Digital Image Processing

<u>Prepared by:</u>
Dr. G. Thirumalaiah
Dept of ECE
Assisitant Professor
Annamacharya University
Rajampet

DIGITAL IMAGE FUNDAMENTALS

Image Sensing and Acquisition :-

The types of images in which we are interested are generated by the combination of an "illumination" source and reflection or absorption of energy from that source by the elements of the 'scene' being imaged. We enclose illumination and scene in quotes to emphasize the fact that they are considerably more general than the familiar situation in which a visible light source illuminates a common everyday 3-D scene. For example, the illumination may originate from a source of electromagnetic energy such as radar, infrared, or x-ray energy. But as noted earlier, it could originate from less traditional sources, such as ultra-sound or even as computer-generated illumination pattern. Similarly, the scene elements could be familiar objects, but they can just as easily be molecules, buried rock formations, or a human brain. we could even image a source, such as acquiring images of the sun. Depending on the nature of the source, illumination energy is reflected from, or transmitted through objects. An example in the first category is light reflected from a planar surface. An example in the second category is when x-rays pass through a patient's body for the purpose of generating a diagnostic x-ray film. In some applications, the reflected or transmitted energy is focused onto a photo convertor which converts the energy into visible light. Electron microscopy and some applications of gamma imaging using this approach. The idea is simple: Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected. The output voltage waveform is the response of the sensors and a digital quantity is obtained from each sensor by digitalizing its response. In this section, we look at the principal modalities for image sensing and generation.

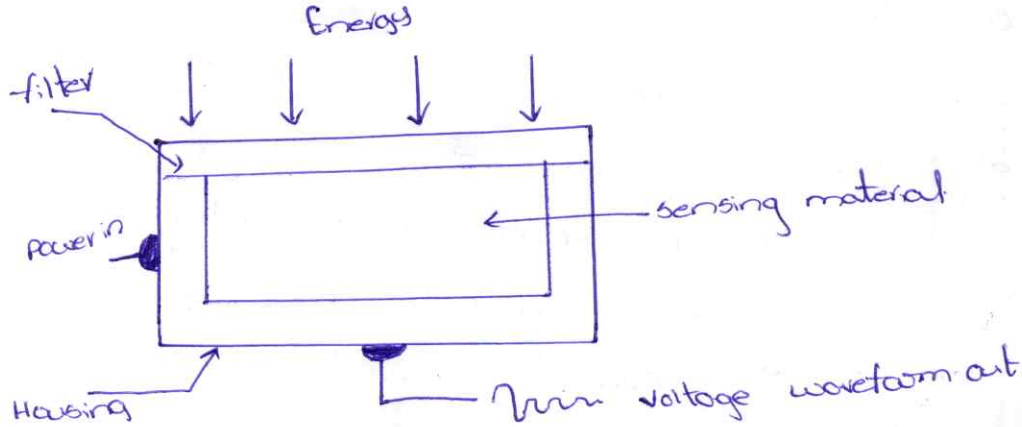


fig: Single Image Sensor

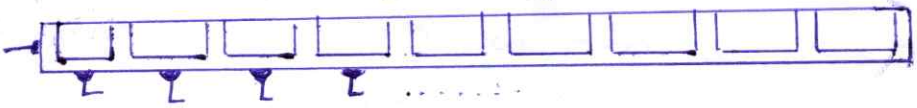


fig:- Line Sensor

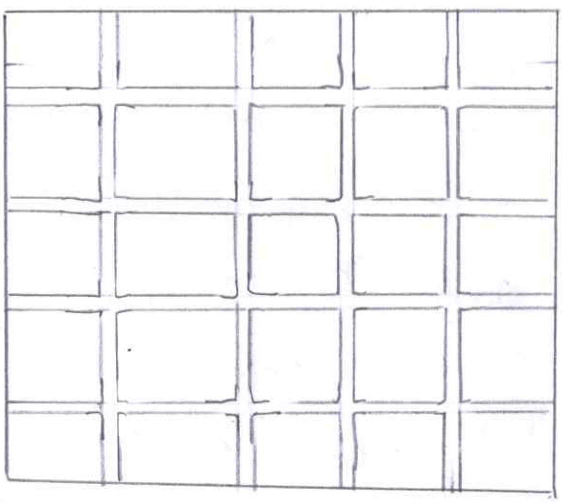
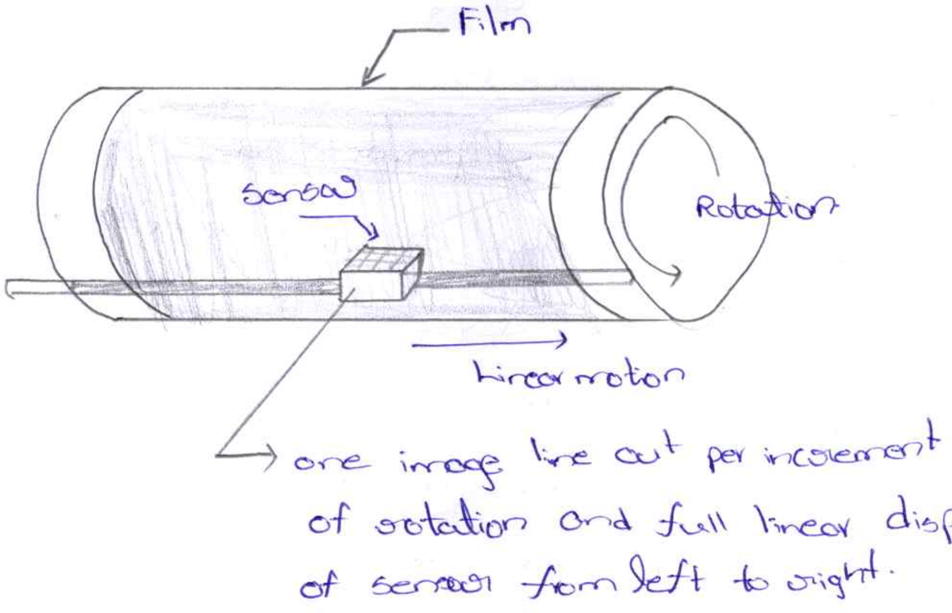


fig:- Array Sensor.

Image Acquisition using a Single sensor:

The components of a single sensor. Perhaps the most ~~similar~~ familiar sensor of this type is the photodiode, which is constructed of silicon materials and whose output voltage waveform is proportional to light. The use of a filter in front of a sensor improves selectively. For example, a green (pass) filter in front of a light sensor favors light in the green band of the color spectrum. As a consequence, the sensor output will be stronger for green light than for other components in the visible spectrum.



In order to generate a 2-D image using a single sensor, there has to be relative displacements in both the x and y directions between the sensor and the area to be imaged. Figure shows an arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Since mechanical motion is can be controlled with high precesion, this method is an inexpensive way to obtain high-resolution images. Other similar mechanical arrangements use a flatbed, with the sensor moving in two linear directions. These types of mechanical digitilizers sometimes are referred to as micro densito-meters.

Image Acquisition using a Sensor strips :-

A geometry that is used much more frequently than single sensors consists of an in-line arganement of sensors in the form of a sensortrip, shows. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction. This is the type of arrangement used in most flat bed scanners. Sensing devices with 4000 or more in-line sensors are possible. In-line sensors are used routinely in airborne imaging applications, in which the imaging system is mounted on an aircraft that flies at a constant altitude and speed over the geographical area to be imaged. One dimensional imaging sensor strips that respond to various band of the electromagnetic spectrum are mounted perpendicular to the direction of flight. The imaging strip gives one line of an image at a time, and the motion of the strip completes the other dimension of a 2-D image. Lenses or other focusing schemes are used to project area to be scanned onto the sensors. Sensor strips mounted in a ring configuration are used to ~~p~~ in medical and industrial imaging to obtain cross-sectional ("slice") images of 3-D objects.

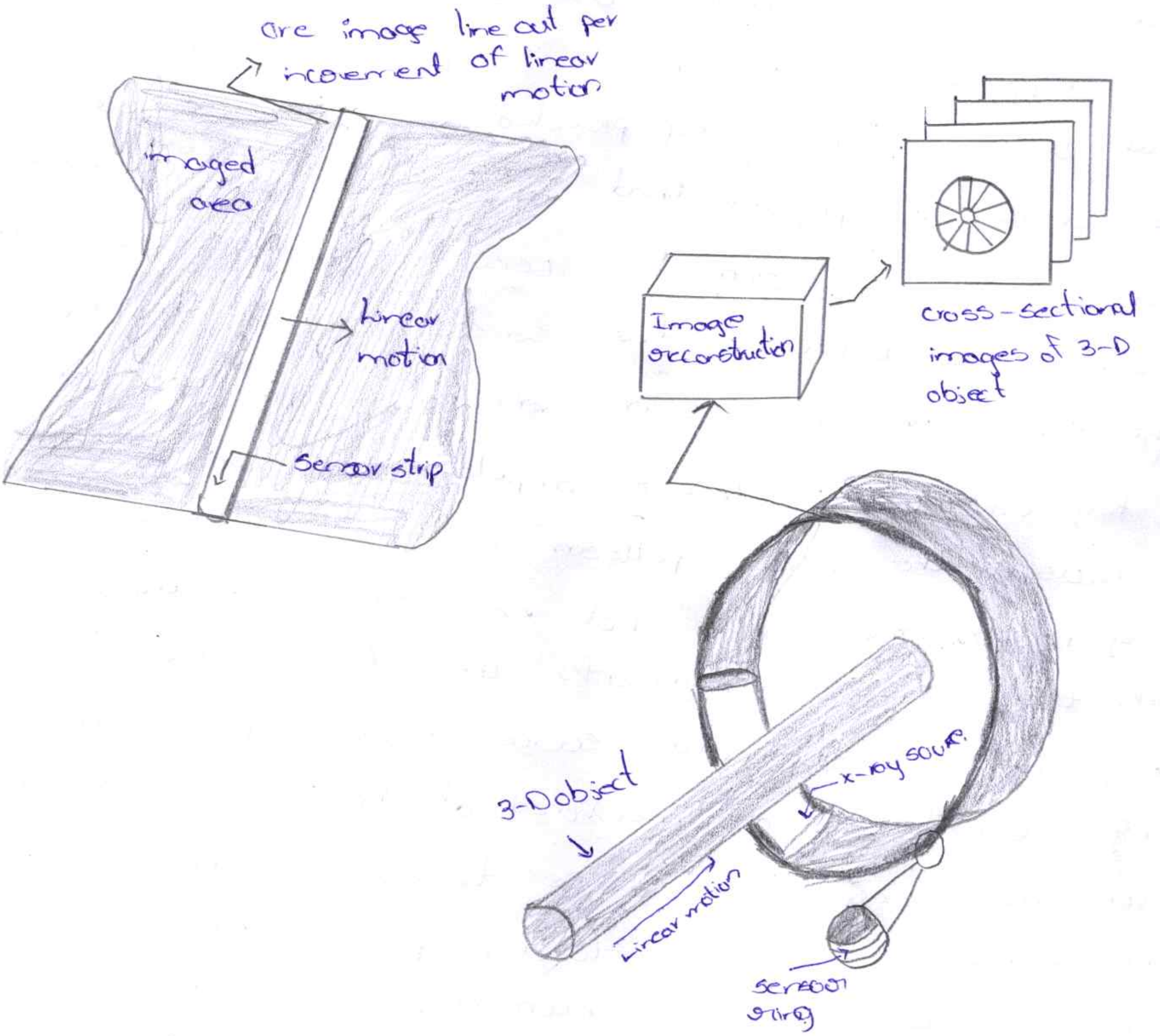
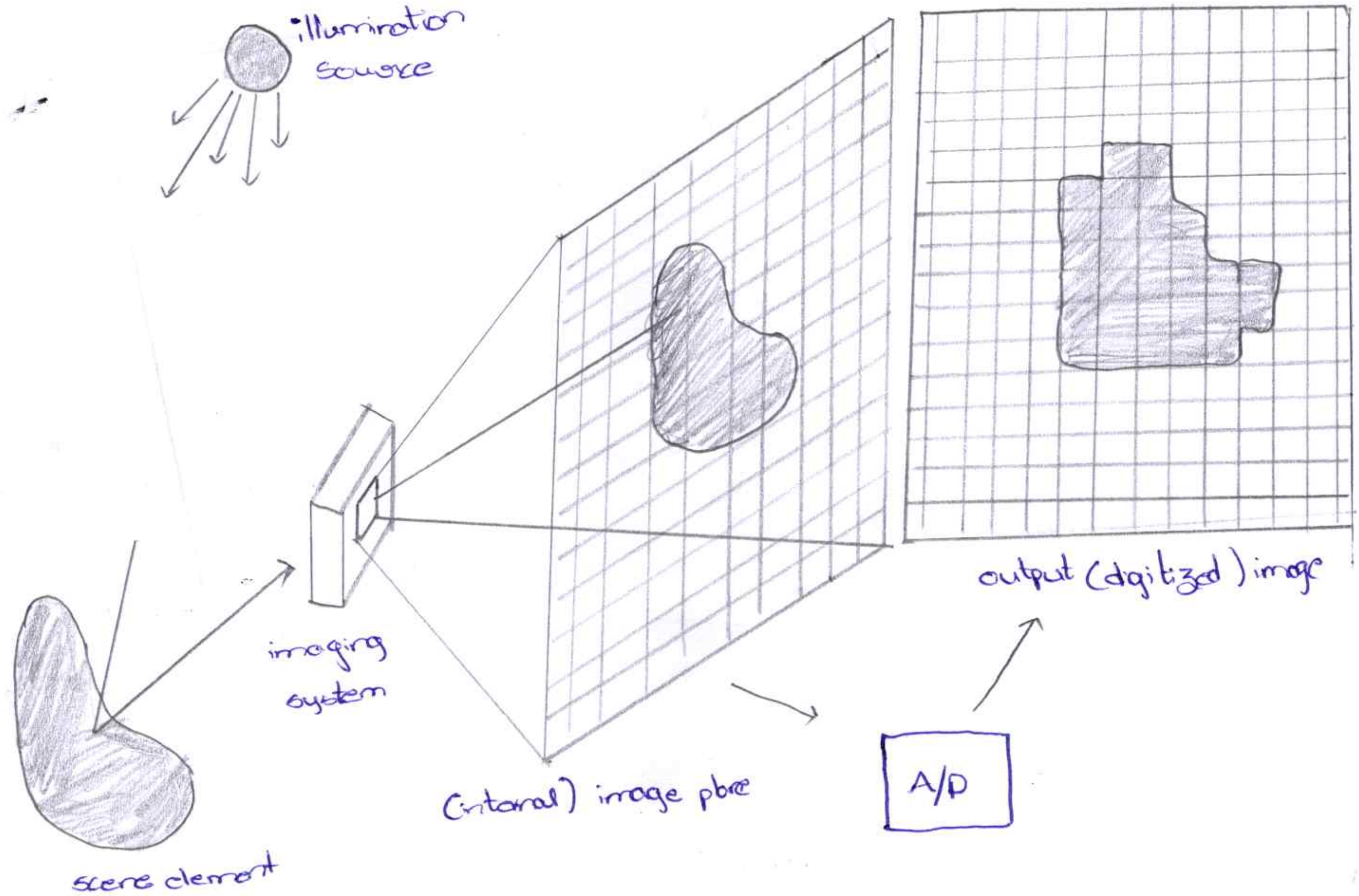


fig:- Image Acquisition using Linear strip and Circular strips

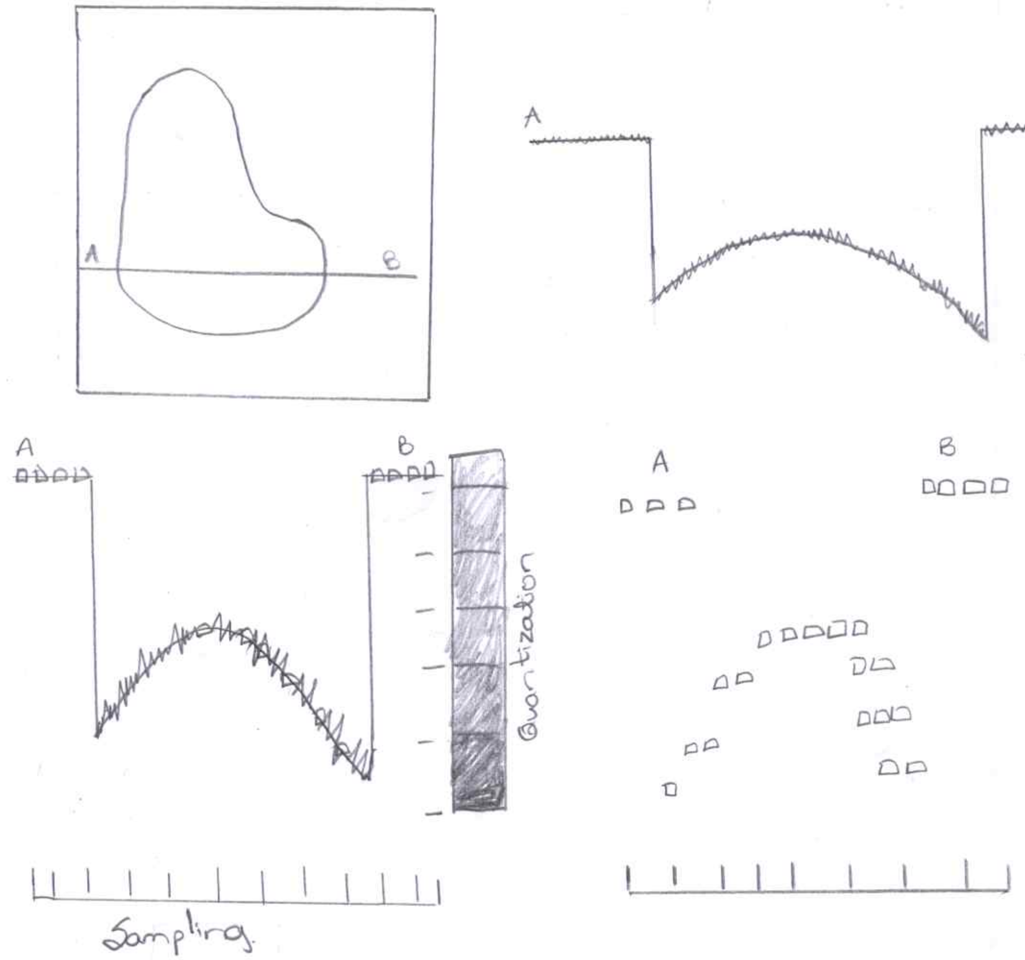
Image Acquisition using a Sensor Arrays:-

The individual sensors arranged in the form of a 2-D array. Numerous electromagnetic and some ultrasonic sensing devices frequently are arranged in an array format. This is also the predominant arrangement found in digital cameras. A typical sensor for these cameras is a CCD array, which can be manufactured with a broad range of sensing properties and can be packaged in rugged arrays of elements or more. CCD sensors are used widely in digital cameras and other light sensing instruments. The response of each sensor is proportional to the integral of the light energy projected onto the surface of the sensor, a properly that is used in astronomical and other applications requiring low noise images. Noise reduction is acheived by letting ~~is~~ the sensor integrate the input light signal over minutes or even hours. The two dimensional, its key advantage is that a complete image can be obtained by focusing the energy pattern onto the surface of the array. Motion obviously is not necessary, as is the case with the sensor arrangements this figure shows the energy from an illumination source being reflected from a scene element, but, as mentioned at the beginning of this section, the energy also could be transmitted through the scene elements. The first function performed by the imaging system is to collect the incoming energy and focus it onto an image plane. If the illumination is light, the front end of the imaging system is a lens, which projects the viewed ~~one~~ scene onto the lens focal plane. The sensor array, which is coincident with the focal plane, produce outputs proportional to the integral of the light received at each sensor. Digital and analog circuitory sweeps these outputs and converts them to a video signal, which is then digitilized by another section of the imaging system

Image Sampling and Quantization :-



To create a digital image, we need to convert the continous sensed data into digital form. This involves two processes: Sampling and quantization. A continous image, f(x,y), that we want to convert into digital form. An image may be continous with respect to the x and y cordinates, and also in amplitude. To convert it into digital form, we have to sample the function in both coordinates and in amplitude. Digitilizing the coordinate values is called Sampling. Digitalizing the amplitude values is called quantization.



