DIGITAL IMAGE PROCESSING

SUBJECT NAME : DIGITAL IMAGE PROCESSING (EC5A1)

UNIT I

PART- A

1. Define Image
An image may be defined as two dimensional light intensity function f(x, y) where x and y denote spatial co-ordinate and the amplitude or value of f at any point (x, y) is called intensity or grayscale or brightness of the image at that point.

2. What is Dynamic Range?
The range of values spanned by the gray scale is called dynamic range of an image. Image will have high contrast, if the dynamic range is high and image will have dull washed out gray look if the dynamic range is low.

3. Define Brightness
Brightness of an object is the perceived luminance of the surround. Two objects with different surroundings would have identical luminance but different brightness.

5. What do you meant by Gray level?
Gray level refers to a scalar measure of intensity that ranges from black to grays and finally to white.

6. What do you meant by Color model?
A Color model is a specification of 3D-coordinates system and a subspace within that system where each color is represented by a single point.

7. List the hardware oriented color models
1. RGB model
2. CMY model
3. YIQ model
4. HSI model

8. What is Hue and saturation?
Hue is a color attribute that describes a pure color where saturation gives a measure of the degree to which a pure color is diluted by white light.

9. List the applications of color models
1. RGB model--- used for color monitors & color video camera
2. CMY model---used for color printing
3. HIS model----used for color image processing
4. YIQ model---used for color picture transmission

10. What is Chromatic Adoption?
` The hue of a perceived color depends on the adoption of the viewer. For example, the American Flag will not immediately appear red, white, and blue of the viewer has

been subjected to high intensity red light before viewing the flag. The color of the flag will appear to shift in hue toward the red component cyan.

11. Define Resolutions

Resolution is defined as the smallest number of discernible detail in an image. Spatial resolution is the smallest discernible detail in an image and gray level resolution refers to the smallest discernible change is gray level.

12. What is meant by pixel?

A digital image is composed of a finite number of elements each of which has a particular location or value. These elements are referred to as pixels or image elements or picture elements or pels elements.

13. Define Digital image

When x, y and the amplitude values of f all are finite discrete quantities , we call the image digital image.

14. What are the steps involved in DIP?

1. Image Acquisition
2. Preprocessing
3. Segmentation
4. Representation and Description
5. Recognition and Interpretation

15. What is recognition and Interpretation?

Recognition means is a process that assigns a label to an object based on the information provided by its descriptors.

Interpretation means assigning meaning to a recognized object.

16. Specify the elements of DIP system

1. Image Acquisition
2. Storage
3. Processing
4. Display

17. List the categories of digital storage

1. Short term storage for use during processing.
2. Online storage for relatively fast recall.
3. Archival storage for infrequent access.

18. What are the types of light receptors?

The two types of light receptors are

- Cones and
- Rods

19. Differentiate photopic and scotopic vision

Photopic vision	Scotopic vision

1. The human being can resolve the fine details with these cones because each one is connected to its own nerve end. 2. This is also known as bright light vision.	Several rods are connected to one nerve end. So it gives the overall picture of the image. This is also known as thin light vision.

20. How cones and rods are distributed in retina?

In each eye, cones are in the range 6-7 million and rods are in the range 75-150 million.

21. Define subjective brightness and brightness adaptation

Subjective brightness means intensity as preserved by the human visual system.

Brightness adaptation means the human visual system can operate only from scotopic to glare limit. It cannot operate over the range simultaneously. It accomplishes this large variation by changes in its overall intensity.

22. Define weber ratio

The ratio of increment of illumination to background of illumination is called as weber ratio.(ie) $\Delta i/i$

If the ratio ($\Delta i/i$) is small, then small percentage of change in intensity is needed (ie) good brightness adaptation.

If the ratio ($\Delta i/i$) is large , then large percentage of change in intensity is needed (ie) poor brightness adaptation.

23. What is meant by machband effect?

Machband effect means the intensity of the stripes is constant. Therefore it preserves the brightness pattern near the boundaries, these bands are called as machband effect.

24. What is simultaneous contrast?

The region reserved brightness not depend on its intensity but also on its background. All centre square have same intensity. However they appear to the eye to become darker as the background becomes lighter.

25. What is meant by illumination and reflectance?

Illumination is the amount of source light incident on the scene. It is represented as i(x, y).

Reflectance is the amount of light reflected by the object in the scene. It is represented by r(x, y).

26. Define sampling and quantization

Sampling means digitizing the co-ordinate value (x, y).

Quantization means digitizing the amplitude value.

27. Find the number of bits required to store a 256 X 256 image with 32 gray levels

32 gray levels = 2^5

= 5 bits

$256 * 256 * 5 = 327680$ bits.

28. Write the expression to find the number of bits to store a digital image?

The number of bits required to store a digital image is

$$b=M \times N \times k$$

When M=N, this equation becomes

$$b=N^2k$$

29. Write short notes on neighbors of a pixel.

The pixel p at co-ordinates (x, y) has 4 neighbors (ie) 2 horizontal and 2 vertical neighbors whose co-ordinates is given by (x+1, y), (x-1,y), (x,y-1), (x, y+1). This is called as direct neighbors. It is denoted by $N_4(P)$

Four diagonal neighbors of p have co-ordinates (x+1, y+1), (x+1,y-1), (x-1, y-1), (x-1, y+1). It is denoted by $N_D(4)$.

Eight neighbors of p denoted by $N_8(P)$ is a combination of 4 direct neighbors and 4 diagonal neighbors.

30. Explain the types of connectivity.

1. 4 connectivity
2. 8 connectivity
3. M connectivity (mixed connectivity)

31. What is meant by path?

Path from pixel p with co-ordinates (x, y) to pixel q with co-ordinates (s,t) is a sequence of distinct pixels with co-ordinates.

32. Give the formula for calculating D_4 and D_8 distance.

D_4 distance (city block distance) is defined by

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

D_8 distance(chess board distance) is defined by

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

33. What is geometric transformation?

Transformation is used to alter the co-ordinate description of image.
The basic geometric transformations are

1. Image translation
2. Scaling
3. Image rotation

34. What is image translation and scaling?

Image translation means reposition the image from one co-ordinate location to another along straight line path.

Scaling is used to alter the size of the object or image (ie) a co-ordinate system is scaled by a factor.

35. Define the term Luminance

Luminance measured in lumens (lm), gives a measure of the amount of energy an observer perceiver from a light source.

PART-B

1. Explain Brightness adaptation and Discrimination

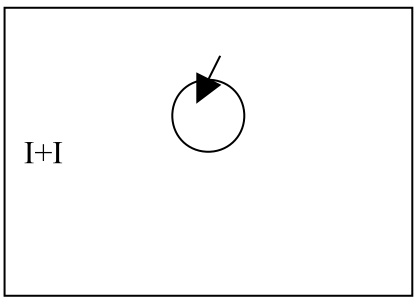
The digital images are displayed as a discrete set of intensities, the eye's ability to discriminate between different intensity levels.

Subjective brightness is a logarithmic function of the light intensity incident on the eye. The long solid curve represents the range of intensities t o which the visual system can adapt.

In photopic vision alone the range is about 10^6.

It accomplishes the large variation by changes in its overall sensitivity phenomenon is known as brightness adaptation.

The eye's ability to discriminate between different intensity levels at any specific adaptation.



The eye is capable of detecting contouring effects in monochrome Image whose overall intensity is represented by fewer than approximately two dozen levels. The second phenomenon called simultaneous contrast is related to the fact that a region's perceived brightness does not depend on its intensity. They app ear to the eye become dark eras the background gets lighter.

2.Explain sampling and quantization:

For computer processing, the image function f(x,y)must be digitized both spatially and in amplitude. Digitization of spatial co-ordinates is called image sampling and amplitude digitization is called grey level quantization.

Sampling:

Consider a digital image of size 1024*1024,256 with a display area used for the image being the same ,the pixels in the lower resolution images where duplicated inorder to fulfill the entire display .the pixel replication produced a checker board effect, which is visible in the image of lower resolution .it is not possible to differentiate a 512*512 images from a1024*1024 under this effect. but a slight increase in graininess and a small decrease in sharpness is noted.

A 256*256 image shows a fine checker board pattern in the edges and more pronounced graininess there out the image .these effect is much more visible in 128*128 images and it becomes quite pronounced in 64*64 and 32*32 images.

Quantization:

It discusses the effects produced when the number of bits used to represent the grey level in an image is decreased .this is illustrated by reducing the grey level required to represent a 1024*1024,512 image.

The 256,128,and 64 level image are visually identical for all practical purposes the 32 level images has developed a set of rigid like structure in areas of smooth grey lines.this effect caused by the user insufficient number of grey levels in smooth areas of digital image is called a false contouring.this is visible in images displayed using 16 or lesser gray level values.

3.Explain about Mach band effect?

Two phenomena demonstrate that perceived brightness is not only a function of intensity. They are mach band pattern and simultaneous contrast.

Mach band pattern:

It states that the visual system tends to undershoot or overshoot around the boundary of regions of different intensities .This is called mach band pattern. Although the width of the stripe is constant, it is perceived as if the brightness pattern is strongly scalloped near the boundaries by darker part.

Simultaneous contrast is related to the fact that a regions perceived brightness does not depend only on its intensity. In the figure all the center square have the same intensity however they appear to the eye as the background gets lighter.

Example: A piece of paper seems white when lying on the desk but can appear when used to shield the eyes while looking at brighter sky.

4. Explain color image fundamentals.

Although the process followed by the human brain in perceiving and interpreting color is a physiopsychological phenomenon that is not yet fully understood, the physical nature of color can be expressed on a formal basis supported by experimental and theoretical results.

Basically, the colors that humans and some other animals perceive in an object are determined by the nature of the light reflected from the object. The visible light is composed of a relatively narrow band of frequencies in the electromagnetic spectrum. A body that reflects light that is balanced in all visible wavelengths appears white to the observer. For example, green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths.

Three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance and brightness. Radiance is the total amount of energy that flows from the light source, and is usually measured in watts(W). Luminance, measured in lumens(lm), gives a measure of the amount of energy an observer perceives from a loght source. Finally, brightness is a subjective descriptor that is practically impossible to measure.

5. Explain CMY model.

This model deals about the cyan,magenta and yellow are the secondary colors of light.When a surface coated with cyan pigment is illuminated with white light no red lihgt is reflected from the surface.Cyan subtracts red light from reflected white light,which itself is composed of equal amounts of red,green and blue light.in this mode cyan data input or perform an RGB to CMY conversion internally.

$C = 1 - R$

$M = 1 - G$

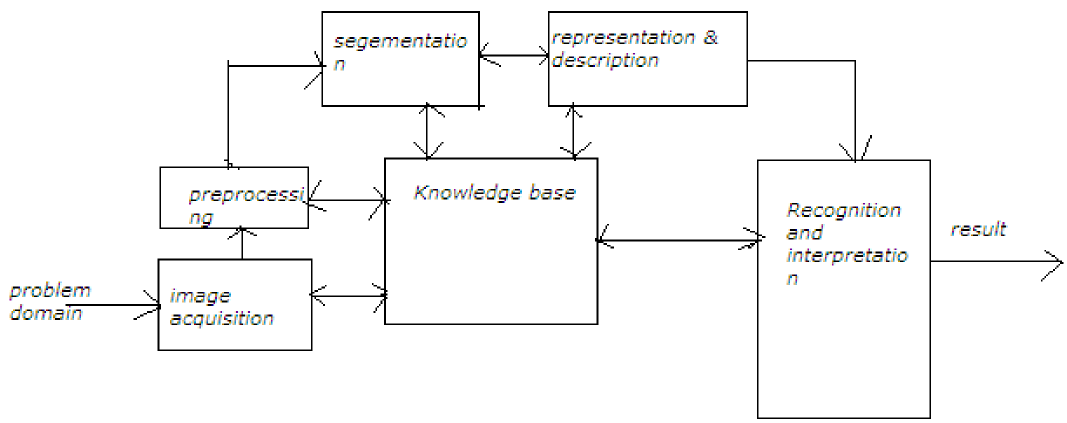
$Y = 1 - B$

All color values have been normalized to the range [0,1].the light reflected from a surface coated with pure cyan does not contain red .RGB values can be obtained easily from a set of CMY values by subtracting the individual Cmy values from 1.Combining these colors prosuces a black .When black is added giving rise to the CMYK color model.This is four coluring printing .

PART-C

1. Describe the fundamental steps in image processing?

Digital image processing encompasses a broad range of hardware, software and theoretical underpinnings.



The problem domain in this example consists of pieces of mail and the objective is to read the address on each piece. Thus the desired output in this case is a stream of alphanumeric characters.

The first step in the process is image acquisition that is acquire a digital image .To do so requires an imaging sensor and the capability to digitize the signal produced by the sensor.

After the digital image has been obtained the next step deals with preprocessing that image. The key function of this is to improve the image in ways that increase the chances for success of the other processes.

The next stage deals with segmentation. Broadly defined segmentation partitions an input image into its constituent parts or objects. The key role of this is to extract individual characters and words from the background,

The output of the segmentation stage usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself.

Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description also called feature selection deals with extracting features that result in some quantitative information of interest that are basic for differentiating one class of object from another.

The last stage involves recognition and interpretation. Recognition is the process that assigns a label to an object based on the information provided by its descriptors. Interpretation involves assigning meaning to an ensemble of recognized objects.

Knowledge about a problem domain is coded into an image processing system in the form of knowledge database. This knowledge may be simple as detailing regions of an image where the information of interest is known to be located thus limiting the search that has to be conducted in seeking that information.

The knowledge base also can be quite complex such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high resolution satellite images of a region in connection with change detection application.

Although we do not discuss image display explicitly at this point it is important to keep in mind that viewing the results of image processing can take place at the output of any step.

2. Explain the basic Elements of digital image processing:

Five elements of digital image processing,

- image acquisitions
- storage
- processing
- communication
- display

1)Image acquisition :

Two devices are required to acquire a digital image ,they are

1)physical device:

Produces an electric signal proportional to the amount of light energy sensed.

2)a digitizer:

Device for converting the electric output into a digital form.

2.storage:

An 8 bit image of size 1024*1024 requires one million bits of storage.three types of storage:

1.short term storage:

It is used during processing. it is provide by computer memory. it consisits of frame buffer which can store one or more images and can be accessed quickly at the video rates.

2.online storage:

It is used for fast recall. It normally uses the magnetic disk,Winchester disk with100s 0f megabits are commonly used .

3.archival storage:

They are passive storage devices and it is used for infrequent access.magnetic tapes and optical disc are the media. High density magnetic tapes can store 1 megabit in about 13 feet of tape .

3)Processing:

Processing of a digital image p involves procedures that are expressedin terms of algorithms .with the exception of image acquisition and display most image processing functions can be implemented in software .the need for a specialized hardware is called increased speed in application. Large scale image processing systems are still being used for massive image application .steps are being merge for general purpose small computer equipped with image processing hardware.

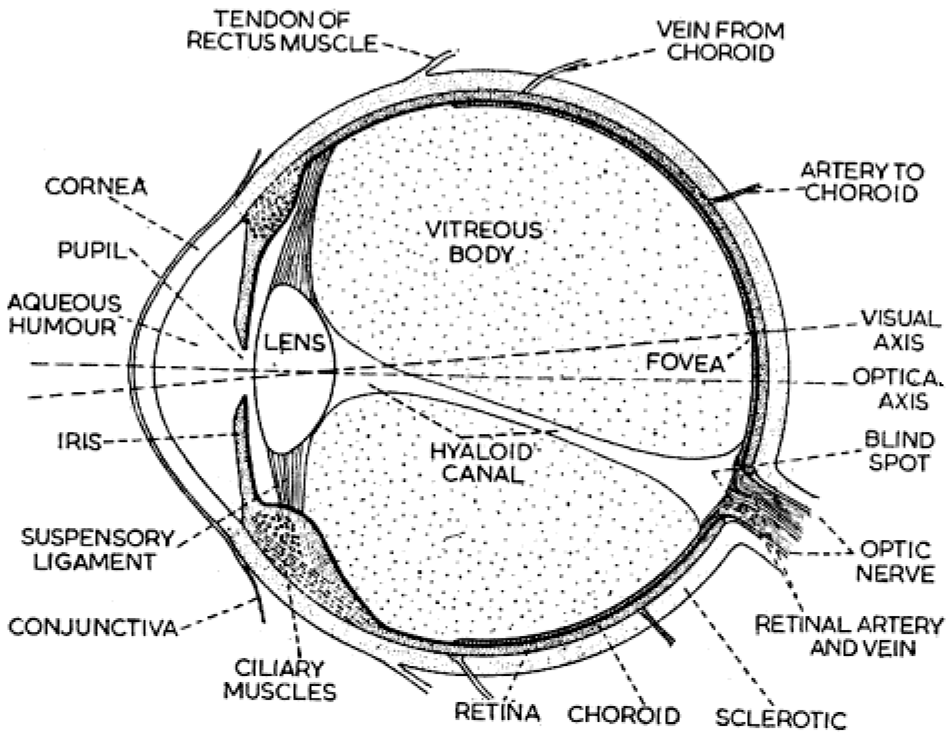
4)communication:

Communication in ip involves local communication between ip systems and remote communication from one point to another in communication with the transmission of image hardware and software are available for most of the computers .the telephone line can transmit a max rate of 9600 bits per second.so to transmit a 512*512,8 bit image at this rate require at last 5 mins.wireless link using intermediate stations such as satellites are much faster but they are costly.

