{min}$ or $Z_{xy} = Z_{max}$
→ In either case, value of pixel is extreme value & o/p Z_{med}, which is from stage A, we know is not impulse.

→ If $A1 > 0$ & $A2 < 0$ fails i.e Zmed is find as impulse, it won't go for stage B.
else
increase the size of window &
repeat stage A

→ This looping continues until it either find a med. value that is not an impulse or the maximum window size is reached

→ If maximum window size is reached, its o/p is Zmed & it can be impulse or not.

Scanned with CamScanner

(I Numerical sure)

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

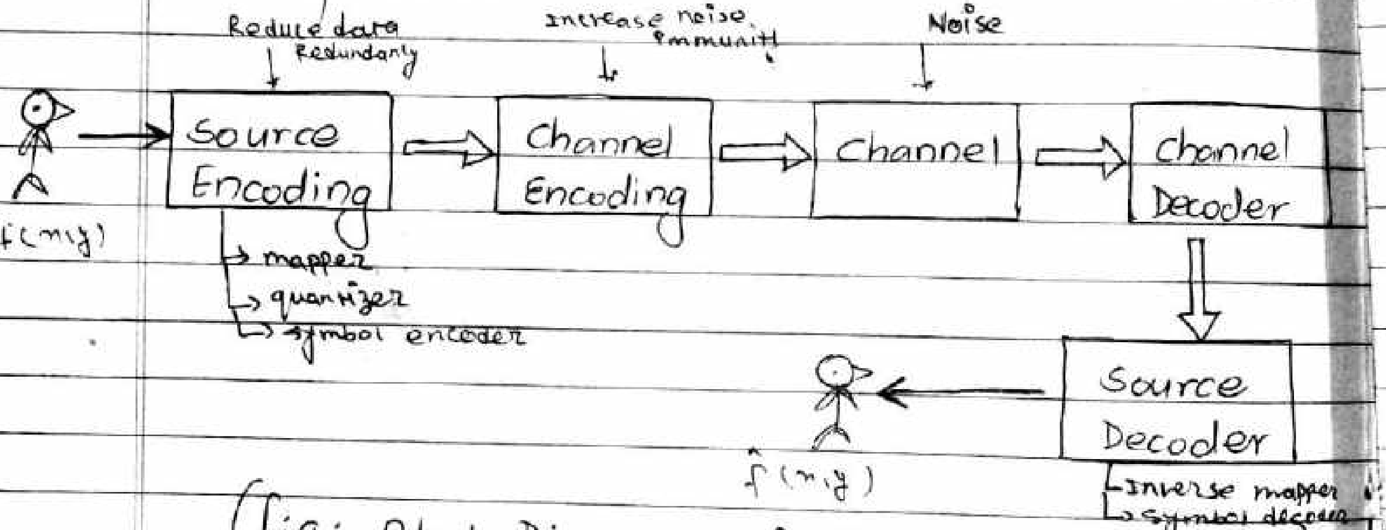
mapper: Reduce Interpixel Redundancy
quantizer: reduce psychovisual Redundancy
Symbol encoder: reduce Coding Redundancy

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

classmate
Date
Page

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.

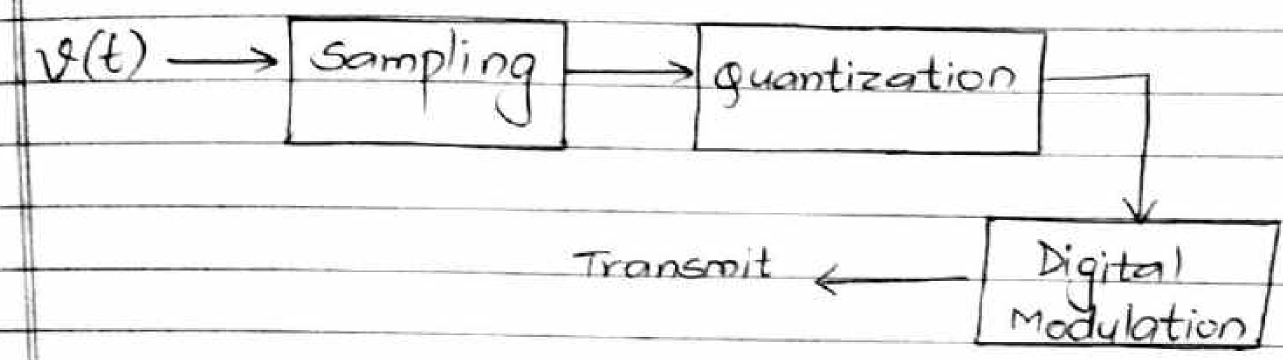
Pixel coding technique is further divided into following three types:

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

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

classmate
Date
Page

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

classmate
Date
Page

(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.

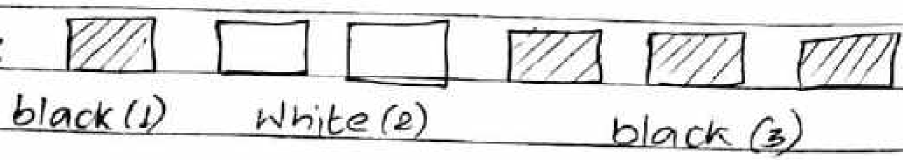
for example, aaabbbaaa is coded as 3a3b3a.

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 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) <u>Entropy Coding</u>: (Replace input string by code word)

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

The most important entropy coding is Huffman Coding. The <u>theoritical minimum average of bits</u> 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.

classmate
Date
Page

step(1):

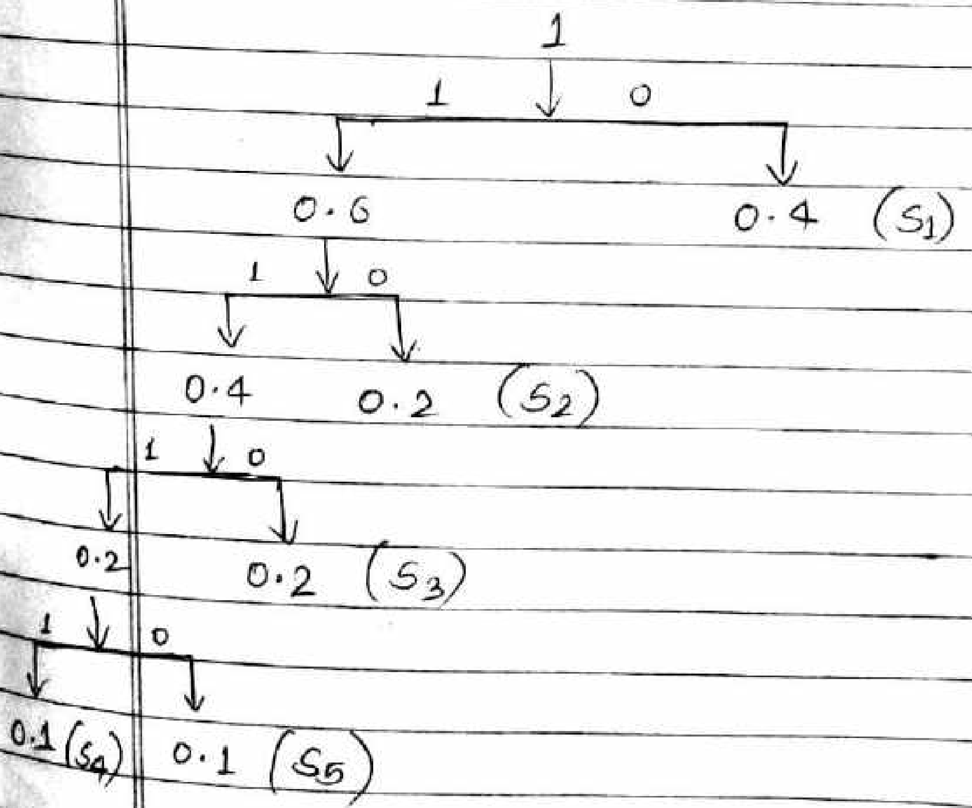
Sol^n:

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)

classmate
Date
Page

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,

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

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

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

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

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

$$= 2.12$$

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

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

$$= 96.36\%$$

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

classmate
Date
Page

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

Gray level (γ)	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 (γ)	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
	$n = 9569$	

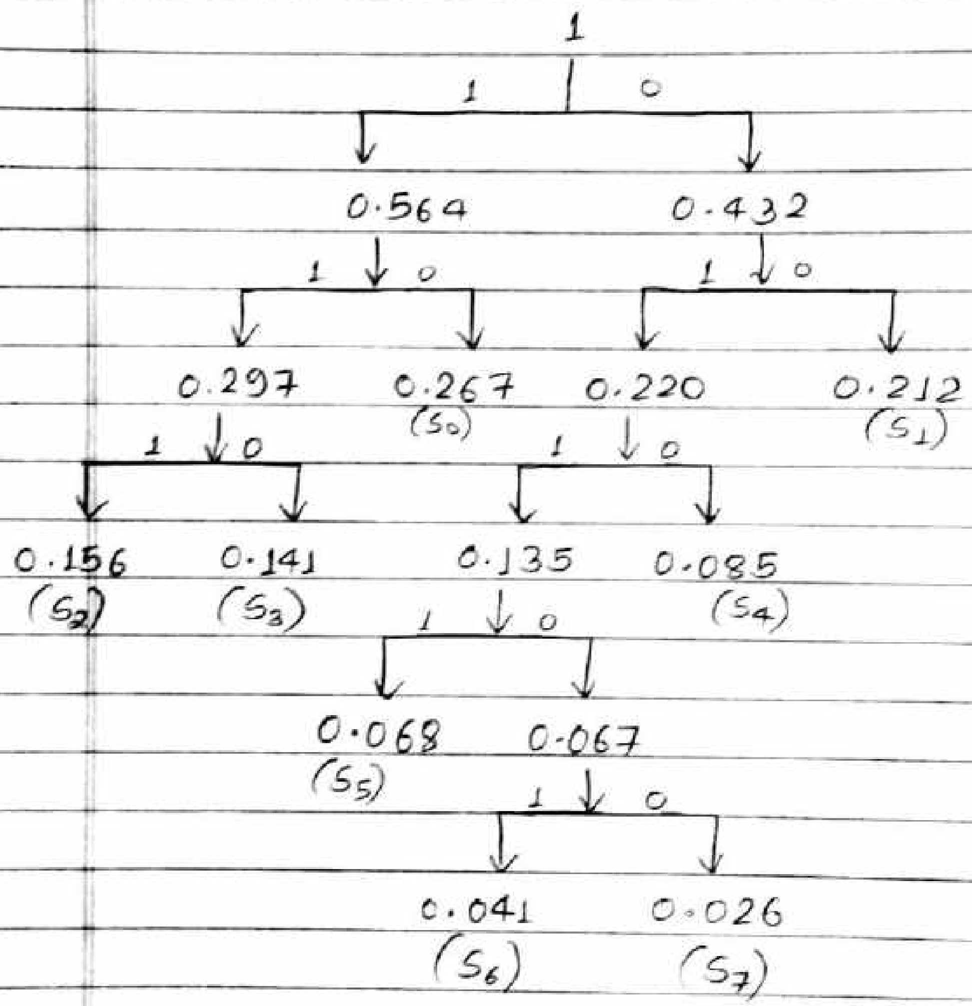
step(2): Assume symbol for each probability in descending order.

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							

classmate
Date
Page

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,

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

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

$$= 2.711$$

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

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

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

$$= 2.687$$

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

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

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

→ 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, i^{th} matrix consist of i^{th} 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.

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

(2) Predictive Coding:

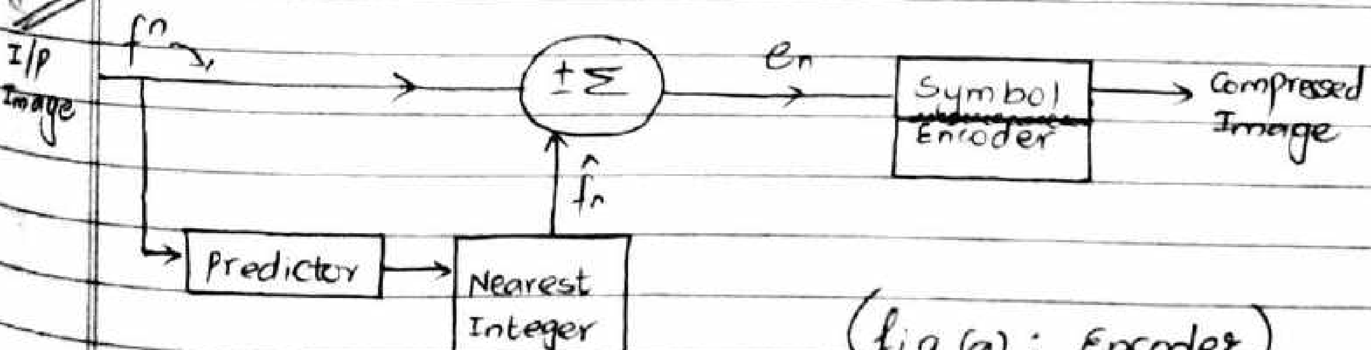
Future value will be predicted on the basis of past values.

