



MOHAMED SATHAK A.J. COLLEGE OF ENGINEERING

(Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai)



EC8093- DIGITAL IMAGE PROCESSING NOTES

EC8093 DIGITAL IMAGE PROCESSING

OBJECTIVES:

1. To become familiar with digital image fundamentals & get exposed to simple image enhancement techniques in Spatial and Frequency domain.
2. To learn concepts of degradation function and restoration techniques & study the image segmentation and representation techniques.
3. To become familiar with image compression and recognition methods

UNIT I - DIGITAL IMAGE FUNDAMENTALS

Steps in Digital Image Processing – Components – Elements of Visual Perception – Image Sensing and Acquisition – Image Sampling and Quantization – Relationships between pixels - Color image fundamentals - RGB, HSI models, Two-dimensional mathematical preliminaries, 2D transforms - DFT, DCT.

UNIT II IMAGE ENHANCEMENT

Spatial Domain: Gray level transformations – Histogram processing – Basics of Spatial Filtering– Smoothing and Sharpening Spatial Filtering, Frequency Domain: Introduction to Fourier Transform– Smoothing and Sharpening frequency domain filters – Ideal, Butterworth and Gaussian filters, Homomorphic filtering, Color image enhancement.

UNIT III IMAGE RESTORATION

Image Restoration - degradation model, Properties, Noise models – Mean Filters – Order Statistics – Adaptive filters – Band reject Filters – Band pass Filters – Notch Filters – Optimum Notch Filtering – Inverse Filtering – Wiener filtering.

UNIT IV IMAGE SEGMENTATION

Edge detection, Edge linking via Hough transform – Thresholding - Region based segmentation – Region growing – Region splitting and merging – Morphological processing- erosion and dilation, Segmentation by morphological watersheds – basic concepts – Dam construction – Watershed segmentation algorithm.

UNIT V IMAGE COMPRESSION AND RECOGNITION

Need for data compression, Huffman, Run Length Encoding, Shift codes, Arithmetic coding, JPEG standard, MPEG. Boundary representation, Boundary description, Fourier Descriptor, Regional Descriptors – Topological feature, Texture - Patterns and Pattern classes - Recognition based on matching.

OUTCOMES: At the end of the course, the students should be able to:

1. Know and understand the basics and fundamentals of digital image processing, such as digitization, sampling, quantization, and 2D-transforms.
2. Operate on images using the techniques of smoothing, sharpening and enhancement & understand the restoration concepts and filtering techniques.
4. Learn the basics of segmentation, features extraction, compression and recognition methods for color models.

TEXT BOOKS:

1. Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Pearson, 3Edition, 2010.
2. Anil K. Jain, Fundamentals of Digital Image Processing, Pearson, 2002.

REFERENCES

1. Kenneth R. Castleman, Digital Image Processing, Pearson, 2006.
2. Rafael C. Gonzalez, et al, Digital Image Processing using MATLAB, Pearson 2011.
3. D.E. Dudgeon et al, Multidimensional Digital Signal Processing, Prentice Hall Professional Technical Reference, 1990.
4. William K. Pratt, Digital Image Processing, John Wiley, New York, 2002
5. Milan Sonka et al, Image processing, analysis and machine vision, Brookes/Cole, Vikas Publishing House, 2nd edition, 1999

LECTURE NOTES

UNIT I - DIGITAL IMAGE FUNDAMENTALS

Steps in Digital Image Processing – Components – Elements of Visual Perception – Image Sensing and Acquisition – Image Sampling and Quantization – Relationships between pixels - Color image fundamentals - RGB, HSI models, Two-dimensional mathematical preliminaries, 2D transforms - DFT, DCT.

PART - A

1. Define Image

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.

4. 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.

5. Define Digital Image.

An image may be defined as a two dimensional function, $f(x,y)$, where x and y are spatial co-ordinates and f is the amplitude. When x,y and the amplitude values of f are all finite, discrete quantities such type of image is called Digital Image.

6. What is Digital Image processing?

The term Digital Image Processing generally refers to processing of a two dimensional picture by a digital computer. i.e., processing of any two dimension data

7. Define Intensity (or) gray level.

An image may be defined as a two-dimensional function, $f(x,y)$. the amplitude of f at any pair off co-ordinates (x,y) is called the intensity (or) gray level off the image at that point.