5)display:

Monochrome and colour tv monitors are the principal display devices used in modern ips.monitors are driven by the outputs of the hardware in the display module of the computer.

3. Explain the Structure of the Human eye



The eye is early a sphere, with an average diameter of approximately 20 mm. Three membrance encloses the eye,

1. Cornea
2. Sclera or Cornea:
3. Retina

. The cornea is a tough, transparent tissue that covers the anterior surface of the eye.

Sclera:

Sclera is an opaque membrance e that encloses the remainder of the optical globe.

Choroid:

-Choroid directly below the sclera. This membrance contains a network of blood vessels that serve as the major source of nutrition to the eye.

-Choroid coat is heavily pigmented and helps to reduce the amount of extraneous light entering the eye.

-The choroid is divided into the ciliary body and the iris diaphragm.

Lens:

The lens is made up of concentric lay ours of fibrous cells and is suspended by fibrous that attach to the ciliary body. It contains 60to 70% of water about 60%fat and m ore protein than any other tissue in the eye.

Retina:

The innermost membrance of the eye is retina, which lines the inside of the wall's entire posterior portion. There are 2 classes of receptors,

1. Cones
2. Rods

Cones:

The cones in each eye between 6and7 million. They are located primarily in the central portion of the retina called the fovea, and highly sensitive to Colour.

Rods:

The number of rods is much larger; some 75 to 150 millions are distributed over the retinal surface.

Fovea as a square sensor array of size 1.5mm*1.5mm.

4. Explain the RGB model

RGB model,each color appears in its primary spectral components of red ,green and blue.This model is based on a Cartesian coordinate system.This color subspace of interest is the cube.RGB values are at three corners cyan.magenta and yellow are at three other corner black is at the origin and white is the at the corner farthest from the origin this model the gray scale extends from black to white along the line joining these two points .The different colors in this model are points on or inside the cube and are defined by vectors extending from the origin.

Images represented in the RGB color model consist of three component images,one for each primary colors.The no of bits used to represented each pixel in which each red,green and blue images is an 8 bit image.Each RGB color pixel of values is said to be 24 bits .The total no of colors in a 24 bit RGB images is 92803=16777,216.
The acquiring a color image is basically the process is shown in fig,. A color image can be acquired by using three filters,sensitive to red,green and blue.When we view a color scene with a monochrome camera equipped with one of these filters the result is a monochrome image whose intensity is proportional to the response of that filter. Repeating this process with each filter produces three monochrome images that are the RGB component images of the color scene.the subset of color is called the set of safe RGB colors or the set of all system safe colors. In inter net applications they are called safe Web colors or safe browser colors.There are 256 colors are obtained from different combination but we are using only 216 colors .

5.Descibe the HSI color image model

The HSI Color Model

The RGB,CMY and other color models are not well suited for describing colors in terms that are practical for human interpretation.For eg,one does not refer to the color of an automobile by giving the percentage of each of the primaries composing its color.

When humans view a color object we describe it by its hue, saturation and brightness.

- Hue is a color attribute that describes a pure color.
- Saturation gives a measure of the degree to which a pure color is diluted by white light.
- Brightness is a subjective descriptor that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation

- Intensity is a most useful descriptor of monochromatic images.

Converting colors from RGB to HSI

Given an image in RGB color format ,

- the H component of each RGB pixel is obtained using the equation

$$H = \{\text{theta} \quad \text{if } B<=G$$
$$360\text{-theta} \quad \text{if } B>G$$

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

- The saturation component is given by

$$S =1-3/(R+G+B)[\min(R,G,B)]$$

- the intensity component is given by

$$I=1/3(R+G+B)$$

Converting colors from HSI to RGB

Given values of HSI in the interval [0,1],we now want to find the corresponding RGB values in the same range .We begin by multiplying H by 360^o,which returns the hue to its original range of $[0^o,360^o]$

RG sector($0^o<=120^o$).when h is in this sector ,the RGB components are given by the equations

$$B = I\,(1 - S)$$
$$R = I\,[1 + S\cos H/\cos(60^o - H)]$$
$$G = 1 - (R + B)$$

GB Sector($120^o <= H < 240^o$).If the given value of H is in this ,we first subtract 120^o from it

$$H = H -120^o$$

Then the RGB components are

$$B = I\,(1 – S)$$
$$G = I\,[1 + S\cos H/\cos(60^o - H)]$$
$$B = 1 - (R + G)$$

BR Sector($240^o <=H<=360^o$).Finally if H is in this range we subtract 240^o from it

$$H = H - 240^o$$

Then the RGB components are

$$G = I\,(1 - S)$$
$$B = I\,[1 + S\cos H/\cos(60^o - H)]$$
$$R = 1 - (G + B)$$

6. Describe the basic relationship between the pixels

<u>2-D Mathematical preliminaries</u>

- Neighbours of a pixel
- Adjacency, Connectivity, Regions and Boundaries
- Distance measures

Neighbours of a pixel

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

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

- This set of pixels, called the 4-neighbours of p, is denoted by $N_4(p)$. Each pixel is a unit distance from (x,y) and some of the neighbours of p lie outside the digital image if (x,y) is on the border of the image.
- The four diagonal neighbours of p have coordinates
  (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)
- And are denoted by $N_D(p)$. These points together with the 4-neighbours are called the 8-neighbours of p, denoted by $N_8(p)$.

Adjacency, Connectivity, Regions and Boundaries

Three types of adjacency:

- **4-adjacency**. Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.
- **8-adjacency**. Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.
- M-adjacency. Two pixels p and q with values from V are m-adjacent if q is in $N_4(p)$, or q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.
- A (digital) path (or curve) from pixel p with coordinates (x,y) to pixel q with coordinates (s,t) is a sequence of distinct pixels with coordinates
  (x_0,y_0), (x_1,y_1),................(x_n,y_n)
  Where (x_0,y_0)= (x,y), (x_n,y_n)=(s,t) and pixels (x_i,y_i) and (x_{i-1},y_{i-1}) are adjacent for 1<=i<=n. N is the length of the path.

Distance measures

- For pixels p,q and z with coordinates (x,y), (s,t) and (v,w) respectively, D is a distance function or metric if
  - D(p,q)>=0 (D(p,q)=0 iff p=q),
  - D(p,q) = D(q,p) and
  - D(p,z) <= D(p,q) + D(q,z)
- The Euclidean distance between p and q is defined as,
  - $D_e(p,q) = [(x-s)^2+(y-t)^2]$
- The D_4 distance (also called city-block distance) between p and q is defined as
  - $D_4(p,q) = |x-s|+|y-t|$
- The D_8 distance (also called chessboard distance) between p and q is defined as
  - $D_8(p,q) = \max(|x-s|+|y-t|)$

UNIT II

1. What is the need for transform?

The need for transform is most of the signals or images are time domain signal (ie) signals can be measured with a function of time. This representation is not always best. For most image processing applications anyone of the mathematical transformation are applied to the signal or images to obtain further information from that signal.

2. What is Image Transform?

An image can be expanded in terms of a discrete set of basis arrays called basis images. These basis images can be generated by unitary matrices. Alternatively, a given NxN image can be viewed as an N^2x1 vectors. An image transform provides a set of coordinates or basis vectors for vector space.

3. What are the applications of transform?

1) To reduce band width
2) To reduce redundancy
3) To extract feature.

4. Give the Conditions for perfect transform

Transpose of matrix = Inverse of a matrix.
Orthoganality.

5. What are the properties of unitary transform?

1) Determinant and the Eigen values of a unitary matrix have unity magnitude
2) the entropy of a random vector is preserved under a unitary Transformation
3) Since the entropy is a measure of average information, this means information is preserved under a unitary transformation.

6. Define Fourier transform pair

The Fourier transform of f(x) denoted by F(u) is defined by

$$F(u)=\int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux}\, dx \text{ ----------------(1)}$$

The inverse fourier transform of f(x) is defined by

$$f(x)=\int_{-\infty}^{\infty} F(u)\, e^{j2\pi ux}\, dx \text{ --------------------(2)}$$

The equations (1) and (2) are known as fourier transform pair.

7. Define Fourier spectrum and spectral density

Fourier spectrum is defined as

$F(u) = |F(u)|\, e^{j\varphi(u)}$

Where

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

$\varphi(u) = \tan^{-1}(I(u)/R(u))$

Spectral density is defined by

$p(u) = |F(u)|^2$

$p(u) = R^2(u)+I^2(u)$

8. Give the relation for 1-D discrete Fourier transform pair

The discrete Fourier transform is defined by

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

The inverse discrete Fourier transform is given by

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

These equations are known as discrete Fourier transform pair.

9. Specify the properties of 2D Fourier transform.

The properties are

- Separability
- Translation
- Periodicity and conjugate symmetry
- Rotation
- Distributivity and scaling
- Average value
- Laplacian
- Convolution and correlation
- sampling

10. Mention the separability property in 2D Fourier transform

The advantage of separable property is that F(u, v) and f(x, y) can be obtained by successive application of 1D Fourier transform or its inverse.

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

Where

$$F(x, v) = N[1/N \sum_{y=0}^{n-1} f(x, y)\, e^{-j2\pi vy/N}$$

11. List the Properties of twiddle factor.

1. Periodicity

   W_N^(K+N)= W_N^K

2. Symmetry

   W_N^(K+N/2)= -W_N^K

12. Give the Properties of one-dimensional DFT

1. The DFT and unitary DFT matrices are symmetric.
2. The extensions of the DFT and unitary DFT of a sequence and their inverse transforms are periodic with period N.
3. The DFT or unitary DFT of a real sequence is conjugate symmetric about N/2.

13. Give the Properties of two-dimensional DFT

1. Symmetric
2. Periodic extensions
3. Sampled Fourier transform
4. Conjugate symmetry.

14. What is meant by convolution?

The convolution of 2 functions is defined by

$$f(x)*g(x) = \int f(\alpha)\, .g(x-\alpha)\, d\alpha$$

where α is the dummy variable

15. State convolution theorem for 1D

- If f(x) has a fourier transform F(u) and g(x) has a fourier transform G(u) then f(x)*g(x) has a fourier transform F(u).G(u).
- Convolution in x domain can be obtained by taking the inverse fourier transform of the product F(u).G(u).
- Convolution in frequency domain reduces the multiplication in the x domain

  $f(x).g(x) \Leftrightarrow F(u)^* G(u)$
- These 2 results are referred to the convolution theorem.

16. What is wrap around error?

The individual periods of the convolution will overlap and referred to as wrap around error

17. Give the formula for correlation of 1D continuous function.

The correlation of 2 continuous functions f(x) and g(x) is defined by

$f(x) \circ g(x) = \int f^*(\alpha) \, g(x+\alpha) \, d\alpha$

18. What are the properties of Haar transform.

1. Haar transform is real and orthogonal.
2. Haar transform is a very fast transform
3. Haar transform has very poor energy compaction for images
4. The basic vectors of Haar matrix sequensly ordered.

19. What are the Properties of Slant transform

1. Slant transform is real and orthogonal.
2. Slant transform is a fast transform
3. Slant transform has very good energy compaction for images
4. The basic vectors of Slant matrix are not sequensely ordered.

20. Specify the properties of forward transformation kernel

The forward transformation kernel is said to be separable if g(x, y, u, v)

g(x, y, u, v) = g1(x, u).g2(y, v)

The forward transformation kernel is symmetric if g1 is functionally equal to g2

g(x, y, u, v) = g1(x, u). g1(y,v)

21. Define fast Walsh transform.

The Walsh transform is defined by

$$w(u) = 1/N \sum_{x=0}^{n-1} f(x) \prod_{i=0}^{x-1} (-1)^{b_i(x).b_{n-1-i}(u)}$$

22. Give the relation for 1-D DCT.

- The 1-D DCT is,

$$C(u)=\alpha(u)\sum_{X=0}^{N-1} f(x) \cos[((2x+1)u\pi)/2N] \text{ where } u=0,1,2,\ldots.N-1$$

- Inverse $f(x)=\sum_{V=0}^{N-1} \alpha(u)\, c(u) \cos[((2x+1)\, u\pi)/2N]$ where $x=0,1,2,\ldots N-1$

23. Write slant transform matrix SN.

$$S_N = 1/\sqrt{2}\left[\begin{array}{c|c|c|c} \begin{matrix}1 & 0\\ a_N & b_N\end{matrix} & 0 & \begin{matrix}1 & 0\\ -a_N & b_N\end{matrix} & 0 \\ \hline 0 & I_{(N/2-1)} & 0 & I_{(N/2-1)} \\ \hline \begin{matrix}1 & 0\\ -b_N & a_N\end{matrix} & 0 & \begin{matrix}0 & -1\\ b_N & a_N\end{matrix} & 0 \\ \hline 0 & I_{(N/2-1)} & 0 & -I_{(N/2-1)} \end{array}\right] \left[\begin{array}{c|c} S_{N/2} & 0 \\ \hline 0 & S_{N/2} \end{array}\right]$$

24. Define Haar transform.

The Haar transform can be expressed in matrix form as,

$T=HFH$

Where F = N X N image matrix
H = N X N transformation matrix
T = resulting N X N transform.

25. Define K-L transform.

Consider a set of n or multi-dimensional discrete signal represented as column vector x1,x2,…xn each having M elements,

$$X=\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_n \end{bmatrix}$$

The mean vector is defined as $M_x=E\{x\}$
Where $E\{x\}$ is the expected value of x.
For M vector samples mean vector is $M_x=1/M \sum_{K=1}^{M} X_k$

The co-variant matrix is, $C_x=E\{(X-M_x)(X-M_x)^T\}$
Where T-vector transposition X->N-D vector
Cx->nxn matrix.
For M samples, $C_x=1/M \sum^{M} (x_k-M_x)(x_k-M_x)^T$.

$K=1$

K-L Transform $Y= A(X- M_X)$

26. Give the equation for singular value decomposition of an image?

$U= {}_{m=1}\sum^{r}\psi\sqrt{\lambda_m}\ \varphi_m^T$

This equation is called as singular value decomposition of an image.

27. Write the properties of Singular value Decomposition(SVD)?

- The SVD transform varies drastically from image to image.
- The SVD transform gives best energy packing efficiency for any given image.
- The SVD transform is useful in the design of filters finding least square,minimum solution of linear equation and finding rank of large matrices.

PART-B

---

1. Write short notes on Discrete Cosine Transform (DCT)

- **I-D DCT**

  I-D DCT is defined as

  $C(u)= \alpha(u)\ {}_{x=0}\sum^{N-1}f(x)\cos[(2x+1)u\pi/2N]$ for u=0,1,2,…….N-1

  Inverse DCT is defined as

  $f(x)= {}_{u=0}\sum^{N-1}\alpha(u)C(u)\cos[(2x+1)u\pi/2N]$ for x=0,1,2,…….N-1

  In both cases $\alpha(u)=1/\sqrt{N}$ for u=0 and $\sqrt{2}/\sqrt{N}$ for u=1,2,…….N-1

- **2-D DCT**

  I-D DCT is defined as

  $C(u,v)= \alpha(u)\,\alpha(v)\ {}_{x=0}\sum^{N-1}\ {}_{y=0}\sum^{N-1}f(x,y)\cos[(2x+1)u\pi/2N]\cos[(2y+1)v\pi/2N]$

  for u,v=0,1,2,…….N-1

  Inverse DCT is defined as

  $f(x,y)={}_{u=0}\sum^{N-1}\ {}_{v=0}\sum^{N-1}\alpha(u)\,\alpha(v)C(u,v)\cos[(2x+1)u\pi/2N]\ \cos[(2y+1)u\pi/2N]$

  for x,y=0,1,2,…….N-1

  In both cases $\alpha(u)=1/\sqrt{N}$ for u=0 and $\sqrt{2}/\sqrt{N}$ for u=1,2,…….N-1

2. Describe Fast Fourier Transform

- The Fourier transform of f(x) denoted by F(u) is defined by

  $$F(u)=\int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux}\, dx \text{ ----------------(1)}$$

- The inverse fourier transform of f(x) is defined by

  $$f(x)= \int_{-\infty}^{\infty}F(u)\ e^{j2\pi ux}\, dx \text{ --------------------(2)}$$

  The equations (1) and (2) are known as fourier transform pair.