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.

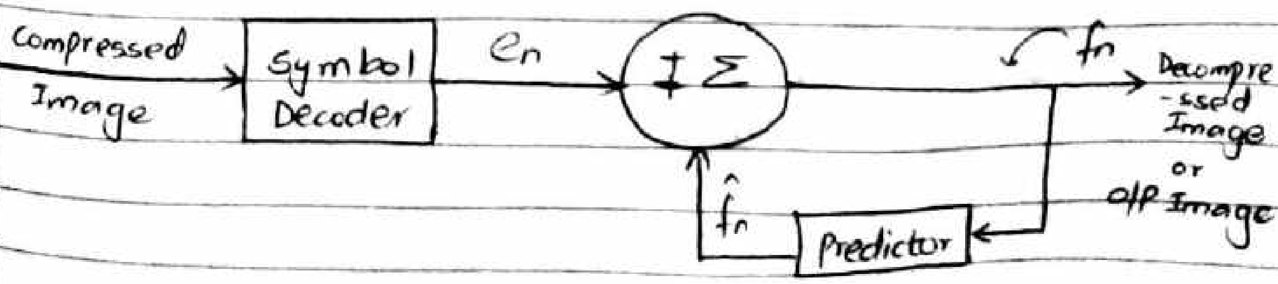
It is of following two types:
(a) lossless predictive coding
(b) lossy Predictive coding

(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 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, symbol encoder uses variable length coding technique to encode symbol.

In case of decoding, 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

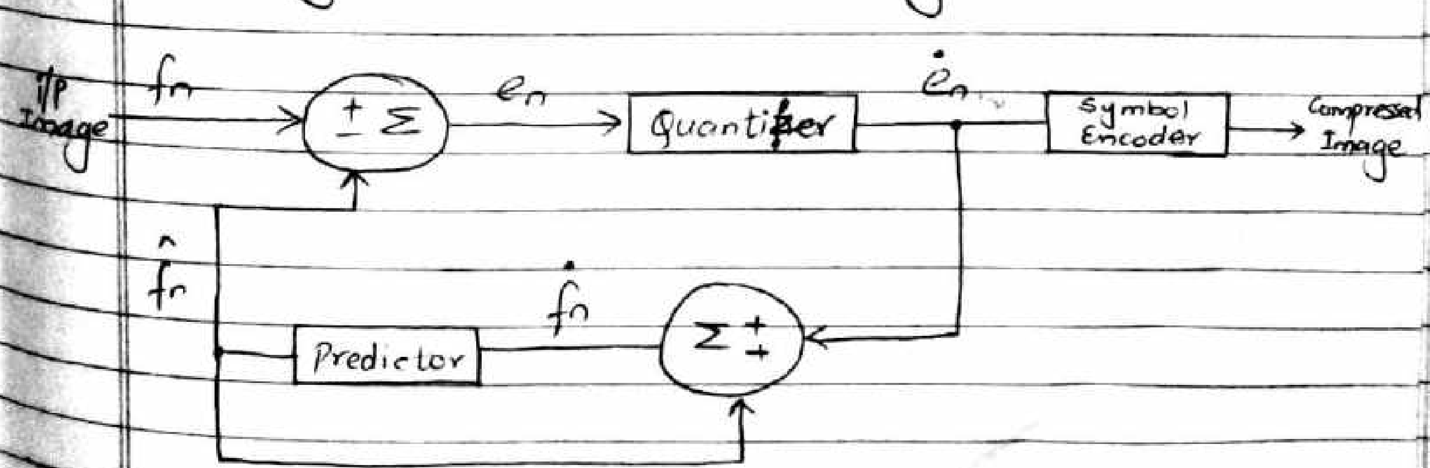
$$\boxed{\hat{f}_n = \text{round}\left[\sum_{i=1}^{m} \alpha_i \, f_{m-i}\right]}$$

Where,

α = Prediction co-efficients

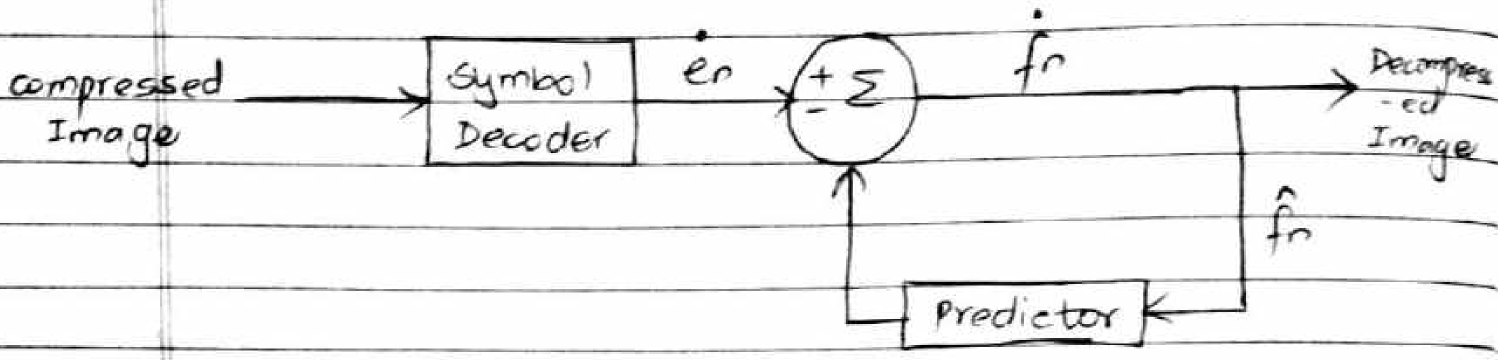
m = order of linear prediction

6) Lossy Predictive Coding:



[fig: Encoder]

classmate
Date
Page



[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 ~~ere~~ encoding & 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. U.K. Raibagi

→ compressing number of frame together

→ compressing individual frame

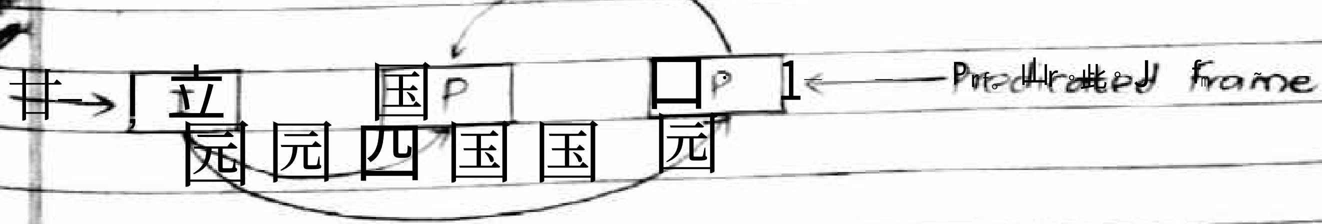
Interframe & Intraframe Coding:

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

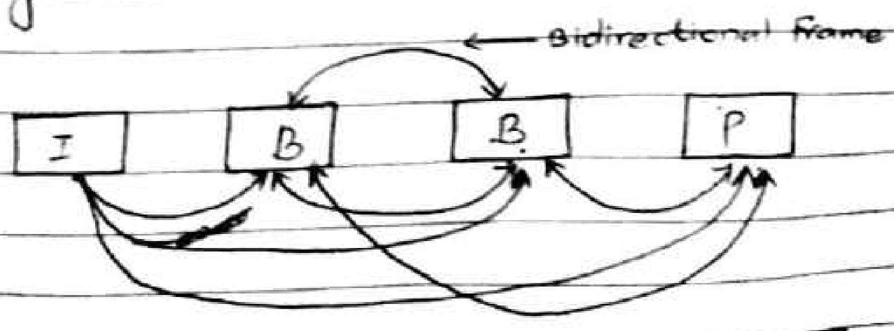
Interframe coding is encoding and compression technique between the video signal. (use number of frame)

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 & next frame for encoder.)



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



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

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 (CIP)
(2) Interpixel Redundancy
(3) Psychovisual Redundancy

(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 pixels 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 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.

PNK

Date
Page

Chapter-6

(3) Morphology Image Processing

Morphology is a branch of biology dealing with form and structure of animals & plants.

- It is a tool for extracting image component that are usefull in the representation & description of region & shape such as boundaries, skeletons and the convex hull.

- Transition from purely image processing methods (whose I/p & O/p are images) to processing in which inputs are images but output are attributes extracted from those images.

→ Reflection of set B donoted by $\hat{B}$, is defined as

$$\hat{B} = \{ w \mid w = -b, \text{ for } b \in B \}$$

→ Transition of a set B by point $z = (z_1, z_2)$, denoted by $(B)_z$ is defined as

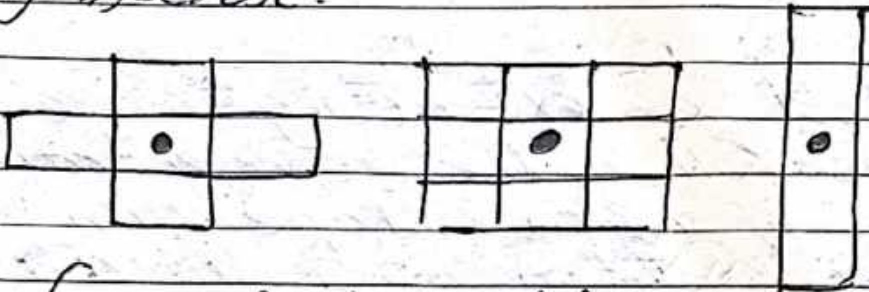
$$(B)_z = \{ c \mid c = b + z \text{ for } b \in B \}$$

6.1

Scanned with CamScanner

Structuring Element

→ small sets or subimages used to probe an image under study for properties of interest.



{symmetry city determine center point}

Dilution

→ A & B are sets in integer 2-D, dilution is denoted by $A \oplus B$ and is defined by

$$A \oplus B = \{ Z \mid (\hat{B})_Z \cap A \neq \phi \}$$

→ If, Reflecting B about its origin & shifting this reflection by Z, then dilution of A by B is the set of all displacements, Z, such that, B and A overlap at least one element.

Scanned with CamScanner

i.e $A \oplus B = \{ z \mid [(\hat{B})_z \cap A] \subseteq A \}$

→ dilution grows or thickens object in binary image.

→ Specific manner and extent of thickening is controlled by the shape of the structuring element used

→ It is used for bridging gaps

Erosion:

→ with A and B as sets in integer 2D, then the erosion of A by B is denoted by $A \ominus B$ and is defined as

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

→ i.e erosion of A by B is the set of all points z such that B, translated by z, is contained in A

6.3

→ B has to be contained in A is equivalent to B not sharing any common element with back ground

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

→ erosion shrinks or thins objects in a binary image

For dilation		For Erosion	
Perfect Match	= 1	Perfect Match	= 1
Some "	= 1	Some "	= 0
No "	= 0	No "	= 0

Scanned with CamScanner

PNC

Date ____ Page ____

Opening and closing

Opening ($O E A$)

→ Opening of set A by structural Element B is denoted by $A \circ B$ and is given by

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

→ The opening of A by B is obtained by taking union of all ~~transition~~ translation of B that fit into A. i.e

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

→ Opening smoothes contour of object, breaks narrow isthmuses and eliminating thin protousions.

Properties of opening

(1) $A \circ B$ is subset/subimage of A

(2) If C is a subset of D, then $C \circ B$ is a subset of $D \circ B$.

6.5

Scanned with CamScanner

③ $(A \circ B) \circ B = A \circ B$

closing (CDE)

→ closing of set A by structuring element B is denoted $A \bullet B$ and is given by

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

→ closing smoothes section of contours but fuses narrow breaks and long thin gulf eliminates small holes & fill gaps in the ~~coun~~ contour.