8. Define Pixel

A digital image is composed of a finite number of elements each of which has a particular location value. The elements are referred to as pixels, picture elements, pels, and image elements. Pixel is the term most widely used to denote the elements of a digital image.

9. 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.

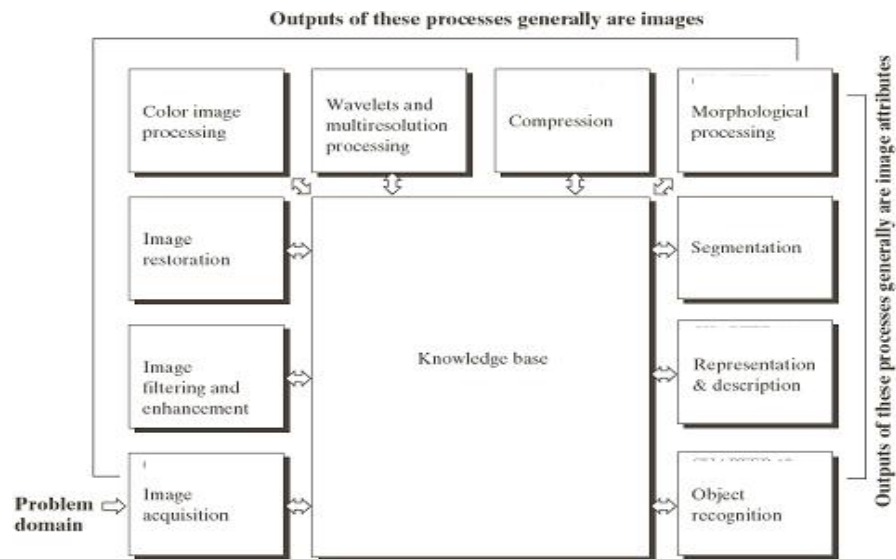
10. List the hardware oriented color models

1. RGB model, 2. CMY model, 3. YIQ model, 4. HSI model

11. 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.

12. Show the block diagram for the steps involved in digital Image processing system?



13. List the applications of color models

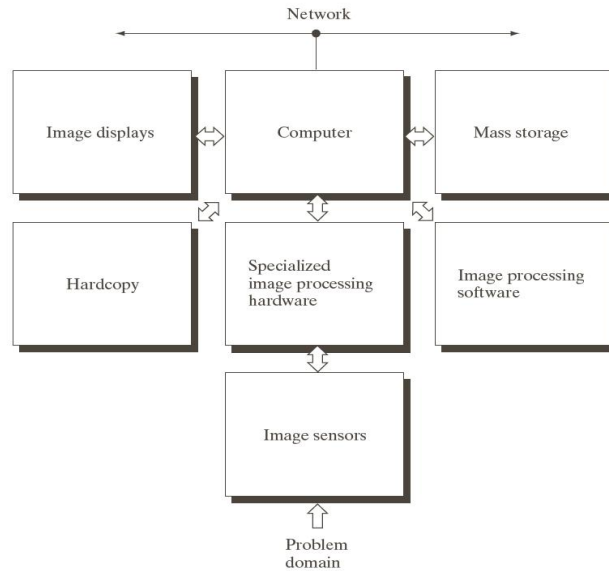
1. RGB model--- used for color monitors & color video camera
2. CMY model---used for color printing
3. HIS model----used for color image processing
4. YIQ model---used for color picture transmission

14. Differentiate Luminance and Chrominance

A. Luminance: received brightness of the light, which is proportional to the total energy in the visible band.

B. Chrominance: describe the perceived color tone of a light, which depends on the wavelength composition of light, chrominance is in turn characterized by two attributes – hue and saturation.

15. Show the block diagram of a general purpose Digital Image Processing system.



16. Write the different combinations of colours in additive color mixing.

Red + Green = Yellow, Red + Blue = Magenta
 Blue + Green = Cyan, Red + Blue + Green = White

17. What are the two ways of mixing of colours?

Mixing of colours can be done by two ways:

(i) Additive mixing, (ii) Subtractive mixing

18. What do you mean by saturation and chrominance?

Saturation: Saturation refers to the spectral purity of the color light.