  $F(u)=1/2M\ \ {}_{x=0}\Sigma^{2M-1}f(x)\ W_{2M}^{ux}$

  $F(u)=1/2\{1/M\ {}_{x=0}\Sigma^{M-1}f(2x)W_{2M}^{u2x} + 1/M\ {}_{x=0}\Sigma^{M-1}f(2x+1)W_{2M}^{u(2x+1)}$

  $F(u)=1/2\{1/M\ {}_{x=0}\Sigma^{M-1}f(2x)W_M^{ux} + 1/M\ {}_{x=0}\Sigma^{M-1}f(2x+1)W_M^{ux}\,.W_{2M}^{u}\ \}$

$F(u)=1/2\{F_{even}(u)+F_{odd}(u).W_{2M}^{u}\}$
$F(u+M)=1/2\{1/M \sum_{x=0}^{M-1} f(2x)W_M^{(u+M)x} + 1/M \sum_{x=0}^{M-1} f(2x+1)W_M^{(u+M)x} .W_{2M}^{u+M}\}$
$F(u+M)=1/2\{1/M \sum_{x=0}^{M-1} f(2x)W_M^{ux} - 1/M \sum_{x=0}^{M-1} f(2x+1)W_M^{ux} .W_{2M}^{u}\}$
$F(u+M)=1/2[F_{even}(u)+F_{odd}(u) .W_{2M}^{u}]$

3.Write short notes on Hotelling transform

Consider a set of n or multi-dimensional discrete signal represented as column vector x1,x2,…xn each having M elements,

$$X=\begin{bmatrix} X1 \\ X2 \\ . \\ . \\ Xn \end{bmatrix}$$

- The mean vector is defined as Mx=E{x},Where E{x} is the expected value of x.

  For M vector samples mean vector is $Mx=1/M \sum_{K=1}^{M} Xk$

- The co-variant matrix is, Cx=E{(X-Mx)(X-Mx)}

  Where T-vector transposition X->N-D vector,Cx->nxn matrix.

  For M samples, $Cx=1/M \sum_{K=1}^{M} (xk-Mx)(xk-Mx)$.

  K-L Transform $Y= A(X- M_X)$

- Features:

  - HT is based on the statistical properties of vector representation
  - HT has several useful properties that makes an important tool for image processing
  - It converts discrete signal s into a sequence of uncorrelated coefficients

---

PART-C

1Explain Discrete Fourier Transform in detail.

- **ID Case**

  $F(u)=1/N \sum_{x=0}^{N-1} f(x)\exp[-j2\pi ux/N]$ for u=0,1,2,…….N-1------------------(1)

  $f(x)=\sum_{u=0}^{N-1} F(u)[j2\pi ux/N]$, for x=0,1,2,…….N-1--------------------------(2)

  Equations (1) and (2) called Discrete Fourier transform pair

The values u=0,1,2,………N-1 in the discrete Fourier transform corresponds to the samples of the continuous transform at values 0, Δu, $2\Delta u$….$(N-1)\Delta u$. In other words F(u) corresponds $F(u\Delta u)$. The terms Δu and Δx related by the expression $\Delta u=1/N\Delta x$

- **2D Case**

  $F(u,v)=1/MN \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi ux/M+vy/N]$

  for u=0,1,2,…….M-1, v=0,1,2,……..N-1

  $f(x,y)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v)\exp[j2\pi ux/M+vy/N]$

  for x=0,1,2,…….M-1, y=0,1,2,……..N-1

  For a square image M=N, FT pair will be

  $F(u, v)=1/N \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi(ux +vy)/N]$

  for u, v=0,1,2,…….N-1

  $f(x, y)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} F(u,v)\exp[j2\pi(ux+vy)/N]$

  for x, y=0,1,2,…….N-1

2.Explain the Properties of 2D discrete Fourier Transform

1. **Separability**

   $F(u, v)=1/N \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi(ux +vy)/N]$ for u, v=0,1,2,…….N-1

   $f(x, y)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} F(u,v)\exp[j2\pi(ux+vy)/N]$ for x, y=0,1,2,…….N-1

   $F(u,v)=1/N \sum_{x=0}^{N-1} F(x,v)\exp[-j2\pi ux/N]$

   where $F(x,v)=N[1/N \sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi vy/N]$

2. **Translation**

The translation properties of the Fourier Transorm pair are

$f(x,y)\exp[-j2\pi(u_0x +v_0y)/N] \Leftrightarrow F(u-u_0,v-v_0)$ are Fourier Transform pair.

And $f(x-x_0,y-y_0) \Leftrightarrow F(u,v)\exp[-j2\pi(ux_0 +vy_0)/N]$

Where the double arrow indicates the correspondence between a function and its Fourier transform.

3. **Periodicity and Conjugate Symmetry**

- Periodicity:

  The Discrete Fourier Transform and its inverse are periodic with period N; that is,

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

- Conjugate symmetry:

  If f(x,y) is real, the Fourier transform also exhibits conjugate symmetry,

  $F(u,v)=F^*(-u,-v)$ or $|F(u,v)|=|F(-u,-v)|$ where $F^*(u,v)$ is the complex conjugate of F(u,v)

4. **Rotation**

Polar Coordinates $x=r\cos\theta$, $y=r\sin\theta$, $u=w\sin\Phi$, $v=w\sin\Phi$ then $f(x,y)$ and $F(u,v)$ become $f(r,\theta)$ and $F(w,\Phi)$ respectively. Rotating $f(x,y)$ by an angle θ_0 rotates $F(u,v)$ by the same angle. Similarly rotating $F(u,v)$ rotates $f(x,y)$ by the same angle.
i.e, $f(r,\theta+\theta_0) \Leftrightarrow F(r,\Phi+\theta_0)$

5. **Distributivity and scaling**

- Distributivity:
  The Discrete Fourier Transform and its inverse are distributive over addition but not over multiplication.
  $F[f1(x,y)+f2(x,y)]=F[f1(x,y)]+F[f2(x,y)]$
  $F[f1(x,y).f2(x,y)]\neq F[f1(x,y)].F[f2(x,y)]$
- Scaling
  For the two scalars a and b,
  $af(x,y) \Leftrightarrow aF(u,v)$ and $f(ax,by) \Leftrightarrow 1/|ab|\,F(u/a,v/b)$

6. **Laplacian**

The Laplacian of a two variable function $f(x,y)$ is defined as $\nabla^2 f(x,y)=\partial^2 f/\partial x^2+\partial^2 f/\partial y^2$

7. **Convolution and Correlation**

- Convolution
  The convolution of two functions f(x) and g(x) denoted by f(x)*g(x) and is defined by the integral, $f(x)*g(x)=\int_{-\infty}^{\infty} f(\alpha)g(x-\alpha)d\alpha$ where α is a dummy variable.
  Convolution of two functions F(u) and G(u) in the frequency domain =multiplication of their inverse f(x) and g(x) respectively.
  Ie, $f(x)*g(x) \Leftrightarrow F(u)G(u)$
- Correlation
  The correlation of two functions f(x) and g(x) denoted by f(x)og(x) and is defined by the integral, $f(x)og(x)=\int_{-\infty}^{\infty} f^*(\alpha)g(x+\alpha)d\alpha$ where α is a dummy variable.
  For the discrete case $f_e(x)og_e(x)= 1/M \sum_{M=0}^{M-1} f^*(m)g(x+m)$

$$f_e(x)= \begin{cases} f(x), & 0\le x\le A-1, \\ 0, & A\le x\le M-1 \end{cases}$$

$$g_e(x)= \begin{cases} g(x), & 0\le x\le B-1, \\ 0, & B\le x\le N-1 \end{cases}$$

3.Discuss Hadamard transform in detail

- ID Hadamard transform

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

$$= \sum_{x=0}^{n-1} f(x)\, g(x,u)$$

where $g(x,u)= 1/N\,(-1)^{\sum_{i=0}^{n-1} bi(x)\,bi(u)}$ which is known as 1D forward Hadamard kernel.

bk(x) is the kth bit binary representation of z
bi (z)=1

- Inverse 1D Hadamard Transform

$$f(x) = \sum_{u=0}^{N-1} H(u) \sum_{i=0}^{n-1} bi(x)\, bi(u)$$

$$f(x) = \sum_{u=0}^{N-1} H(u)\, h(x,u) \qquad x=0,1\ldots.N-1$$

$$h(x,u) = (-1)^{\sum_{i=0}^{n-1} bi(x)\, bi(u)}$$

- 2D Hadamard Transform

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

$$= \sum_{x=0}^{n-1} \sum_{y=0}^{N-1} f(x, y)^{\, g(x,y,u,v)}$$

where $g(x, y, u, v) = 1/N\ (-1)^{\sum_{i=0}^{n-1} b_i(x)\, b_i(u) + b_i(y)\, b_i(v)}$

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

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

Where $h(x, y, u, v) = 1/N\ (-1)^{\sum_{i=0}^{n-1} [b_i(x)\, b_i(u) + b_i(y)\, b_i(v)]}$ which is the inverse kernel Therefore ,forward and reverse kernel are same

- Ordered Hadamard Transform

  1D Ordered Hadamard Transform

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

$$= \sum_{x=0}^{n-1} f(x)\, g(x, u)$$

where $g(x, u) = 1/N\ (-1)^{\sum_{i=0}^{n-1} bi(x)\, pi(u)}$

po(u) = b (n-1) (u)

p1 (u) = b (n-1) (u) + b (n-2) (u)

p2 (u) = b (n-2) (u) + b (n-3) (u)

.

.

.

.

pn-1 (u) = b1(u) + b o (u)

- Inverse Hadamard Transform

$$f(x)= \sum_{u=0}^{N-1} H(u)\, (-1)^{\sum_{i=0}^{n-1} bi(x)\, pi(u)}$$

$$f(x) = \sum_{u=0}^{N-1} H(u)\, h(x,u)$$

$$\text{where } h(x, u) = (-1) \sum_{i=0}^{n-1} bi(x)\, pi(u)$$

- 2D ordered HT Pair

$$H(u,v) = 1/N \sum_{x=0}^{n-1} \sum_{y=0}^{N-1} f(x, y)\, (-1) \sum_{i=0}^{n-1} [bi(x)\, pi(u) + bi(y)\, pi(v)]$$

$$= \sum_{x=0}^{n-1} \sum_{y=0}^{N-1} f(x, y)\, g(x,y,u,v)$$

$$\text{where } g(x, y, u, v) = 1/N\ (-1) \sum_{i=0}^{n-1} [bi(x)\, pi(u) + bi(y)\, pi(v)]$$

Similarly ,$f(x,y) = = 1/N \sum_{u=0}^{n-1} \sum_{v=0}^{N-1} H(u.v)\, (-1) \sum_{i=0}^{n-1} [bi(x)\, pi(u) + bi(y)\, pi(v)]$

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

Where $h(x, y, u, v) = 1/N\ (-1) \sum_{i=0}^{n-1} [bi(x)\, pi(u) + bi(y)\, pi(v)]$

4. Explain Walsh Transform with suitable equations
Walsh Transform

For $N=2^n$, Walsh transform function f(x) is denoted by W(u).W(u) is obtained by substituting the forward transformational kernel,

---

$g(x,u)=1/N \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}$
The Walsh transform of f(x)
$W(u)=1/N \sum_{x=0}^{N-1} f(x) \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}$
$W(u)= \sum_{x=0}^{N-1} f(x)g(x,u)$
Where $b_k(z)$ is the k^{th} bit in the binary representation of z

- **Inverse Walsh Transform**

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

forward and inverse transformational kernel differs by 1/N

When N=8

u \ v	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	+	+	+	-	-	-	-
2	+	+	-	-	+	+	-	-
3	+	+	-	-	-	-	+	+
4	+	-	+	-	+	-	+	-
5	+	-	+	-	-	+	-	+
6	+	-	-	+	+	-	-	+
7	+	-	-	+	-	+	+	-

2D Walsh Transform

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

forward transformational kernel, $g(x,y,u,v)=1/N\{\prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v)}\}$

$W(u,v)=\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)g(x,y,u,v)$

2D Inverse Walsh Transform

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

$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} W(u,v)h(x,y,u,v)$

inverse transformational kernel, $h(x,y,u,v)= \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v)}$

Walsh transformational kernels are separable and symmetric

$g(x,y,u,v)=g_1(x,u)\, g_1(y,v)$

$h(x,y,u,v) =h_1(x,u)\, h_1(y,v)$

$\{1/\sqrt{N} \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)}\}\ \{1/\sqrt{N} \prod_{i=0}^{n-1} (-1)^{b_i(y)b_{n-1-i}(v)}\}$

Walsh Transform can be obtained by successive doubling method

$W(u)=1/2[W_{even}(u)+W_{odd}(v)]$

$W(u+m)=1/2[W_{even}(u)-W_{odd}(v)]$

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

5. Explain Discrete cosine transform in detail

The discrete cosine transform(DCT) gets its name from the fact that the rows of the N*N transform matrix C are obtained as a function of cosines.

$$|C|_{i,j}= \sqrt{1/N}\cos\frac{(2j+1)i\Pi}{2N} \quad i=0,j=0,1,\ldots.N-1$$

$$\sqrt{2/N}\cos\frac{(2j+1)i\Pi}{2N} \quad i=1,2\ldots..N-1,j=0,1,\ldots.N-1$$

The rows of the transform matrix are shown in graphical form. Here the amount of variation increases as we progress down the row; that is the frequency of the rows increases as we go from top to bottom.

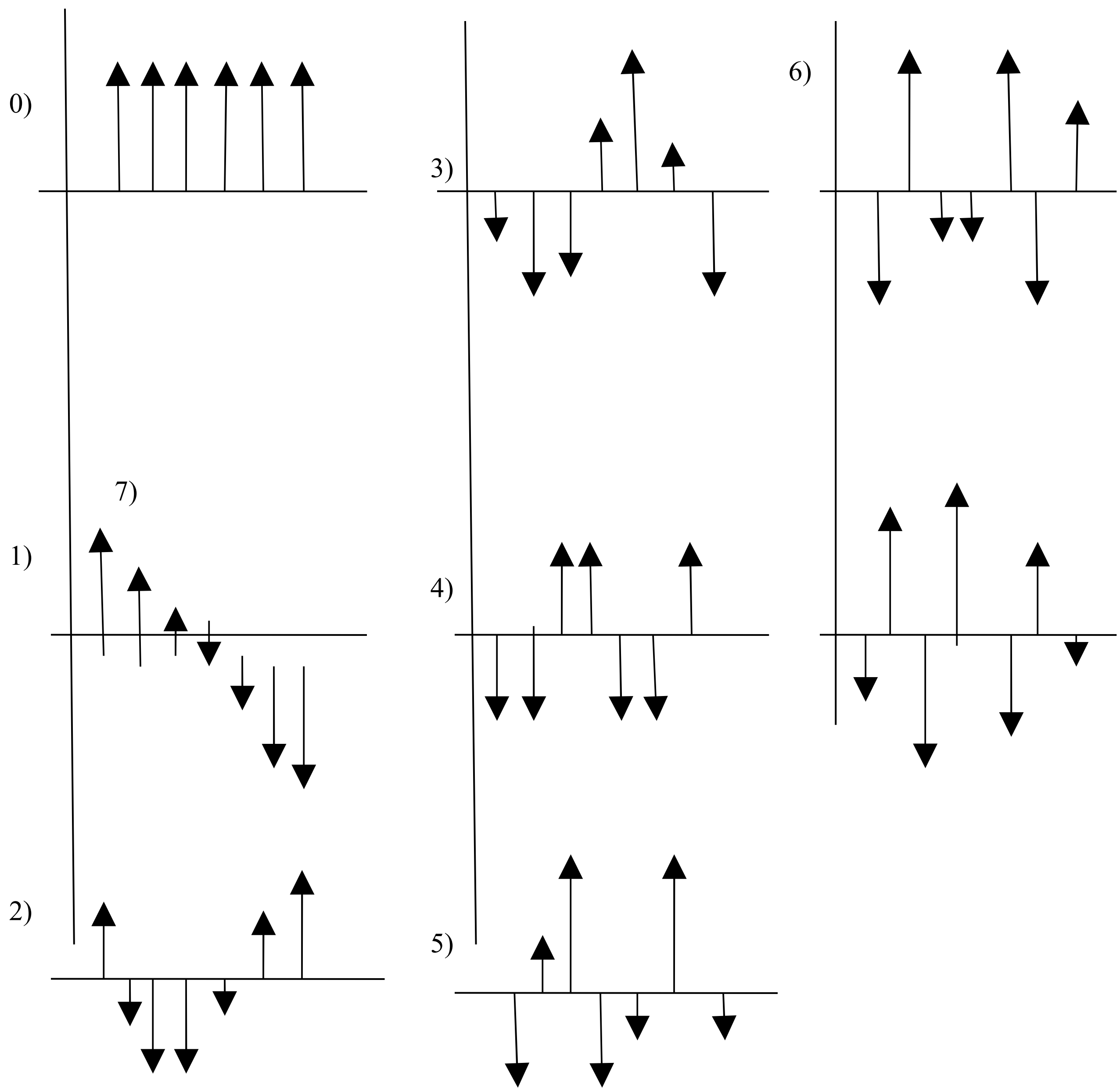


Fig .Basic set of the discrete cosine transform. The numbers correspond to the row of the transform matrix.

Also, the basis matrices show increased variation as we go from the top-left matrix, corresponding to the Θ_{00} coefficient, to the bottom-right matrix, corresponding to the $\Theta_{(N-1)(N-1)}$ coefficient.