Properties:

①. A is a subset of $A \bullet B$

② If C is a subset of D then $C \bullet B$ is a subset of $D \bullet B$

③ $(A \bullet B) \bullet B = A \bullet B$.

Q. Given an image A and a ~~stur~~ structural element B below, calulate erosion $A \ominus B$, dilution $(A \oplus B)$ and closing $(A \bullet B)$

$A =$

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

$B =$

0	1	0
1	1	1
0	1	0

$A_e =$

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

(b) $A \oplus B = A(d)$

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

6.7

Scanned with CamScanner

P.M=1
S.M=1
N=0.

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

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

$(A \ominus B)$

0	1	0
1	1	1
0	1	0

S.E

$(A \ominus B) \oplus B$.

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

6.B

P.N.C

P.M = 1
S.M = 0
N.O = 0

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

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

$A \oplus B$

0	1	0
1	1	1
0	1	0

S.E

$(A \oplus B) \ominus B$

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

6-9

Scanned with CamScanner

```
%Program for Erosion and Dilution in Binary Images
close all
clear
clc
%Reading Input Image
A=imread('demo1.bmp');
mask=[0 1 0;1 1 1; 0 1 0];
%Image Dilation
A_Dilation=imdilate(A,mask);
%Image Erosion
A_Erode=imerode(A,mask);
subplot(221);imshow(A),title ('Original Image');
subplot(222);imshow(A_Dilation), title('Dilated Image');
subplot(223);imshow(A),title ('Original Image');
subplot(224);imshow(A_Erode), title('Eroded Image');
```

Original Image

Dilated Image

Original Image

Eroded Image

```
%Program for Opening and Closing in Binary Images
close all
clear
clc
%Reading Input Image for Opening and Closing Operation
A=imread('demo1.bmp');
mask=[1 1 1; 1 1 1; 1 1 1];
%Image Opening
A_Opened=imopen(A,mask);
%Image Closing
A_Closed=imclose(A,mask);
%Display Of Input and Output Images
imshow(A),title ('Original Image');
figure,imshow(A_Opened), title('Opened Image After Morpholocial Processing');
figure,imshow(A_Closed), title('Closed Image After Morpholocial Processing');
```

19

+ 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

0 → background pixel, 1 → 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 bivalued 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 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)

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,

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 5-pixels arranged in a diagonal line.

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

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)

classmate
Date
Page

overlap 1-valued pixel in the original Image

```
                * * * * * [*] * * *
              * * * * * * * * *
            1 1 1 1 1 1 1 * *
          * 1 1 1 1 1 1 1 *
        * * 1 1 1 1 1 1 1
      * * * * * * * * *
    * * * * * * * * *
```

Donot overlap 1-valued pixel in the original Image

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

```
            1 1 1 1 1 1 1 1 1
          1 1 1 1 1 1 1 1 1
        1 1 1 1 1 1 1 1 1
      1 1 1 1 1 1 1 1 1
    1 1 1 1 1 1 1 1 1
  1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1
```

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

classmate
Date
Page

reflected & translated is $\boxed{\text{overlaps}}$ 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 $\boxed{\text{convention}}$ 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

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

Date
Page

(Imp) (S.N)

Erosion:

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.

1
$\boxed{1}$ ← Origin of the structuring element is shown with dark border.
1

③ Structuring element translated into several location in the image.

classmate
Date
Page

$$\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & \boxed{1} \\ 1 & 1 & 1 & \boxed{1} & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}$$

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.

$$\begin{array}{ccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

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

classmate
Date
Page

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

Opening & Closing:

imp L.Q

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:

classmate
Date
Page

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

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:

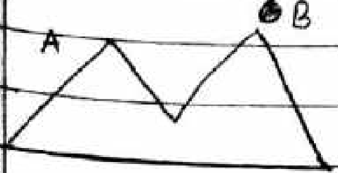
The morphological closing of A by B is denoted as $A \bullet B$ and defined as the dilation of A by B and followed by a erosion of result by B.

i.e

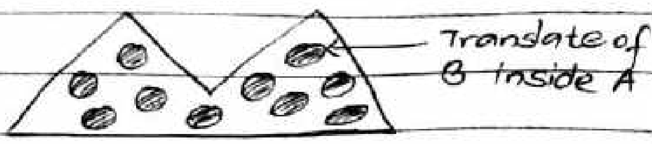
$$A \bullet B = (A \oplus B) \ominus B$$

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:

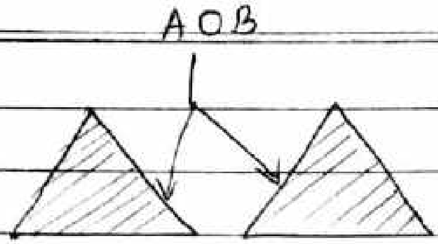


fig(a): set A and structing element B

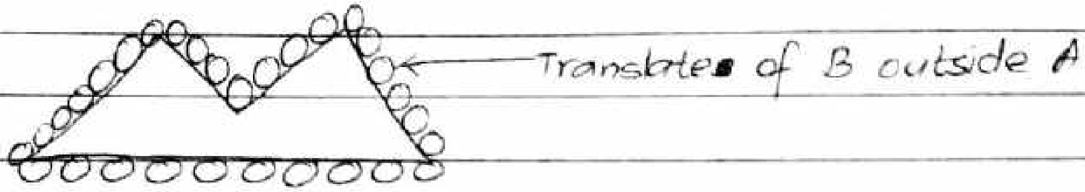


fig(b): Translation of B that fit entirely within set A.

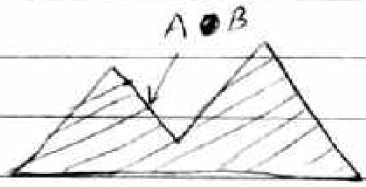
classmate
Date
Page



(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.

(29)

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.

classmate
Date
Page

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

classmate
Date
Page

(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) $P(R_i) = TRUE$

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 a 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 $2m * 2n$.

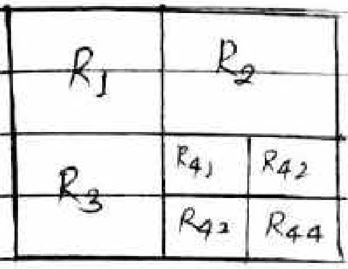
So, the procedure is summarized as:

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

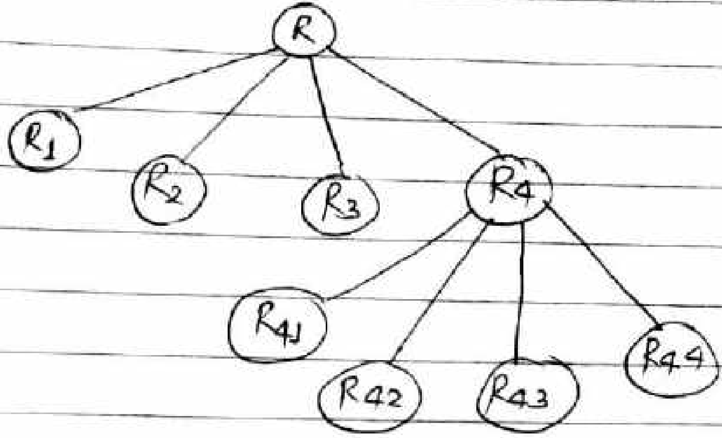
(ii) Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$ = TRUE

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

Example:



fig(a): Partitioned Image



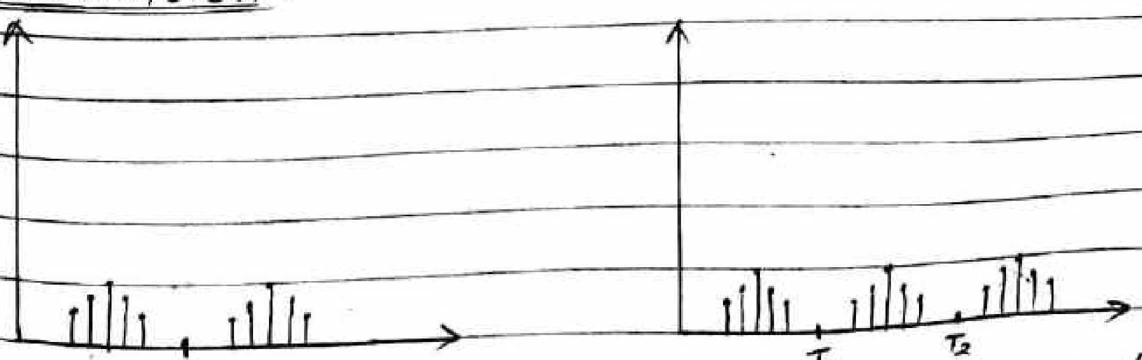
fig(b): Corresponding quadtree

(IMP) (4) Segmentation By Thresholding: (2012 fall) (2016 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 by as 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,

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

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 thresholding

Fix (all) value

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 in...
③ mean μ_1 for group G_1, μ_2 for G_2
Repeat 4. $T = \frac{\mu_1 + \mu_2}{2}$
5. $|T_i - T_{i+1}| < T_i$

Classmate
Date ______
Page ______

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.

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

CLASSMATE
Date
Page

with correlation

(IMP)

Template Matching : 2012 fall) (1 Q)

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 the 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 through the use of correlation Theorem.

(P.T.O)

* using co-relation find the matching betn template pixel & original image pixel.

Part 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

Edge Linking & boundary ~~detection~~

Edge or Boundary Linking & Detection:

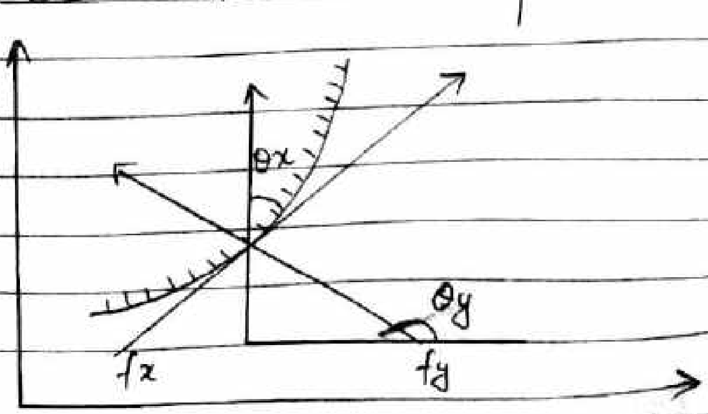
① Using Template Matching:

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 discontinuties 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:

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

Now,

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

$$\therefore \frac{df}{dr} = fx \cdot \cos\theta + fy \cdot \sin\theta$$

1

classmate
Date
Page

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):**

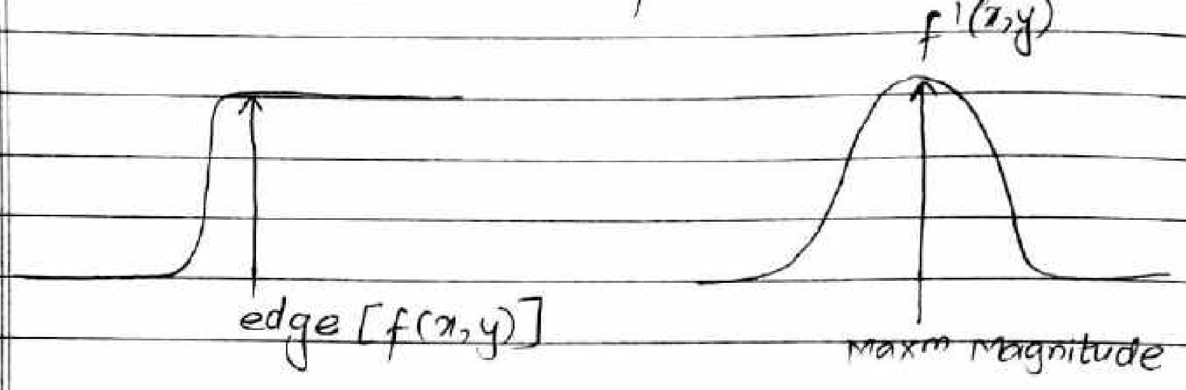
(A) First Derivative:

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

(B) Second Order Derivative

(C) ~~Compass Operator~~

Ⓐ First Derivative :
[i.e Difference operator]



fig(a) : Cross section of an image edge

fig(b) : Cross section of an image edge using first derivative

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 x 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 x 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:

classmate
Date
Page

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

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 3rd & 1st column approximates the derivates in y-direction.