Chrominance: Hue and Saturation of a color put together is known as chrominance.

19. State Grassman's law.

The brightness impression produced by the three primary colours that constitute the single light. This property of the eye generating a response which depends on the algebraic sum of the Blue, Red and Green. This forms the basis of color signal generation is called "Grassman's law".

20. List out the factors that affect tonal gradation of the reproduced picture.

a) Brightness b) Contrast c) Viewing distance d) Luminance e) Hue and Saturation.

21. Write the basic principle of Television camera tube.

The camera pickup and converts optical information of the scene into electrical energy form may be called as eye of the television system.

22. Define Contrast and Hue.

Contrast: This is the difference in light intensity between black and white parts of the picture over and above the average brightness level.

Hue: This is the predominant spectral color of the received light. Thus the color of any object is distinguished by its hue or tint.

23. Define Image Quantization.

An image may be continuous with respect to the x- and y- co-ordinates and also in amplitude digitizing the amplitude value is called Quantization.

24. Define Image Sampling.

The process of converting continuous spatial co-ordinates into its digitized form is called as sampling.

25. 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 in gray level.

26. What are the steps involved in DIP?

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

27. What is recognition and Interpretation?

Recognition means 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.

28. Specify the elements of DIP system

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

29. What are the types of light receptors?

The two types of light receptors are

1. Cones and 2. Rods

30. 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.	Several rods are connected to one nerve end. So it gives the overall picture of the image.
2. This is also known as bright light vision.	This is also known as Dim light vision.

31. 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 75-150million.

32. 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.

33. 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.

34. What is meant by connectivity? Explain the types of connectivity.

Connectivity between pixels is a fundamental concept that simplifies the definition of numerous digital image concepts, two pixels are said to be connected if they are neighbors and if their gray levels satisfy a condition i.e. the grey levels are equal.

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

35. What are the types of adjacency?

There are three types of adjacency. They are

a. 4-adjacency, b. 8-adjacency, c. m-adjacency

36. Name some application of Digital Image Processing.

1. Remote sensing
2. Image transmission and storage for business applications.
3. Medical processing
4. Radar and Sonar
5. Acoustic Image Processing
6. Robotics
7. Automated inspection of industrial parts.

37. Define Euclidean distance.

The Euclidean distance between p and q is defined as

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

38. Define City – block distance.

The D_4 distance, also called City-block distance between p and q is defined as

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

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

40. 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

41. 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.

42. What is Image Transform? And what is the need for 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 $N \times N$ image can be viewed as an $N^2 \times 1$ vectors. An image transform provides a set of coordinates or basis vectors for vector space. The term image transfer usually refers to a class of unitary matrices used for representing images. 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.

43. Define the term Luminance

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

44. What are the applications of transform?

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

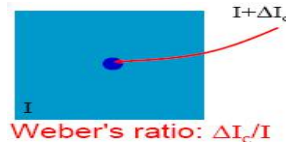
45. What are the properties of unitary transform?

- 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.

46. Define Weber ratio

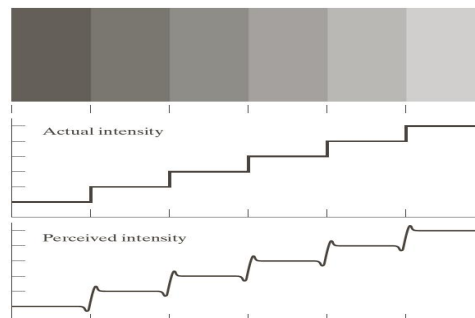
The ratio of increment of illumination to background of illumination is called as Weber ratio.(ie) $\Delta I/I$

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



47. What is meant by mach band effect?

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



48. What is simultaneous contrast?

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



49. 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 $r(x, y)$.

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

$$32 \text{ gray levels} = 2^5 = 5 \text{ bits}, \quad 256 * 256 * 5 = 327680 \text{ bits.}$$

51. 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

$$b = M * N * k,$$

When $M=N$, this equation becomes
 $=N^2k$

52. What do you mean by Zooming and shrinking of digital images?

Zooming may be viewed as over sampling. It involves the creation of new pixel locations and the assignment of gray levels to those new locations.