The DCT is closely related to the discrete Fourier transform(DFT) and DCT can be obtained from DFT. In terms of compression, the DCT performs better than the DFT.

In DFT, to find the Fourier coefficients for s sequence of length N, we assume that the sequence is periodic with period N. The DFT assumes that the sequence outside the interval behaves in a different manner. This introduces sharp discontinuities, at the beginning and the end of the sequence. In order to represent these sharp discontinuities, the DFT needs nonzero coefficients for the high-frequency components. Because these components are needed only at the two end points of the sequence, their effect needs to be cancelled out at other points in the sequence. Thus, the DFT adjusts other coefficients accordingly. When we discard the high frequency coefficients during the compression process, the coefficients that were cancelling out the high-frequency effect in other parts of the sequence result in the introduction of additional distortion.

The DCT can be obtained using the DFT by mirroring the original N-point sequence to obtain a 2N-point sequence. The DCT is simply the first N points of the resulting 2N-point DFT. When we take the DFT of the 2N-point mirrored sequence, we again have to assume periodicity. Here it does not introduce any sharp discontinuities at the edges.

The DCT is better at energy compaction for most correlated sources when compared to the DFT. For Markov sources with high correlation coefficient ρ,

$$\rho = \frac{E[x_n x_{n+1}]}{E[x_n 2]},$$

The compaction ability of the DCT is very close to that of the KLT. As many sources can be modelled as Markov sources with high values for ρ, this superior compaction ability has made the DCT the most popular transform. It is a part of many international standards, including JPEG,MPEG and CCITT H.261.

UNIT III

1. **Specify the objective of image enhancement technique.**

The objective of enhancement technique is to process an image so that the result is more suitable than the original image for a particular application.

2. **List the 2 categories of image enhancement.**
   - Spatial domain refers to image plane itself & approaches in this category are based on direct manipulation of picture image.
   - Frequency domain methods based on modifying the image by fourier transform.

3. What is the purpose of image averaging?

An important application of image averaging is in the field of astronomy, where imaging with very low light levels is routine, causing sensor noise frequently to render single images virtually useless for analysis.

4. What is meant by masking?
- Mask is the small 2-D array in which the values of mask co-efficient determines the nature of process.

- The enhancement technique based on this type of approach is referred to as mask processing.

5. Define histogram.

The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function $h(r_k)=n_k$.

r_k-kth gray level

n_k-number of pixels in the image having gray level r_k.

6.What is meant by histogram equalization?

$$S_k = T(r_k) = \sum_{j=0}^{k} P_r(r_j) = \sum_{j=0}^{k} n_j/n \qquad \text{where } k=0,1,2,\ldots,L-1$$

This transformation is called histogram equalization.

7.Differentiate linear spatial filter and non-linear spatial filter.

s.no.	Linear spatial filter	Non-linear spatial filter
1.	Response is a sum of products of the filter co-efficient.	They do not explicitly use co-efficients in the sum-of-products.
2.	$R = w(-1,-1)\,f(x-1,y-1) + w(-1,0)\,f(x-1,y) + \ldots + w(0,0)\,f(x,y) + \ldots + w(1,0)\,f(x+1,y) + w(1,1)\,f(x+1,y+1).$	$R = w_1z_1 + w_2z_2 + \ldots + w_9z_9 = \sum_{i=1}^{9} w_iz_i$

8. Give the mask used for high boost filtering.

0	-1	0
-1	A+4	-1
0	-1	0

-1	-1	-1
-1	A+8	-1
-1	-1	-1

9. What is meant by laplacian filter?

The laplacian for a function f(x,y) of 2 variables is defined as,

2 2 2 2 2

$\nabla f = \partial f / \partial x + \partial f / \partial y$

10. Write the steps involved in frequency domain filtering.

1. Multiply the input image by (-1) to center the transform.
2. Compute F(u,v), the DFT of the image from (1).
3. Multiply F(u,v) by a filter function H(u,v).
4. Compute the inverse DFT of the result in (3).
5. Obtain the real part of the result in (4).
6. Multiply the result in (5) by (-1)

11. What do you mean by Point processing?

Image enhancement at any Point in an image depends only on the gray level at that point is often referred to as Point processing.

12. Define Derivative filter?

For a function f (x, y), the gradient f at co-ordinate (x, y) is defined as the vector

$$\Delta f = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix}$$

$$\Delta f = \text{mag}\ (\Delta f) = \{[(\partial f/\partial x)^2 + (\partial f/\partial y)^2]\}^{1/2}$$

13. Define spatial filtering

Spatial filtering is the process of moving the filter mask from point to point in an image. For linear spatial filter, the response is given by a sum of products of the filter coefficients, and the corresponding image pixels in the area spanned by the filter mask.

14. What is a Median filter?

The median filter replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel.

15. What is maximum filter and minimum filter?

The 100^{th} percentile is maximum filter is used in finding brightest points in an image. The 0^{th} percentile filter is minimum filter used for finding darkest points in an image.

16. Write the application of sharpening filters

1. Electronic printing and medical imaging to industrial application
2. Autonomous target detection in smart weapons.

17. Name the different types of derivative filters

1. Perwitt operators
2. Roberts cross gradient operators
3. Sobel operators

18. What is meant by Image Restoration?

Restoration attempts to reconstruct or recover an image that has been degraded by using a clear knowledge of the degrading phenomenon.

19. What are the two properties in Linear Operator?

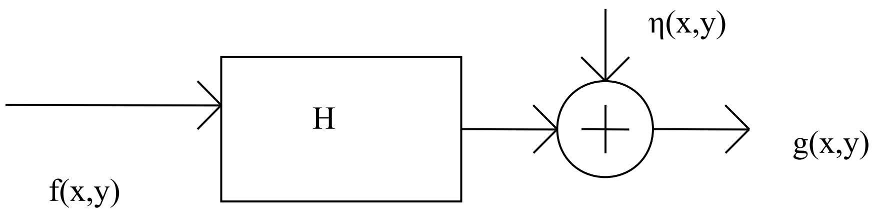
- Additivity
- Homogenity

20. Give the additivity property of Linear Operator

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

The additive property says that if H is the linear operator,the response to a sum of two is equal to the sum of the two responses.

21. How a degradation process is modeled?



A system operator H, which together with an additive white noise term η(x,y) a operates on an input image f(x,y) to produce a degraded image g(x,y).

22.Give the homogenity property in Linear Operator

$$H[k_1 f_1(x,y)]=k_1\, H[f_1(x,y)]$$

The homogeneity property says that,the response to a constant multiple of any input is equal to the response to that input multiplied by the same constant.

23. Give the relation for degradation model for continuous function

$$g(x,y) = {}_{-\infty}\int^{\infty}\int f(\alpha,\beta)\S(x-\alpha,y-\beta).d\alpha d\beta+\eta(x,y)$$

24. What is Fredholm integral of first kind?

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

which is called the superposition or convolution or fredholm integral of first kind. It states that if the response of H to an impulse is known, the response to any input f(α,β) can be calculated by means of fredholm integral.

25. Define circulant matrix

A square matrix, in which each row is a circular shift of the preceding row and the first row is a circular shift of the last row, is called circulant matrix.

$$He = \begin{bmatrix} h_e(o) & h_e(M-1) & h_e(M-2)\ldots\ldots\ldots\ldots & h_e(1) \\ h_e(1) & h_e(0) & h_e(M-1)\ldots\ldots\ldots & h_e(2) \\ . & & & \\ . & & & \end{bmatrix}$$

.

$h_e(M-1)$ $h_e(M-2)$ $h_e(M-3)$………. $h_e(0)$

26. What is the concept algebraic approach?

The concept of algebraic approach is to estimate the original image which minimizes a predefined criterion of performances

27. What are the two methods of algebraic approach?

- Unconstrained restoration approach
- Constrained restoration approach

28. Define Gray-level interpolation

Gray-level interpolation deals with the assignment of gray levels to pixels in the spatially transformed image

29. What is meant by Noise probability density function?

The spatial noise descriptor is the statistical behavior of gray level values in the noise component of the model.

30. Why the restoration is called as unconstrained restoration?

In the absence of any knowledge about the noise 'n', a meaningful criterion function is to seek an f^ such that H f^ approximates of in a least square sense by assuming the noise term is as small as possible.

Where H = system operator.

f^ = estimated input image.

g = degraded image.

31. Which is the most frequent method to overcome the difficulty to formulate the spatial relocation of pixels?

The point is the most frequent method, which are subsets of pixels whose location in the input (distorted) and output (corrected) imaged is known precisely.

32. What are the three methods of estimating the degradation function?

1. Observation
2. Experimentation
3. Mathematical modeling.

33. What are the types of noise models?

- Guassian noise
- Rayleigh noise
- Erlang noise
- Exponential noise
- Uniform noise
- Impulse noise

34. Give the relation for guassian noise

Guassian noise:

The PDF guassian random variable Z is given by

$P(Z)=e^{-(Z-\mu)2/2\sigma 2}/\sqrt{2\pi\sigma}$

Z->Gray level value

σ->standard deviation

σ^2->varianze of Z

μ->mean of the graylevel value Z

35. Give the relation for rayleigh noise

Rayleigh noise:

The PDF is

$P(Z)= 2(z-a)e^{-(z-a)2/b}/b$ for Z>=a

0 for Z<a

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

standard deviation $\sigma^2=b(4-\pi)/4$

36. Give the relation for Gamma noise

Gamma noise:

The PDF is

$P(Z)=a^b z^{b-1} ae^{-az}/(b-1)$ for Z>=0

0 for Z<0

mean $\mu=b/a$

standard deviation $\sigma^2=b/a^2$

37. Give the relation for Exponential noise

Exponential noise

The PDF is

$P(Z)= ae^{-az}$ Z>=0

0 Z<0

mean $\mu=1/a$

standard deviation $\sigma^2=1/a^2$

38. Give the relation for Uniform noise

Uniform noise:

The PDF is

$P(Z)=1/(b-a)$ if a<=Z<=b

0 otherwise

mean $\mu=a+b/2$

standard deviation $\sigma^2=(b-a)^2/12$

39. Give the relation for Impulse noise

Impulse noise:

The PDF is

$P(Z) =P_a$ for z=a

P_b for z=b

0 Otherwise

40. What is inverse filtering?

The simplest approach to restoration is direct inverse filtering, an estimate F^(u,v) of the transform of the original image simply by dividing the transform of the degraded image G^(u,v) by the degradation function.

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

41. What is pseudo inverse filter?

It is the stabilized version of the inverse filter.For a linear shift invariant system with frequency response H(u,v) the pseudo inverse filter is defined as

$$H^{-}(u,v)=\begin{cases} 1/(H(u,v) & H\neq 0 \\ 0 & H=0 \end{cases}$$

42. What is meant by least mean square filter?

The limitation of inverse and pseudo inverse filter is very sensitive noise. The wiener filtering is a method of restoring images in the presence of blurr as well as noise.

43. Give the difference between Enhancement and Restoration

- Enhancement technique is based primarily on the pleasing aspects it might present to the viewer. For example: Contrast Stretching.
- Where as Removal of image blur by applying a deblurrings function is considered a restoration technique

PART-B

1.Discuss different mean filters

Arithmetic mean filter

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

Geometric mean filter

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

$$\hat{f}(x,y) = \Big[\prod_{(s,t)\in S_{xy}} g(s,t) \Big]$$

here ,each restored pixel is given by the product of the pixels in the subimage window , raised to the power 1/mn

- **Harmonic filters**

  The harmonic mean filtering operation is given by the expression

  $$\hat{f}(x,y) = mn/\sum(1/g(s,t))$$

- **Contra harmonic mean filter**

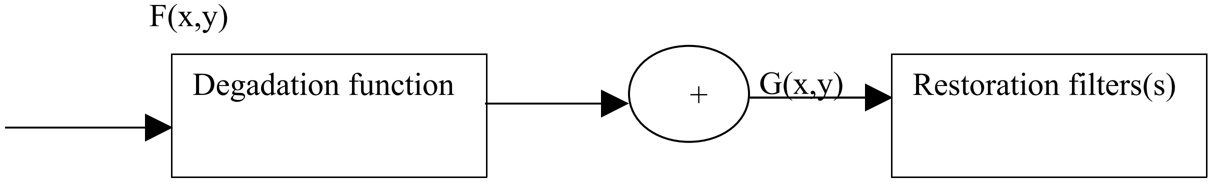
Contra harmonic mean filtering operation yields a restored image based on the expression

$$\hat{f}(x,y)=\sum g(x,t)^{Q+1} / \sum g(s,t)^{Q}$$

where Q is called the order of the filter.This filter is well suited for reducing or virtually eliminating the effect of salt and pepper noise.

2. Draw the degradation model and explain.

F(x,y)



Noise

- H is a linear positive invariant process the the degraded image is given in the spatial domain by

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

where h(x,y) is the spatial representation of the degradation function and the symbol "*" indicates spatial convolution.

- The convolution in the spatial domain is equal to multiplication in the frequency domain.
- The equivalent frequency domain representation is

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

Where the terms in capital letters are the Fourier transforms of the corresponding terms in the previous equation.

3.Write short notes on Median Filters

Introduction:

-Median Filter is one of the part of the smoothing filter.
-No mask is used in the median filters.
-We choose 3x3 sub-image arranged in ascending order and leave first four values.

3	5	7
2	10	20
30	9	4

2,3,4,5,7,9,10,20,30
-Take the median value.
-This median filter is the non-linear spatial filtering.
1)median filtering smoothing
2)Max filter
3)Min filter

Max filter:

R=Max
-Max filter gives the brightest points

Min filter:

R=Min
-It helps to get the largest point in the image.

4. Write short notes on Wiener Filtering.

- The inverse filtering approach makes no explicit provision for handling noise.
- An approach that incorporate both the degradation function and statistical characteristics of noise into the restoration process.
- The method is founded on considering images and noise as random processes,and the objective is to find an estimate f of the uncorrupted image f such that the mean square error between them is minimized.
- This error measure is given by

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

where E{.}is the expected value of the argument.

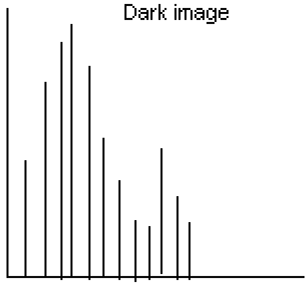
- It is assumed that the noise and the image are uncorrelated,that one or other has zero mean:and that the gray levels in the estimate are a linear function of the levels in the degradated image.
- Based on these conditions,the minimum of the error function in Eq is given in the frequency domain by the expression
    - This result is known as the wiener filter after N.Wiener,who proposed the concepts in the year shown.the filter which consists of the term inside the brackets also is commonly referred to as the minimum mean square error filter or the least square error filter.
- We include references at the end of sources containing detailed derivations of the wiener filter.
    - The restored image in the spatial domain is given by the inverse Fourier transform of the frequency domain estimate F(u,v).
- If the noise is zero,then the noise power spectrum vanishes and the wiener filter reduces to the inverse filter.
    - However the power spectrum of the undegraded image seldom is known. Where k is a specified constant.
    - Example illustrates the power spectrum of wiener filtering over direct inverse filtering.the value of K was chosen interactively to yield the best visual results.
  - It illustrates the full inverse filtered result similarly is the radially limited inverse filter .
- These images are duplicated here for convenience in making comparisons.
    - As expected ,the inverse filter produced an unusable image.The noise in the inring filter.
- The wiener filter result is by no means perfect,but it does give us a hint as to image content.
    - The noise is still quite visible, but the text can be seen through a “curtain” of noise.

PART-C

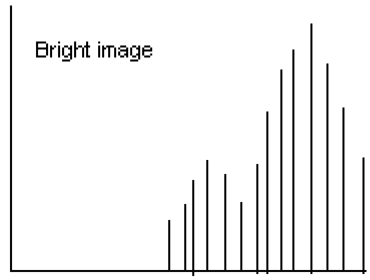
1.Explain Histogram processing

- The Histogram of the digital image with gray levels in the range [0,L-1]is the discrete function $p(r_k)=n_k/n$ where r_k is the kth gray level, n_k is the number of pixel,n is the total number of pixel in the image and k=0,1,2,…….L-1.
- $P(r_k)$ gives the an estimate probability of occurrence of gray-level r_k.. Figure show the the histogram of four basic types of images.