classmate
Date
Page

(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)

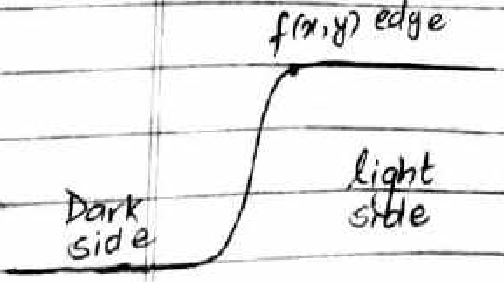


fig (a) : Cross Section of ramp edge

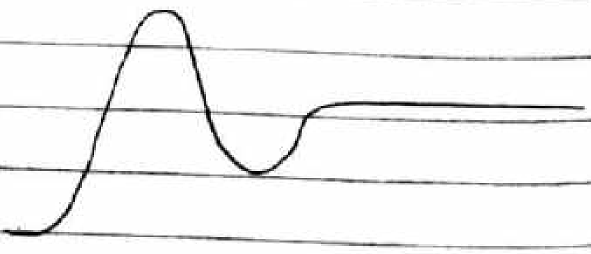


fig (b) : Second derivative of ramp edge

The second derivative is $\boxed{\text{positive}}$ at the transition into ramp 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.

classmate
Date
Page

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:

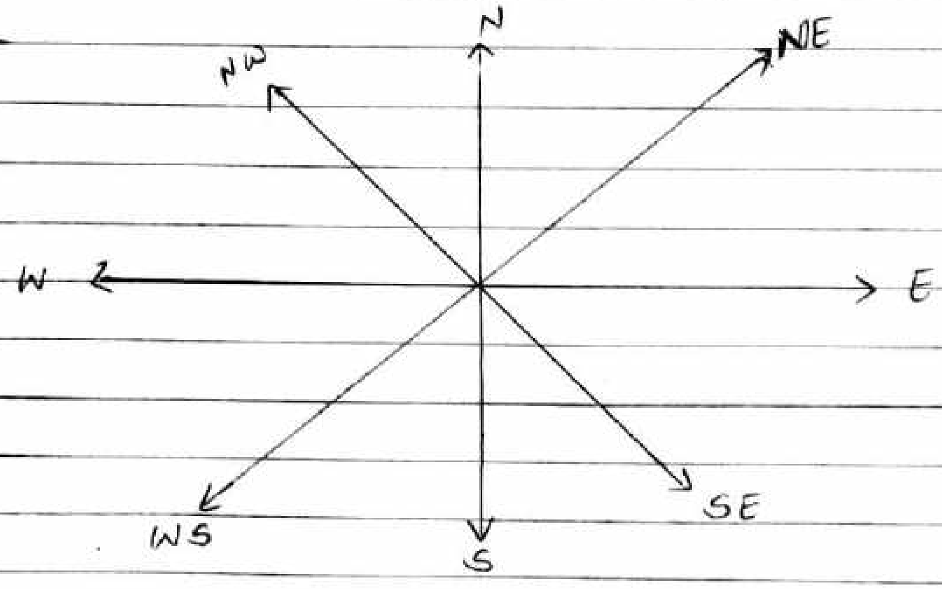
$\Delta^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$

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 → NE

(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) $\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$ ↗ +45°

→ Horizontal

→ +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.

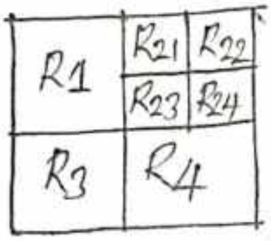
7. Image Segmentation

→ Segmentation is the process of partitioning a digital image into multiple regions and extracting the meaningful regions is known as Region of Interest (RoI).

→ Image Enhancement Algorithms are based on 2 basic properties of intensity values

ⓐ Discontinuity → objective is to extract region differ in properties like intensity, color, texture. Also known as Boundary Approach.

ⓑ Similarity (Continuity):- objective is to ~~extract~~ group pixels based on common property to extract a coherent region. Also known as region based approach.



Let R represent entire image, region & segmentation is partitioning R into n-subgroup R_i

→ characteristics of Segmentation Process:-

ⓐ $\bigcup_{i=1}^{n} R_i = R$

ⓑ R_i should be connected region $i = 1, 2, \ldots n$

ⓒ $R_i \cap R_j = \phi$ for all i & j $i \neq j$

ⓓ $P(R_i) =$ True for $i = 1, 2, 3 \ldots n$

$P(R_i \cup R_j) =$ false for $i \neq j$

where P is predicate that some property over the regions.

Scanned with CamScanner

* Detection of Discontinuities.

① Point
→ Isolated Point

② Line
→ Horizontal
→ Vertical
→ Slanted

③ Edge
→ objelt outline

① Point Detection

- An isolated point is a point whose gray level is significantly different from its background in a homogeneous Area.

→ It run mask through the image.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

mask

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

image.

$$\text{Response } (R) = w_1 z_1 + w_2 z_2 + \ldots\ldots w_9 z_9$$

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

If $|R| \geq T$, a point is detected. where T is threshold non negative value

Sample mask for point detection is

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

Scanned with CamScanner

Line Detection

Four types of mask are used to get the response i.e. R_1, R_2, R_3 & R_4 for directions; horizontal, vertical, $+45°$ & $-45°$ respectively.

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Horizontal vertical $+45°$ $-45°$

$$R_K = \sum_{K=1}^{4} w_K z_K$$

$R_1 \rightarrow LR$
$R_2 - TB$
R_3, R_4

→ If at certain point in an image $|R_i| > |R_j|$ for all $j \neq i$, that point is said to be more likely associated with a line in the direction of mask i.

③ Edge Detection:-

→ An Edge is a set of connected pixels that lies on the boundary between two regions which differ in gray value.
→ Pixel on edge is known as edge points
→ It provides outline of the object.
→ In physical plane, edge corresponds to discontinuities in depth, surface orientation, changes in material properties, light variation, etc.
→ It locates sharp changes in the intensity function
→ Edges are pixels where brightness changes abruptly.

Scanned with CamScanner

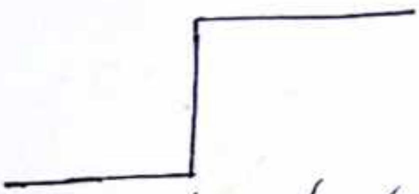
→ An edge can be extracted by computing the derivative of image

→ Magnitude of derivative → indicates strength of the edge

→ Direction of derivative vector indicates edge orientation.

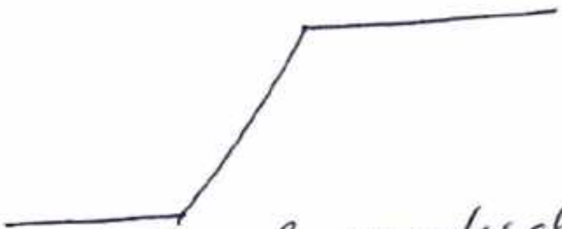
Types of Edges:-

① Step Edges



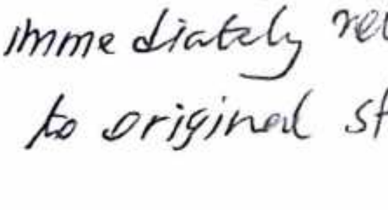
An abrupt changes in intensity

② Ramp Edge



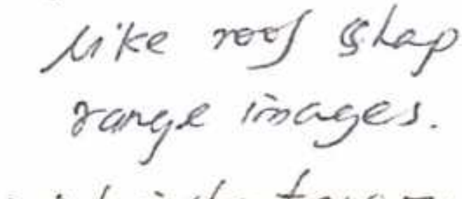
A slow & gradual changes in intensity

③ Spike



Quick change or immediately return to original state

④ Roof Edge.



gradual changes Forming like roof shape / range images.

→ not instantaneous over short distance.

Scanned with CamScanner

Stages in Edge Detection

Start → Input Image → Filtering → Differentiation → Localization → Display → End.

Differentiation

$$\text{Vector } \nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix}$$

$$\text{magnitude } \nabla f = \text{mag}(\nabla f)$$

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

$$\text{Direction of Gradient } (\alpha(x,y)) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

→ The direction of an edge at an orbitary point (x,y) is orthogonal to the direction, $\alpha(x,y)$, of the gradient vector of the point.

→ Direction angle of edge $= \alpha - 90^\circ$

Scanned with CamScanner

First order Edge Detection Operator

- → Local transitions among different image intensities constitutes an edge.
- → Therefore, the objective is to measure intensity gradient
- → Edge detectors can be viewed as gradient calulators

$$\Delta f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \delta f / \delta x \\ \delta f / \delta y \end{bmatrix}$$

$$mag\,(\Delta f) = \left[G_x^2 + G_y^2 \right]^{1/2}$$

$$\approx |G_x| + |G_y|$$

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

Backward Difference

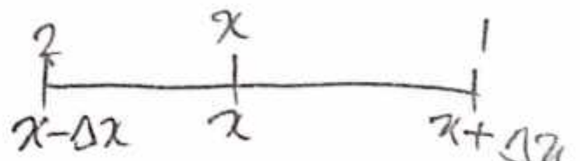
$$[f(x) - f(x-\Delta x)] / \Delta x$$

Forward Difference

$$[f(x+\Delta x) + f(x)] / \Delta x$$

Centeral Distance

$$[f(x+\Delta x) - f(x-\Delta x)] / 2\Delta x$$



Scanned with CamScanner

→ These differences can be obtained by applying following masks, assuming $\Delta x = 1$

⇒ Backward Difference = $f(x) - f(x-1) = [1 \quad -1]$

⇒ Forward Difference = $f(x+1) - f(x) = [-1 \quad +1]$

* Robert operator :- First order operator

→ Robert kernel are derivatives with respective to diagonal element

→ Also known as cross gradient operator.

→ are based on cross-diagonal differences.

Let $f(x,y)$ & $f(x+1,y)$ be neighbouring pixels, then

$$\frac{\delta f}{\delta x} = f(x+1,y) - f(x,y)$$

$$G_x \begin{vmatrix} -1 & 0 \\ 0 & +1 \end{vmatrix} \quad \Big| \quad G_y \begin{bmatrix} 0 & -1 \\ +1 & 0 \end{bmatrix}$$

Robert mask.

Generic Gradient Based Algorithm :-

1. Read the image and smooth it.
2. Convolve the image f with g_x $\hat{f}(x) = f * g_x$
3. Convolve the image f with g_y $\hat{f}(y) = f * g_y$
4. Compute the edge magnitude & edge oritation
5. Compute the edge magnitude with a threshold value *If edge magnitude is higher, assign it as a possible edge point

Scanned with CamScanner

Prewit Operator :-

→ It takes the central difference of the neighbouring pixels, this difference can be represented mathematically as

$$\frac{\delta f}{\delta x} = \left[\frac{f(x+1) - f(x-1)}{2}\right] \text{ as 1D.}$$

$$\left[\frac{f\,x+1,y - f(x+1,y)}{2}\right] \text{ as 2D}$$

central difference is obtained by marks

$[-1, 0, +1]$

$$G_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

$$G_x = \frac{\delta f}{\delta x} = (Z_7 + Z_8 + Z_9) - (Z_1 + Z_2 + Z_3)$$

$$G_y = \frac{\delta f}{\delta y} = (Z_3 + Z_6 + Z_9) - (Z_1 + Z_4 + Z_7)$$

Scanned with CamScanner

Sobel Operator

→ provides both a differentiation and smoothing effect

→ relies on central difference

→ can be viewed as approximation of first Gaussian Derivative

→ Here convolution is both commutative & associative

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

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

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

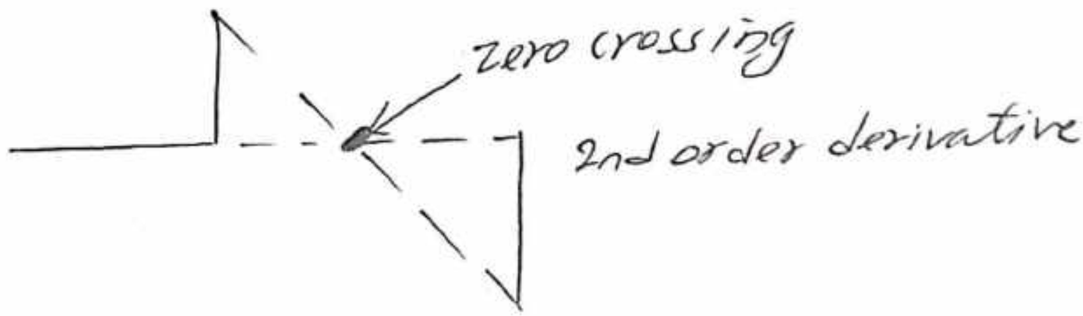
Scanned with CamScanner

Second Order derivative Filter

→ In first derivative, edges are consider to be present when edge magnitude is larger compared to threshold value

→ In case of 2nd derivative, edge is present at that location where 2nd derivative is zero.