Shrinking may be viewed as under sampling. To shrink an image by one half, we delete every row and column. To reduce possible aliasing effect, it is a good idea to blur an image slightly before shrinking it.

53. List out the properties of a cosine transforms

- a. It is real and orthogonal, b. is a fast transform
- c. It is not the real part of the unitary DFT, d. it has excellent energy compaction for highly correlated data

54. Define Cosine Transform.

The $N \times N$ cosine transform matrix $C=\{c(k,n)\}$ also called the discrete cosine transform (DCT) is defined as

$$C(k,n) = \begin{cases} \frac{1}{\sqrt{N}}, & k=0, 0 \leq n \leq N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N}, & 1 \leq k \leq N-1, 0 \leq n \leq N-1 \end{cases}$$

55. List out the significance of Karhunen- Loeve Transform.

- 1. Has not fast algorithm
- 2. Useful for small size vectors
- 3. Useful in performance evaluation and for finding performance bounds
- 4. Has the best energy compaction.

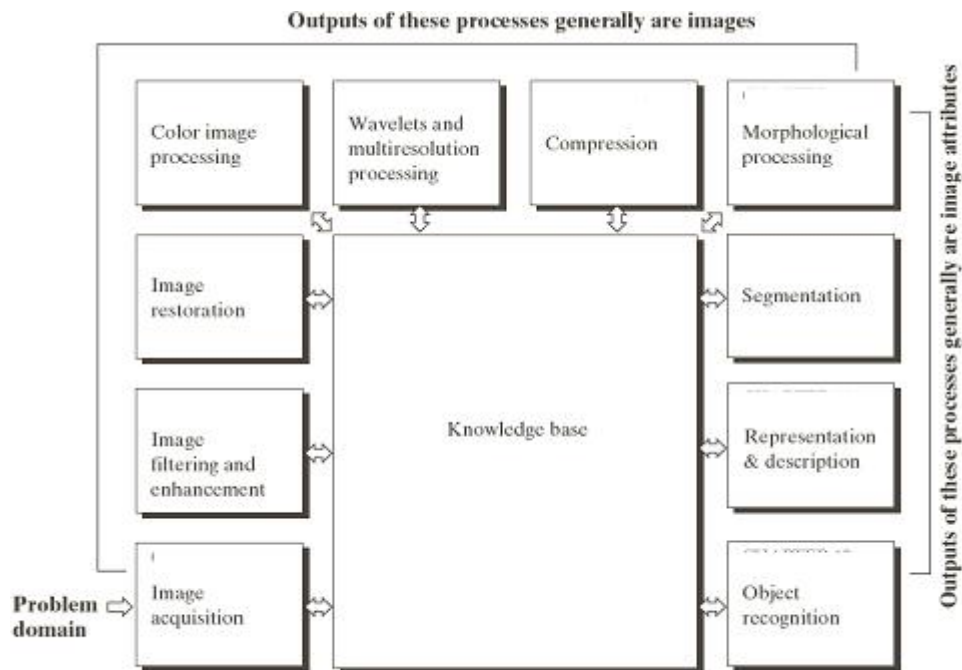
56. Write any two properties of 2D Fourier transform.

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

Power Spectrum $\rightarrow p(u,v) = |F(u,v)|^2$

PART – B

1. Explain the fundamental steps in digital image processing



1. Image acquisition is the first process in the digital image processing. Note that acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves pre-processing, such as scaling.

2. The next step is image enhancement, which is one among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. A familiar example of enhancement is when we increase the contrast of an image because “it looks better.” It is important to keep in mind that enhancement is a very subjective area of image processing.

3. Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation. Enhancement, on the other hand, is based on human subjective preferences regarding what constitutes a “good” enhancement result.i.e remove noise and restores the original image

4. Color image processing is an area that has been gaining in importance because of the significant increase in the use of digital images over the Internet. Color image processing involves the study of fundamental concepts in color models and basic color processing in a digital domain. Image color can be used as the basis for extracting features of interest in an image.

5. Wavelets are the foundation for representing images in various degrees of resolution. In particular, wavelets can be used for image data compression and for pyramidal representation, in which images are subdivided successively into smaller regions.