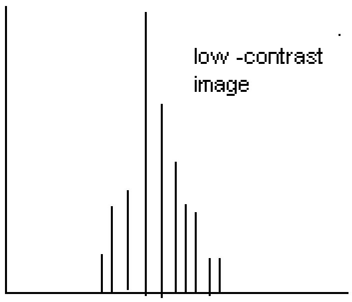
Figure: Histogram corresponding to four basic image types



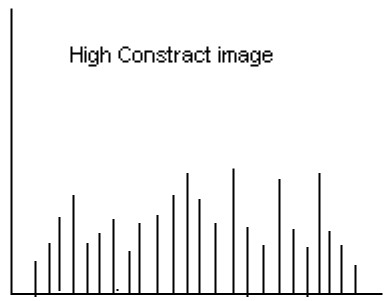
(a)



(b)



(c)



(d)

Histrogram Equalization

- Let the variable r represent the gray levels in the image to be enhanced. The pixel value are continous quantities normalized that lie in the interval [0,1] with r=0 represent black with r=1 represent white.
- The transformation of the form
    - $S=T(r)$ ……………………………………(1)
- Which produce a level s for every pixel value r in the original image.it satisfy condition:
    - o $T(r)$ is the single-valued and monotonically increasing in the interval $0 \leq r \leq 1$ and
    - o $0 \leq T(r) \leq 1$ for $0 \leq r \leq 1$
        - ▪ Condition 1 preserves the order from black to white in the gray scale.
        - ▪ Condition 2 guarantees a mapping that is consistent with the allowed range of pixel values.

          $R=T^{-1}(s) \quad 0 \leq s \leq 1$ ………………………..(2)
- The probability density function of the transformed graylevel is

  $$Ps(s)=[pr(r)dr/ds] \; r=T^{-1}(s) \qquad \text{…………………….(3)}$$
- Consider the transformation function

  $$S=T(r)= \int Pr(w)dw \qquad 0 \leq r \leq 1 \qquad \text{…………………….(4)}$$

  Where w is the dummy variable of integration .

  From Eqn(4) the derivative of s with respect to r is

  $$ds/dr=pr(r)$$

  Substituting dr/ds into eqn(3) yields

  $$Ps(s)=[1] \quad 0 \leq s \leq 1$$

Histogram Specfication

- Histogram equalization method does not lent itself to interactive application .
- Let Pr(r) and Pz(z) be the original and desired probability function. Suppose the histogram equalization is utilized on the original image

$$S=T(r)=\int Pr(w)\,dw \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(5)$$

- Desired image levels could be equalized using the transformation function

$$V=G(z)=\int Pr(w)dw \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(6)$$

- The inverse process is, $z=G^{-1}(v)$. Here Ps(s) and Pv(v) are identical uniform densities

$$Z=G^{-1}(s)$$

Assume that $G^{-1}(s)$ is single-valued, the procedure can be summarized as follow

1. Equalize the level of the original image using eqn(4)
2. Specify the desired density function and obtain the transformation function G(z) using eqn(6)
3. Apply the inverse transformation function $Z=G^{-1}(s)$ to the level obtained in step 1.

we can obtain result in combined transformation function

$$z=G^{-1}[T(r)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(7)$$

Histogram specification for digital image is limited one

1. First specify a particular histogram by digitizing the given function.
2. Specifying a histogram shape by means of a graphic device whose output is fed into the processor executing the histogram specification algorithm.

2. Explain Spatial Filtering

- The use of spatial mask for the image processing usually called spatial filtering and spatial mask are spatial filters.
- The linear filter classified into
    - Low pass
    - High pass
    - Band pass filtering
- Consider 3*3 mask

W1	W2	W3
W4	W5	W6
W7	W8	W9

- Denoting the gray level of pixels under the mask at any location by z1,z2,z3……,z9, the response of a linear mask is

$$R=w1z1+ w2z2 +\ldots\ldots..+w9z9$$

Smoothing Filters

- Lowpass Spatial filtering:
  - The filter has to have positive coefficient.
  - The response would be the sum of gray levels of nine pixels which could cause R to be out of the gray level range.
  - The solution is to scale the sum by dividing R by 9.The use of the form of mask are called neighborhood averaging

1/9

1	1	1
1	1	1
1	1	1

- Median filtering:
  - To achive noise reduction rather than blurring.
  - The gray level of each pixel is replaced by the median of the gray level in the neighbourhood of that pixel

Sharpening Filters

- Basic highpass spatial filtering:
  - The filter should be positive ecoefficient near the center and negative in the outer periphery.
  - The sum of the coefficient are 0.
  - This eliminate the zero- frequency term reducing significantly the global contrast of the image

1/9*

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

- High_boost filtering:

  The definition is

  High-boost=(A)(Original)-Lowpass
  =(A-1) (Original)+ Original –Lowpass
  =(A-1) (Original)+Hignpass

- Derivative Filters:
  - Averaging is anlog to integration , differentiation can be expected to have opposite effect and thus sharpen the image

3. Explain the Geometric Transformations used in image restoration.

- Geometric transformations are used for image restoration, modify the spatial relationship between the pixels in an image.
- Geometric transformations are often called rubber sheet transformations, because they are may be viewed as the process of printing an image on a sheet of rubber.
- The geometric transformations consists of two basic operations:

  (1) Spatial transformation
  (2) Gray level interpolation

1.Spatial transformations:-

An image f of pixel coordinates(x,y) undergoes geometric distortion to produce an image g with coordinates(x',y').this transformation may be expressed as

$$x' = r(x,y)$$
$$y' = s(x,y)$$

- where r(x,y) and s(x',y') are the spatial transformations that produced the geometrically distorted image g(x',y').
- If r(x,y) and s(x,y) were known analytically recovering f(x,y) from the distorted image g(x',y') by applying the transformations in reverse might possible theoretically.
- The method used most frequently to formulate the spatial relocation of pixels by the use of tiepoints,which are a subset of pixels whose location in the input and output image is known precisely.
- The vertices of the quadrilaterals are corresponding tiepoints.
    - $r(x,y)=c1x+c2y+c3xy+c4$
    - $S(x,y)=c5x+c6y+c7xy+c8$
    - $x'=c1x+c2y+c3xy+c4$
    - $y'=c5x+c6y+c7xy+c8$
- Since there are a total of eight known tiepoints these equations can be solved for eight coefficients $c_i, i=1,2,\ldots 8$.
- The coefficient constitute the geometric distortion model used to transform all pixels within the quadrilateral region defined by the tiepoints used to obtain the coefficients.
- Tiepoints are established by a number of different techniques depending on the application.

2. Gray level Interpolation:-

- Depending on the values of coefficients ci equations can yield noninteger values for x' and y'.
- Because the distorted image g is digital ,its pixel values are defined only at integer co ordinates .
- Thus using non integer values for x', y' causes a mapping into locations of g for which no gray levels are defined.
- The technique is used to accomplish this is called gray level interpolation.

4.Describe homomorphic filtering

- The illumination – reflectance model can be used to develop a frequency domain procedure for improving the appearance of an image by simultaneous gray – level compression and contrast enhancement.
- An image can be expressed as the product of illumination and reflectance components.

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

Where F i(u,v)) and F(r(u,v)) are the Fourier transformation of i(x,y)and r(x,y) respectively.

- The inverse (exponential) operation yields the desird enhanced image, denoted by g(x,y); that is,

$$\mathrm{Ln}[f(x,y)] = \ln[i(x,y)\, r(x,y)]$$
$$F[\ln(f(x,y))] = F[\ln(i(x,y)]+F[\ln(\, r(x,y))]$$

- This method is based on a special case of a class of systems know as homomorphism systems.
- In this particular application,
    - The key to the approach is the separation of the illumination and reflectance components achieved in the from.
    - The homomorphism filter function can then operate on these on these component separately.
    - The illumination components of an image generally is characterized by slow spatial variations.
    - While the reflectance component tends to vary abruptly, particularly at the junction, while the reflectance component tends to vary abruptly, particularly at the junctions of dissimilar objects.
    - A good deal of control can be gained over the illumination and reflectance components with a homomorphic filter.
    - This control requires specification of a filter function H(u.v) that affects the low - and high – frequency components of the Fourier transform in different ways.

5.Explain the different Noise Distribution in detail.

Introduction:

- Noise are unwanted signal which corrupts the original signal.
- Origin of noise source is during image acquisition and/or transmission and digitization.
- During capturing ,performance of imaging sensors are affected by the environmental conditions due to the quality of sensors.
- Image acquisition are the principle source of noise.
- Due to the interference in the transmission it will affect the transmission of the image.
- **<u>Types:</u>**

Rayleigh noise:

The PDF is

$P(Z)= 2(z-a)e^{-(z—a)2/b}/b$ for $Z>=a$

0 for $Z<a$

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

standard deviation $\sigma^2=b(4-\pi)/4$

Gamma noise:

The PDF is

$P(Z)=a^b z^{b-1} ae^{-az}/(b-1)$ for $Z>=0$
0 for $Z<0$
mean $\mu=b/a$
standard deviation $\sigma^2=b/a^2$

Exponential noise

The PDF is
$P(Z)= ae^{-az}$ $Z>=0$
0 $Z<0$
mean $\mu=1/a$
standard deviation $\sigma^2=1/a^2$

Uniform noise:

The PDF is
$P(Z)=1/(b-a)$ if $a<=Z<=b$
0 otherwise
mean $\mu=a+b/2$
standard deviation $\sigma^2=(b-a)^2/12$

Impulse noise:

The PDF is
$P(Z) =P_a$ for z=a
P_b for z=b
0 Otherwise

UNIT I V

1. What is segmentation?

Segmentation subdivides on image in to its constitute regions or objects. The level to which the subdivides is carried depends on the problem being solved .That is segmentation should when the objects of interest in application have been isolated.

2. Write the applications of segmentation.

- Detection of isolated points.
- Detection of lines and edges in an image.

3. What are the three types of discontinuity in digital image?

Points, lines and edges.

4. How the derivatives are obtained in edge detection during formulation?

The first derivative at any point in an image is obtained by using the magnitude of the gradient at that point. Similarly the second derivatives are obtained by using the laplacian.

5. Write about linking edge points.

The approach for linking edge points is to analyze the characteristics of pixels in a small neighborhood (3x3 or 5x5) about every point (x,y)in an image that has undergone edge detection. All points that are similar are linked, forming a boundary of pixels that share some common properties.

6. What are the two properties used for establishing similarity of edge pixels?

(1) The strength of the response of the gradient operator used to produce the edge pixel.

(2) The direction of the gradient.

7. What is edge?

An edge isa set of connected pixels that lie on the boundary between two regions edges are more closely modeled as having a ramplike profile. The slope of the ramp is inversely proportional to the degree of blurring in the edge.

8. Give the properties of the second derivative around an edge?

- The sign of the second derivative can be used to determine whether an edge pixel lies on the dark or light side of an edge.
- It produces two values for every edge in an image.
- An imaginary straightline joining the extreme positive and negative values of the second derivative would cross zero near the midpoint of the edge.

9. Define Gradient Operator?

First order derivatives of a digital image are based on various approximation of the 2-D gradient. The gradient of an image f(x,y) at location(x,y) is defined as the vector

Magnitude of the vector is

$\Delta f = mag(\Delta f) = [Gx^2 + Gy^2]^{1/2}$

$\infty(x,y) = \tan^{-1}(Gy/Gx)$

$\infty(x,y)$ is the direction angle of vector Δf

10. What is meant by object point and background point?

To execute the objects from the background is to select a threshold T that separate these modes. Then any point (x,y) for which f(x,y)>T is called an object point. Otherwise the point is called background point.

11. What is global, Local and dynamic or adaptive threshold?

When Threshold T depends only on f(x,y) then the threshold is called global . If T depends both on f(x,y) and p(x,y) is called local. If T depends on the spatial coordinates x and y the threshold is called dynamic or adaptive where f(x,y) is the original image.

12. Define region growing?

Region growing is a procedure that groups pixels or subregions in to layer regions based on predefined criteria. The basic approach is to start with a set of seed points and from there grow regions by appending to each seed these neighbouring pixels that have properties similar to the seed.

13. Specify the steps involved in splitting and merging?

Split into 4 disjoint quadrants any region Ri for which P(Ri)=FALSE.

Merge any adjacent regions Rj and Rk for which P(RjURk)=TRUE.

Stop when no further merging or splitting is positive.

14. Define pattern.

A pattern is a quantitative or structural description of an objective or some other entity of interest in an image,

15. Define pattern class.

A pattern class is a family of patterns that share some common properties .Pattern classes are denoted w1,w2,----wm, where M is the number of classes .

16.List the three pattern arrangements.

Vectors
Strings
Treestching

17. Give the decision theoretic methods.

Matching-Matching by minimum distance classifier
Matching by correlation

18. Define training pattern and training set.

The patterns used to estimate the parameters are called training patterns,anda set of such patterns from each class is called a training set.

19. Define training

The process by which a training set is used to obtain decision functions is called learning or training.

20. What are the layers in back propagation network?

Input layer, Hidden layer and out put layer

PART –B

1. **Write short notes on image segmentation.**

- Segmentation subdivides on image in to its constitute regions or objects. The level to which the subdivides is carried depends on the problem being solved .
- Examples: In autonomous air to ground target acquisition applications identifying vehicles on a road is of interest.
- The first step is to segment the road from the image and then to segment the elements of the road down to objects of a range of sizes that correspond potential vehicles.
- In target acquistition ,the system designer has n control of the environment.
- So the usual approach is to focus on selecting the types of sensors most likely to enhance the objects of interest .
- Example is the use of infrared imaging to detect objects with a strong heat signature,such as tanks in motion.
- Segmentation algorithms for monochrome images are based on one of the two basic properties of gray level values . They are discontinuity and similarity.

- Based on the first category ,the approach is based on abrupt changes in gray level and the areas of interest based on this category are detection of isolated points and detection of lines and edges in an image.
- Based on the second category the approach is based on thresholding, region growing and region splitting and merging .
- The concept of segmenting an image based on discontinuity or similarity of the gray level values of its pixels is applicable to both static and dynamic images.

2. Write short notes on edge detection

Edge Detection:

- Edge detection is "local" image processing methods designed to detect edge pixels.
- Concept that is based on a measure of intensity-level discontinuity at a point.
- It is possible to link edge points into edge segments, and sometimes these segments are linked in such a way that they correspond to boundaries, but this is not always the case.

The image gradient and its properties:

- The tool of choice for finding edge strength and direction at location (x,y) of an image, f, is the gradient, denoted by ∇f, a and defined as the vector

$$\nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix}$$

- The magnitude length of vector ∇f, denoted as $M(x,y)$

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

  Is the value of the rate of change in the direction of the gradient vector.
- The direction of the gradient vector is given by the angle

$$\alpha(x,y)=\tan^{-1} \; g_y/g_x$$

  measured with respect to the x-axis.
- Follows, using these differences as our estimates of the partials, that $\partial f/\partial x=-2$ and $\partial f/\partial y=2$ at the point in equation. Then

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} = \begin{matrix} -2 \\ 2 \end{matrix}$$

  from which we obtain $M(x,y)=2\sqrt{2}$ at that point.

Gradient operators:

- Obtaining the gradient of an image requires computing the partials derivatives $\partial f/\partial x$ and $\partial f/\partial y$ at every pixel location in the image.

$$g_x=\partial f(x,y)/\partial x= f(x+1,y)- f(x,y)$$
$$g_y=\partial f(x,y)/\partial y= f(x,y+1)- f(x,y)$$

- An approach used frequently is to approximate the gradient by absolute value:

$$\nabla f \approx |G_x| + |G_y|$$

The Laplacian

- The laplacian of a 2-D function f(x,y) is a second order derivatives defined as

$$\nabla^2 f=\partial^2 f/\partial^2 x+\partial^2 f/\partial^2 y$$

- The first laplacian is combined with smoothing as a precursor to finding edges via zero crossings. Consider the function.

$$\nabla^2 f=8z5-(z1+z2+z3+z4+z6+z7+z8+z9)$$

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

3.Write Short notes on edge linking by local processing.

- One of the simplest approaches f or linking edge points is to analyze the characteristics of the pixels in a small neighborhood about every point in an image that has undergone edge detection.
- Two properties used for establishing similarity of edge pixels in the analysis are
  - The strength of the response of the gradient operator used to produce the edge pixel,
  - The direction of the gradient.

The first property is given by the value of ∇f.
Thus an edge pixel with coordinates (x',y') and in the predefined neighborhood of (x,y) is similar in magnitude to the pixel at (x,y) if $|\nabla f(x,y) - \nabla(x',y')|<=T$ where T is a nonnegative threshold.
The direction of the gradient vector is given by

$$\alpha(x,y)=\tan^{-1}\ gy/gx$$

Then an edge pixel at (x',y') in the predefined neighborhood of (x,y) has an angle similar to the pixel at (x,y) if $|\ \alpha(x,y)= \alpha(x',y')|<A$ where A is an angle threshold. Note that the direction of the edge at (x,y) in reality is perpendicular to the direction of the gradient vector at that point.
A point in the predefined neighborhood of (x,y) is linked to the pixel at (x',y') if both magnitude and direction criteria are satisfied. This process is repeated for every location in the image. A record must be kept of linked points as the center of the neighborhood is moved from pixel to pixel. A simple bookkeeping procedure is to assign a different gray level to each set of linked edge pixels.

4.Write short notes on the applications of artificial neural networks in image processing.