Digital Image Representation:

Digital image is a finite collection of discrete samples of any observable object. The pixels represent a two-or higher dimensional 'view' of the object, each pixel having its own discrete value in a finite range. The pixel values may represent the amount of visible light, infra red light, absorption of x-rays, electrons, or any other measureable value such as ultrasound wave impulses. The image does not need to have any visual sense; it is sufficient that the samples from a 2D spatial structure that may be illustrated as an image. The images may be obtained by a digital camera, scanner, electron microscope, ultrasound stethoscope, or any other optical non optical sensor.

Examples of digital image are :

- Digital photographs
- Satellite images
- radiological images
- binary images, fax images, engineering drawings.

Computer graphics, CAD drawings, and vector graphics in general are not considered in this course even though their reproduction is a possible source of an image. In fact, one goal of intermediate level image processing may be to reconstruct a model (eg. vector representation) for a given digital image.

Some basic Relationship between pixels :-

We consider several important relationships between pixels in a digital image.

Neighbors of a Pixel:-

- A pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by: (x+1, y), (x-1, y) (x, y+1), (x, y-1)

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

This set of pixels, called the 4-neighbors or p, is denoted by $N_4(p)$. Each pixel is one unit distance from (x, y) and some of the neighbors of p lie outside the digital image if (x, y) is on the border of the image. The four diagnd neighbors of p have coordinates and are denoted by $N_D(p)$.

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

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

These points, together with the 4-neighbors, are called the 8-neighbors of P, denoted by $N_8(P)$.

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

As before, some of the points in $N_D(P)$ and $N_8(P)$ fall outside the image if $f(x,y)$ is on the border of the image.

Adjacency and Connectivity :-

Let V be the set of gray-level values used to define adjacency, in a binary image, $V = \{1\}$. In a gray-scale image, the idea is the same, but V typically contains more elements.

For example, $V = \{180, 181, 182, \ldots 200\}$.

If the possible intensity values 0-255, V set can be any subset of these 256 values.

If we are reference to adjacency of pixel with value.

Three types of adjacency

- 4-adjacency - two pixel P and Q with value from V are 4-adjacency if A is in the set $N_4(P)$.
- 8-adjacency - two pixel p and Q with value from V are 8-adjacency if A is in the set $N_8(P)$.
- M-adjacency - two pixel p and Q with value from V are m-adjacency if (i) Q is in $N_4(P)$ (or) (ii) Q is in $N_D(Q)$ and the

set $N_4(P) \cap N_4(q)$ has no pixel whose values are from V.

- Mixed adjacency is a modification of 8-adjacency. It is introduced to eliminate the ambiguities that often arise when 8-adjacency is used.

- For example (1) :-

$$\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \qquad \begin{array}{ccc} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \qquad \begin{array}{ccc} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}$$

(a) (b) (c)

fig(a): Arrangement of pixels
fig(b):- pixels that are 8 adjacent to the centre pixel
fig(c):- m-adjacency.

Types of adjacency:-

- In this example, we can note that to connect between two pixels:
  - in 8-adjacency way, you can find multiple paths between two pixels.
  - while, in m-adjacency, you can find only one path between two pixels.

- So, m-adjacency has eliminated the multiple path connection that has been generated by the 8-adjacency.

- Two subsets S_1 and S_2 are adjacent, if some pixel in S_1 is adjacent to some pixel in S_2.

  Adjacent means, either 4-, 8- or m-adjacency.

A Digital Path :-

- A digital path from pixel p with coordinate (x, y) to pixel q with coordinate (s, t) is a sequence of distinct pixels with

with coordinates $(x_0, y_0), (x_1, y_1), \dots (x_n, y_n)$ where $(x_0, y_0) = (x, y)$ and $(x_n, y_n) = (s, t)$ and pixels (x_i, y_i) and ~~for~~ (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$

- n is the length of the path.
- If $(x_0, y_0) = (x_n, y_n)$, the path is closed.

We can specify 4-, 8- or m-paths depending on the type of adjacency specified.

* Return to previous example ①:
  i) In figure (b) the paths between the top right and bottom right pixels are 8-paths. And the path between the same 2 pixels in figure (c) ~~in~~ is m-path

Connectivity:-

- Let S represent a subset of pixels in an image, two pixels p and q are said to be connected in S if there exists a path between them consisting entirely of pixels in S.
- For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S. If it only has one connected component, then set S is called a connected set.

Region and Boundary:-

- Region: Let R be a subset of pixels in an image, we call R is a connected set.
- Boundary: The boundary of a region R is the set of pixels in the region that have one or more neighbors that are not in R.

If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns in the image. This extra defination is required because an image has no neighbors beyond its borders. Normally, when we refer to a region, we are referring to subset of an image, and any pixels in the boundary of the region that happen to coincide with the border of the image are included implicitly as part of the region boundary.

Distance Measures :-

For pixel p, q and z with coordinate (x, y), (s, t) and (v, w) respectively D is a distance function or metric

if $D[p,q] \geq 0$ $\{D[p,q] = 0$ iff $p = q\}$

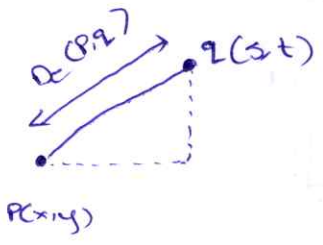
$D[p.q] = D[p.q]$ and

$D[p.q] \geq 0 \{D[p.q] + D[q, z)$

• The Euclidean Distance between p and q is defined as:

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

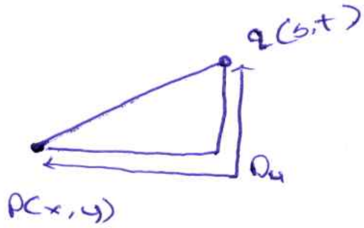
pixels having a distance less than or equal to some value r from (x, y) are the points contained in a disk of radius 'r' centred at (x, y)



• The D_4 distance between p and q is defined as:

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

pixels having a D_4 distance from (x,y), less than or equal to some value r from a Diamond centred at (x,y).



Example :

The pixels with distance $D_4 \le 2$ from (x,y) from the following contours of constant distance.

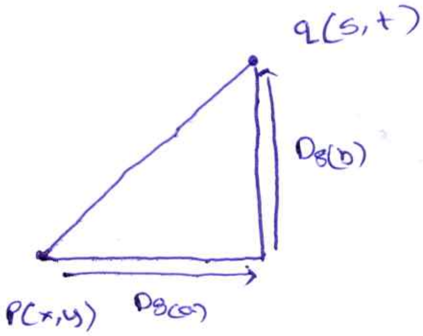
The pixels with $D_4 = 1$ are the 4-neighbors of (x,y)

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

- The D_8 distance between p and q is defined as:

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

pixels having a D_8 distance from (x,y) less than or equal to some value r from a square centred at (x,y).



$$D_8 = \max(D_{8(a)}, D_{8(b)})$$

Example:-

D_8 distance ≤ 2 from (x,y) form the following contours of constant distance.

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

- D_m distance:-

  * It is defined as the shortest m-path between the points.
  * In this case, the distance between two pixels will depend on the values of the pixels along the path. as well as the values of their neighbors.

- Example:-

  Consider the following arrangement of pixels and assume that P_1, P_2, and P_4 have value 1 and that P_1, and P_3 can have a value of 0 or 1 suppose that we consider the adjacency of pixels values 1 (i.e $V=\{1\}$)

	P_3	P_4
P_1	P_2	
P		

Now, to compute the D_m between points p and P_4

Here we have 4 cases:

Case 1:- If $P_1=0$ and $P_3=0$

The length of the shortest m-path is 2 (P, P_2, P_4)

	0	1
0	1	
1		

Case 2:- If $P_1 = 1$ and $P_3 = 0$

Now, P_1 and P will no longer be adjacent.

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

	0	1
1	1	
1		

Case-3:- If $P_1 = 0$ and $P_3 = 1$

The same applies here, and the shortest m-path will be 3 (P, P_2, P_3, P_4)

	1	1
0	1	
1		

Case-4:- If $P_1 = 1$ and $P_3 = 1$

The length of the shortest m-path will be 4 (P, P_1, P_2, P_3, P_4)

	1	1
1	1	
1		

An Introduction to Mathematical tools used in Digital Image Processing :-

To process the digital images, the mathematical tools are required. Those mathematical tools are

1. Array versus matrix operation.
2. Linear versus non linear.
3. Arithmatic operations
4. Set and logical operations
5. spatial operation
6. Vector and matrix operation
7. probabilistic methods.

1. Array Versus Matrix operation :-

The array operation involves two (or) more images whose operation is performed pixel by pixel.

Consider 2x2 images, the array product is written as

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$\begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

matrix product is written as

$$\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

2. Linear versus non-linear:-

One of the most important classification of image processing method is whether it is the linear (or) non-linear.

Consider a general operator 'H' generates an output image $g(x,y)$ with input image $f(x,y)$ is

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

if 'H' is said to be linear, it is written as

$$H[a_i f_i(x,y) + a_j f_j(x,y)] = a_i H[f_i(x,y)] + a_j H[f_j(x,y)]$$

Here a_i and a_j are arbitrary constants.

The above condition is satisfied, then H is said to be linear, otherwise non-linear summation operation is always a linear operation.

maximum operation is always a non-linear operation.

Ex:- Consider f_1 and f_2 images whose intensity values are.

$$f_1 = \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}$$

$$\Sigma[a_i f_i(x,y) + a_2 f_2(x,y)] = a_i \Sigma[f_i(x,y)] + a_2 \Sigma[f_2(x,y)]$$

Here $a_1 = 1, \; a_2 = -1$

$$\Sigma\left[1\begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1)\begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}\right] = \Sigma\left[1\begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} - 1\begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}\right]$$

$$= \Sigma\left[\begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} -6 & -5 \\ -4 & -7 \end{bmatrix}\right]$$

$$= -15$$

$a_1 \Sigma [f_1(x,y)] + a_2 \Sigma [f_2(x,y)]$

$$= 1\Sigma \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} - 1\Sigma \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}$$

$$= -15$$

$\therefore$ summation operation is always linear.

Since L.H.S = R.H.S

Ex:- Consider the operator as maximum operator.

$$\max [a_1 f_1(x,y) + a_2 f_2(x,y)] = a_1 \max [f_1(x,y)] + a_2 \max [f_2(x,y)]$$

$$\max \left\{ 1 \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1) \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix} \right\}$$

$$\max \left\{ \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} -6 & -5 \\ -4 & -7 \end{bmatrix} \right\}$$

$$\max \begin{bmatrix} -6 & -3 \\ -2 & -4 \end{bmatrix} = -2$$

$\Rightarrow a_1 \max [f_1(x,y)] + a_2 \max [f_2(x,y)]$

$\Rightarrow 1 \max \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1) \max \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}$

$\Rightarrow 1 \times 3 - 1 \times 7$

$\Rightarrow 3 - 7$

$\Rightarrow -4$

$\therefore$ maximum operator is always non-linear

since L.H.S $\neq$ R.H.S.

The system is said to be linear if it satisfies superposition and homogenity

3. Arithmatic Operations :-

The arithmatic operations are performed on two or more images (or) it may perform on the corresponding pixel pair in an image.

The operations involve in Arithmatic operations are addition, substraction, multiplication and division.

$$s(x,y) = f(x,y) + g(x,y) \quad (\text{sum})$$
$$d(x,y) = f(x,y) - g(x,y) \quad (\text{difference})$$
$$p(x,y) = f(x,y) * g(x,y) \quad (\text{Product})$$
$$v(x,y) = f(x,y) \div g(x,y) \quad (\text{division})$$

The addition operation is performed on image averaging which is to reduce the noise.

Assume the $g(x,y)$ be corrupted image with noise written as

$$\boxed{g(x,y) = f(x,y) + \eta(x,y)}$$

To remove the noise component add the corrupted images in the original image

ie written as $\bar{g}(x,y) = \frac{1}{K} \sum_{i=0}^{K} g_i(x,y)$

where $g_i(x,y)$ are corrupted images.

The mean (or) expected value of $\bar{g}(x,y)$ is $E[\bar{g}(x,y)] = f(x,y)$

variance $\boxed{\sigma^2 \bar{g}(x,y) = \frac{1}{K} \sigma^2 \eta(x,y)}$

Image substraction is used in enhancement which enhances the difference between the two images.

Image multiplication and division is used for shading correction.

4. Set and logical operations :-

Set operations :-

Let A be the set composed of ordered elements (or) group of elements $a = (a_1, a_2)$. If 'a' is in set A, then it is said to be $\boxed{a \in A}$.

If no elements are presents within the set A, it is said to be null set.

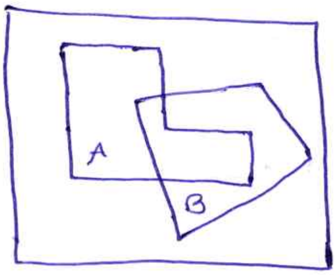
$$\boxed{A = \phi}$$

The another set operations used in image processing are $A \cup B$, $A \cap B$, A^C and $A - B$

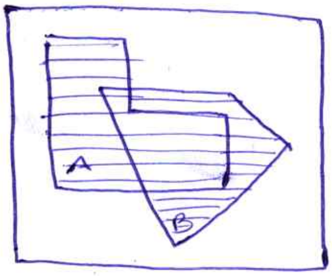
$A \cup B$:- represents the set of elements belongs to A (or) B (or) both.

$A \cap B$:- The set of elements belongs to both set A and set B.

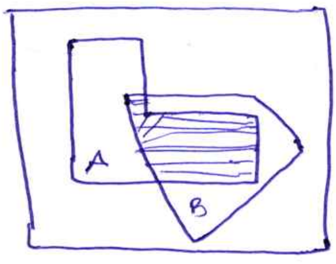
The operations performed in image procesing is as shown in figure.



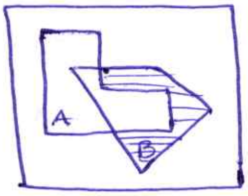
(i) $A \cup B$

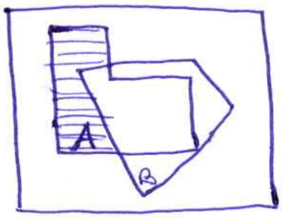


(ii) $A \cap B$



iii) A'   (iv) A-B



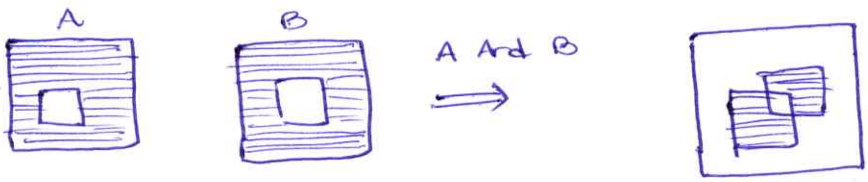


Logical Operations :-

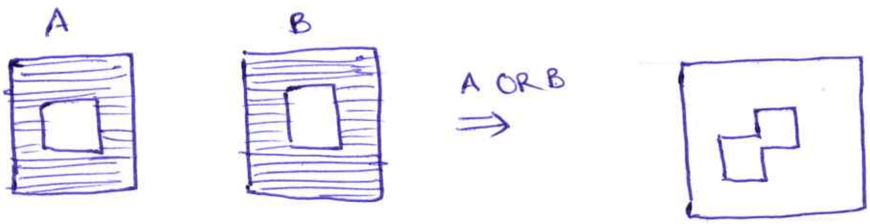
Different types of logical operations are used in Image processing such as AND, NOT, OR and XOR. The operations performed in image processing is shown in figure.