→ It is like zero crossing, which can be ~~oby~~ observed as sign change



* Laplacian Operator

→ It is one of the zero crossing algorithm

→ These mask are very sensitive to noise because there is no magnitude checking, even a small ripple looks like edge

→ Therefore image must be filter first and then edge detection process is applied.

→ Advantage is that they are rotationally invariant.

Scanned with CamScanner

Algorithm

1. Generate the mask
2. Apply the mask
3. Detect zero crossing
4. Zero crossing is a situation where pixels in neighbour differ from each other pixel in sign.

Laplacian Operator

→ laplacian of a 2D function $f(x,y)$ is a 2^{nd} order derivative & is defined as

$$\nabla^2 f(x,y) = \frac{\delta^2 f(x,y)}{\delta x^2} + \frac{\delta^2 f(x,y)}{\delta y^2}$$

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

$x-1,y$ $\quad f(x,y)$ $\quad x+1,y$

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

$(x,y+1)$

$f(x,y)$

$(x,y-1)$

Scanned with CamScanner

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

Different Laplacian masks

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

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

3X3 mask,

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

for other mask

$$\nabla^2 f = 8z_5 - [z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9]$$

Scanned with CamScanner

Edge Linking & Boundry Detection

1. Local Processing

→ one of the simplest approach for linking edge point is to analyze the charalteristics of pixel in a small neighbourhood about every pixel point (x,y) that has been declared as edge point

→ All points that are similar according to predefined criteria are linked

2 principle propperties for similarity of edge pixels :-

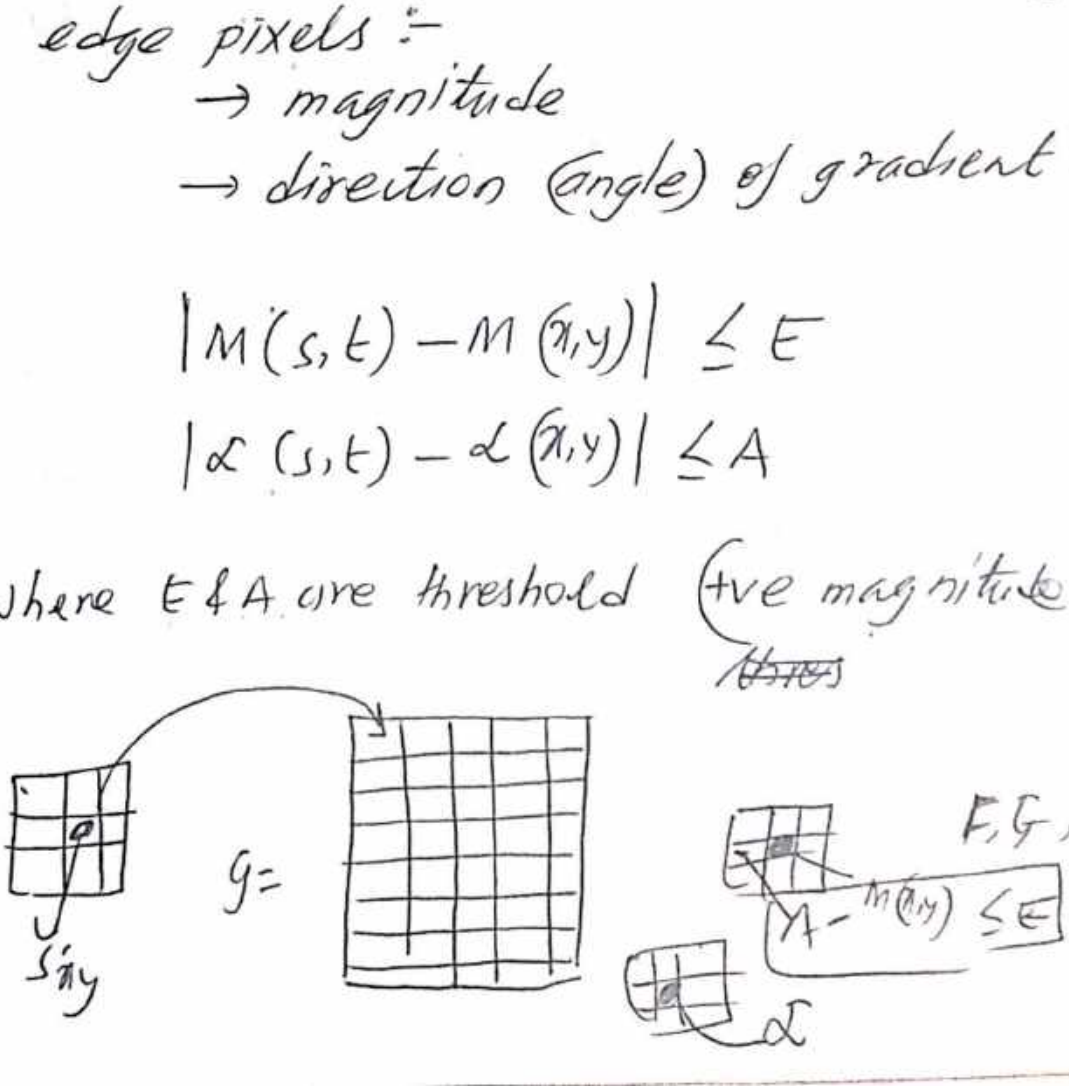
→ magnitude

→ direction (angle) of gradient operator

$$|M(s,t) - M(x,y)| \leq E$$

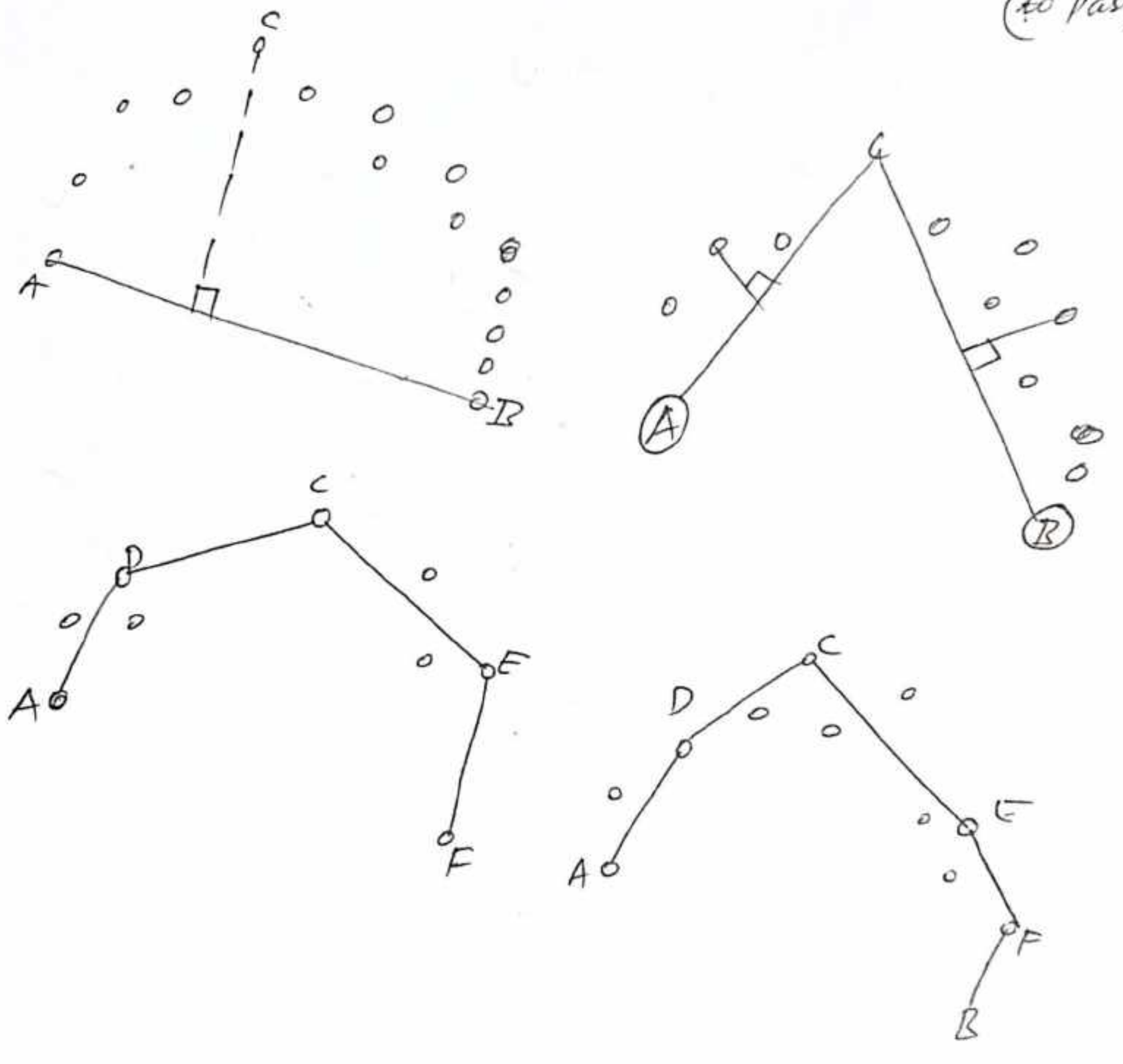
$$|\alpha(s,t) - \alpha(x,y)| \leq A$$

Where E & A are threshold (+ve magnitude & angle)



② Regional Processing :-

(perpendicular) distance
Threshold to check.
(to pass)



Scanned with CamScanner

Hough Transform :-

→ It is a feature extraction method for detecting simple ~~edges~~ shapes such as circle, lines etc in an image

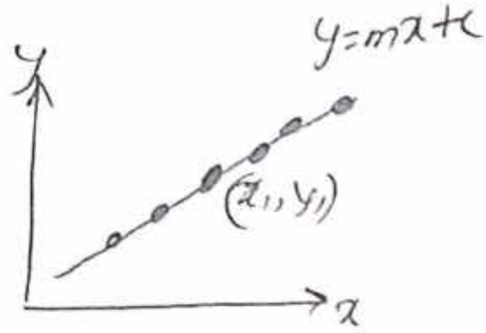
→ takes images created by edge detection operators but must of the time, edge map is disconnected.

→ Therefore Hough Transform is used to connect the disjoint edge points.

→ The eqⁿ of line

$$y = mx + c$$

→ problem is that, infinite lines can be drawn connecting these points



∴ an edge point in x-y plane is transformed to the c-m plane.

Let point (x_i, y_i), then

$$y_i = mx_i + c$$

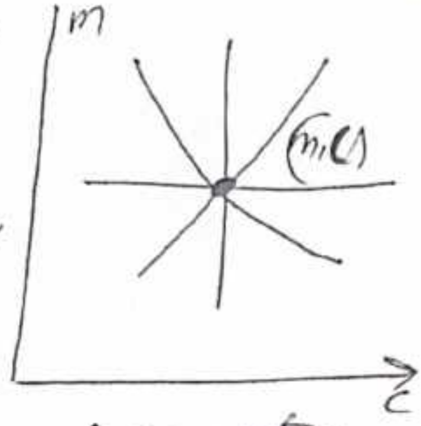
$$\boxed{\text{or } c = -x_i m + y_i}$$

Now, any edge point in x-y plane can be written in c-m plane

- A common intersection point indicates that the edge point are the part of the same line

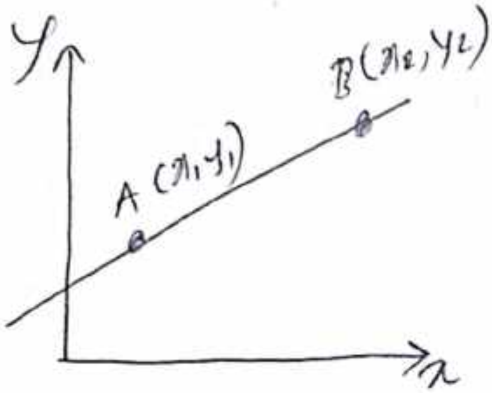
Scanned with CamScanner

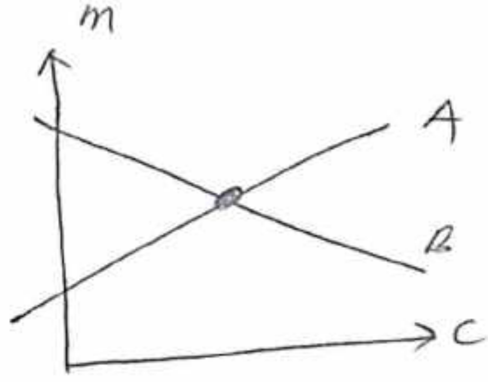
→ Parameter space is also called Hough space



parameter space

→ If A&B are two point connected by a line in spatial domain, They will be intersecting line in Hough space.