The real-time automatic images processing and pattern recognition are very important for many problems in medicine, physics, geology, space research, military applications and so on. For example, it is necessary for pilots and drivers for immediate decision-making in poor visibility conditions. An approach to image enhancement through artificial neural network's (ANN) processing is proposed.ANN is for images enhancement through approximation of image transform function *T*. This function is approximated with use of ANN which is trained evolutionary in the time of test images processing. Each ANN is genetically encoded as the list of its connections. Truncation selection is used for parental subpopulation formation. Original crossover and mutation operators, which respect structures of the ANNs undergoing recombination and mutation, are used. Nodes with sigmoid activation functions are considered. The population size adapts to the properties of evolution during the algorithm run using simple resizing strategy. In this application pixel-by-pixel brightness processing with use of ANN paradigm is adopted. The topology of ANN is tuned simultaneously with connections weights. The ANN approximating *T* function should have three input nodes and one output node. During the training we evaluate each ANN with respect to the visual quality of the processed images.

The three-step procedure for image enhancement is proposed:
(1) multiplicative adjustment of image brightness
(2) local level processing using ANN;
(3) global level auto smoothing algorithm.

The artificial neural network training stage with use of single 128x128 pixels image takes about 70 seconds on the Intel Pentium IV 3 GHz processor. After completion of the learning process the obtained artificial neural network is ready to process
arbitrary images that were not presented during the training. The processing time for 512x512 pixels image is about 0.25 second. The ANN, as a rule, included 3 input nodes, one or more hidden nodes and one output node.

PART-C

1.Discuss region oriented segmentation in detail

The objective of segmentation is to partition an image into regions. We approached this problem by finding boundaries between based on discontinuities in gray levels, segmentation was accomplished via thresholds based on the distribution of pixels properties, such as gray level values or color.

Basic Formulation:

Let Represent the region of image. We may view segmentation as a process that partition R into n subregions,$R_1, R_2, \ldots\ldots\ldots\ldots$,such that

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

(b) R_i is a connected region, $i=1,2,\ldots\ldots\ldots\ldots n$.

(c) $R_i \cap R_j = \Phi$ for all i and $j, i \neq j$.

(d) $P(R_i) = \text{TRUE}$ for $i=1,2,\ldots\ldots\ldots\ldots\ldots\ldots n$.

(e) $P(R_i \cup R_j) = \text{FALSE}$ for $i \neq j$.

Here, $P(R_i)$ is a logical predicate defined over the points in set R_i and Φ is the null set.

- Condition (a) indicates that the segmentation must be complete that is every pixel must be in a region.
- Condition (b) requires that points in a region must be connected in some predefined sense.
- Condition(c) indicates that the regions must be disjoint.
- Condition(d) deals with the properties that must be satisfied by the pixels in a segmented region.

Region Growing:

As its name implies region growing is a procedure that groups pixel or subregions into larger regions based on predefined criteria. The basic approach is to start with a set of "seed" points and from these grow regions.

- If the result of these computation shows clusters of values, the pixels whose properties place them near the centroid of these clusters can be used as seeds.
- Descriptors alone can yield misleading results if connectivity or adjacency information is not used in the region growing process.

Region Splitting and Merging:

The procedure just discussed grows regions from a set of seed points. An alternative into subdivided an image initially into a set of arbitrary, disjointed regions and then merge and/or split the regions in an attempt to satisfy the conditions.

<table>
<tr><td>R1</td><td colspan="2">R2</td></tr>
<tr><td rowspan="2">R3</td><td>R41</td><td>R42</td></tr>
<tr><td>R43</td><td>R44</td></tr>
</table>

1. Split into four disjoint quadrants any region R_i for which $P(R_i)=\text{FALSE}$.
2. Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)=\text{TRUE}$.
3. Stop when no further merging or splitting is possible.

Mean and standard deviation of pixels in a region to quantify the texture of region.

Role of thresholding:

We introduced a simple model in which an image f(x,y) is formed as the product of a reflectance component r(x,y) and an illumination components i(x,y). consider the computer generated reflectance function.

- The histogram of this function is clearly bimodal and could be portioned easily by placing a single global threshold, T, in the histogram valley.
- Multiplying the reflectance function by the illumination function.
- Original valley was virtually eliminated, making segmentation by a single threshold an impossible task.
- Although we seldom have the reflectance function by itself to work with, this simple illustration shows that the reflective nature of objects and background can be such that they are separable.

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

Taking the natural logarithm of this equation yields a sum:

$$\begin{aligned} z(x,y) &= \ln f(x,y) \\ &= \ln i(x,y)+\ln r(x,y) \\ &= i'(x,y)+r'(x,y) \end{aligned}$$

- If $i'(x,y)$ and $r'(x,y)$ are independent random variable, the histogram of z(x,y) is given by the convolution of the histogram of $i'(x,y)$ and $r'(x,y)$.
- But if $i'(x,y)$ had a border histogram the convolution process would smear the histogram of $r'(x,y)$, yielding a histogram for z(x,y) whose shape could be quite different from that of the histogram of $r'(x,y)$.
- The degree of distortion depends on the broadness of the histogram of $i'(x,y)$, which in turn depends on the nonuniformity of the illumination function.
- We have dealt with the logarithm of $f(x,y)$, instead of dealing with the image function directly.
- When access to the illumination source is available, a solution frequently used in practice to compensate for nonuniformity is to project the illumination pattern onto a constant, white reflective surface.
- This yields an image g(x,y)=ki(x,y), where k is a constant that depends on the surface and i(x,y) is the illumination pattern.
- For any image $f(x,y)=i(x,y)r(x,y)$ obtained from the same illumination function, simply dividing $f(x,y)$ by g(x,y) yields a normalized function h(x,y)= $f(x,y)/g(x,y)= r(x,y)/k$.
- Thus, if r(x,y) can be segmented by using a single threshold T, then h(x,y) can be segmented by using single threshold of value T/k.

2.**Explain Back propogation neural networks in detail**

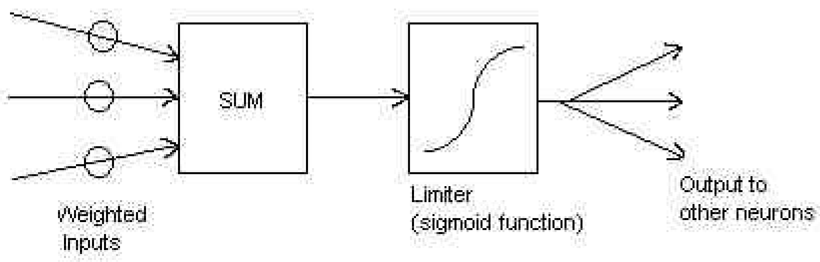


Fig: Structure of a basic neuron

Back propagation training algorithm:

Basic neuron: Figure shows the Structure of a basic neuron. A set of inputs applied either from the outside of from a previous layer. Each of these is multiplied by a weight and the products are summed. This summation of this products is turned NET and must be calculated for each neuron in the network. After NET is calculated an activation function F is applied to modify it, thereby producing the signal out

Where,

$$NET = x1w1 + x2w2 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots xnwn$$

$$NET = \sum_{i=1}^{n} xiwi$$

and out=F(NET)

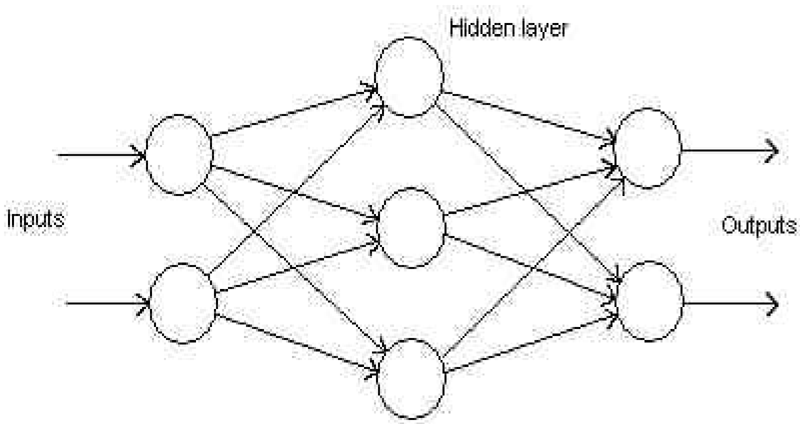


Fig: Back propagation neural network

Multilayer Back Propagation Network:

Figure shows the Back propagation neural network .The first set of neurons serve only as distribution points. They perform no input summation. The input signal is simply passed into the weight on there outputs. Each neuron in subsequent layers produces NET and OUT signals as described above. A neuron is associated with the set of weights that connects to its input. The input of distribution layer is designated as layer 0. Backpropagation can be applied to network with a number of layers. However only two layers of weights are needed to demonstrate the algorithm.

Steps involved in backpropagation training:

- Select the next training pair from the training set. Apply input vector to the network input.
- Calculate the output of the network.
- Calculate the error between the output network and the desired output.
- Adjust the weight of the network in a way that minimizes the error.
- Repeat steps1 through 4 for each vector in the training until the error for the entire set acceptably low.

Forward pass:

Step 1 and 2 can be expressed in vector form as follows: an input vector x is applied and the output vector y is produced. Calculations in multilayer network are done layer by layer starting at the layer nearest to the inputs. The NET value of each neuron in the first layer is calculated as the weight sum of its neurons input. The activation function F then 'squashes' NET to produce the OUT value for each neuron in that layer. Once a set a output for a layer is found, it serves as input to the next layer. This process is repeated layer by layer until the final set of output network is produced.

Reverse bias:

Because a target value is available for each neuron in the output layer, adjusting ht associated weights id easily accomplished as a modification of the delta rule. Interior layers are referred to as hidden layers as their outputs has no target values for comparison. Hence the training is more complicated. Hidden layers have no target vectors so the ordinary training process described above can not be used.

3.Disscuss in detail on pattern and pattern classes.

A pattern is a quantitative or structural description of an objective or some other entity of interest in an image,

A pattern class is a family of patterns that share some common properties .Pattern classes are denoted w1,w2,----wm, where M is the number of classes

Three principle pattern arrangements used in practice are vectors(for quantitative descriptors) and strings and trees (for structural descriptions) .

Patternvectors are represented by bold lower case letters such as **x,y**, and **z,**where

Each component x represent the ith descriptors.Pattern vectors are represented in coloumns (i.e. n x 10 marices) or in the equilant form x=9x1,x2,------xn)T,T-transpose.

The nature of the pattern vector depends on the measurement technique used to describe the physical pattern itself.

Ex. If we want to describe the three types of iris floers(iris setosa,virginica,and versicolor)by measuring the width and length of the petals.It is represented in the vector form x=[x1,x2]T;x1,x2 correspond to width length respectively.Three pattern classes are w1,w2,w3 corresponding to the three verities.

Because the petals of all flowers vary in width and length to some degree the pattern vectors describing three flowers also will vary, not only between different classes ,but also with in a class.
The result of this classic feature selection problem shows that the degree of class seperability depends strongly on the choice of pattern measurements selected for an application.

UNIT V

1. What is image compression?

Image compression refers to the process of redundancy amount of data required to represent the given quantity of information for digital image. The basis of reduction process is removal of redundant data.

2. What is Data Compression?

Data compression requires the identification and extraction of source redundancy. In other words, data compression seeks to reduce the number of bits used to store or transmit information.

3. What are two main types of Data compression?

- Lossless compression can recover the exact original data after compression. It is used mainly for compressing database records, spreadsheets or word processing files, where exact replication of the original is essential.
- Lossy compression will result in a certain loss of accuracy in exchange for a substantial increase in compression. Lossy compression is more effective when used to compress graphic images and digitised voice where losses outside visual or aural perception can be tolerated.

4. What is the need for Compression?

In terms of storage, the capacity of a storage device can be effectively increased with methods that compress a body of data on its way to a storage device and decompresses it when it is retrieved.

In terms of communications, the bandwidth of a digital communication link can be effectively increased by compressing data at the sending end and decompressing data at the receiving end.

At any given time, the ability of the Internet to transfer data is fixed. Thus, if data can effectively be compressed wherever possible, significant improvements of data throughput can be achieved. Many files can be combined into one compressed document making sending easier.

5. What are different Compression Methods?

Run Length Encoding (RLE)
Arithmetic coding
Huffman coding and
Transform coding

6. Define is coding redundancy?

If the gray level of an image is coded in a way that uses more code words than necessary to represent each gray level, then the resulting image is said to contain coding redundancy.

7. Define interpixel redundancy?

The value of any given pixel can be predicted from the values of its neighbors. The information carried by is small. Therefore the visual contribution of a single pixel to an image is redundant. Otherwise called as spatial redundant geometric redundant or interpixel redundant.

Eg: Run length coding

8. What is run length coding?

Run-length Encoding, or RLE is a technique used to reduce the size of a repeating string of characters. This repeating string is called a *run*; typically RLE encodes a run of symbols into two bytes, a count and a symbol. RLE can compress any type of data regardless of its information content, but the content of data to be compressed affects the compression ratio. Compression is normally measured with the compression ratio:

9. Define compression ratio.

Compression Ratio = original size / compressed size: 1

10. Define psycho visual redundancy?

In normal visual processing certain information has less importance than other information. So this information is said to be psycho visual redundant.

11. Define encoder

Source encoder is responsible for removing the coding and interpixel redundancy and psycho visual redundancy.

There are two components

A) Source Encoder

B) Channel Encoder

12. Define source encoder

Source encoder performs three operations

1) Mapper -this transforms the input data into non-visual format. It reduces the interpixel redundancy.

2) Quantizer - It reduces the psycho visual redundancy of the input images .This step is omitted if the system is error free.

3) Symbol encoder- This reduces the coding redundancy .This is the final stage of encoding process.

13. Define channel encoder

The channel encoder reduces reduces the impact of the channel noise by inserting redundant bits into the source encoded data.

Eg: Hamming code

14. What are the types of decoder?

Source decoder- has two components

a) Symbol decoder- This performs inverse operation of symbol encoder.

b) Inverse mapping- This performs inverse operation of mapper.

Channel decoder-this is omitted if the system is error free.

15. What are the operations performed by error free compression?

1) Devising an alternative representation of the image in which its interpixel redundant are reduced.

2) Coding the representation to eliminate coding redundancy

16. What is Variable Length Coding?

Variable Length Coding is the simplest approach to error free compression. It reduces only the coding redundancy. It assigns the shortest possible codeword to the most probable gray levels.

17. Define Huffman coding

- Huffman coding is a popular technique for removing coding redundancy.
- When coding the symbols of an information source the Huffman code yields the smallest possible number of code words, code symbols per source symbol.

18. Define Block code

Each source symbol is mapped into fixed sequence of code symbols or code words. So it is called as block code.

19. Define instantaneous code

A code word that is not a prefix of any other code word is called instantaneous or prefix codeword.

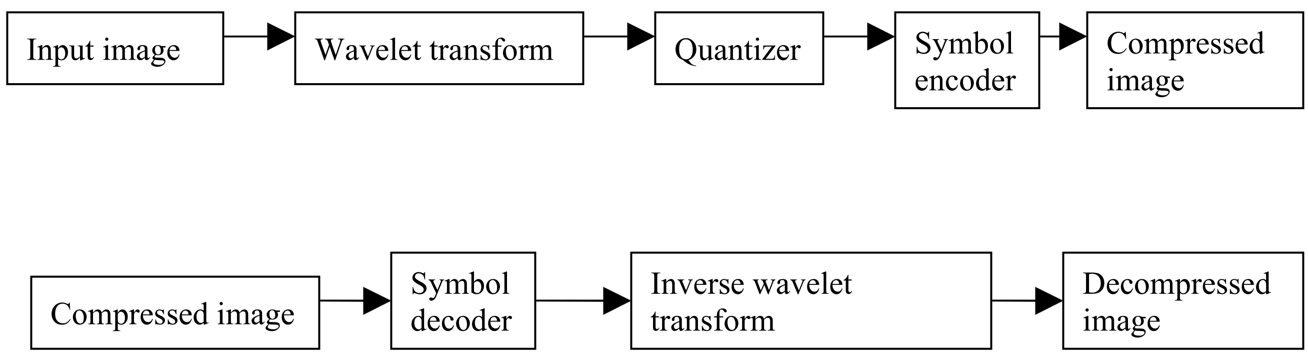
20. Define arithmetic coding

In arithmetic coding one to one corresponds between source symbols and code word doesn't exist where as the single arithmetic code word assigned for a sequence of source symbols. A code word defines an interval of number between 0 and 1.

21. What is bit plane Decomposition?

An effective technique for reducing an image's interpixel redundancies is to process the image's bit plane individually. This technique is based on the concept of decomposing multilevel images into a series of binary images and compressing each binary image via one of several well-known binary compression methods.

22. Draw the block diagram of transform coding system



23. How effectiveness of quantization can be improved?

- Introducing an enlarged quantization interval around zero, called a dead zero.
- Adapting the size of the quantization intervals from scale to scale. In either case, the selected quantization intervals must be transmitted to the decoder with the encoded image bit stream.

24. What are the coding systems in JPEG?

1. A lossy baseline coding system, which is based on the DCT and is adequate for most compression application.
2. An extended coding system for greater compression, higher precision or progressive reconstruction applications.
3. a lossless independent coding system for reversible compression.