Consider the foreground portion as '1' and background portion as '0'.

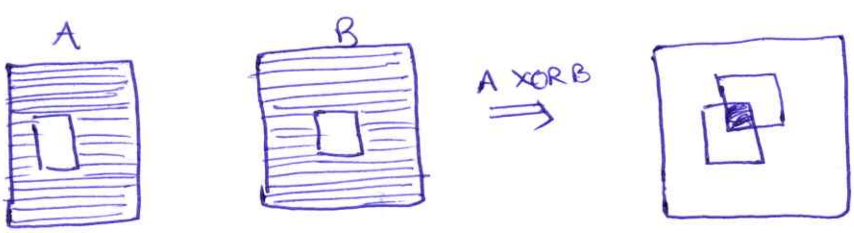
A	
0	→ 1
1	→ 0



A	B	
0	0	1
0	1	1
1	0	1
1	1	0



A	B	
0	0	0
0	1	1
1	0	1
1	1	1



A	B	
0	0	0
0	1	1
1	0	1
1	1	0

5) Spatial Operations :-

The spatial operations are performed directly on pixels in an image. These are 3 spatial operations used in image processing systems.

i) Single - pixel operation
ii) neighbourhood operation
iii) Geometrical spatial operation.

i) Single - pixel operation :-

This operation is performed directly on pixels which modifies the values based on light intensity values.

To perform this operation, it employs the transform operation. The single pixel operation is defined as

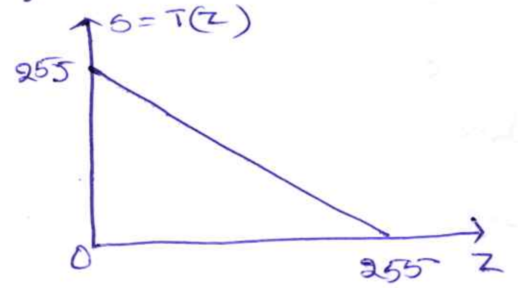
$$\boxed{s = T(z)}$$

where z represents intensity value of input image (Pixel).

T represents transformation operation

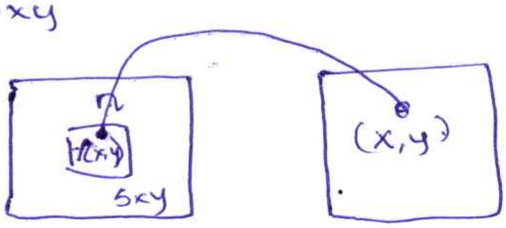
s represents the processed output image (Pixel) intensity value.

Graphically it is represented as

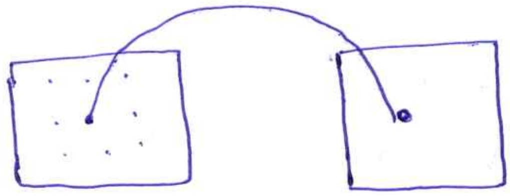


ii) Neighbourhood Operation :-

Let S_{xy} represents the set of co-ordinate values centred at a arbitrary point (x, y). The arbitrary point is transformed after performing the neighbourhood operation at some co-ordinate such that its pixel is determined by averaging of all the pixels in neighbourhood S_{xy}



Ex:-



mathematically, the operation is expressed as

$$g(x,y) = \frac{1}{MN} \sum f(r,c)$$

r = co-ordinates of row pixels.

c = co-ordinates of column pixels.

iii) Geometrical Spatial Operation (or) Transformation :-

This transformation modify the spatial relationship between pixels in an image. This transformation is also known as Rubber sheet transformation because it may viewed as analogous (or) similar to printing of image on a sheet of rubber and stretched the image according to predefined set of rules.

Geometrical transformation performs two operations.

1. Spacial co-ordinate transformation.
2. Intensity Interpolation.

The operation of ~~geom~~ geometrical transformation is given as $(x,y) = T(v,w)$

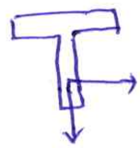
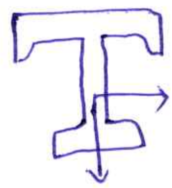
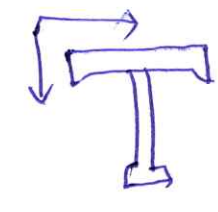
(x, y) are co-ordinates of processed image.
v, w are co-ordinates of original image.

To perform the transformation operation the method utilizes "Affine transform" is given as

$$[x \quad y \quad 1] = T\,[v \quad w \quad 1] = [v \quad w \quad 1]\begin{bmatrix} T_{11} & T_{12} & 0 \\ T_{21} & T_{22} & 0 \\ T_{31} & T_{32} & 1 \end{bmatrix}$$

The Affine transform is used to scale, rotate and shear the image.

The operations of Affine transform is as shown in figure.

Operation name	Affine matrix	co-ordinate equation	Example.
i) Identity operation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = w$	
ii) Scale	$\begin{bmatrix} C_x & 0 & 0 \\ 0 & C_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = C_x v$ $y = C_y w$	
iii) Rotation	$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v\cos\theta - w\sin\theta$ $y = v\sin\theta + w\cos\theta$	
iv) Translation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ T_x & T_y & 1 \end{bmatrix}$	$x = v + T_x$ $y = w + T_y$	

G. Vector and matrix operation:-

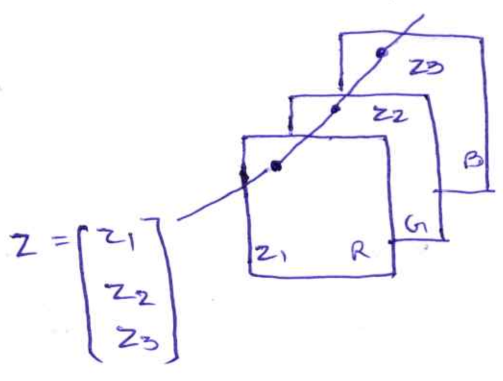
These operations are commonly used in multi spectoral image processing. Generally the colour images are formed by RGB (Red Green Blue) colour spaces. Each pixel in colour image is composed of three components. i.e, intensity values of Red, Green and Blue.

i.e $z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$

where z_1 is intensity values of Red

z_2 is intensity values of green.

z_3 is intensity values of blue.



The euclidian distance between the pixel vector z and the orbitorary point a is given as

$$D(z,a) = [(z-a)^T (z-a)]^{1/2}$$

$$= [(z_1-a_1)^2 + (z_2-a_2)^2 + (z_n-a_n)^2]^{1/2}$$

$$D(z,a) = \|z-a\|$$

7. Probabilistic methods :-

The probability is used in many ways in image processing to findout the random values of intensities. The probability method is used generally, this method is applicable to image Restoration.

Assume z_i be the intensity values of image

where $i = 0, 1, 2, \ldots\ldots L-1$

The probability of intensity level occuring in an image is written as

$$\boxed{P(z_k) = \frac{n_k}{MN}}$$

where

n_k indicates n times occuring

m - rows

N - columns

$P(z_k)$ is probability of k^{th} intensity value.

As we know, the probobity of all the pixel values is '1'.

i.e $$\boxed{\sum_{k=0}^{L-1} P(z_k) = 0}$$

To find out the mean and variance values related to the probability is given as

$$\boxed{\text{mean } (m) = \sum_{K=0}^{L-1} Z_K \, P(Z_K)}$$

$$\boxed{\text{variance } (\sigma^2) = \sum_{K=0}^{L-1} (Z_K - m)^2 \, P(Z_K)}$$

Image Transforms :-

2-D FFT :-

2D Discrete Fourier Transform

The independent variable (t, x, y) is discrete

$$F_r = \sum_{k=0}^{N_0-1} f[k] e^{-jr\Omega_0 k}$$

$$f_{N_0}[k] = \frac{1}{N_0} \sum_{r=0}^{N_0-1} F_r e^{jr\Omega_0 k}$$

$$\Omega_0 = \frac{2\pi}{N_0}$$

$\Longrightarrow$

$$F[u,v] = \sum_{i=0}^{N_0-1} \sum_{k=0}^{N_0-1} f[i,k] e^{-j\Omega_0(ui+vk)}$$

$$f_{N_0}[i,k] = \frac{1}{N_0^2} \sum_{u=0}^{N_0-1} \sum_{v=0}^{N_0-1} F[u,v] e^{j\Omega_0(ui+vk)}$$

$$\Omega_0 = \frac{2\pi}{N_0}$$

Properties

- Linearity $\rightarrow$ $af(x,y) + bg(x,y) \Leftrightarrow aF(u,v) + bG(u,v)$
- shifting $\rightarrow$ $f(x-x_0, y-x_0) \Leftrightarrow e^{-j2\pi(ux_0+vy_0)} F(u,v)$
- Modulation $\rightarrow$ $e^{j2\pi(u_0x+v_0y)} f(x,y) \Leftrightarrow F(u-u_0, v-v_0)$
- Convolution $\rightarrow$ $f(x,y) * g(x,y) \Leftrightarrow F(u,v) G(u,v)$
- Multiplication $\rightarrow$ $f(x,y) g(x,y) \Leftrightarrow F(u,v) * G(u,v)$
- Separability $\rightarrow$ $f(x,y) = f(x) f(y) \Leftrightarrow F(u,v) = F(u) F(v)$

Separability

1. Separability of the 2D Fourier Transform -

- 2D Fourier Transforms can be implemented as a sequence of 1D Fourier Transform operations performed independently along the two axis.

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

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-j2\pi ux}\, e^{-j2\pi vy}\, dx\, dy = \int_{-\infty}^{\infty} e^{-2\pi vy}\, dy \int_{-\infty}^{\infty} f(x,y)\, e^{-j2\pi ux}\, dx =$$

$$= \int_{-\infty}^{\infty} F(u,y)\, e^{-j2\pi vy}\, dy \equiv F(u,v)$$

2D DFT → 1D DFT along the rows → 1D DFT along the cols

Separability

- Separable functions can be written as $f(x,y) = f(x) \cdot g(y)$

2. The FT of a separable function is the product of the FTs of the two functions.

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)\, g(y)\, e^{-j2\pi ux}\, e^{-j2\pi vy}\, dx\, dy$$

$$= \int_{-\infty}^{\infty} g(y)\, e^{-j2\pi vy}\, dy \int_{-\infty}^{\infty} h(x)\, e^{-j2\pi ux}\, dx = H(u)\, G(v)$$

$$f(x,y) = h(x)g(y) \Rightarrow F(u,v) = H(u)\,G(v)$$

Walsh Transform :-

We define now the 1-D walsh Transform as follows:

$$W(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\left[\prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)}\right]$$

The above is equivalent to:

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

The transform kernal values are obtained from:

$$T(u,x) = T(x,u) = \frac{1}{N}\left[\prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)}\right] = \frac{1}{N}(-1)^{\sum_{i=1}^{n-1} b_i(x)b_{n-1-i}(u)}$$

Therefore, the array formed by the walsh matrix is a real symmetric matrix. It is easily shown that it has orthogonal columns and rows.

1-D Inverse Walsh Transform

$$f(x) = \sum_{x=0}^{N-1} W(u)\left[\prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)}\right]$$

The above is again equivalent to

$$f(x) = \sum_{x=0}^{N-1} W(u)\,(-1)^{\sum_{i=1}^{n-1} b_i(x)\,b_{n-1-i}(u)}$$

The array formed by the Inverse walsh matrix is identical to the one formed by the forward Walsh matrix apart from a multiplicative factor N.

2-D Walsh Transform :-

We define now the 2-D Walsh Transform as a straight forward extension of the 1-D Transform:

$$W(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\left[\prod_{i=0}^{n-1}(-1)^{b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v)}\right]$$

- The above is equivalent to:

$$W(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\,(-1)^{\sum_{i=1}^{n-1}(b_i(x)b_{n-1-i}(u)+b_i(x)b_{n-1-i}(u))}$$

Inverse Walsh Transform :-

We define now the Inverse 2-D Walsh transform. It is identical to the forward 2-D walsh transform

$$f(x,y) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} W(u,v)\left[\prod_{i=0}^{n-1}(-1)^{b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v)}\right]$$

- The above is equivalent to:

$$f(x,y) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} W(u,v)\,(-1)^{\sum_{i=1}^{n-1}(b_i(x)b_{n-1-i}(u)+b_i(x)b_{n-1-i}(u))}$$

Hadamard Transform :

We define now the 2-D Hadamard transform. It is similar to the 2-D Walsh transform.

$$H(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \left[\prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_i(u) + b_i(y)\, b_i(v)} \right]$$

The above is equivalent to:

$$H(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\, (-1)^{\sum_{i=1}^{n-1} (b_i(x) b_i(u) + b_i(x) . b_i(u))}$$

We define now the Inverse 2-D Hadamard Transform. It is identical to the forward 2-D Hadamard transform.

$$f(x,y) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} H(u,v) \left[\prod_{i=0}^{n-1} (-1)^{b_i(x)\, b_i(u) + b_i(y)\, b_i(v)} \right]$$

The above is equivalent to:

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

2. Image Enhancement

Image Enhancement in Spatial Domain:

Image enhancement approaches fall into two broad categories:

- Spatial domain methods
- Frequency domain methods

The term spatial domain refers to the image plane itself, and approaches in this category are based on direct manipulation of pixels in an image.

Frequency domain processing techniques are based on modifying the Fourier transform of an image. It is used to enhance Medical images, images captured in remote sensing, images from satellite. Image enhancement has very good Applications. The term spatial domain refers to the aggregate of pixels composing an image.

Spatial Domain processes will be denoted by the expression.

$$\boxed{g(x,y) = T[f(x,y)]}$$

Where $f(x,y)$ is the input image, $g(x,y)$ is the processed image, and T is an operator on f, defined over some neighborhood of (x,y). The principal approach in defining a neighborhood about a point (x,y) is to use a square or rectangular subimage area centered at (x,y). The center of the subimage is moved from pixel to pixel starting, say, at the top left corner. The operator T is applied at each location (x,y) to yield the output, g, at that location.

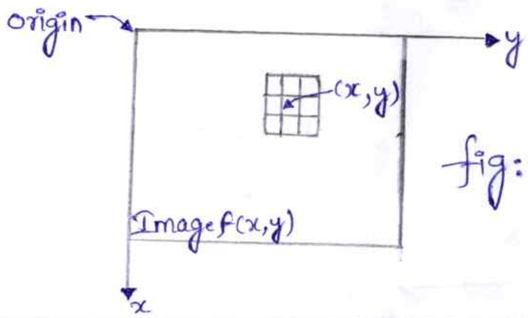


fig: 3x3 neighborhood about a point (x,y) in an image

The simplest form of T is when the neighborhood is of size 1*1. In this case, g depends only on the value of f at (x,y), and T becomes a gray-level (also called an intensity or mapping) transformation function of the form.

$$\boxed{S = T(r)}$$

Where r is the pixels of the input image and S is the pixels of the output image. T is a transformation function that maps each value of 'r' to each value of 's'.

Basic Intensity transformation / Basic Gray Level Transformations:

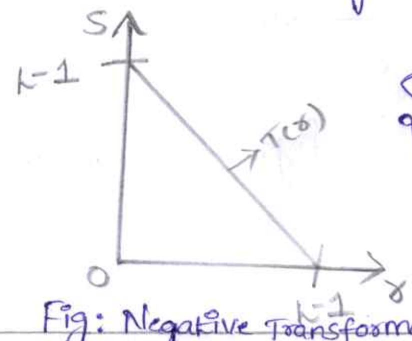
- Image Negative
- Log Transformations
- power Law Transformations
- piecewise-Linear Transformation functions

Image Negative:

The image negative with gray level value in the range of [0, L-1] is obtained by negative transformation given by S=T(r) or

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

Where r = gray level value at pixel (x,y).
L is the largest level consist in the image.
→ It results in getting photograph negative. It is useful for enhancing white details embedded in dark regions of the image.



{conversion of Black pixels to white pixels (vice versa)}

Fig: Negative Transformation

Logarithmic Transformations:

The Log transformation can be defined by the formula

$$\boxed{S = C\log(r+1)}$$

where s and r are the pixel values of the output and the input image and c is constant. Value 1 is added to each of the pixel value of the input image becoze if there is a pixel intensity of 0 in the image, then log(0) is equal to infinity. So 1 is added, to make the minimum value at least 1.

It maps Narrow Range of pixels to wide range pixels with in the same range.

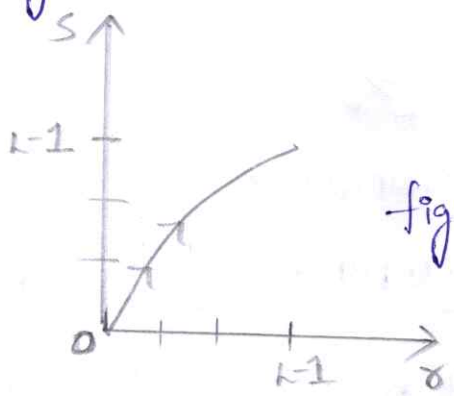


fig: Log transformation curve i/p v/s o/p

Power-Law Transformations (Gamma Correction):

There are further two transformation is power law transformations, that include nth power and nth root transformation. These transformation can be given by the expression.

$$\boxed{S = (Cr)^{\gamma}}$$

Here, $\gamma = 1 \rightsquigarrow$ Linear

$\gamma < 1 \rightsquigarrow$ nth root

$\gamma > 1 \rightsquigarrow$ nth power

This Symbol γ is called gamma, due to which this transformation is also known as gamma transformation.

Where C and g are positive constants. Sometimes Equation $S = C(r+\varepsilon)^{\gamma}$ for an offset.