Algorithm

1. Load the image
2. Determine the image edge using any edge detector
3. Quantize the parameter space P.
4. Repeat the process for all pixels of image

   If the pixel is an edge pixel, then,

$$c = -x(m) + y$$

$$P(c,m) = P(c,m) + 1$$

5. Show the Hough space
6. Find the local maxima in parameter Space
7. Draw the line using local maxima.

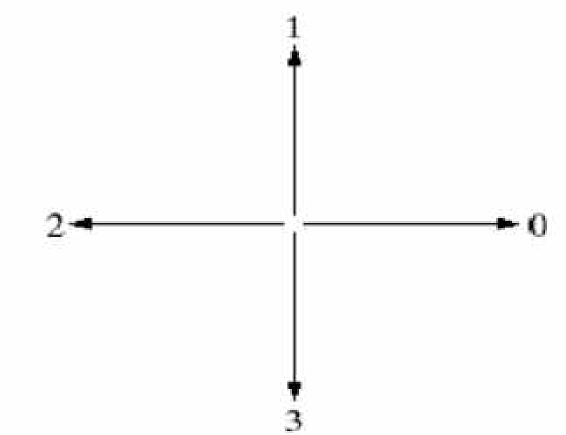
Scanned with CamScanner

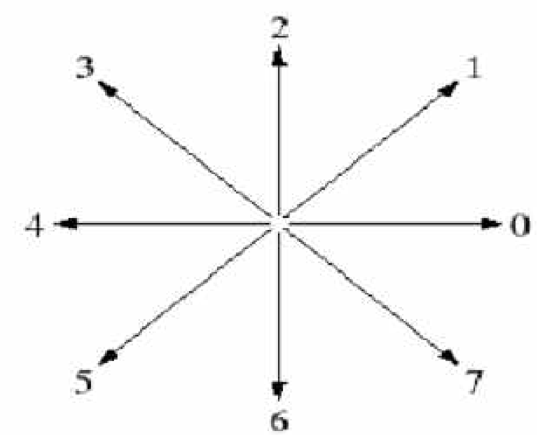
8. Representation and Description

- After an image is segmented into regions, the resulting aggregate of segmented pixels usually is represented and described in a form suitable for further computer processing.
- Representation a region can be done in 2 ways: -
  - Represent the region in terms of its external characteristics (boundary)
  - In term of internal characteristics (region)
- The feature selected as descriptors should be as insensitive as possible to variations is size translation and rotation.

1. **Chain Codes:**

- It is used to represent a boundary by a connected sequence of straight-line segments of specified length and direction.
- Representation is based on 4 – or – 8 connectivity of the segments.
- Direction of each segment is coded by using a number sequence.





4-direction Chain Code

8-direction Chain Code

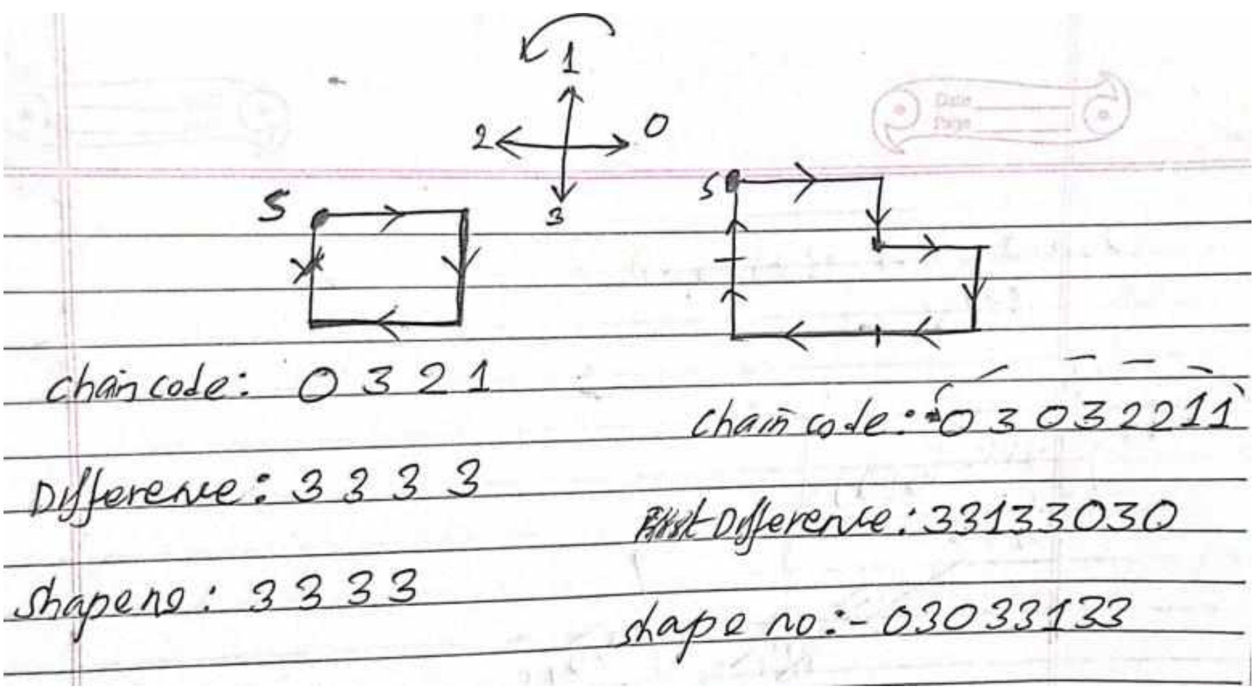
- Boundary of code is formed as a sequence of such directional numbers which often called as **Freeman chain code**.
- Chain code of a boundary depends on the starting point.
- It can be normalized with respect to starting point by treating chain code as a circular sequence of directional number and redefine the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude.
- Normalize the rotation by using first difference of the chain code instead of code itself.

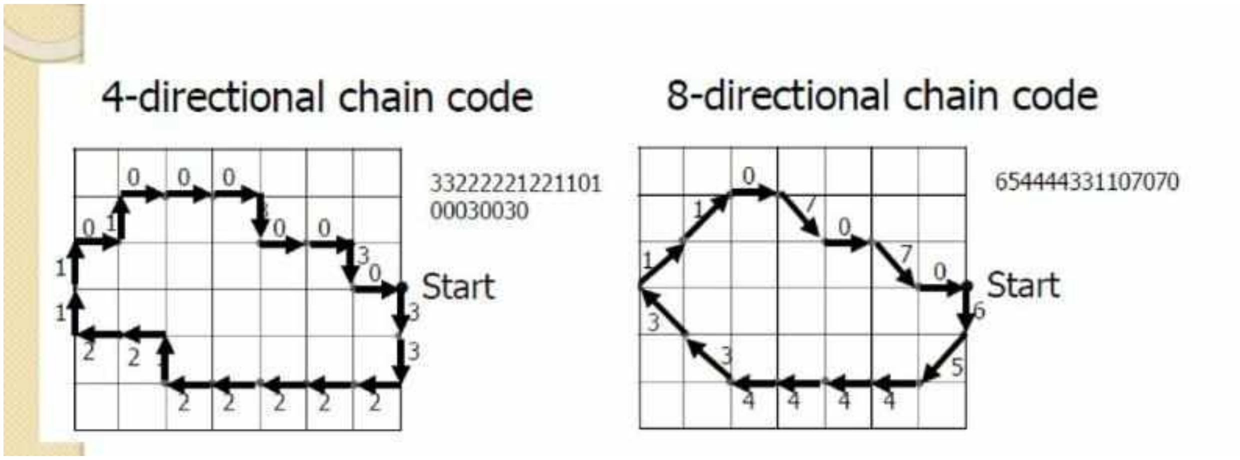
First Difference

- It is obtained by counting the number of direction changes (in counter clockwise direction) that separate two adjacent elements of the code.
  Eg. 4-Direction Chain Code: 1 0 1 0 3 3 2 2
  First Difference: 3 1 3 3 0 3 0
- If we treat the code as a circular sequence to normalize with respect to starting point, the first element of difference is computed by using the transition between last and first component of the chain code.
- Starting point normalized difference: 3 3 1 3 3 0 3 0

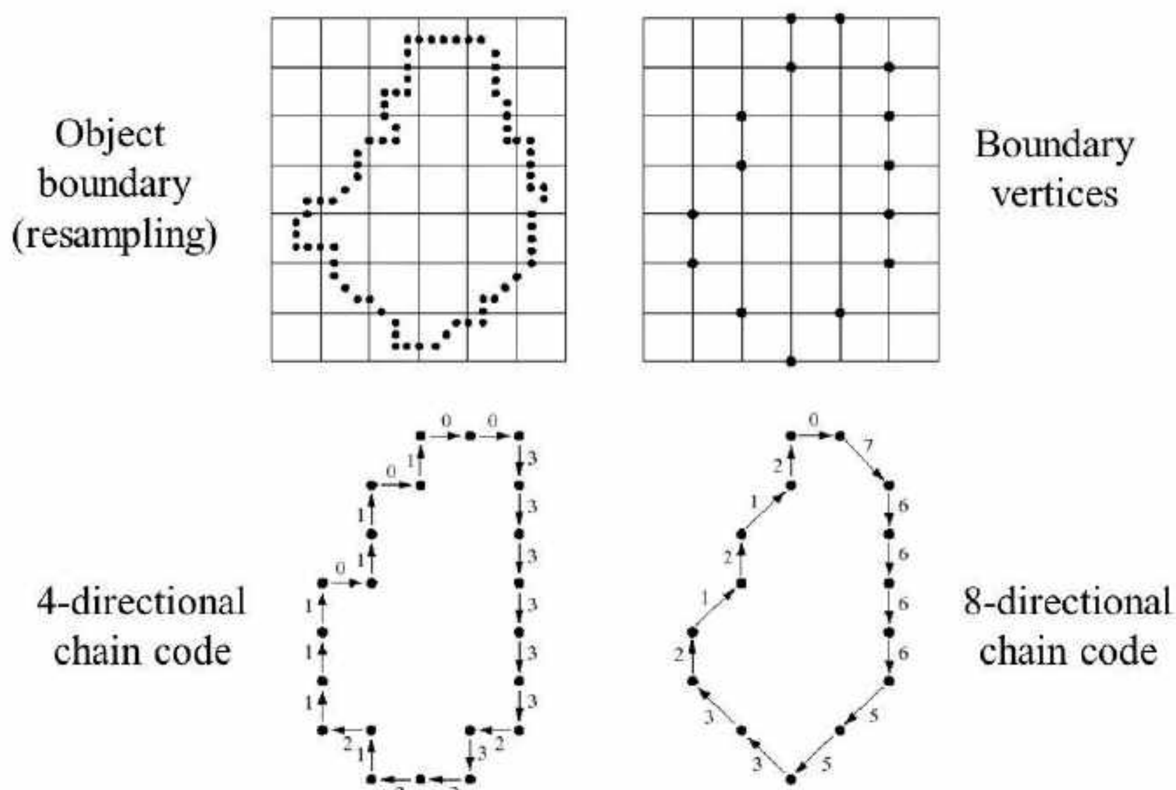
Shape Number:

- The first difference of a Chain code depends on the starting point.
- Shape number is defined as the first difference of smallest magnitude.
- The order n of a shape number is defined as a number of digits in its representation.
- n is even for closed boundary.
- First difference is computed by treating the chain code as a circular sequence.





Examples of Chain Codes



(Images from Rafael C. Gonzalez and Richard E. Wood, Digital Image Processing, 2nd Edition.