6. Compression, as the name implies, deals with techniques for reducing the storage required saving an image, or the bandwidth required transmitting it. Although storage technology has improved significantly over the past decade, the same cannot be said for transmission capacity. This is true particularly in uses of the Internet, which are characterized by significant pictorial content. Image compression is familiar (perhaps

inadvertently) to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

7. Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape. The morphological image processing is the beginning of transition from processes that output images to processes that output image attributes.

8. Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually. On the other hand, weak or erratic segmentation algorithms almost always guarantee eventual failure. In general, the more accurate the segmentation, the more likely recognition is to succeed.

9. Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region (i.e., the set of pixels separating one image region from another) or all the points in the region itself.

10. Recognition is the process that assigns a label (e.g., “vehicle”) to an object based on its descriptors. Recognition topic deals with the methods for recognition of individual objects in an image.

2. Explain the Components of an Image Processing System

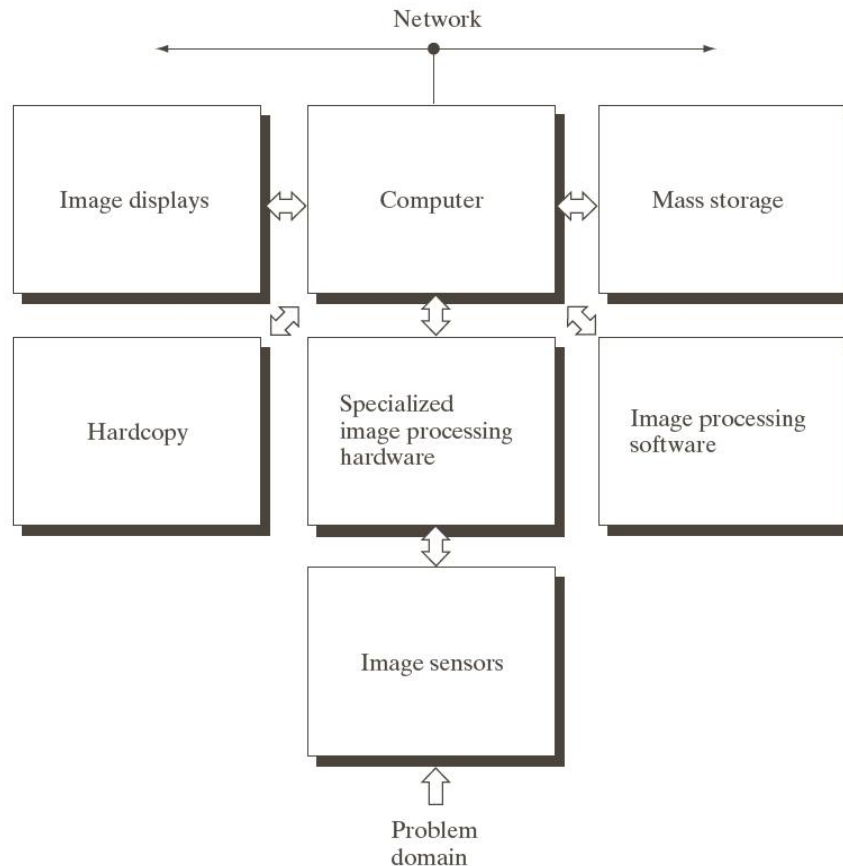
Figure shows the basic components comprising a typical *general-purpose* system used for digital image processing. The function of each component is discussed in the following paragraphs, starting with image sensing.

1. With reference to sensing, two elements are required to acquire digital images. The first is a physical device that is sensitive to the energy radiated by the object we wish to image. The second, called a digitizer, is a device for converting the output of the physical sensing device into digital form. For instance, in a digital video camera, the sensors produce an electrical output proportional to light intensity. The digitizer converts these outputs to digital data.