In the below figure that curves generated with values of $\gamma > 1$ have exactly. The opposite being true for generated with values of $\gamma < 1$. Here, to the identify transformation when $C = \gamma = 1$

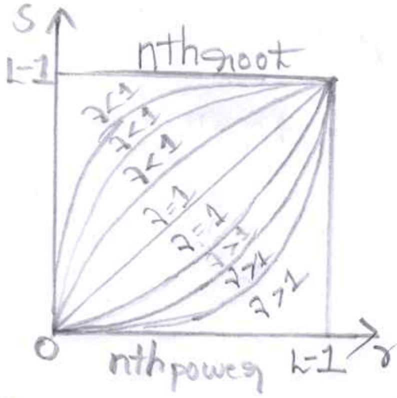


fig: power Law Transformation

Piecewise-Linear Transformation Functions/Contrast Stretching:

A piecewise linear transformation function is a type of mathematics function is defined by multiple linear segments, each applicable over a specific interval of the i/p domain. This means that the function's rule changes depending on the value of the input variable. Consider two pixels located at (r_1, s_1) (r_2, s_2) as shown in below fig.

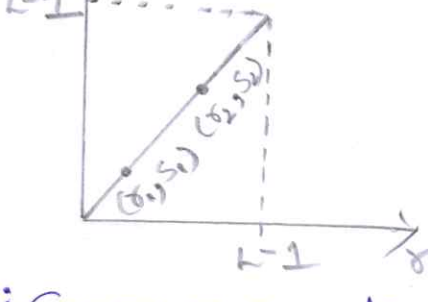


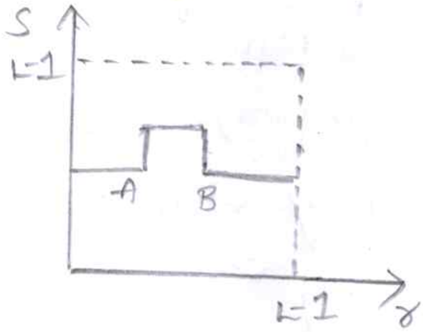
fig: Constrast Stretching

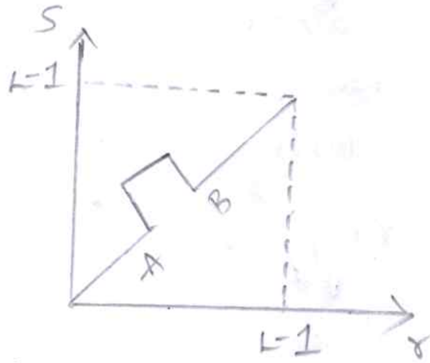
Gray-Level Slicing:

Highlighting a specific range of gray levels in an image often is desired. Applications include enhancing features such as masses of water in satellite imagery and enhancing flaws in X-ray images.

There are several ways of doing level slicing, but most of them are variations of two basic themes. One approach is to display a high value for all gray levels in the range of interest and a low value for all other gray levels.

Highlights the specified region as '1' and unspecified as '0' (or) same.





Bit-plane slicing:

Instead of highlighting the specific region highlight the specific bits of enhances the image such phenomenon known as Bit plane slicing.

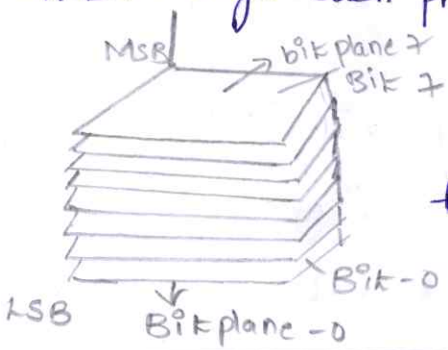


fig: Bit-plane slicing

Histogram processing:

The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function of the form.

$$H(r_K) = n_K$$

Where n_K is the k^{th} gray level and n_K is the number of pixels in the image having the level. A normalized histogram is given by the equation.

$$P(r_K) = n_K/MN \qquad K = 0, 1, 2 \ldots, L-1$$

$P(r_K)$ gives the estimate of the probability of occurrence of gray level r_K. The sum of all components of a normalized histogram is equal to 1. The histogram plots are simple plots of $P(r_K) = r_K$ versus r_K. In the dark image the components of the histogram are concentrated on the low(dark) side of the gray scale. In case of bright image the histogram components are baised towards the high side of the gray scale. The histogram of a low constrast image will be narrow and will be centered towards the middle of the grayscale.

Generally the images are four types. They are:

i. Dark Image

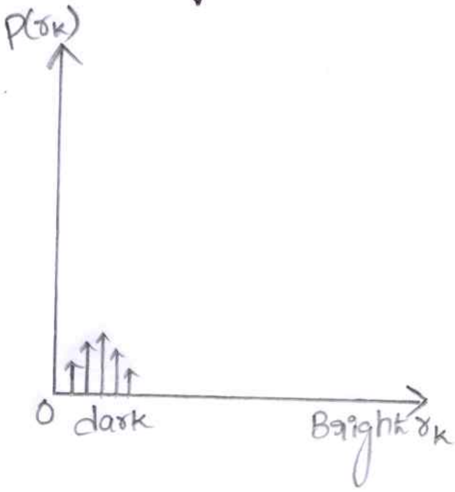
ii. Bright Image

iii. Low Constrast Image

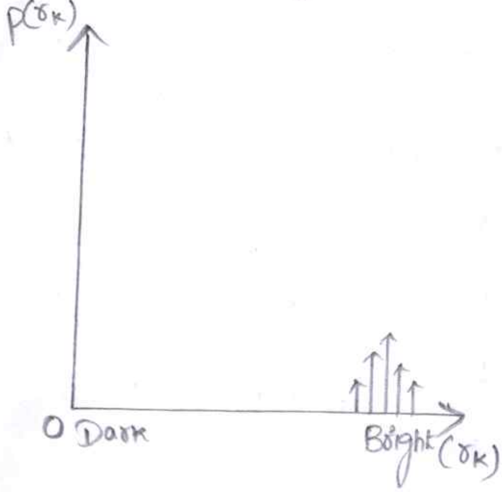
iv. High contrast Image

The components of the histogram in the high contrast image cover a board range of the gray scale. The net effect of this will be an image that shows a great deal of gray level high dynamic range.

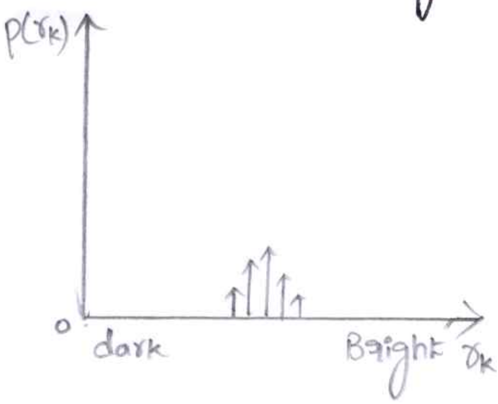
1. Dark Image



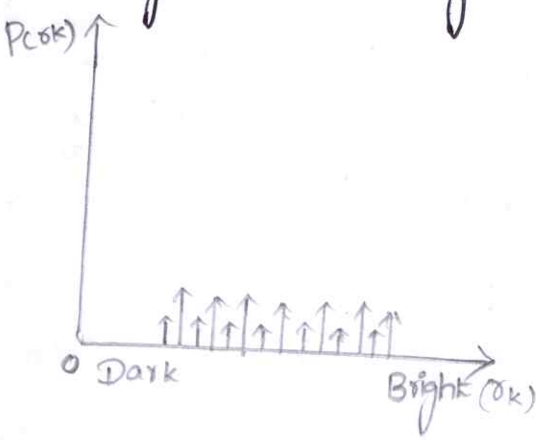
2. Bright Image



3. Low Contrast Image



4. High Contrast image



Histogram Equalization:

Histogram equalization is a common technique for enhancing the appearance of images. Suppose we have an image which is predominantly dark. Then its histogram would be skewed towards the lower end of the grey scale and all the image detail are compressed into the dark end of the histogram. If we could 'stretch out' the grey levels at the dark end to produce a more uniformly distributed then the image would become much clearer.

We know $S = T(r) \longrightarrow (1)$

$$P_s(S) = P_r(r)\left|\frac{dr}{ds}\right| \longrightarrow (2)$$

$$T(r) = L-1 \int_0^\infty P_r(r) \rightsquigarrow (3)$$

To Equalize the image in spatial domain

diff equ(1) w.r.to (r) then

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

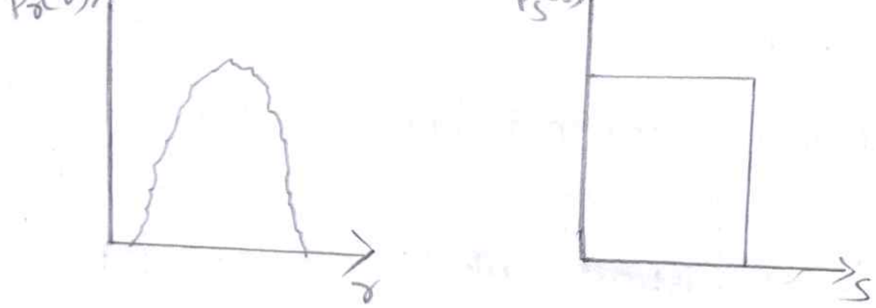
$$\frac{ds}{dr} = (L-1)\, P_r(r) \rightsquigarrow (4)$$

Subs above equ (4) in equ (2)

$$P_s(S) = \cancel{P_r(r)} \frac{1}{(L-1)\cancel{P_r(r)}}$$

$$P_s(S) = \frac{1}{L-1}$$

$$\boxed{P_s(S) = 1}$$



Smoothing and Sharpening Spatial filters:

The use of masks in image processing System is known as Spatial filtering and masks themselves are known as filters. For Image enhancement, the linear and non-linear filters are used usually a Lowpass filters, High pass filters, Band pass filters.

Low pass filter allows low frequency Components (Smoothering).
High pass filter allows high frequency Components.
Bandpass filter allows a band of frequency Components.

Generally the Spatial filtering is performed by the use of mask, which consists of coefficients and the response of filtering is the Sum of product b/w Coefficients of mask and gray levels of original Image.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

3x3

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

3x3

$R = w_1 z_1 + w_2 z_2 + w_3 z_3 + \ldots\ldots w_9 z_9.$

The Spatial filtering Can be Classified into two types

1) Image Smoothering
2) Image Sharpening

1) Image Smoothering:

Smoothering performs noise reduction and blurring. Blurring is used in pre processing steps to remove the Small details from an image prior to object extraction.

Usually the Smoothering is performed by two types of filters

i) Low pass filter
ii) Median filter

i) **Low pass filter:**

The shape of impulse response needed to be implemented by Low pass filter which allows low frequency components and attenuates high frequency components.

It performs averaging of an image i.e., each and every pixel in gets average by sum of all gray levels in an Image.

The LPF mask contain all positive coefficients.

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

The main drawback of LPF is it removes edges (or) some other sharp details.

ii) **Median filter:**

The principle objective of the median filter is it performs noise reduction rather than blurring.

It also works (or) operates to preserve fine sharp details.

It performs by replacing each and every gray level of the pixel by its median value which is obtained by sorting all the pixels in 3x3 image are 10, 20, 20, 20, 15, 20, 20, 25, 100

10	20	20
20	15	20
20	25	100

Sorting 10, 15, 20, 20, 20, 20, 20, 25, 100

All the pixels are replaced by 20 in next step.

2) **Image Sharpening:**

The sharpening is used to highlight the Sharp details (edges). The Sharpening is performed by '3' types of filters.

ⓐ Basic High pass filter

ⓑ High Boost filter

ⓒ Derivative filter

ⓐ **Basic High pass filter:**

The shape of impulse response needed to be implemented by HPF. Which allows high frequency Components and attenuates LF Components. In high pass mask, the Centre Coefficient is +ve and outer peripherals is -ve.

The mask is shown in figure.

In HPF the Sum of all Co-efficient is zero.

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

ⓑ **High Boost filter:**

A High pass filtered image can be computed interms of low pass filtered image and original images.

i.e., original image = HPF + LPF

A high boost filter is designed by multiply a gain parameter to the original image.

So, HPF = original image − LPF

High boost filter can be formed

HBF = A(original) − LPF

HBF = (A−1) original + original − LPF

HBF = (A−1) original + HPF

If A = 1; the HBF act as HPF.

(C) Derivative filter:→

Averaging of pixels in a region tend to blur, details in an Image. usually the averaging is performed by integration (or) differentiation. In image processing systems the differentiation is performed by derivative Filters, usually a gradient filter denoted as ∇f.

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

The magnitude of gradient is

$$mag(\nabla f) = \sqrt{\left[\frac{\partial f}{\partial x}\right]^2 + \left[\frac{\partial f}{\partial y}\right]^2}$$

Image Smoothing and Sharpening using Frequency Domain Filters:

Improving the quality of an Image is known as enhancement. The enhancement in frequency domain is computed as applying the fourier transform on original Image and multiply with filter transfer function & then apply inverse fourier transform to get enhanced Image.

The enhancement in frequency domain given as

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

In spatial domain, the product becomes as convolution.

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

Filtering in Frequency Domain:

The image enhancement is performed by two types of filtering.

1) Image Smoothening
2) Image Sharpening

1) Image Smoothening:

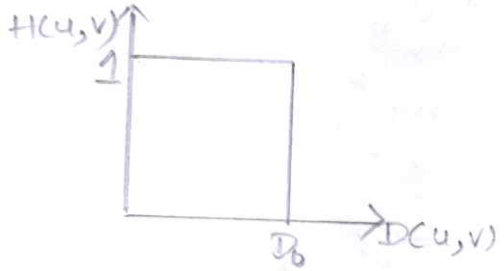
Smoothening filters are used to remove blurness and noise reduction in frequency domain.

The Smoothening is achieved by the three types of low pass filters which is used to reduce the noise at center of an image. i.e., blurness is achieved by attenuating the high frequency components.

i) Ideal Low pass filter:-

Low pass filter allows low frequency components and attenuates high frequency components with a cutoff frequency D_0 located at a distance $D(u,v)$ from the origin.

The transfer characteristics of ideal low pass filter is as shown in fig



The transfer function of 2-dimensional ideal LPF is

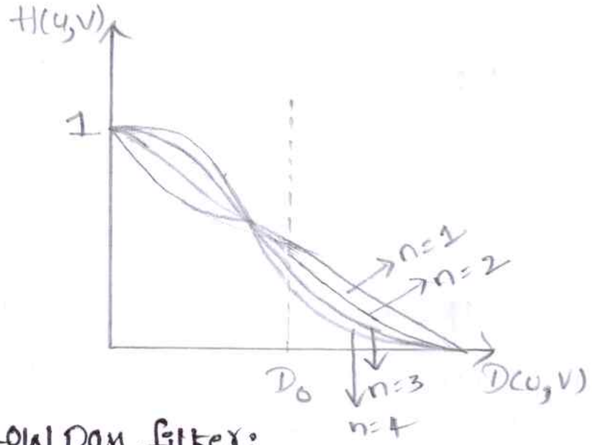
$$H(u,v) = \begin{cases} 1, & D(u,v) \le D_0 \\ 0, & D(u,v) > D_0 \end{cases}$$

ii) Butterworth low pass filter:-

The transfer function of 2-dimensional butterworth LPF of order 'n' with the cut-off frequency D_0 is given as

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

The transfer characteristics of butterworth LPF is as shown in fig.



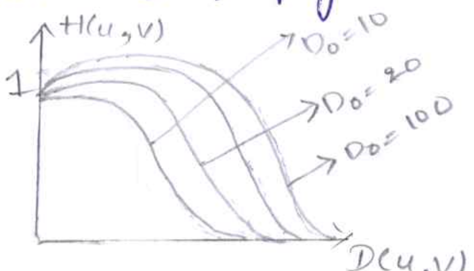
iii) Gaussian low pass filter:

The transfer function of 2-dimensional gaussian LPF with cut-off frequency D_0 is given as

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

where σ is measure of spread at the centre of an image

The transfer characteristics of gaussian LPF is



2) Image Sharpening:→

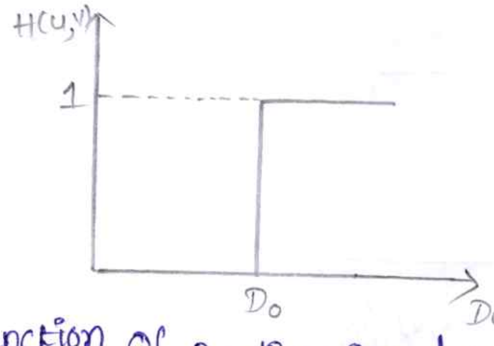
The Sharpening filters are used to highlight (or) enhance a fine details in an Image.

At edges (or) Sharp transitions have significant high frequency components. To enhance Such such components, three types of high pass filters are used which removes noise at edges (or) boundaries of an Image.

i) Ideal High pass filter:→

High pass filter allows high frequency components and attenuates low frequency components with a cut-off frequency 'D_0' located at a distance of $D(u,v)$ from the origin.