25. What is JPEG?

The acronym is expanded as "Joint Photographic Expert Group". It is an international standard in 1992. It perfectly Works with color and grayscale images, Many applications e.g., satellite, medical,...

26. What are the basic steps in JPEG?

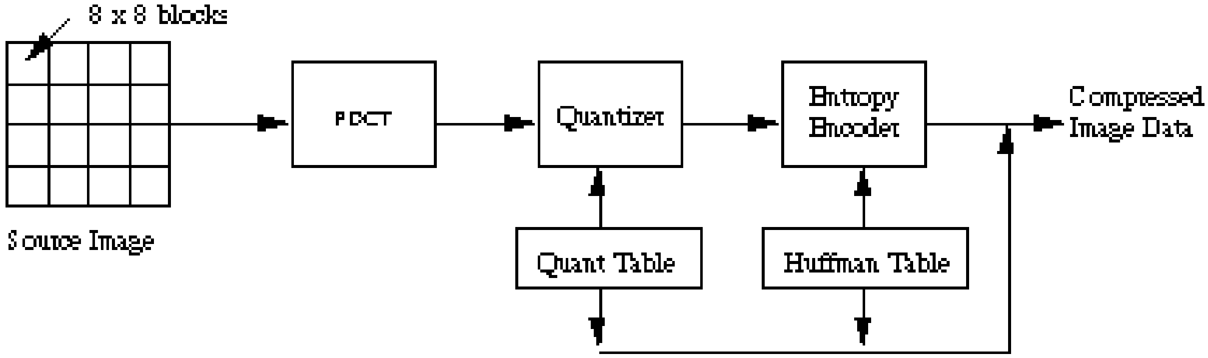
The Major Steps in JPEG Coding involve:

- DCT (Discrete Cosine Transformation)
- Quantization
- Zigzag Scan
- DPCM on DC component
- RLE on AC Components
- Entropy Coding

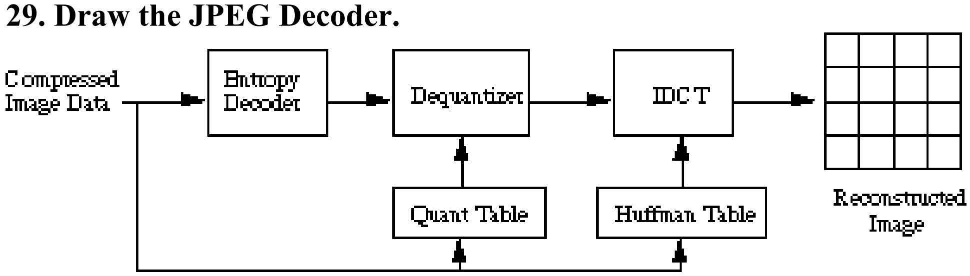
27. What is MPEG?

The acronym is expanded as "Moving Picture Expert Group". It is an international standard in 1992. It perfectly Works with video and also used in teleconferencing

28. Draw the JPEG Encoder.



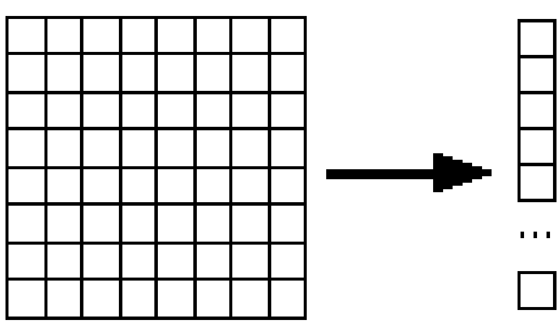
29. Draw the JPEG Decoder.

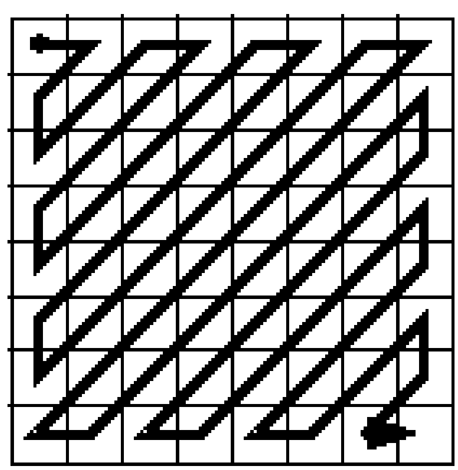


30. What is zig zag sequence?

The purpose of the Zig-zag Scan:

- To group low frequency coefficients in top of vector.
- Maps 8 x 8 to a 1 x 64 vector





31. Define I-frame

I-frame is Intraframe or Independent frame. An I-frame is compressed independently of all frames. It resembles a JPEG encoded image. It is the reference point for the motion estimation needed to generate subsequent P and P-frame.

32. Define P-frame

P-frame is called predictive frame. A P-frame is the compressed difference between the current frame and a prediction of it based on the previous I or P-frame

33. Define B-frame

B-frame is the bidirectional frame. A B-frame is the compressed difference between the current frame and a prediction of it based on the previous I or P-frame or next P-frame. Accordingly the decoder must have access to both past and future reference frames.

PART B

1. **Define Compression and explain data Redundancy in image compression**

Compression: It is the process of reducing the size of the given data or an image. It will help us to reduce the storage space required to store an image or File.

Data Redundancy:

The data or words that either provide no relevant information or simply restate that which is already known .It is said to be data redundancy.

Consider N1 and N2 number of information carrying units in two data sets that represent the same information

Data Redundancy Rd = 1-1/Cr

Where Cr is called the Compression Ratio.

` Cr=N1/N2.

Types of Redundancy

There are three basic Redundancy and they are classified as

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

1. Coding Redundancy :

We developed this technique for image enhancement by histogram processing on the assumption that the grey levels of an image are random quantities. Here the grey level histogram of the image also can provide a great deal of insight in the construction of codes to reduce the amount of data used to represent it.

2. Interpixel Redundancy :

Inorder to reduce the interpixel redundancy in an image, the 2-D pixel array normally used for human viewing and interpretation must be transformed in to more efficient form.

3. Psychovisual Redundancy:

Certain information simply has less relative importance than other information in the normal visual processing. This information is called Psycovisual Redundant.

2. Explain the Coding phase in JPEG

In this approach the lable for the DC and AC coefficient are coded differently using Huffman codes. The DC coefficient values partitioned into categories. The categories are then Huffman coded. The AC coefficient is generated in slightly different manner. There are two special codes: End-of-block(EOF) and ZRL

Table: Coding of the differences of the DC labels

0				0			
1			-1		1		
2		-3	-2		2		3
3	-7	……………………	-4		4	………	7

………………………………………

Table: sample table for obtaining the Huffman code for a given label value and run length

Z/C	Codeword	Z/c	Codeword	…..	Z/C	Codeword
0/0	1010				F/0	11111111001
0/1	00	1/1	1100		F/1	1111111111110101

……………………

To encode the AC coefficient First using Zigzag scan. We obtain

-9 3 0 0 0 0 0 ……… 0

The first value belong to category 1. transmit the code corresponding to 0/1 follow by a single bit 1 to indicate that the value being transmitted is 1 and not -1 .Simillarly other AC coefficient code are transmited.

To obtain the reconstruction of the original block Dequantization is performed and taking inverse transform of the coefficient we get the reconstructed block

3.Explain Huffman coding with an example.

- This technique was developed by David Huffman.

- The codes generated using this technique or procedure are called Huffman codes.
- These codes are prefix codes and are optimum for a given model.

The Huffman procedure is based on two observations regarding optimum prefix codes

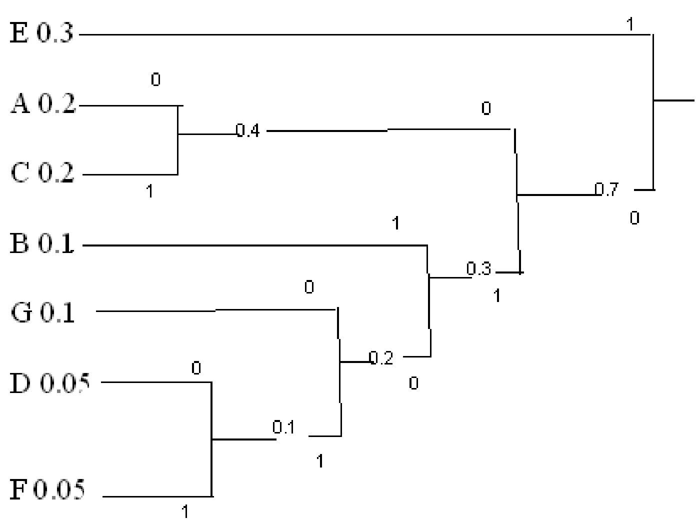
1.In an optimum code, symbols that occur more frequently will have shorter code words than symbols that occur less frequently.

2.in an optimum code ,the two symbols that occur least frequently will have the same length

Design of a Huffman Code

To design a Huffman code ,we first sort the letters in descending probability

Find the Huffman code for the following:
P(A)=0.2, p(B)=0.1, p(C)=0.2,p(D)=0.05,p(E)=0.3,p(F)=0.05,p(G)=0.1



Find the average length and entropy

Average length =I= $\sum_{k=1}^{M} p(a_k)l(a_k)$

$I=3(0.2)+3(0.1)+3(0.2)+5(0.05)+1(0.3)+5(0.05)+4(0.1)$
$=2.7$ bits/symbol

Entropy =H(a$_k$)= - $\sum_{k=1}^{M} p(a_k)\log_2 p(a_k)$
$=0.7667$

Find Efficiency

Efficiency =η = Entropy/average length
=0.284%

Find redundancy

Redundancy =Re =1-η =0.716

PART -C

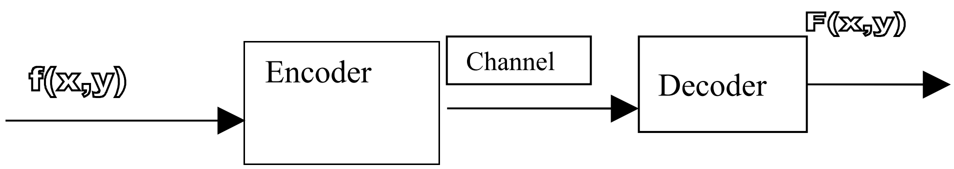
1) **Define Compression and Explain the general compression system model**

Compression: It is the process of reducing the size of the given data or an image. It will help us to reduce the storage space required to store an image or File.

Image Compression Model:

There are two Structural model and they are broadly Classified as follows

1. An Encoder
2. A Decoder.

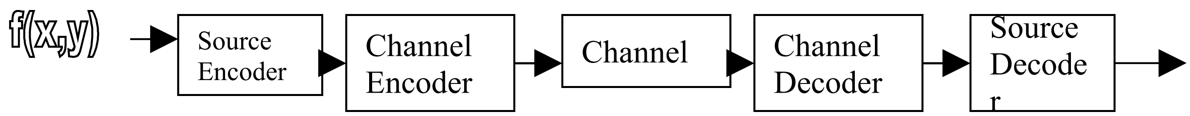


An Input image f(x,y) is fed in to encoder and create a set of symbols and after transmission over the channel ,the encoded representation is fed in to the decoder.

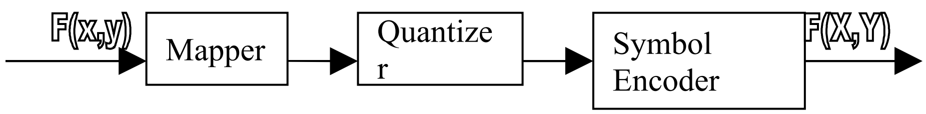
A General Compression system model:

The General system model consist of the following components,They are broadly classified as

1. Source Encoder
2. Channel Encoder
3. Channel
4. Channel Decoder
5. Souce Decoder



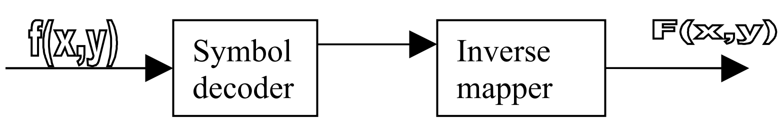
The Source Encoder Will removes the input redundancies. The channel encoder will increase the noise immunity of the source encoder's output. If the channel between encoder and decoder is noise free then the channel encoder and decoder can be omitted.



MAPPER:
It transforms the input data in to a format designed to reduce the interpixel redundancy in the input image.

QUANTIZER:
It reduce the accuracy of the mapper's output.

SYMBOL ENCODER:
It creates a fixed or variable length code to represent the quantizer's output and maps the output in accordance with the code.

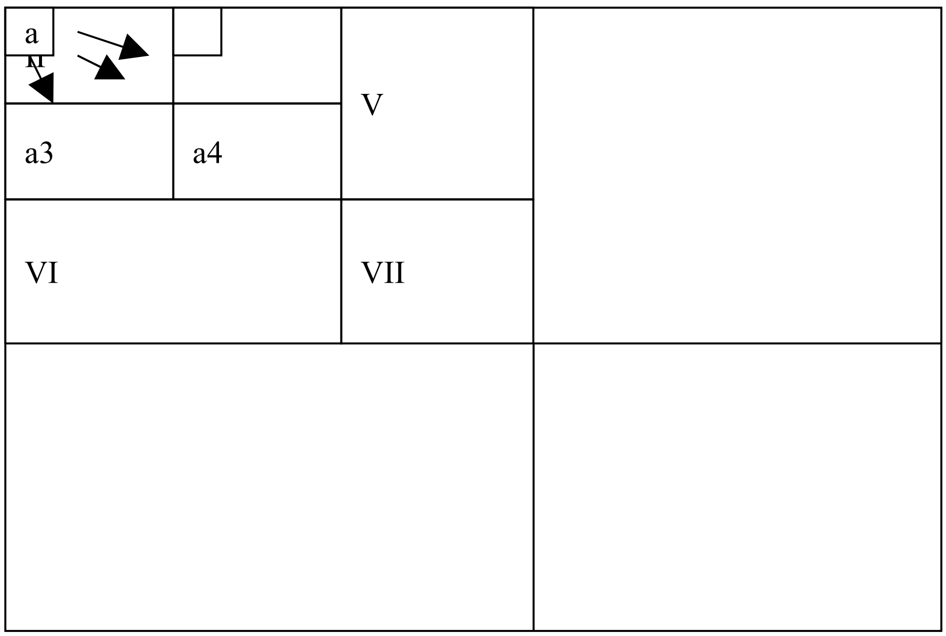


SYMBOL DECODER:
The inverse operation of the source encoder's symbol will be performed and maps the blocks.

2.Explain the concepts of Embedded Zero Tree coding

EZW coder was introduced by Shapiro. It is a quantization and coding strategy that characteristics of the wavelet decomposition.The particular characteristic used by the EZW algorithm is that there are wavelet coefficients in different subbands that represent the same spatial location in the image.

In 10-band decomposition ,the coefficient a in the upper-left corner of band I represents the same spatial location as coefficient a1 represents the same spatial location as coefficients a11,a12,a13,a14 in band V. Each of these pixel represents the same spatial location as four pixels in band VIII.



We can visualize the relationships of these coefficients in form of tree:The coefficient a forms the root of the tree with three descendants a1,a2,a3.

EZW algorithm is a multiple pass algorithm,with each pass consisting of two steps.

1.significance map encoding or the dominant pass
2.refinement or the subordinate pass

If Cmax is the value of the largest coefficient,the initial value of the threshold T0 is given by

$$T0=2^{[\text{Log Cmax}]}$$

This selection guarantees that the largest coefficient will lie in the interval[T0,2T0].In each pass,the threshold Ti is reduced to half the value it had in the previous pass:

$$Ti=1/2(Ti-1)$$

For given value of Ti,we assign one of four possible labels to the coefficients:

1.significance positive(sp)
2.significant negative(sn)
3.zerotree root(zr)
4.isolated zero(iz)

The coefficients labeled significant are simply those that fall in the outer levels of the quantized and are assigned an initial reconstructed value of 1.5Ti or -1.5Ti,depending on whether the coefficient is positive or negative.