2. Specialized image processing hardware usually consists of the digitizer just mentioned, plus hardware that performs other primitive operations, such as an arithmetic logic unit (ALU) that performs arithmetic and logical operations in parallel on entire images. ALU is used in averaging images as quickly as they are digitized, for the purpose of noise reduction. This unit performs functions that require fast data throughputs (e.g., digitizing and averaging video images at 30 frames/s)

3. The computer in an image processing system is a general-purpose computer and can range from a PC to a supercomputer. For dedicated applications,

4. Software for image processing consists of specialized modules that perform specific tasks.



5. Mass storage capability is a must in image processing applications. An image of size pixels, in which the intensity of each pixel is an 8-bit quantity, Digital storage for image processing applications falls into three principal categories:

(1) Short-term storage for use during processing, (2) on-line storage for relatively fast recall, and (3) archival storage, characterized by infrequent access. Storage is measured in bytes (eight bits), Kbytes (one thousand bytes), Mbytes (one million bytes), Gbytes (meaning giga, or one billion, bytes), and Tbytes (meaning tera, or one trillion, bytes). On-line storage generally takes the form of magnetic disks or optical-media storage. The key factor characterizing on-line storage is frequent access to the stored data.

6. Image displays in use today are mainly color (preferably flat screen) TV monitors. Monitors are driven by the outputs of image and graphics display cards that are an integral part of the computer system.

7. Hardcopy devices for recording images include laser printers, film cameras, heat-sensitive devices, inkjet units, and digital units, such as optical and CDROM disks. Film provides the highest possible resolution,

8. Networking is almost a default function in any computer system in use today. Because of the large amount of data inherent in image processing applications, the key consideration in image transmission is bandwidth. Communications with remote sites via the Internet are not always as efficient. Fortunately, this situation is improving quickly as a result of optical fiber and other broadband technologies.

Purpose of Image processing:

The purpose of image processing is divided into 5 groups. They are:

1. Visualization - Observe the objects that are not visible.
2. Image sharpening and restoration - To create a better image.

3. Image retrieval - Seek for the image of interest.
4. Measurement of pattern – Measures various objects in an image.
5. Image Recognition – Distinguish the objects in an image.

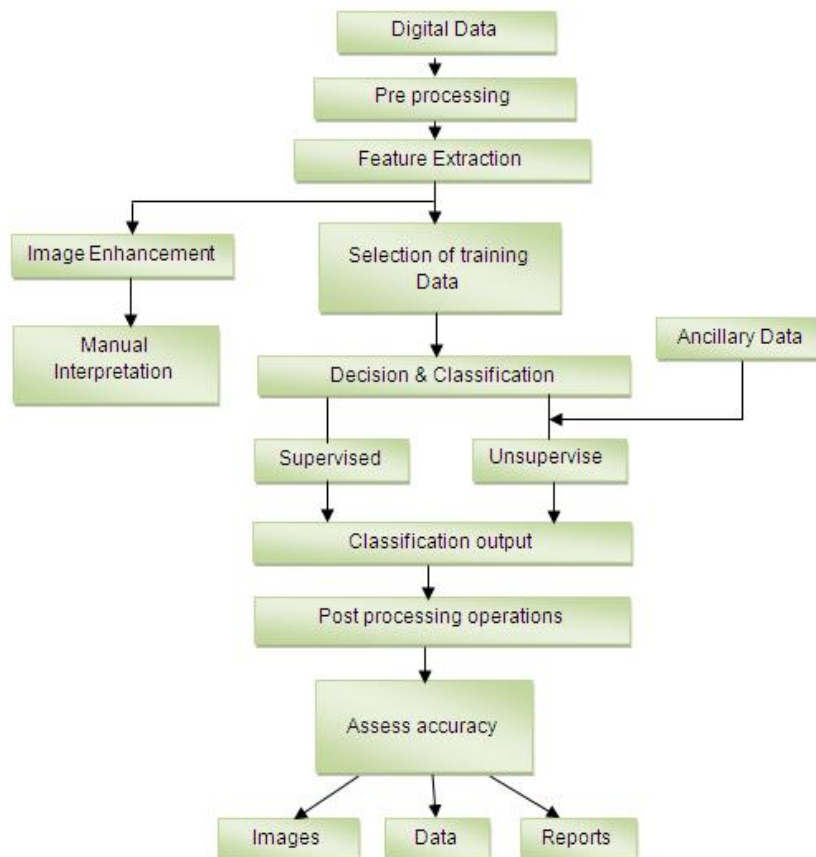
Image Processing Types:

1. Analog Image Processing:

Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

2. Digital Image Processing:

Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre- processing, enhancement and display, information extraction.



Applications of Image Processing:

1. Intelligent Transportation Systems -

This technique can be used in Automatic number plate recognition and Traffic sign recognition.