The transfer characteristics of Ideal high pass filter is



The transfer function of 2-dimensional HPF with a cut-off frequency D_0 is given as

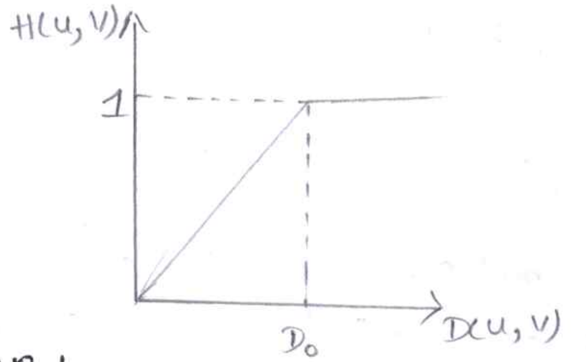
$$H(u,v) = \begin{cases} 0, & D(u,v) \le D_0 \\ 1, & D(u,v) > D_0 \end{cases}$$

ii) Butterworth HPF:→

The transfer function of 2-dimensional butterworth high pass filter of order n with the cut-off frequency D_0 is

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

The transfer characteristics of butterworth HPF is

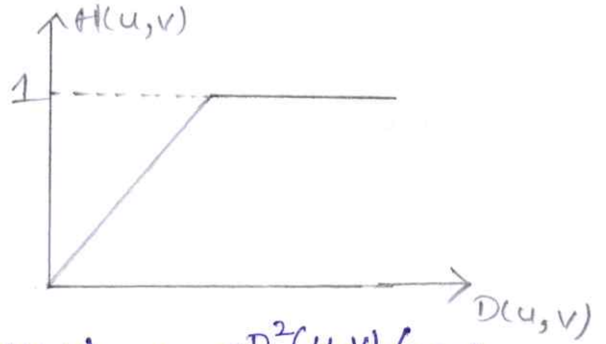


iii) Gaussian High pass Filter :-

The transfer function of 2-dimensional gaussian High pass filter with cut-off frequency D_0 is given as

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

The transfer characteristics of Gaussian high pass filter is



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

UNIT-III
IMAGE RESTORATION

A Model Of The Image Degradation:

Degradation Process operates on a degradation function that operates on an input image with an additive Noise term. Input image is represented by using the notation $f(x,y)$, noise term can be represented as $\eta(x,y)$. Those two terms when combined gives the result as $g(x,y)$. If we are given $g(x,y)$, some knowledge about the degradation function H or J and some knowledge about the additive noise term $\eta(x,y)$, the objective of restoration is to obtain an estimate $\hat{f}(x,y)$ of the original image. We want estimate to be as close as possible to the original image. The more we know about h and η, the closer $\hat{f}(x,y)$ will be to $f(x,y)$. If it is a linear Position invariant Process, then degraded image is given in the spatial domain by

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

$h(x,y)$ is spatial representation of degradation function and symbol * represents convolution. IN frequency domain we may write this equation as

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

The terms in the capital letters are the fourier Transform of the corresponding terms in the spatial domain.

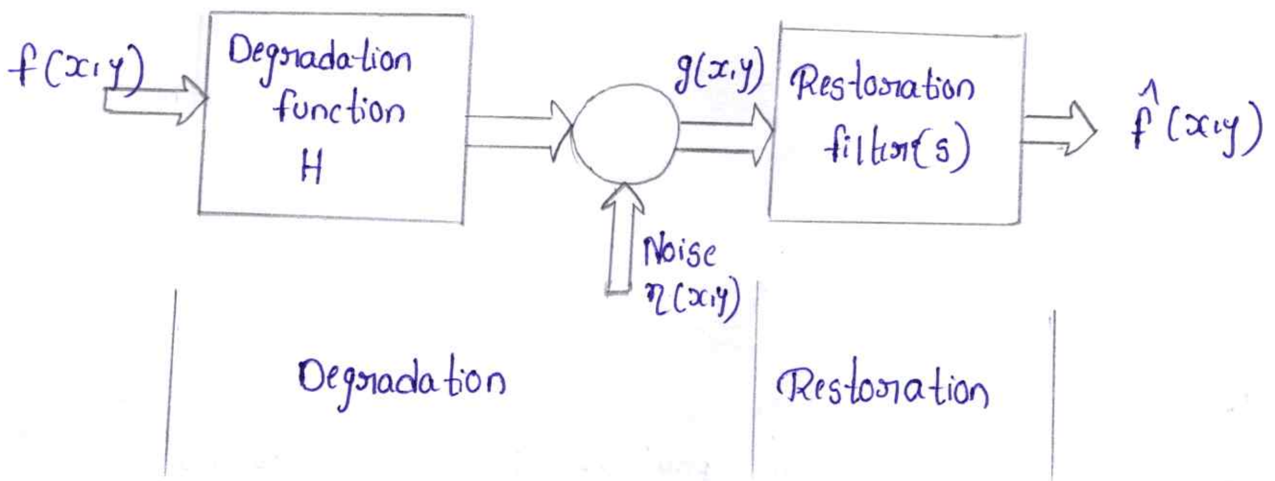


fig:- A Model of the image Degradation / Restoration Process.

Noise Models :

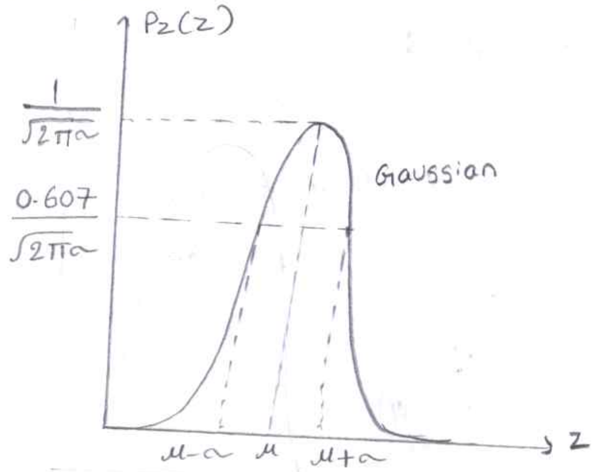
The Principal source of Noise in digital images arises during image acquisition and /or transmission. The Performance of imaging sensors is affected by a variety of factors, such as environmental conditions during image acquisition and by the quality of the sensing elements themselves. Image are corrupted during transmission principally due to interference in the channels used for transmission. Since main sources of noise presented in digital images are resulted from atmospheric disturbance and image sensor circuitry. following assumptions can be made i.e the Noise model is spatial invariant (independent of spatial location) . the Noise Model is uncorrelated with the object function.

Gaussian Noise :

The Noise Models is are used frequently in Practices because of its tractability in both spatial and frequency domain. The Pdf of gaussian random variable is

$$P_z(z) = \begin{cases} \frac{1}{b-a} & \text{if } a \le z \le b \\ 0 & \text{otherwise} \end{cases}$$

Where z represents the gray level , μ = mean of average value of z, σ = standard deviation



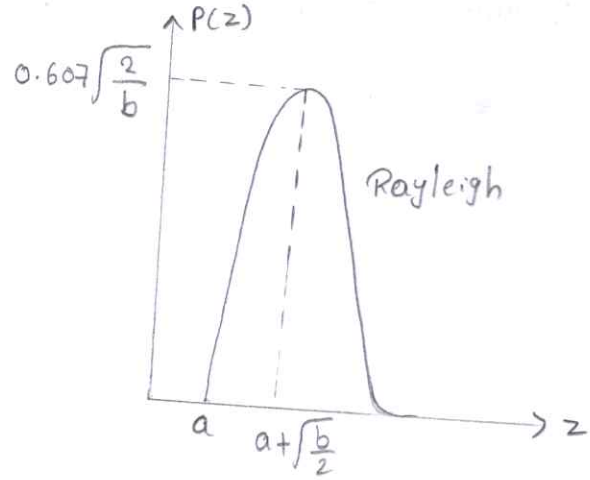
Rayleigh Noise :-

Unlike Gaussian distribution , the Rayleigh distribution is no symmetric. it is given by the formula

$$P_z(z) = \begin{cases} \frac{2}{b}(z-a)^{2/b} & z \geq a \\ 0 & z \leq a \end{cases}$$

The Mean and variance of this density is

$$m = a + \sqrt{\pi b/4}\ , \sigma^2 = \frac{b(4-\pi)}{4}$$



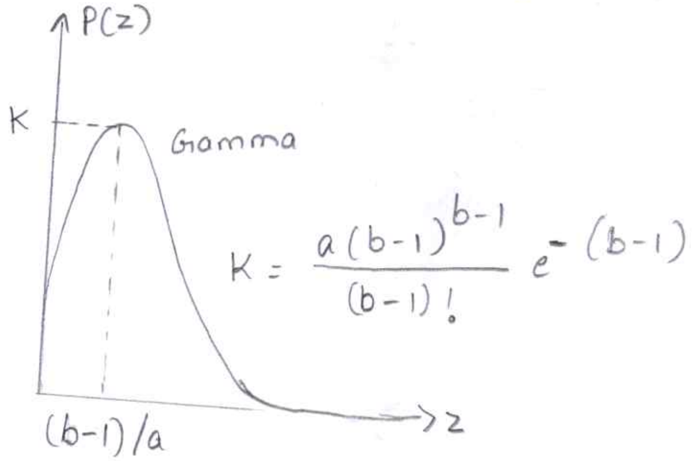
Gamma Noise :

The PDF of erlang noise is given by

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

The mean and variance of this density are given by

$$\text{mean} : \mu = \frac{b}{a} \quad \text{variance} : \sigma^2 = \frac{b}{a^2}$$



it's shape is similar to Rayleigh disruption. this Equation is referred to as gamma density it is correct only when the denominator is the gamma function.

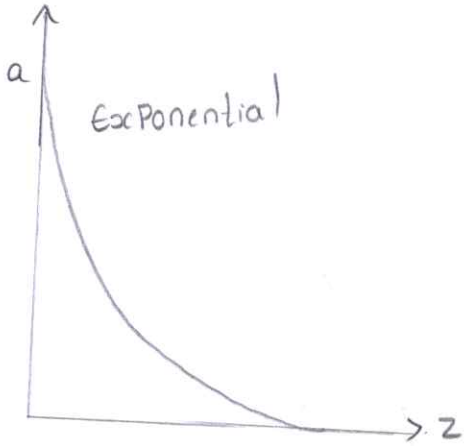
Exponential Noise:

Exponential distribution has an exponential shape. The PDF of exponential Noise is given as

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

Where $a>0$. The Mean and variance of this density are given by

$$m = \frac{1}{a}, \quad \sigma^2 = \frac{1}{a^2}$$



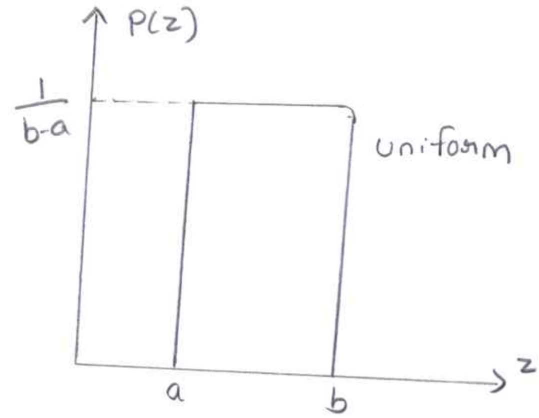
Uniform Noise:

The PDF of uniform noise is given by

$$p_z(z) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq z \leq b \\ 0 & \text{otherwise} \end{cases}$$

The Mean and variance of the Noise is

$$m = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$$



Impulse (salt & Pepper) Noise:

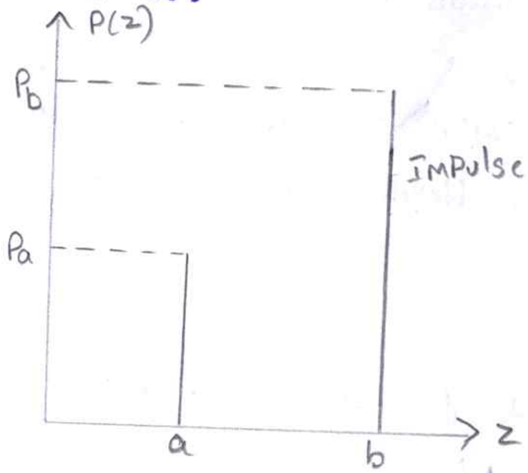
IN this case, the Noise is signal dependent, And is multiplied to the image

The PDF of bipolar (impulse) Noise is given by

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

$b > a$

If $b>a$, gray level b will appear as a light dot in image. level a will appear like a dark dot.



Restoration in the Presence of Noise Only - Spatial filtering:

When the only degradation Present in an image is Noise, i.e

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

(or)

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

The Noise terms are unknown so subtracting then from $g(x,y)$ or $G(u,v)$ is not a realistic approach. IN the case of Periodic Noise it is Possible to Estimate $N(u,v)$ from the Spectrum $G(u,v)$ So $N(u,v)$ can be subtracted from $G(u,v)$ to obtain an Estimate of original image. Spatial filtering can be done when only additive Noise is Present. The following techniques can be used to reduce the noise effect.

(i) Mean Filter:

(ii) (a) Arithmetic Mean filter:

It is the simplest mean filter. let S_{xy} represents the set of coordinates in the sub image of size $m*n$ centered at Point (x,y). The arithmetic mean filter computes the average value of the corrupted image $g(x,y)$ in the area defined by S_{xy}. The value of the restored image f at any Point (x, y) is the arithmetic mean computed using the Pixels in the region defined by S_{xy}

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

The operation can be using a convolution mask in which all coefficients have value $1/mn$ A mean filter smooths local variations in image. Noise is reduced as a result of blurring. for every Pixel in the image, the Pixel value is replaced by the mean value of its neighboring Pixels with a weight. This will resulted in a smoothing effect of the image.

(b) Geometric Mean filter:

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

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

Here, each restored Pixel is given by the Product of the Pixel in the sub image window. raised to the Power $1/mn$. A geometric mean filters but it to loose image details in the Process.

(c) Harmonic Mean filter:

The Harmonic mean filtering operations is given by the Expression.

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

The harmonic mean filter works well for salt noise but fails for Pepper noise. it does well with gaussian noise also.

(d) Order statistics filter:

(It is the Best order statistic filter. it replaces the value of a Pixel by the)

Order statistics filters are spatial filters whose response is based on ordering the Pixel contained in the image area encompassed by the filter. the response of the filter at any Point is determined by the ranking result.

(e) Median filter:

It is the best order statistic filter. it replaces the value of a Pixel by the Meadian of gray levels in the Neighborhood of the Pixel.

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

The original of The Pixel is included in the computation of the median of the filter are quite Possible because for certain types of random Noise. the Provide excellent noise reduction capabilities With considerably less blurring then Smoothning filters of similar size. These are effective for bipolar and unipolar impulse noise.

(f) Max and Min filter:

using the 100th Percentile of ranked set of numbers is called the max filter and is given by the Equation.

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

It is used for finding the brightest Point in an image. Pepper Noise in the image has very low values. it is reduced by max filter using the max selection Process in the Sublimated area sky. the 0th Percentile filter is min filter.

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

This filter is useful for flinging the darkest Point in image. Also, it reduces salt noise of the main operation

(g) MidPoint filter:-

The MidPoint filter simply computes the midPoints between the maximum and minimum values in the area encompassed by

$$\hat{f}(x,y) = \left(\underset{(s,t)\in S_{xy}}{\max} \{g(s,t)\} + \underset{(s,t)\in S_{xy}}{\min} \{g(s,t)\}\right) / 2$$

It comelines the order statistics and averaging. this filter Works best for randomly distributed noise like gaussian or uniform noise.

Periodic Noise by frequency domain filtering:

These types of filters are used for this Purpose.

Band Reject filters:

it removes a band of frequencies about the origin of the fourier transformer.

Ideal Band Reject filter:

An ideal band reject filter is given by the Expression.

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

$D(u,v)$ - the distance from the origin of ther centered frequency Rectangle

w - the width of the band

D_0 - the radial center of the frequency rectangle

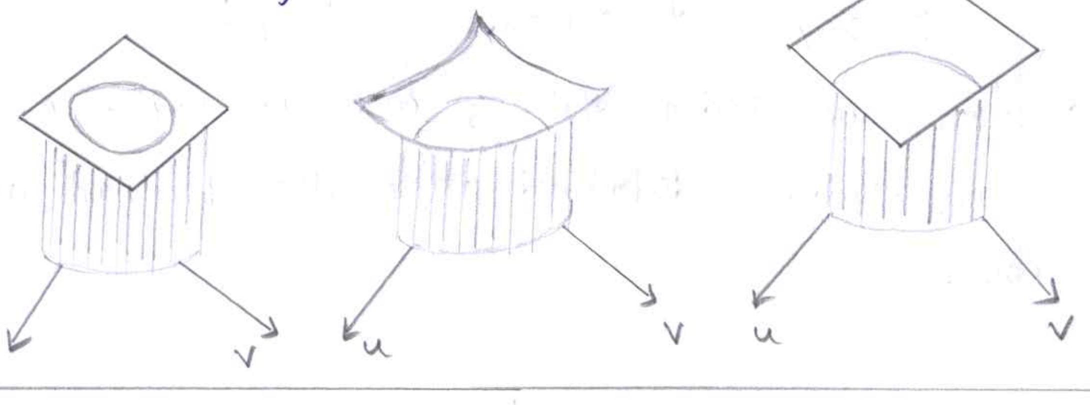
Butter Worth Band Reject filter:

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

Gaussian Band Reject filter:

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

These filters are mostly used with when the location of noise component in the frequency domain. is known. sinusoidal noise can be Easily removed by using these kinds of filters because it, show two impulses that are mirror images of each other about the origin. of the frequency transform.



Band Pass filter:

The function of a band pass filter is opposite to that of a band reject filter it allows a specific frequency band of the image to be passed and blocks the rest of frequencies. The transfer function of a band pass filter can be obtained from a corresponding band reject filter with transfer function $H_{br}(u,v)$ by using the Equation.

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

These filters cannot be applied directly on an image because it may remove too much details of an image but these are effective in isolating the effect of an image of selected frequency bands.

Notch Filters:

A notch filter rejects (or passes) frequencies is predefined neighbourhoods about a center frequency.

Due to the symmetry of the fourier transform notch filter must appear in symmetric pairs about the origin.