(19)

classmate
Date
Page

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

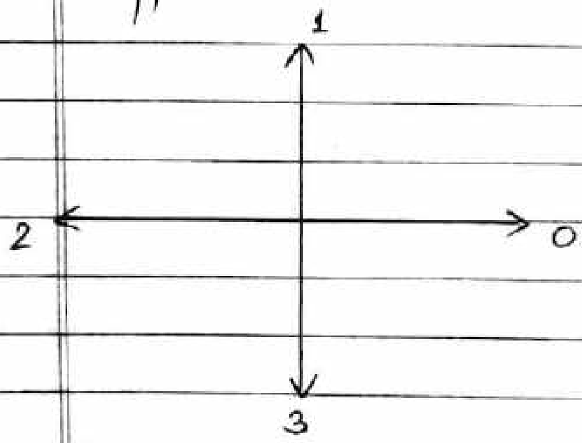
classmate
Date
Page

① Chain Codes :

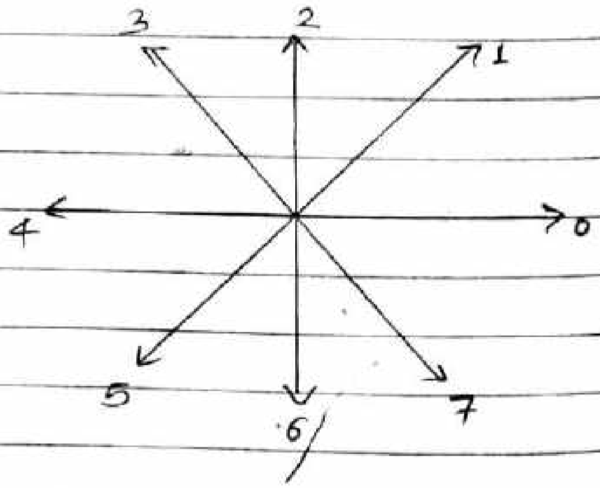
(2015 fall)

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 : (re-university)

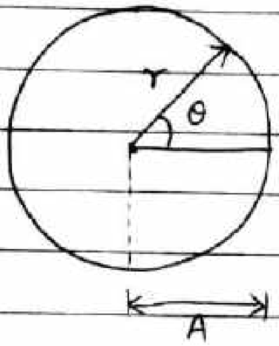
Signature is a 1-Dimensional functional representation of a boundary and may be generated in various ways.

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

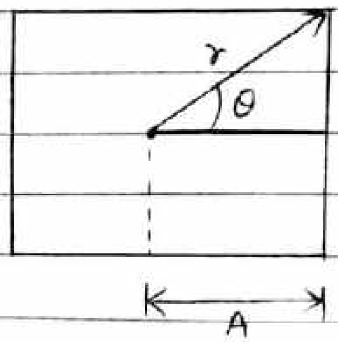


fig (b): square object

The corresponding angle signature can be defined as,

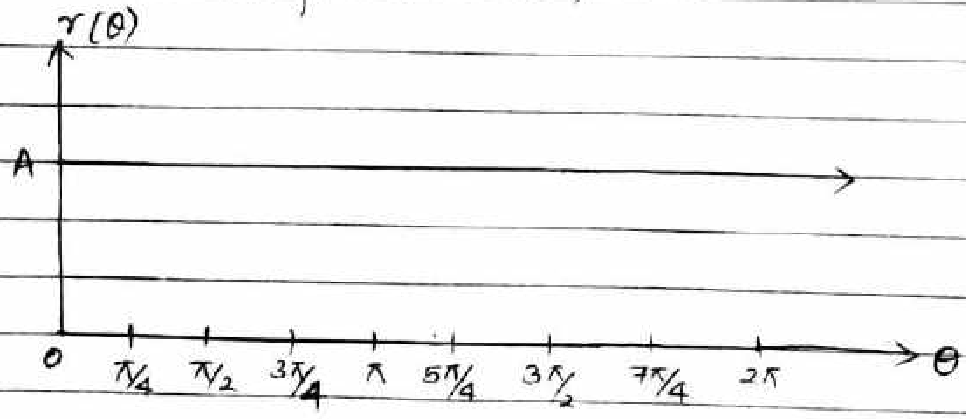


fig (c): Corresponding distance Vs Angle Signatures of circular objects

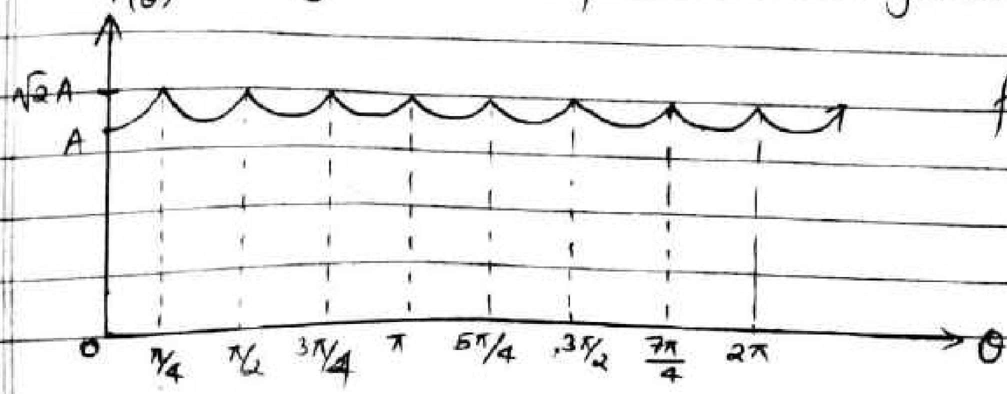


fig (d): Corresponding distance Vs Angle signature of square object

(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:

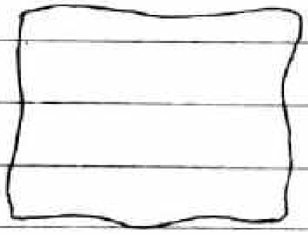


fig (a): Boundaries of an irregulare square

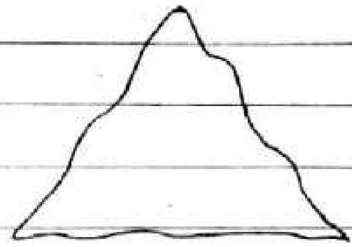


fig (b): Boundaries of irregular triangle

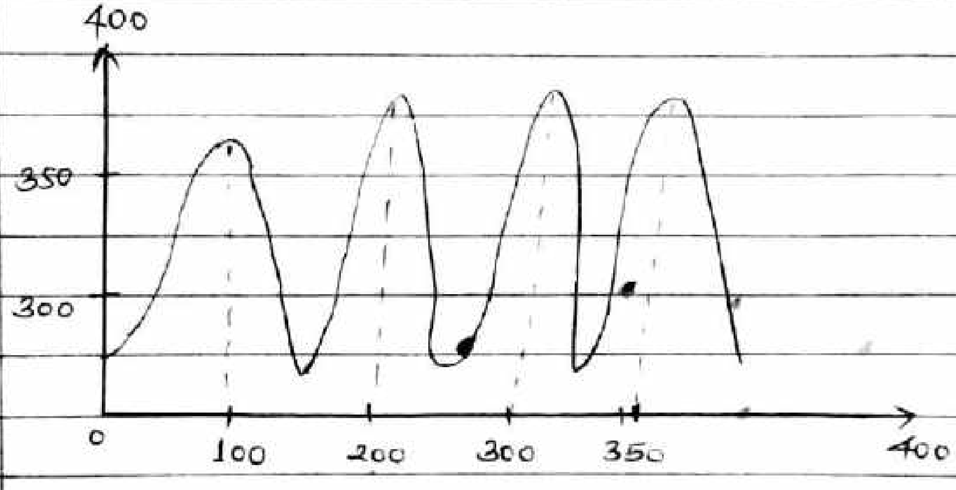


fig (c): corresponding signature of an irregular square

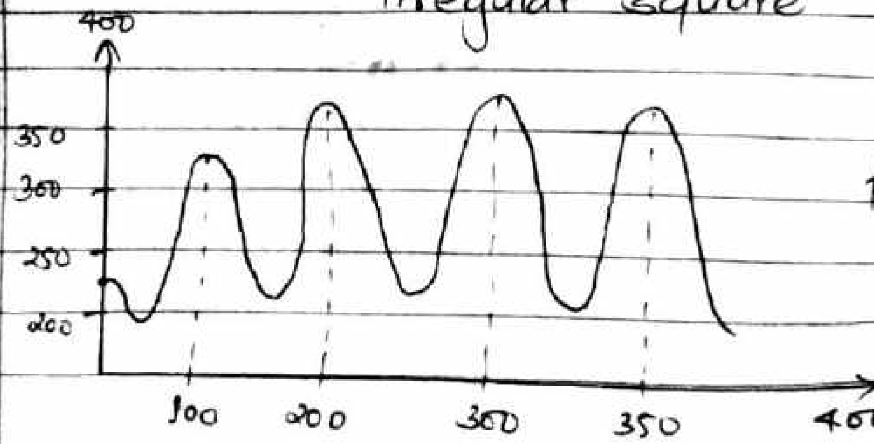


fig (d): corresponding signature of irregular triangle

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.

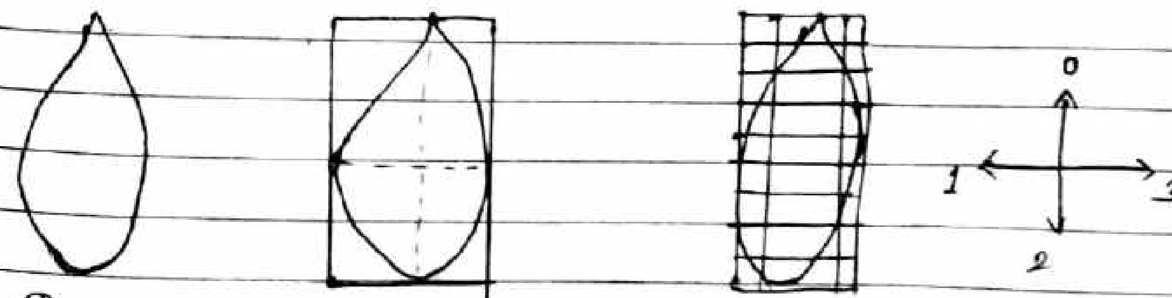
Moreover, 'n' is even for a closed boundary and its value limits the number of possible different shapes.

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

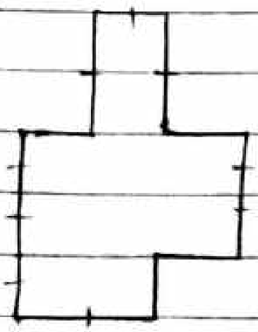
The four directional chain codes can be made insensitive to the starting point by using the integer of minimum magnitude and made rotation that are multiples of 90° by using the first difference of the code.

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

∴ Chain code = 0000 300 3223 2221211
Difference = 3000 31311 311 00310310

Shape no. = 000310330130031303

(fig: steps in the generation of shape number)

(2) Fourier Descriptor:

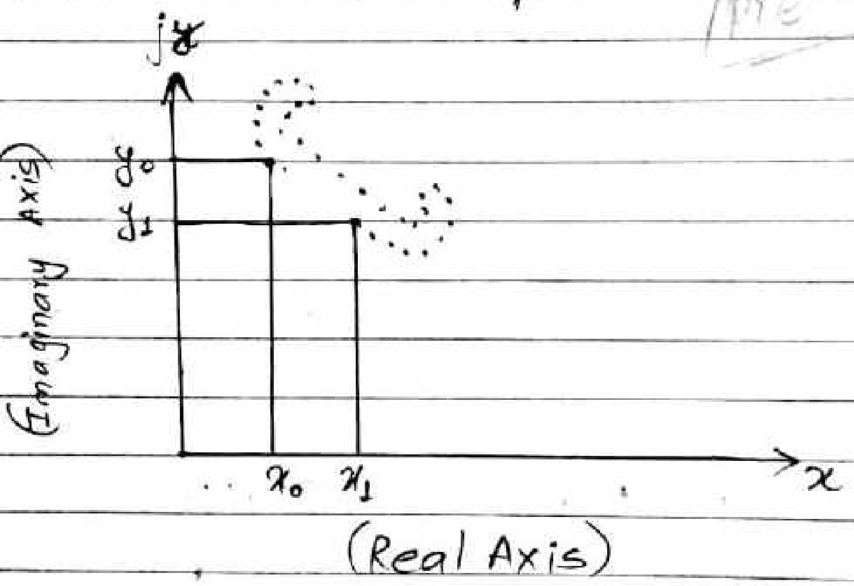


fig: A digital boundary and its representation as a complex sequence. point (x_0, y_0) is selected arbitary or starting point. Point (x_1, y_1) is the next counter clockwise point in a sequence.