3.Discuss MPEGcompression standard

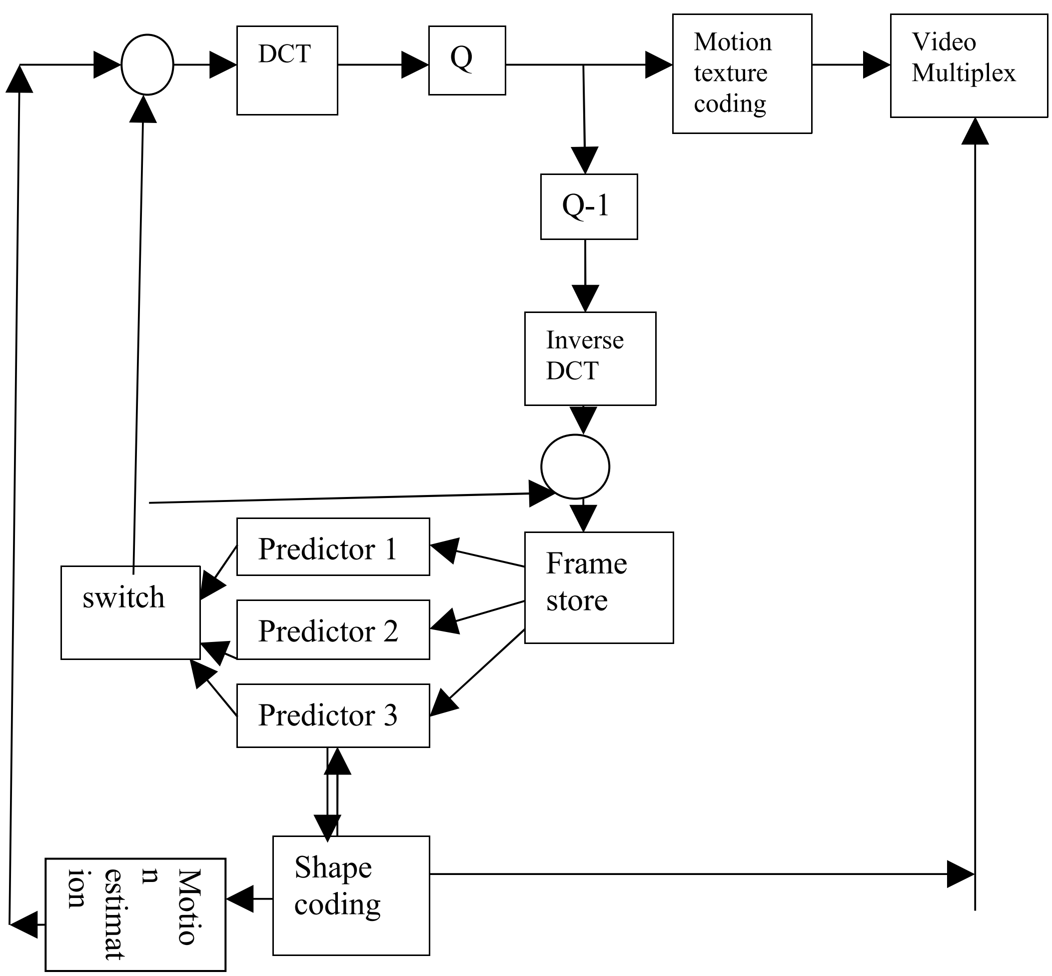
Introduction:
-The basic structure of the compression algorithm proposed by mpeg is very similar to that of ITU-T H.261
-In mpeg the blocks are organized in macro blocks which are defined in the same manner as that of H.261 algorithm
-The mpeg standard initially had applications that require digital storage and retrieval as a major focus

Frames
I-Frames
-Mpeg includes some frames periodically that are coded without any reference to the past Frames. These frames are called I-frames
-I-frames do not use temporal correlation for prediction.Thus the number of frames between two consecutive Iframes is a trade-off between compression efficiency and convenience.
P and B frames
-In order to improve the compression efficiency mpeg1 algorithm contains two other types of frames: predictive coded and bidirectionally predictive coded frames
-Generally the compression efficiency of P-frames is substantially higher than Iframes
Anchor frames
-The I and P frames are sometimes are anchor frames
-To compensate for the reduction in the amount of compression due to the frequent use of Iframes the mpeg standard introduced Bframes
Group of pictures(GOP)
-GOP is a small random access unit in the video sequence
-The GOP structure is set up as a tradeoff between the high compression efficiency of
-Motion compensated coding and the coding and the fast picture acquisition capability of periodic intra-only processing
-The format for mpeg is very flexible however the mpeg committee has provided some suggested value for the various parameters
-For Mpeg 1 these suggested values are called the constraint parameter bitstream
<u>MPEG2</u>
-It takes a toolkit approach providing a number of subsets each containing different options
-For a particular application the user can select from a set of profiles and levels
Types of profiles

-Simple
-Main
-Snr-scalable
-Spatially scalable
-High
-Simple profile uses the Bframes.but removal of the Bframes makes the requirements simpler.

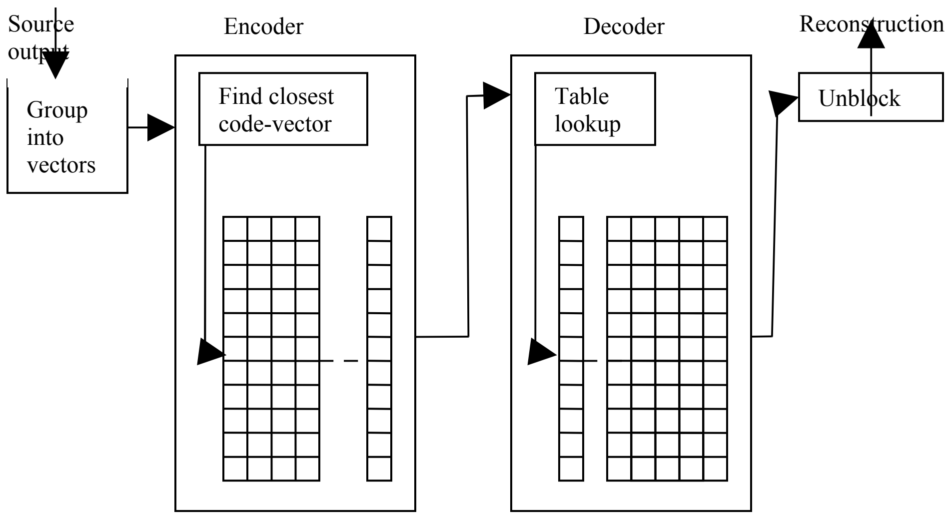
MPEG 4

-Provides a more abstract approach to the coding of multimedia.The standard views the multimedia scene as a collection of objects.These objects can be coded independently.
-Language called the binary format for scenes based on the virtual reality modeling language has been developed by Mpeg.
-The protocol for managing the elementary streams and their multiplexed version called the delivery multimedia integration framework is a part of Mpeg4
-The different objects that makeup the scene are coded and sent to the multiplexer
-The information about the presence of these objects is also provided to the motion compensator predictor
-It is also used in facial animation controlled by facial definition parameter
-It allows for object scalability.



<u>MPEG7</u>:
-Focus on the development of a multimedia content description interface seems to be somewhat removed from the study of data compression
-These activities relate to the core principles of data compression which is the development of compact descriptions of information

4. Discuss about Vector quantization procedure in detail



In vector quantization we group the source output into blocks or vectors. This vector of source outputs forms the input to the vector quantizer. At both the encoder and decoder of the vector quantizer, we have a set of L-dimensional vectors called the codebook of the vector quantizer. The vectors in this codebook are known as code-vectors. Each code vector is assigned a binary index.

At the encoder, the input vector is compared to each code-vector in order to find the code vector closest to the input vector

In order to inform the decoder about which code vector was found to be the closest to the input vector, we transmit or store the binary index of the code-vector. Because the decoder has exactly the same codebook, it can retrieve the code vector

Although the encoder have to perform considerable amount of computations in order to find the closest reproduction vector to the vector of source outputs, the decoding consists of a table lookup. This makes vector quantization a very attractive encoding scheme for applications in which the resources available for decoding are considerably less than the resources available for encoding

Advantages of vector quantization over scalar quantization
For a given rate (bits per sample), use of vector quantization results in lower distortion than when scalar quantization is used at the same rate

Vectors of source output values tend to fall in clusters. By selecting the quantizer output points to lie in these clusters, we have more accurate representation of the source output

Use:

One application for which vector quantizer has been extremely popular is image compression.

Disadvantage of vector quantization:

Vector quantization applications operate at low rates. For applications such as high-quality video coding, which requires higher rates this is definitely a problem.

To solve these problems, there are several approaches which entails some structure in the quantization process

Tree structures vector quantizers:

This structure organizes codebook in such a way that it is easy to pick which part contains the desired output vector

Structured vector quantizers:

Tree-structured vector quantizer solves the complexity problem, but acerbates the storage problem

We now take entirely different tacks and develop vector quantize that do not have these storage problems; however we pay for this relief in other ways

5. Explain Arithmetic coding with an example

Algorithm Implementation

Sequence being encoded as:

$l^n=l^{n-1}+(u^{n-1}-l^{n-1})f_x(x_n-1)$

$u^n= l^{n-1}+(u^{n-1}-l^{n-1})f_x(x_n)$

n becomes larger values gets closer and closer together. The intervals becomes narrower, there are 3 possibilities

1. the interval is entirely confined to the lower half of the unit interval [0,0.5)
2. the interval is entirely confined to the upper half of the unit interval [0.5,1)
3. the interval is in the midpoint of the unit interval

We want to have the sub interval (tag) in the full [0,1) interval

E_1:[0,0.5) $E_1(x)=2x$

E_1:[0.5,1) $E_1(x)=2(x-0.5)$

This process of generating the bits if the tag without waiting to see the entire sequence is called incremental encoding

Tag generation with scaling

Eg: A={a1,a2,a3} P(a1)=0.8 P(a2)=0.02 P(a3)=0.18 Fx(1)=0.8 Fx(2)=0.82 Fx(3)=1

Encode the sequence 1321

Solution:

- first element is 1

Initialize $u^0=1$ $l^0=0$

$l^{1=}0+(1-0)0=0$

$u^1=0+(1-0)0.8=0.8$

The interval [0,0.8) is either in the upper or the lower half of unit interval so proceed

- Second element 3

$l^2=0+(0.8-0)0.82=0.656$
$u^2=0+(0.8-0)0.1=0.8$

interval [0.656,0.8) is in the upper limit. Send the binary code 1 and scale

$l^2=2(0.656-0.5)=0.312$
$u^2=2(0.8-0.5)=0.6$

- Third element 2

$l^3=0.312+(0.6-0.312)0.8=0.5424$
$u^3=0.312+(0.6-0.312)0.82=0.54816$

interval [0.5424,0.54816) is in the upper limit. Send the binary code 1 and scale

$l^3=2(0.5424-0.5)=0.0848$
$u^3=2(0.54816-0.5)=0.09632$

interval [0.0848,0.09632) is in the lower limit. Send the binary code 0 and scale

$l^3=2*0.0848=0.1696$
$u^3=2*0.09632=0.19264$

interval [0.1696,0.19264) is in the lower limit. Send the binary code 0 and scale

$l^3=2*0.1696=0.3392$
$u^3=2*0.19264=0.38528$

interval [0.3392,0.38528) is in the lower limit. Send the binary code 0 and scale

$l^3=2*0.3392=0.6784$
$u^3=2*0.38528=0.77056$

interval [0.6784,0.77056) is in the upper limit. Send the binary code 1 and scale

$l^3=2(0.6784-0.5)=0.3568$
$u^3=2(0.77056-0.5)=0.54112$

The interval [0.3598,0.54112) is either in the upper or the lower half of unit interval so proceed

- Fourth element 1

$l^4=0.3568+(0.54112-0.3568)0=0.3568$
$u^4=0.3568+(0.54112-0.3568)0.8=0.504256$

Stop the encoding.
Binary sequence generated is 110001. Transmit 1 followed by many 0 required by the word length

6. Explain about JPEG compression scheme in detail.

JPEG is a transform coding approach using DCT. Consider 8*8 block of the image as shown in table

Table : an 8*8 block of an image

124	125	122	120	122	119	117	118
121	121	120	119	119	120	120	118
125	124	123	122	121	121	120	120
124	124	125	125	126	125	124	124
127	127	128	129	130	128	127	125
143	142	143	142	140	139	139	139
150	148	152	152	152	152	150	151
156	159	158	155	158	158	157	156

The Transform

The transform used in the Jpeg scheme is the DCT .The input image is first "level shifted by 2p-1 ie) subtract 2p-1 from each pixel value. Then the image is divided into blockes of size 8*8 , which are transformed using an 8*8 forward DCT .The table show the DCT coefficient.

Table: The DCT coefficient

39.88	6.56	-2.24	1.22	-0.37	-1.08	0.79	1.13
-102.43	4.56	2.26	1.12	0.35	-0.63	-1.05	-0.48
37.77	1.31	1.77	0.25	-1.50	-2.21	-0.10	0.23
-5.67	2.24	-1.32	-0.81	1.41	0.22	-0.13	0.17
-3.37	-0.74	-1.75	0.77	-0.62	-2.65	-1.30	0.76
5.98	-0.13	-0.45	-0.77	1.99	-0.26	1.46	0.00
3.97	5.52	2.39	-0.55	-0.051	-0.84	-0.52	-0.13
-3.43	0.51	-1.07	0.87	0.96	0.09	0.33	0.01

Quantization

The JPEG algorithm uses uniform midthread quantization to quantize the various coefficient. The quantizer step sizes are organized in a table called the quantization table as shown in table

Table: Sample Quantization table

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	12	100	103	99

The lable corresponding to the quantized value of the transform coefficient θij is obtained as

$$L_{ij}=\theta_{ij}/Q_{ij}+0.5$$

Where Qij is the (i,j)th element of the quantization table. The reconstructed value is obtained by multiplying the lable with corresponding entry in the quantization table

Table: The quantizer lable

2	1	0	0	0	0	0	0
-9	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Coding

In this approach the lable for the DC and AC coefficient are coded differently using Huffman codes. The DC coefficient values partitioned into categories. The categories are then Huffman coded. The AC coefficient is generated in slightly different manner. There are two special codes: End-of-block(EOF) and ZRL

Table: Coding of the differences of the DC labels

1				0			
1			-1		1		
2		-3	-2		2		3
3	-7	…………………….	-4		4	………	7

Table: sample table for obtaining the Huffman code for a given label value and run length

Z/C	Codeword	Z/c	Codeword	…..	Z/C	Codeword
0/0	1010				F/0	11111111001
0/1	00	1/1	1100		F/1	1111111111110101
……………………						

To encode the AC coefficient First using Zigzag scan. We obtain

-9 3 0 0 0 0 0 ……… 0

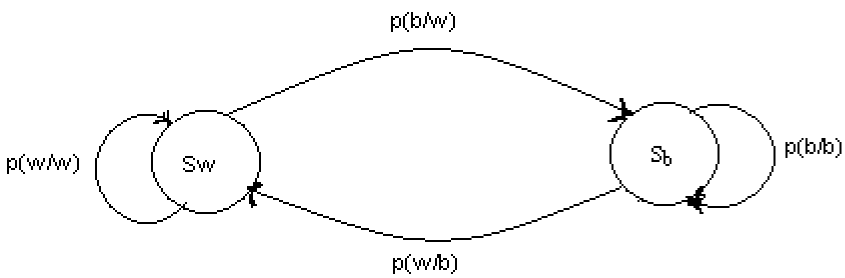
The first value belong to category 1. transmit the code corresponding to 0/1 follow by a single bit 1 to indicate that the value being transmitted is 1 and not -1 .Simillarly other AC coefficient code are transmited.

To obtain the reconstruction of the original block Dequantization is performed and taking inverse transform of the coefficient we get the reconstructed block

7. Describe the concepts of run length coding

- The model that gives rise to run-length coding is the capon model[40], a two-state markov model with state s_w and s_b
- The transition probabilities p(w/b) and p(b/w), and the probability of being in each state $p(s_w)$ and $p(s_b)$, completely specify this model .

- For facsimile images, p(w/w) and p(w/b) are generally significantly higher than p(b/w) and p(b/b)
- The markov model is represented by the state diagram
- The entropy using a probability model and the iid assumption was significantly more than the entropy using the markov model
- Let us try to interpret what the model says about the structure of the data .
- The highly skewed nature of the probabilities p(b/w) and p(w/w),and to a lesser extent p(w/b) and p(b/b), says that once a pixel takes on a particular color, it is highly likely that the following pixels will also be of the same color



- So, rather than code the color of each pixel separately , we can simply code the length of the runs of each color .
- For example, if we had 190 white pixels followed by 30 black pixels , followed by another 210 white pixels , instead of coding the 430 pixels individually, we would code the sequence 190, 30, 210, along with an indication of the color of the first string of pixels .
- Coding the lengths of runs instead of coding individual values is called run-length coding

CCITT GROUP 3 AND 4-RECOMMENDATIONS T.4 AND T.6

- The one dimensional coding scheme is a run-length coding scheme in which each line is represented as a series of alternating white runs and black runs. The first run is always a white run. If the first pixel is a black pixel, then we assume that we have a white run of length zero.
- Runs of different lengths occur with different probabilities, therefore they are coded using a variable length code..
- The number of possible lengths of runs is extremely large and it is not simply feasible to build a codebook that large.
- Therefore instead of generating a Huffman code for each run length r1, the the run length is expressed in the form

  R1=64*m+t for t=0.1….63 and m=1,2…..27
- When we have to represent a run length r1, instead of finding a code for r1,we use the corresponding codes for m and t.
- The codes for t are called the terminating codes and the codes for m are called make up codes.

- Except for the optional codes, there are separate codes for black and white run lengths.
- This coding scheme is generally referred to as a modified Huffman scheme
- In the two dimensional scheme ,instead of reporting the run lengths ,which in terms our Markov model is the length of time we remain in one state ,we report the transition times when we move from one state to another state.