The Transfer function of an ideal notch filter reject of radius D_0 with centers a (u_0, v_0) and by symmetry at $(-u_0, v_0)$ is

$$D_1(u,v) = \sqrt{(u - M/2 - u_0)^2 + (v - N/2 - v_0)^2}$$

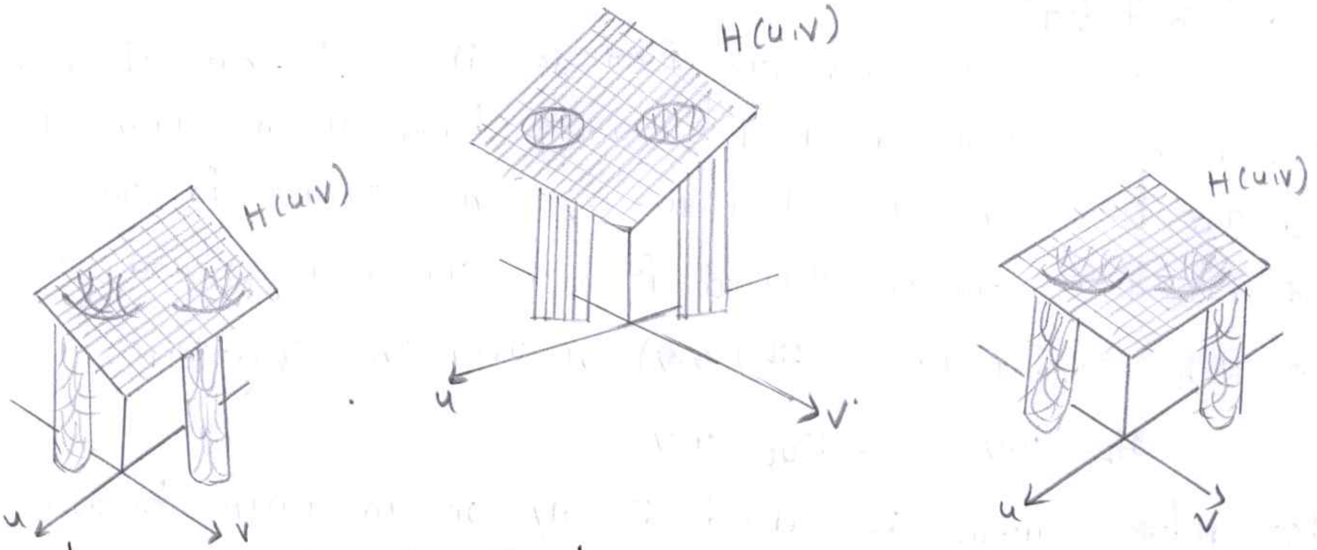
$$D_2(u,v) = \sqrt{(u - M/2 + u_0)^2 + (v - N/2 + v_0)^2}$$

ideal, butterworth, gaussian notch filters.

$$H(u,v) = \begin{cases} 0 & \text{if } D_1(u,v) \le D_0 \text{ or } D_2(u,v) \le D_0 \\ 1 & \text{otherwise.} \end{cases}$$

$$H(u,v) = 1 \Big/ \left[1 + \left[\frac{D_0^2}{D_1(u,v)\, D_2(u,v)} \right]^n \right]$$

$$H(u,v) = 1 - \mathrm{Exp}\left[-\frac{1}{2} \left[\frac{D_1(u,v)\, D_2(u,v)}{D_0^2} \right] \right]$$



Estimating the degradation function:

There are Three Principle ways to Estimate the degradation function H(u,v) they are

1. Estimating by image observation
2. Estimating by Experementation
3. Estimating by Mathematical Modelling.

1. Estimating by Image Observation:

Consider a degraded image without knowing the degraded function. Based on the Assumptions, the image is degraded by linear, Position invarient Process.

The degradation function in the degrades image is Estimated by gathering the information on image itself. i.e simply by observing the degraded image we can Estimate the degradation function.

By observing in the sectional area of their image, the degradation function is Estimated.

Assume the sub image is $G_s(u,v)$

and the image Estimated is $F^{\wedge}(u,v)$

the degraded function is $H_s(u,v)$

$\therefore$ the degradation function is estimated by

$$G_s(u,v) = H_s(u,v) \cdot F^{\wedge}(u,v)$$

$$H_s(u,v) = \frac{G_s(u,v)}{F^{\wedge}(u,v)}$$

The blurness (degradation) is removed by sharpening filters.

2. Estimating by Experimentation:

If the equipment is available same as the equipment used to acquire degraded image. it is possible to obtain accurate estimation of degradation function.

Generally the degraded image are acquired similar to that of the original image by various system settings.

To estimate the degradation image with the same system setting, the impulse response is required.

The fourier transform of impulse response is a constant i.e 'A'

With the presence of the impulse response the degradation function is estimated as

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

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

3. Estimating by Mathematical Modelling:

The degradation function in the degraded image is estimated by mathematical modelling procedure.

The degradation is obtained by the linear motion between image and sensors

By using the degradation model, the degradation function estimates by the equation.

$$H(u,v) = e^{-k(u^2+v^2)^{5/6}}$$

Except the exponent power term 5/6 the equation is similar to that of

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

Gaussian low Pass filter.

The above equation is Proposed by Hufnagel by using the Physical characteristics of atmospheric turbulence.

consider $x_0(t)$ and $y_0(t)$ be the varying intensity values with respect to t'. The total Exposure at any Point in the recording Medium can be obtained as the instantaneous Exposures of all the Points over the interval 0 to T as a function of integration

The degradation function is given as

$$g(x,y) = \int_0^T f(x - x_0(t),\ y - y_0(t))\, dt \quad - (1)$$

Apply the fourier transform to above Equation

$$G(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, e^{-j2\pi(ux+vy)}\, dx\, dy \quad - (2)$$

Substitute Eq (1) and (2) are rearranging the integration, then

$$G(u,v) = \int_0^T \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x - x_0(t), (y - y_0(t))\, e^{-j2\pi(ux+vy)}\, dx\, dy\right]$$

The above Equation can be Written as

$$G(u,v) = \int_0^T F(u,v)\, e^{-j2\pi(ux_0(t) + vy_0(t)}\, dt$$

$$G(u,v) = F(u,v) \int_0^T e^{-j2\pi(ux_0(t) + vy_0(t))}\, dt$$

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

Where $H(u,v) = \int_0^T e^{-j2\pi(ux_0(t) + vy_0(t))}\, dt$

consider $x_0(t) = \frac{at}{T}$

$y_0(t) = \frac{bt}{T}$

then

$$H(u,v) = \int_0^T e^{\frac{-j2\pi uat}{T}}\, dt$$

$$H(u,v) = \frac{T}{\pi u a} \sin(\pi u a)\, e^{-j\pi u a}$$

$$H(u,v) = \frac{T}{\pi(ua+vb)} \sin[\pi(ua+vb)]\, e^{-j\pi(ua+vb)}$$

Inverse filtering:

The simplest approach to restoration is direct inverse filtering. Where we complete an estimate $\hat{F}(u,v)$ of the transform of the original image simply by dividing the transform of the degraded image $G(u,v)$ by degradation function $H(u,v)$

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

We know that

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

Therefore

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

from the above equation We observe that We cannot recover the undegraded image Exactly because $N(u,v)$ is a random function whose fourier transform is not known.

One approach to get around the zero or small-value Problem is to limit the filter frequencies to values near the origin.

We know that $H(0,0)$ is Equal to the average values of $h(x,y)$

By limitting the analysis of frequencies near the origin We reduce the Probability of encountering zero values.

Wiener Filtering:

The inverse filtering approach has Poor Performance. the Wiener filtering approach uses the degradation function and statistical characteristics of noise into the restoration Process.

The objective is to find an Estimate $\hat{f}$ of the uncorrupted image f such that Mean square Error between them is minimized

The Error Measure is given by

$$e^2 = E\{[f(x) - f^1(x)]^2\}$$

Where $E\{\cdot\}$ is the expected value of the argument.

We assume that the noise and the image are uncorrelated one or the other has zero mean.

The gray levels in the Estimate are a linear function of the levels in the degraded image.

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

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

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

Where $H(u,v)$ = degradation function

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

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

$S_\eta(u,v) = |N(u,v)|^2$ = Power Spectrum of the Noise.

$S_f(u,v) = |F(u,v)|^2$ = Power consumption of the undcorrelated image.

The Power Spectrum of the under graded image is rarely known. An Approach used frequently when these quantities are not known or cannot be Estimated then the Expression used is

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

Where K is a specific constant.

Constrained least squares filtering:

The Wiener filter has a disadvantage that We need to know the Power Spectra of the undegraded image and Noise. The Constrained least square filtering required only the knowledge of only the Mean and variance of the Noise. these Parameters usually can be calculated from a given degraded image this is the advantage With this Method. this Method Produces a optimal results. this Method require the optimal criteria Which is important We express the

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

in vector Matrix form

$$g = Hf + \eta$$

The optimally criteria for restoration is based on a Measure of Smoothness, such as the second derivative of an image (Laplacian)

The Minimum of a criterion function c defined as

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

Subject to the constraint.

$$\|g - H\hat{f}\|^2 = \|\eta\|^2$$

Where $\|w\|^2 \triangleq w^T w$ is a Euclidean vector norm $\hat{f}$ is Estimate of the undegraded image. ∇^2 is laplacian operator.

The frequency domain solution to this optimization Problem is given by

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

Where γ is a Parameter that must be adjusted so that the constraint is satisfied.

$P(u,v)$ is the fourier transform of the laplacian operator.

$$P(x,y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Unit 4. Color Image Processing

Color Fundamentals:-

Color is a perceptual property that results from interaction of light with the human visual system.
This depends on both physical wavelength of light and physiological response of eye and brain

a. Visible Spectrum:

- The Human eye can perceive light in wavelength range of 400 nm (violet) to 700 nm (red).
- The spectrum is continous, with no sharp boundaries between colors.
- An object appears coloured because it reflect some wavelengths and absorbs others.
  * white objects reflect nearly all visible wavelengths.
  * black objects absorb all visible wavelengths.

b. Color representation and standards:-

* In the additive color system, the primary are red, green and blue (RGB). Their mixtures produce secondary colors cyan, magenta and yellow (CMY).
* To provide a standard for color specification the CIE 1931 system was introduced using tristimulus values (x, y, z). These values are plotted on CIE chromaticity diagram, where the boundary represents pure spectral colors and the interior show mixtures, with white near the center.

Color Models:-

The purpose of color model is to facilitate the specification of colors in some standard, generally accepted way.

- In essence, a color model is a specification of coordinate system and a subspace within the system where each color is represent by a single point.

- Most color models in use today are oriented either toward hardware or toward applications where color manipulation is goal.

Main types of Color Models:-

a) The RGB color Model:

In RGB model, each color represents in primary spectral components of red, green and blue. This model is based on cartesian coordinate system.

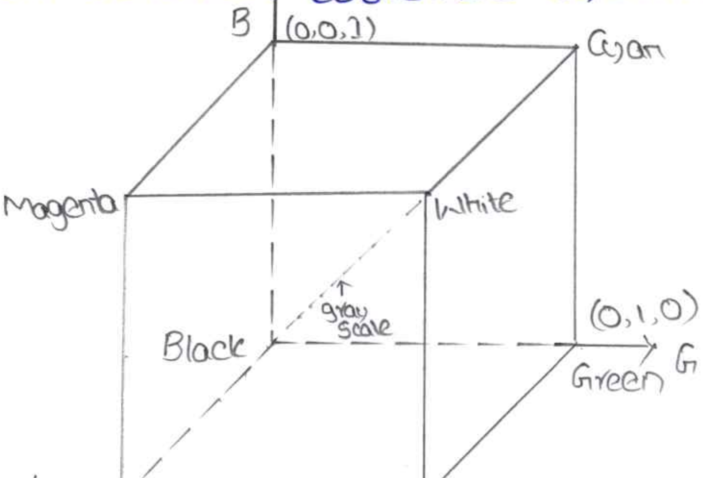


Fig:- Schematic of RGB color cube.

Structure and representation of RGB model:-

- The RGB model is represented as three-dimensional cartesian coordinate system forming a color cube where each axis corresponds to one of primary colors.
- Each color is defined by triplet (R,G,B) where each value typically ranges from 0 to 255 in standard 8 bit images, yielding $256^3 = 16,777,216$ possible colors in 24 bit RGB system.

The origin of cube (0,0,0) is black, while the opposite corner (255,255,255) is white.

The secondary colors are cyan, Magenta and yellow.

Functionality in digital imaging:-

- Images in RGB model are composed of pixels, each described by a set of three values of defining its color.
- This model supports faithful color reproduction on display screens, monitors and in image processing software.
- Color is generated through additive mixing

* The RGB model is standard for monitors, TVs, cameras and web graphics.
* It underlies image processing task, display renders and basis for defining other colors like HSI, HSV.

The detailed geometric, numeric & practical aspect of RGB color model allow accurate representation and transformation.

b) The CMY and CMYK Color Model:-

The CMY and CMYK color models are subtractive color systems used primarily in color printing, where colors are formed by mixing inks on paper rather than mixing light.

* CMY (Cyan, Magenta, Yellow) uses three secondary colors which are components of RGB primaries.

Mathematically, the conversion of RGB to CMY is

$$C = 1 - R$$
$$M = 1 - G$$
$$Y = 1 - B$$

* Combining all three CMY links ideally should produce black, but imperfections in actual ink pigments con gray instead.

This issue can be solved by CMYK where K is pure black component is added to CMY.

Applications:-

1) The CMYK model is worldwide standard for color printing in paper and printing press.

2) Graphic Design use CMYK for designing

3) Photographic Reproduction

4) Packaging & Product labels

5) Proofing and color management.

C) The HSI Color Model:-

The HSI color model (Hue, Saturation, Intensity) is a perceptual color space designed to closely match how humans describe and interpret colors, making it useful in image processing and computer vision application

Hue:- represents the actual color type and measured as an angle on color wheel.

Saturation:- describes the purity of color, expressed in percentage.

Intensity:- indicates the brightness or lightness of the color.

Applications:

- The HSI model is preferred for tasks requiring color segmentation, enhancement, object recognition and pattern analysis.
- Widely used in remote sensing, medical imaging.

Conversion of RGB to HSI:-

Consider an image in RGB color format,

Then the H component of each RGB pixel is obtained using

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

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

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

Finally Intensity component is given by

$$I = \frac{1}{3}(R+G+B)$$

Pseudocolor Image processing:

Pseudocolor (false color) image processing consits of assigning colors to gray values based on specified criterion. The term pseudo is used to differentiate process of assigning colors to mono-chrome images from processes associated with true color images.

The common pseudo color techniques are:

a. Intensity Slicing:

Intensity slicing, also known as density slicing or thresholding with color coding, is a technique used to map specific ranges of intensity value in grayscale image to distinct colors to enhance visualization and interpretation.

This method is useful to highlight the boundaries.

Principle:

The gray scale intensity range $[0, L-1]$ ($L = 256$ for 8 bits) is divided into several intervals by defining P slicing planes

These P planes split intensity scale into P+1 intervals

$V_1 = [0, l_1]$, $V_2 = [l_1, l_2]$, ------ $V_P = [l_{P-1}, l_P]$, $V_{P+1} = [l_P, L-1]$

* Each Intensity interval V_K is assigned a unique color C_K.
* For any pixel with the intensity $f(x,y)$, the pseudocolor assigned follows $f(x,y) = C_k$ if $f(x,y) \in V_K$.

b) Intensity to color transformations:-

Intensity to color transformations are advanced pseudo color image processing techniques that assign colors to grayscale intensity values by applying transformation to intensity data before mapping these values to RGB color channels.

This approach is more flexible and powerful than simple intensity slicing

Principle:

Instead of slicing intensity values to discrete values and assigning fixed colors, this method applies three seperate transformations to intensity values $f(x,y)$

-Each transformed value is fed independently into the red, green, blue color channels

$$R = f_r(f(x,y)), \quad G = f_g(f(x,y)), \quad B = f_b(f(x,y))$$

* The grayscale input passes in parallel through the three transformation functions.

* The outputs control intensity of Red, green, blue on RGB display, producing a composite pseudocolor image

Advantages:-

1, Allows smooth transition of colors over intensity range

2, Designed to highlight particular features in image by tailoring transformations.

3, Enables emphasis of specific intensity ranges by adjusting transformation parameters (phase, frq, amplitude).

Basics of Full-Color Image Processing

Full-color images consits of multiple components, typically three (RGB) in RGB color model.

Processing full-color images involving handle these components either individually or as color vectors.

Color-image representation:-

A color images pixel at position (x,y) is represented as a vector with components corresponding to channels

$$C(x,y) = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

For an image of size MxN there are MxN such vectors representing image

Processing Approaches:-

1) Component-wise processing

- Each color channel (RGB) is treated as individual grayscale image.
- Standard grayscale processing techniques are applied to each channel independently.
- After processing, the channels are combined again to form final color image.
- This approach is straightforward and leverges existing gray scale methods, but may not capture inter-channel color relationships fully

2) Vector-based processing:-

- Processes the color pixels as 3-dimensional vectors directly.
- Operations consider color vector as a whole, which can preserve relationships among colors better.
- This approach more complex and sometimes requires specialized algorithms.

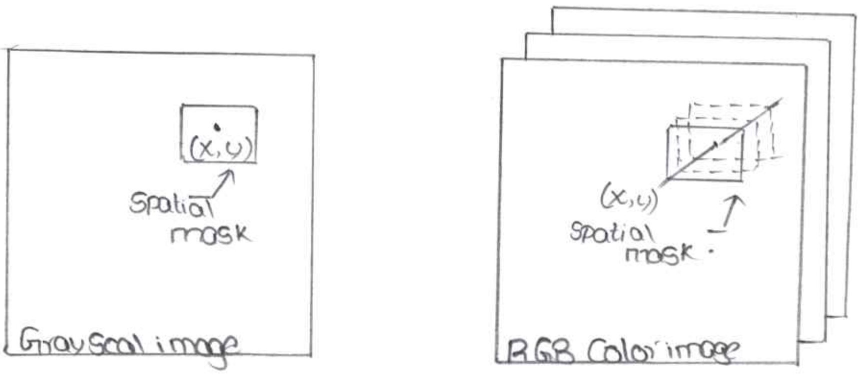


Fig:- Spatial masks for gray scale and RGB color images

A spatial mask (kernel) is a small grid that slides over a image to process each pixel based on neighborhood.