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 $s(k)$

classmate
Date
Page

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 classificatio

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

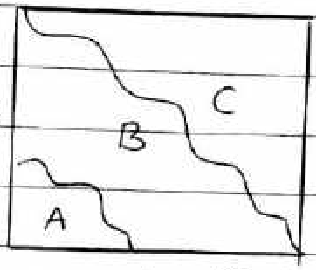
101	105	105	105
101	101	105	105
103	101	101	101
103	103	101	101

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

To justify this classification, we use Bayesian classifier.

ॐ

- one conditional probability
- two unconditional probability

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

(Imp) $P(B/A) = P(A \cap B) / P(A)$

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

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

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.

classmate
Date
Page

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

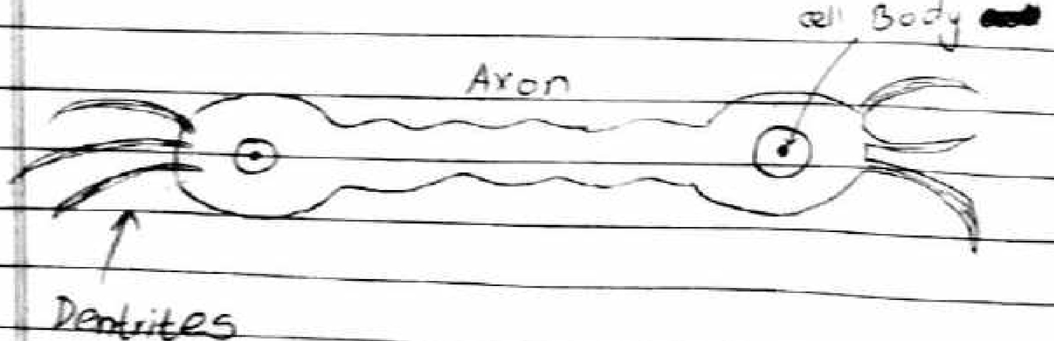
(fig: ~~Recording~~ Recognizing digit using Neighbourhood classifier)

(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)

classmate
Date
Page

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 Netwok 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) ~~Dep~~ Different facial recognition

(v) Manufacturating product design (CAD)

(vi) Telecommunication automated information

(vii) Robotics control system

classmate
Date
Page

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.

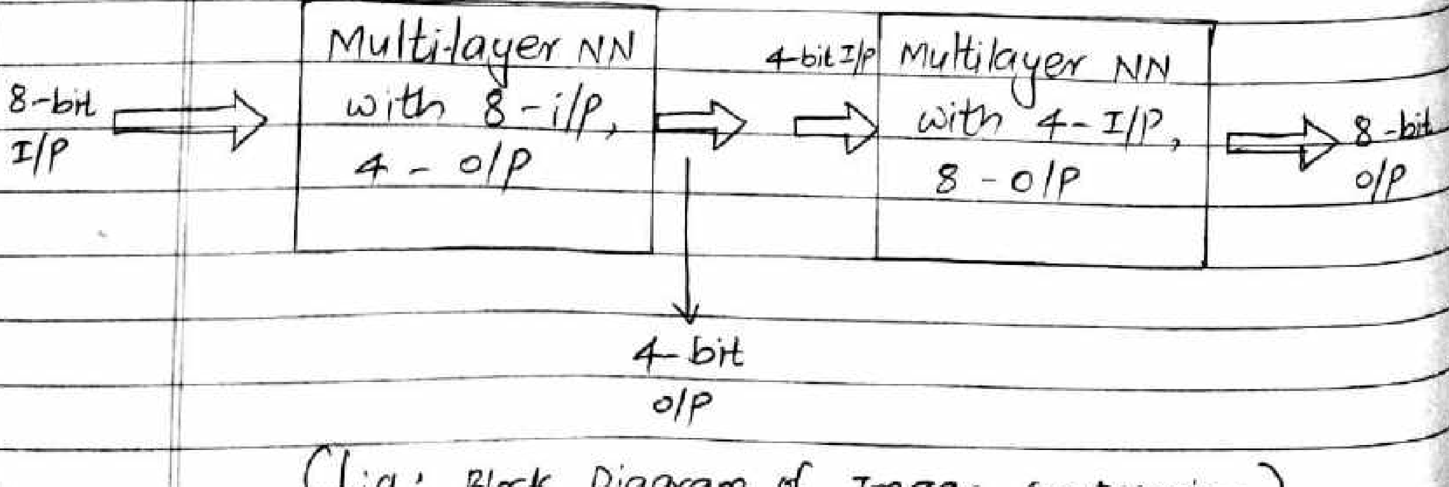
(v) lack of well manpower.

(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)

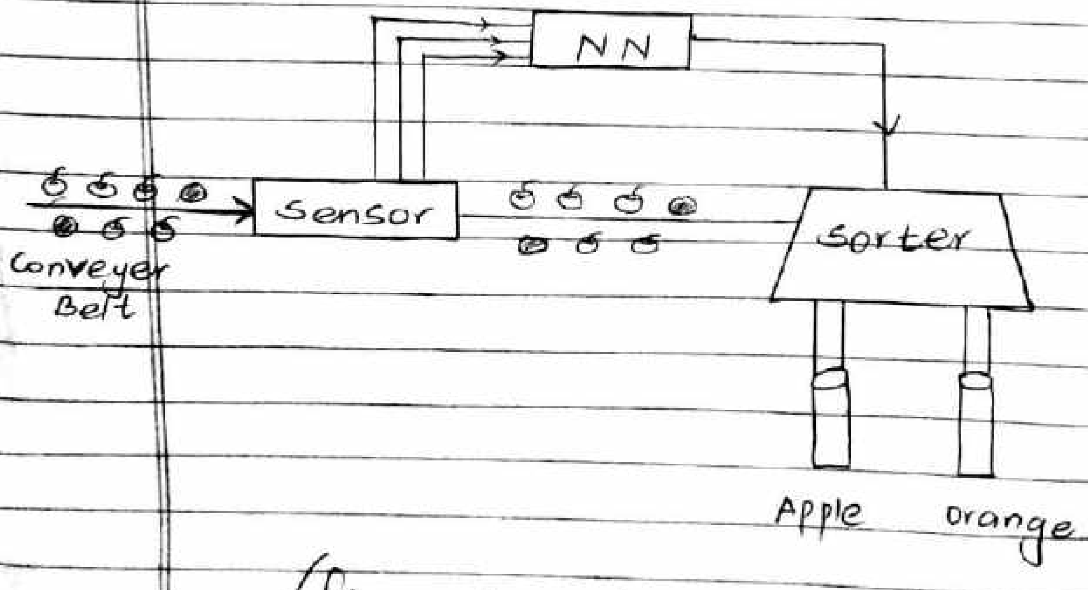
NN model have received much attention in many fields where high compression ratio is required.

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 8×4 multi-layer NN.

Neural network (NN) on the other side then process the compressed image by 4×8 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.

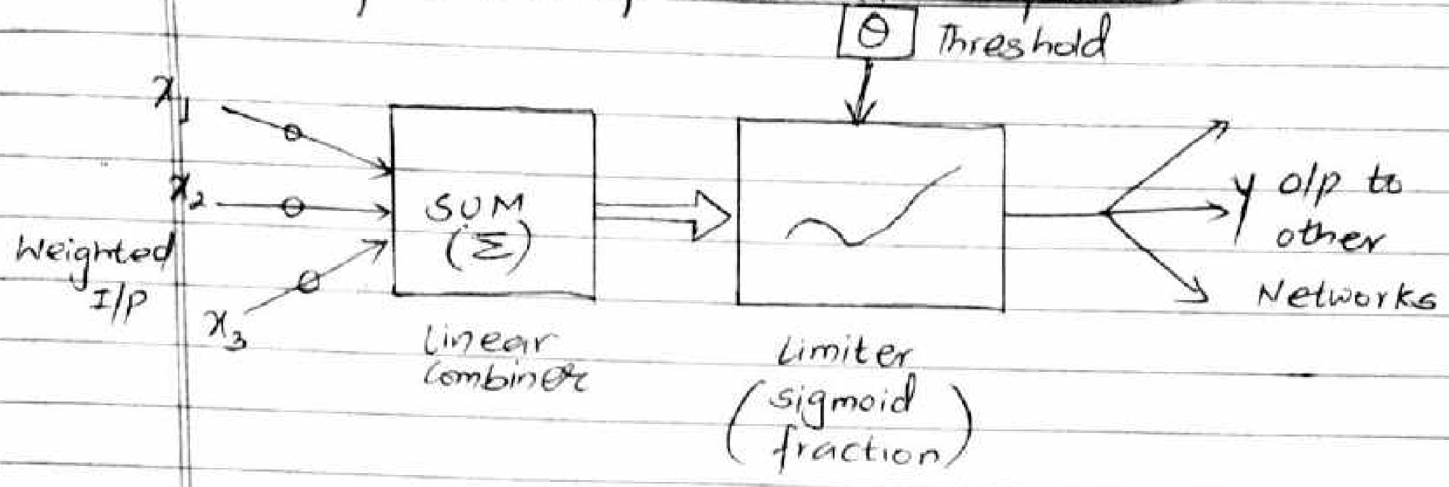
The dealer wants a machine that will sort the fruits according to their type.

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

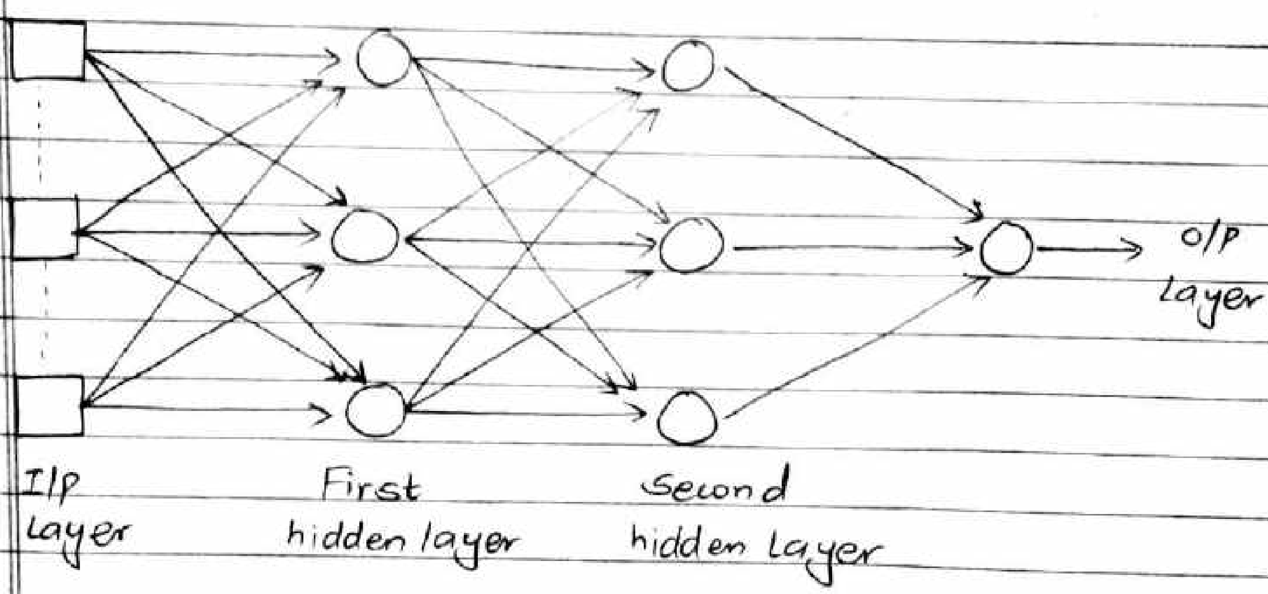
The output will be input to the purpose of NN ~~to~~ decided what kind of fruit is on the conveyer belt. So, fruit can be directed to the correct storer.

(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-linearlly

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.

(P)

(Assignment - I)

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

pramodpatarule@gmail.com

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

(3.83)

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.