* In gray-scale, the mask cover several pixels. For each pixel location, you apply some operation (like average) to pixel values inside mask.

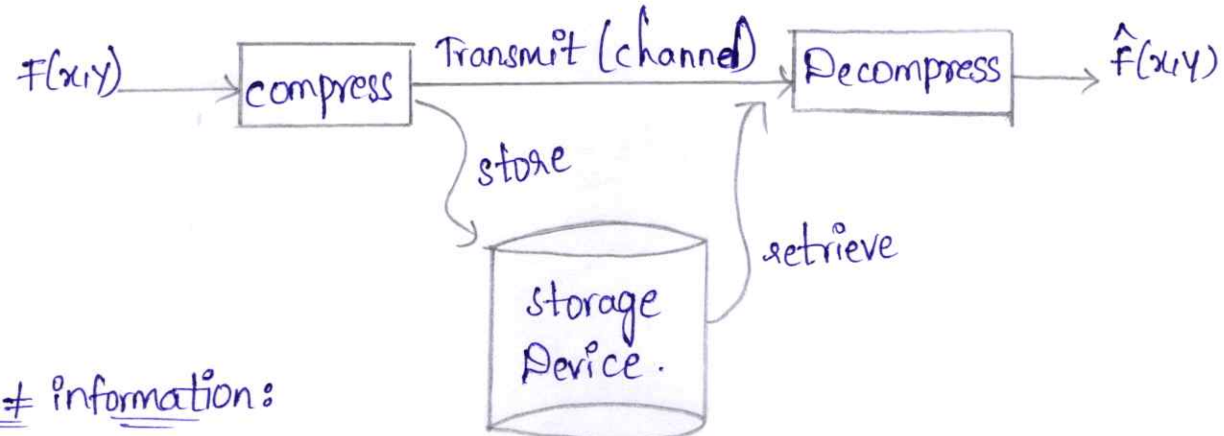
* In color image, sum the R,G,B values separately inside the mask, divide each sum by no of pixels and combine these averages into new color vector

Digital Image Processing

UNIT 5: IMAGE SEGMENTATION & COMPRESSION

*Definition: Image compression deals with reducing the amount of data required to present a digital image by removing of redundant data.

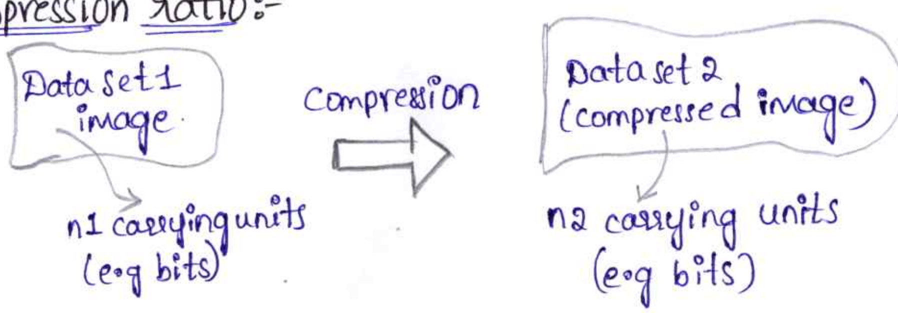
*Goal of Image Compression: The goal of image compression is to reduce the amount of data required a digital image.



Data ≠ information:

- Data and information are not synonymous terms!
- Data is the mean by which information is conveyed.
- Data compression aim storeduce the amount of data required to represent a given quality of information while preserving as much information as possible
- The same amount of information can be represent by various amount of data Ex1: you have an extra class after completion of 3.50 PM.
  Ex2: Extra class have been scheduled after 7th hour for you.

*Definition of compression ratio:-



Compression ratio $C_R = \frac{n_1}{n_2}$

* Definitions of Data Redundancy :-

→ Relative data redundancy : $\boxed{R_D = 1 - \frac{1}{C_R}}$

Example :

if $C_R = \frac{10}{1}$, then $R_D = 1 - \frac{1}{10} = 0.9$
(90% of the data in dataset 1 is redundant

if $n_2 = n_1$, then $C_R = 1$, $R_D = 0$
if $n_2 << n_1$, then $C_R \to \infty$, $R_D \to 1$

* Coding redundancy :-

- code : a list of symbols (Letters, numbers, bits etc..)
- code word : a sequence of symbol used to represent a piece of information or an event (e.g, gray levels)
- code word length : number of symbols in each codeword.

example : (binary code, symbols : 0, 1, Length : 3)

0 : 000 4 : 100
1 : 001 5 : 101
2 : 010 6 : 110
3 : 011 7 : 111

N x M image
r_k : k-th gray level
$p(r_k)$: probability of r_k.
$I(r_k)$: # of bits for r_k

Expected value :
$E(X) = \sum_x x p(X = x)$

Average # of bits : $L_{avg} = E(I(r_k)) = \sum_{k=0}^{L-1} I(r_k) p(r_k)$
Total # of bits : NML_{avg}

Coding Redundancy (con'd)

case 1 : $I(r_k)$ = constant length.

Assume an image with $L = 8$

Assume $I(r_k) = 3$, $L_{avg} = \sum_{k=0}^{7} 3 p(r_k) = 3 \sum_{k=0}^{7} p(r_k) = 3$ bits

Total number of bits : 3NM

case 2 : $I(r_k)$ = variable length.

$L_{avg} = \sum_{k=0}^{7} I(r_k) p(r_k) = 2.7$ bits $\qquad R_D = 1 - \frac{1}{1.11} = 0.099$

$C_R = \frac{3}{2.7} = 1.11$ (about 10%)

Interpixel redundancy

→ interpixel redundancy implies that pixel values are correlated pixel value can be reasonably predicted by its neighbors.

$$f(x) a g(x) = \int_{-\infty}^{\infty} f(x) g(x+a)\,da.$$

Interpixel redundancy (cont'd)

→ To reduce interpixel redundancy, the data must be transformed in another format (using a transformation)

- thresholding, DFT, DWT, etc.

Psychovisual redundancy

→ the human eye does not respond with equal sensitivity to all visual information.

It is more sensitive to the lower frequencies than to the higher frequency in the visual spectrum.

idea: discard data that is perceptually insignificant.

Fidelity criteria.

$$f(x,y) \longrightarrow \boxed{\text{compress}} \longrightarrow g(x,y) \longrightarrow \boxed{\text{Decompress}} \longrightarrow \hat{f}(x,y)$$

$$\hat{f}(x,y) = f(x,y) + e(x,y)$$

→ How close $f(x,y)$ $\hat{f}(x,y)$?

→ criteria.

- subjective:- based on human observers
- objective: Mathematically defined criteria.

Subjective fidelity criteria.

Value	Rating	Description.
1.	Excellent	An image of extremely high quality, as good as you could desire.
2.	Fine	An image of high quality, providing enjoyable viewing Interference is not objectionable.
3.	passable	An image of acceptable quality Interference is not objectionable.

Value	Rating	Description
4.	Marginal	An image of poor quality: you wish you could improve it.
5.	Inferior	A very poor image but you could watch it objectionable interference is definitely present.
6.	unusable	An image so bad that you could not watch it.

Objective fidelity criteria.

→ Root mean square error (RMS)

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

→ Mean-square

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

Compression Methods of images:—

Compression methods can be lossy

→ when a tolerable degree of deterioration in the visual quality of the resulting image is acceptable or Lossless.

→ As a general guideline, lossy compression should be used for general purpose photographic images.

→ whereas lossless compression should be preferred when dealing with line art, technical drawings, cartoons etc.

Fundamentals of visual data compression.

The general problem of image compression is to reduce the amount of data required to represent a digital image or video and the underlying basis of the reduction process is the removal of redundant data. Mathematically, visual data compression typically involves transforming (encoding) a 2-D pixel array into a statistically uncorrelated data set.

this transformation is applied prior to storage a transmission. At some later time the compressed image is decompressed to reconstruct the original image information (preserving or lossless techniques)

Redundancy:-

If the same information can be represented using different amounts of data, it is reasonable to believe that the representation that the representation that requires more data contains what is technically called data redundancy.

- coding redundancy: consists in using variable-length code words selected as to match the statistics of the original source in this case the image itself or a processed version of its pixel values.
- interpixel redundancy: this type of redundancy - sometimes called spatial redundancy inter frame redundancy or geometric redundancy - exploits the fact that an image very often contains strongly correlated pixels. Examples of compression techniques that explore the interpixel redundancy include: constant Area coding (CAC), (1-D or 2-D) Run-Length Encoding (RLE) techniques, and many predictive coding algorithms such as Different pulse code modulation (DPCM)
- psycho visual redundancy: Many experiments on the psychophysical aspects of human vision have proven that the human eye does not respond with equal sensitivity to all incoming visual information the loss of quality that ensues as a byproduct of such techniques is frequently called quantization. Most of the image coding algorithms in use today exploit this type of redundancy such as the Discrete cosine transform (DCT) based algorithm at the heart of the JPEG encoding standard.

Image compression and coding models:-

Figure shows a general image compression model. It consists of source encoder a channel encoder, the storage or transmission media (also referred to as channel) a channel decoder and a source decoder.

At the receiver's side the channel and source decoder perform the opposite functions and ultimately recover (an approximation of) the original image.

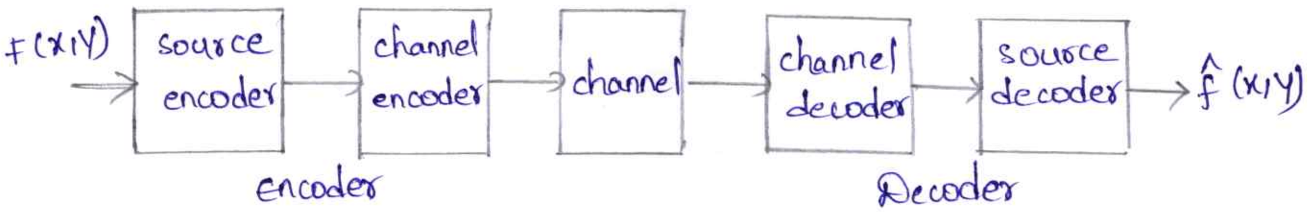


Figure: image compression model

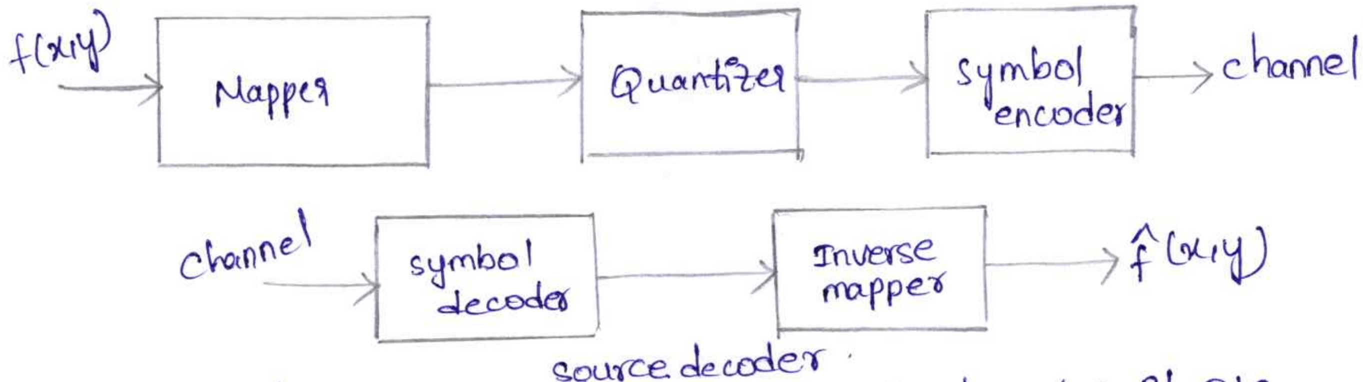


figure:- Shows that source encoder in further detail. Its main components are.

- Mapper :- transforms the input data into a (usually nonviscual) format designed to reduce interpixel redundanices in the input image.
- Quantizer :- reduces the accuracy of the mapper's output in accordance with same preestablished fidelity criterion. Reduces the psychovisual redundancies of the input image.
- Symbol (entropy) encoder : Creates a fixed-or-variable length code to represent the quantizer's output and maps the output in accorda-nce with code.

*Error-free compression:-

Error-free compression techniques usually rely on entropy-based encoding algorithms. the concept of entropy is mathematically described in equation (1)

where
a_j is a symbol produced by the information source.
$P(a_j)$ is the probability of that symbol.

J is the total number of different symbols.

H(z) is the entropy of the source

The concept of entropy provides an upper bound on how much compression can be achieved given the probability distribution of the source. In other words, it establishes a theoretical limit on the amount of losses compression that can be achieved using entropy encoding techniques alone.

Variable Length coding (VLC)

Most entropy-based encoding techniques rely on assigned variable-length codeworks to each symbol whereas the most likely symbols are assigned shorter codewords.

It is described in more detail in a separate short article.

Run-length encoding (RLE)

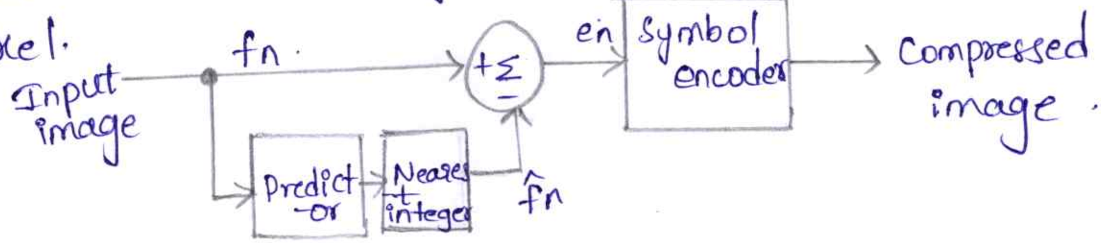
RLE is one of the simplest data compression techniques. It consists of replacing a sequence (run) of identical symbols by a pair containing the symbol and the run length. It is used as the primary compression technique in the 1-D CCITT Group 3 fax standard and in conjunction with other techniques in the JPEG image

Differential coding:-

Differential coding techniques explore the interpixel redundancy in digital images. As a consequence of this narrower distribution - and consequently reduced entropy - Huffman coding or other VLC schems will produce shorter code words for the difference image.

Predictive coding

predictive coding techniques constitute another example of exploration of interpixel redundancy, in which in which the basic idea is to encode only the new information in each pixel.



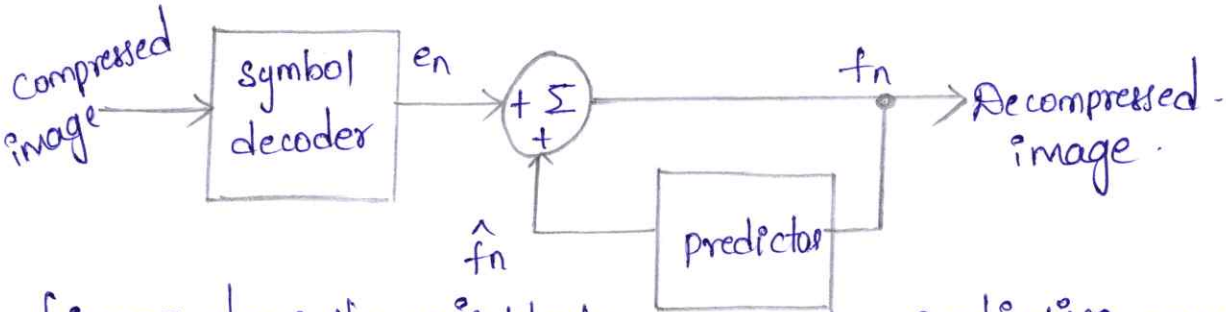


figure 3 shows the main blocks of a lossless predictive encoder the key component is the predictor, whose function is to generate an estimated (predicated) value for each pixel from the input image based on previous pixel values.

Dictionary-based coding

Dictionary-based coding techniques are based on the idea of incrementally building a dictionary (table) there is no need to calculate, store or transit the probability the best-known variant of dictionary-based coding algorithms is the LZW (Lempel Ziv-Welch)

*Lossy compression:—

Lossy compression techniques deliberately introduce a certain amount of distortion to the encoded image. the amount of error (loss) and the resulting bit savings.

Quantization.

the quantitation stage is at the core of any lossy image encoding algorithm. Quantization in at the encoder side means partitioning of the input data range into a smaller set of values. Because of its fast lookup capabilities at the decoder side VQ-based coding schems are particularly attractive to multimedia applications.

Transform coding

the techniques discussed so far work directly on the pixel values and are usually called spatial domain techniques the Discrete cosine Transform (DCT) has become the most widely used transform coding technique.

Wavelet coding

Wavelet coding techniques are also based on the idea that the coefficients of a transform that decorrelates the pixels of an image can be coded more efficiently than the original pixels themselves. wavelet coding has been at the core of the latest image compression standards most notably JPEG 2000, which is discussed in a separate short article.

* Image compression standards:-

Work on international standards for image compression started in the late 1970s with the CCITT (currently ITU-T) need to standardize binary image compression algorithms for Group 3 facsimile communications. (such as JPEG) 1. easier exchage of image files between different devices and applications.

2. existence of benchmarks and reference data sets for new and alternative development.

Binary Image compression standarts.

Work on binary image compression standards was initially motivated by CCITT Group 3 and 4 facsimile standards. the Group 3 standard uses a non-adaptive, 1-D RLE technique in which the last K-1 lines. the Joint Bilevel Image Group (JBIG) - a joint committee of the ITU-T and ISO - has addressed these limitations and proposed two new standards. (JBIG and JBIG2)

Continuous-tone still image compression standards.

1. Entropy coding:- Every block of an image is entropy encoded based upon the pk's within a block. the produces variable length code for each block depending on spatial activities within the blocks

2. Run-Length Encoding:- scan the image horizontally or vertically and while scanning assign a group of pixel with same intensity into a pair (gi li) where gi is the intensity and li is the length of the "run"

Example 1:- consider the following 8X8 image.

4	4	4	4	4	4	4	0
4	5	5	5	5	5	4	0
4	5	6	6	6	5	4	0
4	5	6	7	6	5	4	0
4	5	6	6	6	5	4	0
4	5	5	5	5	5	4	0
4	4	4	4	4	4	4	0
4	4	4	4	4	4	4	0

the run-length codes using vertical (continuous-top-down) scanning mode are:

(4,9) (5,5) (4,3) (5,1) (6,3)
(5,1) (4,3) (5,1) (6,1) (7,1)
(6,1) (5,1) (4,3) (5,1) (6,3)
(5,1) (4,3) (5,5) (4,10) (0,8)

total of 20 pairs = 40 numbers. the horizontal scanning would lead to 34 pairs = 68 numbers, which is more than the actual number of pixels (64)

Example 3:- For the same image in the previous example which requires 3 bits/pixel using standard PCM we can arrange the table.

Gray levels	# occurrences	P_k	C_k	B_k	$P_k B_k$	$-P_k \log_2 P_k$
0	8	0.125	0000	4	0.5	0.375
1	0	0	-	0	-	-
2	0	0	-	0	-	-
3	0	0	-	0	-	-
4	31	0.484	1	1	0.484	0.507
5	16	0.25	01	2	0.5	0.5
6	8	0.125	001	3	0.375	0.375
7	1	0.016	0001	4	0.64	0.095
	64	1			R	H

codewords C_ks are obtained by constructing the binary tree as in fig 5

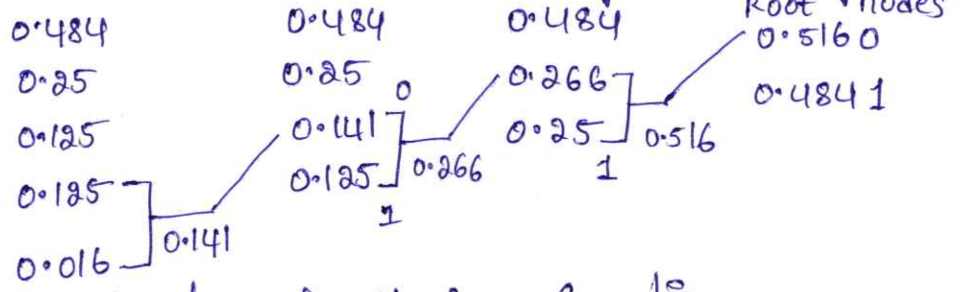


fig:- Tree structure for Huffman Encoding

Note that in this case we have.

$$R=\sum_{K=1}^{8} B_k P_K = 1.923 \text{ bits/pixel}.$$

$$H=-\sum_{K=1}^{8} P_k \log_2 P_K = 1.852 \text{ bits/pixel}$$

thus.

$$1.852 \le R = 1.923 \le H+\frac{1}{L}=1.977$$

i.e an average of 2 bits/pixel (instead of 3 bits/pixel using PCM) can be used to code the image.

*Predictive Encoding:-

Idea: Remove mutual redundancy among successive pixels in a region of support (ROS) or neighborhood and encode only the new informatio-n. this method is based upon linear predication. Let us start with 1-D linear predicators. An N^{th} order linear prediction of $x(n)$ based on N previous samples is generated using a 1-D autoregressive (AR) model

$$\hat{x}(n)=a_1x(n-1)+a_2x(n-2)+\cdots+a_N x(n-N)$$

Now instead of encoding $x(n)$ the predication error.

$$e(n)=x(n)-\hat{x}(n)$$

$x(n)$ using the previous encoded values $x(n-k)$ and the encoded error signal.

$$x(n)=\hat{x}(n)+e(n)$$

this method is also referred to as differential PCM (DPCM)

Minimum variance prediction

The predictor

$$\hat{x}(n)=\sum_{i=1}^{N} a_i x(n-i)$$

is the best N^{th} order linear mean-squared predictor of $x(n)$, which minimizes the MSE

$$e=E\left[(x(n)-\hat{x}(n))^2\right]$$

this minimization wrt a_k's results in the following "orthogonal property"

$$\frac{\partial e}{\partial a_k}=-2E\left[(x(n)-\hat{x}(n))x(n-k)\right]=0 \quad 1\le K\le N$$

which leads to the normal equation

$$r_{xx}(K)-\sum_{i=1}^{N} a_i r_{xx}(K-i)=\sigma_e^2\delta(k), \quad 0\le K\le N$$

where $r_{xx}(k)$ is the autocorrelation of the data $x(n)$ and σ_e^2 is the variance of the driving process $e(n)$

→ To understand the need for compact image representation consider the amount of data required to represent a 2 hour standard Definition (SD) using 720X480X24 pixel arrays.

→ A video is a sequence of video frames where each frame is full color still image Because video player must display the frames sequentially at rates near 30 fps. standard definition data must be accessed 30fps X (720X480) ppf X 3bpp = 31,104,00 bps

-fps: frames per second ppf: pixels per frame, bpp: bytes per pixel, bps: bytes per second.

→ Thus a 2 hour movie consists of : = 31,104,000bps X (60^2) sph X 2hrs where sph is second per hour = 2.24×10^{11} bytes = 224 GB of data.

TWENTY SEVEN 8.5 GB dual layer DVD's are needed to store it

Time required to transmit a small 128X128X24 bit full color image over this range of speed is from 7.0 to 0.03 sec.

Data vs information:

Data and information is not the same thing; data are the means by which information is conveyed Because various amounts of data can be used to represent the same amount of information. A comparative study of performance of a variety of different image transforms is done base on compression ratio, entropy and time factor.

* the flow of image compression coding:-

Image compression coding is store the image into bit-stream as compact as possible and to display the decoded image in the monitor as exact as possible. Now consider an encoder and a decoder as shown in fig 1.3

If the total data quantity of the bit stream is less than the total data quantity of the original image then this is called image Compression. the full compression flow is as shown in fig 1.3.

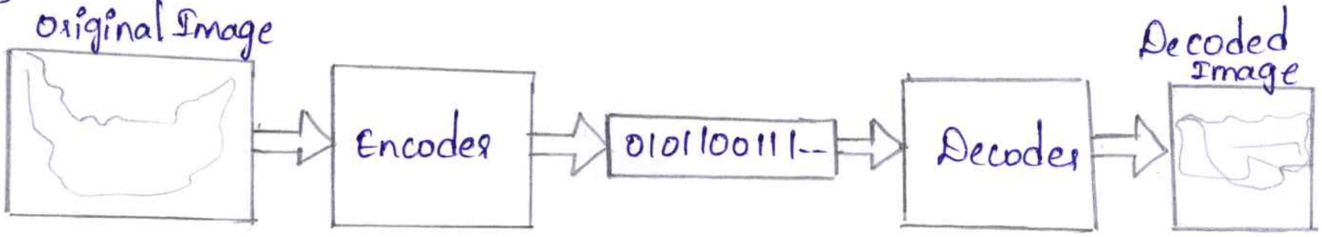


Fig 1.3 The basic flow of image compression coding.

The compression ratio is defined as follows:

$$Cr = \frac{n1}{n2}$$

Where n1 is the data rate of original image and n2 is that of the encoded bit-stream.

In order to evaluate the performance of the image compression coding. it is necessary to define a measurement that can estimate the difference between the original image and the decoded image.

$$MSE = \sqrt{\sum_{x=0}^{W-1}\sum_{y=0}^{H-1}\left[f(x,y)-f(x,y)\right]^2}$$

$$PSNR = 20\log 10 \frac{255}{MSE}$$

the general encoding architecture of image compression system is shaon is fig 1.4 the fundamental theory and concept of each functional block will introduced in the following sections

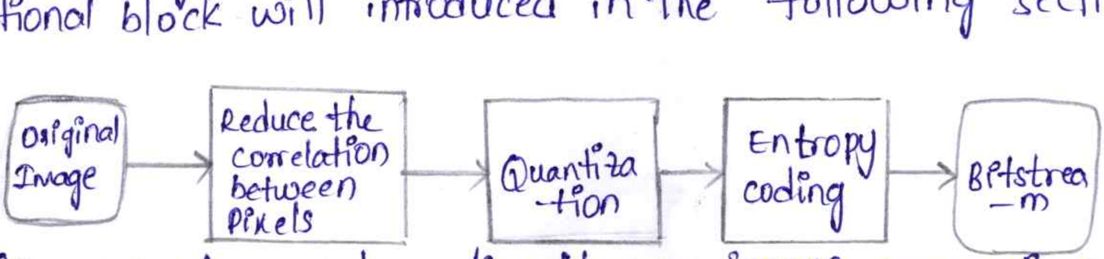


fig:- 1.4 the general encoding flow of image compression

Reduce the correlation between pixels.

the reason is that the correlation between one pixel and its neighour pixels is very high there are a lot of relevant processing methods being proposed.

the best known methods are as follows.

- Predictive coding: predictive coding such as DPCM (Differential pulse code Modulation) is a lossless coding method which means that the decoded image.
- Orthogonal Transform: Karhunen-Loeve Transform (KLT) and Discrete cosine Transform (DCT) are the two most well-known orthogonal transforms. the DCT-based image compression standard such as JPEG is a lossy coding method.
- subband coding: subband coding such as Discrete wavelet Transform (DWT) is also a lossy coding method.

Quantization:

the objective of quantization is to reduce the precision and to achieve higher compression ratio. For instance the original image uses 8 bits to store one element for every pixel quantity the image compression standards such as JPEG and JPEG 2000 have their own quantization methods, and the details of relevant theory will be introduced in the chapter 2.

Entropy coding:

the main objective of entropy coding is to achieve less average length of the image. Entropy coding assigns codewords to the corresponding symbols according to the probability of the symbols. the entropy encoder of JPEG and JPEG 2000 will also be introduced.

*Image compression standard:—

In this chapter we will introduce the fundamental theory of two well-known image compression standards - JPEG and JPEG 2000

JPEG - joint picture expert Group

fig. 2.1 and 2.2 shows the Encoder and Decoder model of JPEG. We will introduce the operation and fundamental theory of each block in the following sections.

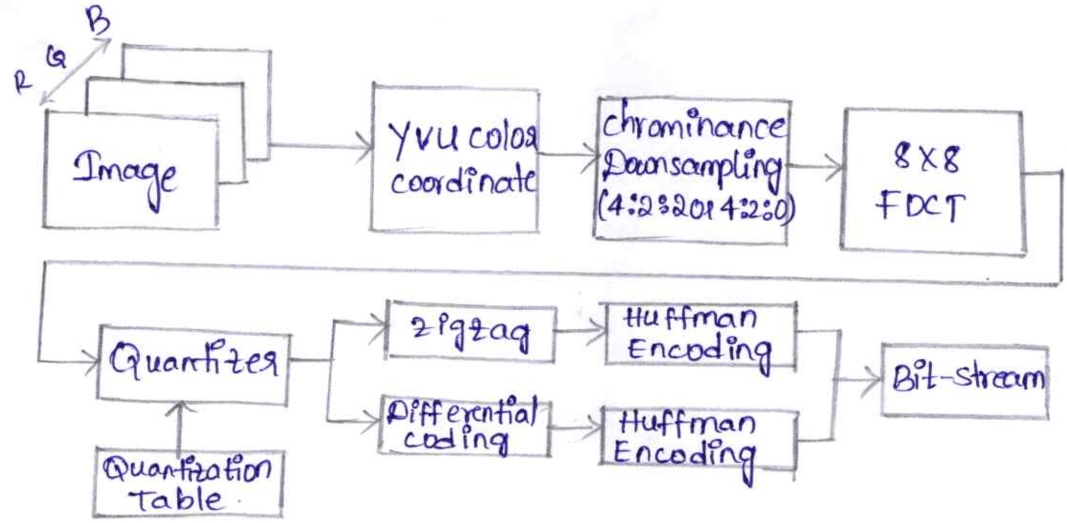


fig: 2.1 Encoder model of JPEG compression standard.

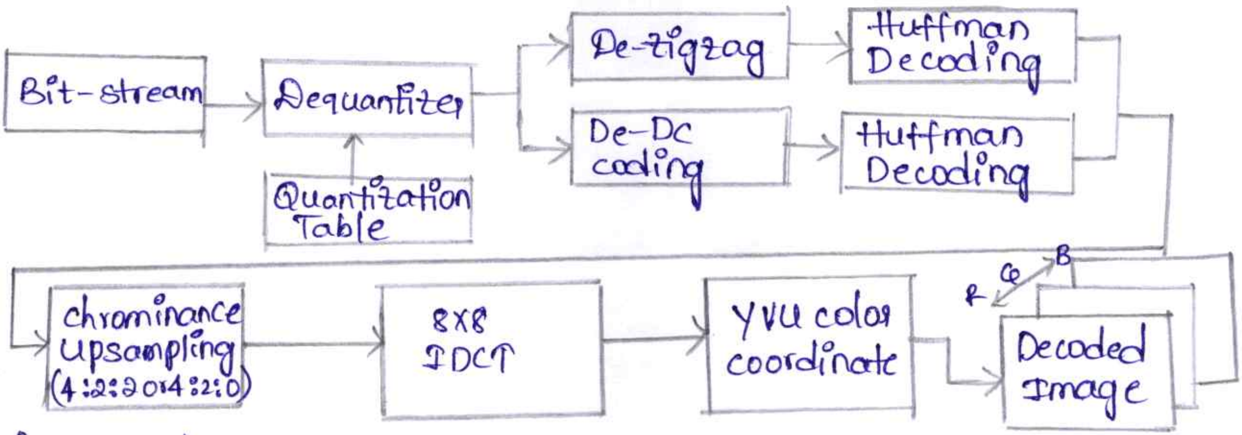


fig: 2.2 The Decoder model of JPEG compression standard

Dicrete cosine transform

the next step after color coordinate conversion is to divide the three colour components of the image into many 8x8 blocks the mathematical definition of the Forward DCT and the Inverse DCT are as follows

Forward DCT

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

for $u=0,\ldots,N-1$ and $v=0,\ldots,N-1$

where $N=8$ and $c(k)=\begin{cases}1/\sqrt{2} & \text{for } k=0\\ 1 & \text{otherwise}\end{cases}$

Inverse DCT

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

for $x=0,\ldots N-1$ and $y=0\ldots N-1$ where $N=8$

the $f(x,y)$ is the value of each pixel in the selected 8x8 block and the $F(u,v)$ is the DCT coefficient after transformation

1. The DCT can concentrate the energy of the transformed signal in Low frequency
2. For image compression the DCT can reduce the blocking effect than the DFT

* Quantization in JPEG:—

Quantization is the step where we actually throw away data the DCT is a Lossless procedure. the quantized coefficient is defined in (2.3) and the reverse the process can be achieved by the (2.4)

$$F(u,v)_{quantization} = \text{round}\left(\frac{F(u,v)}{Q(u,v)}\right)$$

$$F(u,v)_{deQ} = F(u,v)_{quantization} \times Q(u,v)$$

the JPEG Committee has recommended certain Q matrix for luminance and chrominance components is defined in (2.5) and (2.6)

$$Q_Y=\begin{bmatrix}16 & 11 & 10 & 16 & 24 & 40 & 51 & 61\\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55\\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56\\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62\\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77\\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92\\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101\\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99\end{bmatrix}$$

condition the Q Matrix for luminace and chrominance components is defined in (2.5) and (2.6)

$$
Q_C = \begin{bmatrix}
17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\
18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\
24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\
47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\
99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\
99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\
99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\
99 & 99 & 99 & 99 & 99 & 99 & 99 & 99
\end{bmatrix}
$$

* **Zigzag scan:-**

After quantization the DC coefficient is treated separately from the 63 AC coefficients the DC coefficient is a measure of the average value of the original 64 image samples. the other 63 entries are the AC components They are treated separately from the DC coefficients in the entropy coding Process

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

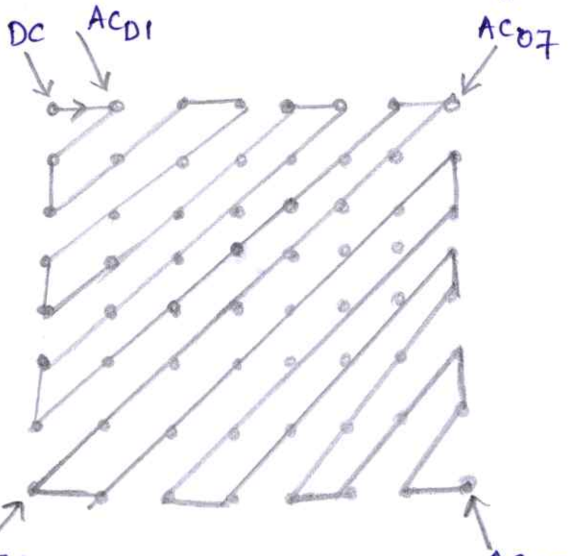


fig 2.3 the zigzag scan order

* **Entropy coding in JPEG Differential coding:—**

The Mathematical representation of the differential coding is

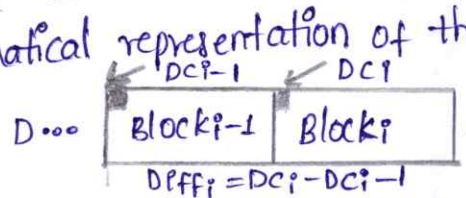


fig:-2.4 Differential coding.

We set $DC_0 = 0$. DC of the current block DC_i will be equal to $DC_{i-1} + Diff_i$ therefore in the JPEG file then the difference is encoded with Huffman coding alogorithm together with the encoding of AC coefficients.