2. Remote Sensing -

For this application, sensors capture the pictures of the earth's surface in remote sensing satellites or multi – spectral scanner which is mounted on an aircraft. These pictures are processed by transmitting it to the Earth station. Techniques used to interpret the objects and regions are used in flood control, city planning, resource mobilization, agricultural production monitoring, etc.

3. Moving object tracking -

This application enables to measure motion parameters and acquire visual record of the moving object. The different types of approach to track an object are:

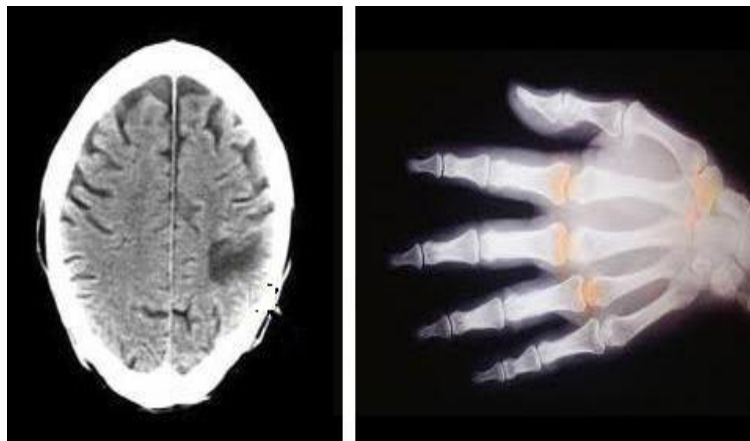
- Motion based tracking
- Recognition based tracking

4. Defense surveillance –

Aerial surveillance methods are used to continuously keep an eye on the land and oceans. This application is also used to locate the types and formation of naval vessels of the ocean surface. The important duty is to divide the various objects present in the water body part of the image. The different parameters such as length, breadth, area, perimeter, compactness are set up to classify each of divided objects. It is important to recognize the distribution of these objects in different directions that are east, west, north, south, northeast, northwest, southeast and south west to explain all possible formations of the vessels. We can interpret the entire oceanic scenario from the spatial distribution of these objects.

5. Biomedical Imaging techniques -

For medical diagnosis, different types of imaging tools such as X- ray, Ultrasound, computer aided tomography (CT) etc are used. The diagrams of X- ray, MRI, and computer aided tomography (CT) are given below.



Some of the applications of Biomedical imaging applications are as follows:

- Heart disease identification– The important diagnostic features such as size of the heart and its shape are required to know in order to classify the heart diseases. To improve the diagnosis of heart diseases, image analysis techniques are employed to radiographic images.
- Lung disease identification – In X- rays, the regions that appear dark contain air while region that appears lighter are solid tissues. Bones are more radio opaque than tissues. The ribs, the heart, thoracic spine, and the diaphragm that separates the chest cavity from the abdominal cavity are clearly seen on the X-ray film.

· Digital mammograms – This is used to detect the breast tumour. Mammograms can be analyzed using Image processing techniques such as segmentation, shape analysis, contrast enhancement, feature extraction, etc.

6. Automatic Visual Inspection System –

This application improves the quality and productivity of the product in the industries.

· Automatic inspection of incandescent lamp filaments – This involves examination of the bulb manufacturing process. Due to no uniformity in the pitch of the wiring in the lamp, the filament of the bulb gets fused within a short duration. In this application, a binary image slice of the filament is created from which the silhouette of the filament is fabricated. Silhouettes are analyzed to recognize the non uniformity in the pitch of the wiring in the lamp. This system is being used by the General Electric Corporation.

· Automatic surface inspection systems – In metal industries it is essential to detect the flaws on the surfaces. For instance, it is essential to detect any kind of aberration on the rolled metal surface in the hot or cold rolling mills in a steel plant. Image processing techniques such as texture identification, edge detection, fractal analysis etc are used for the detection.

· Faulty component identification – This application identifies the faulty components in electronic or electromechanical systems. Higher amount of thermal energy is generated by these faulty components. The Infra-red images are produced from the distribution of thermal energies in the assembly. The faulty components can be identified by analyzing the Infra-red images.

Image Modalities:

Medical image Modalities are

1. X-ray
2. MRI
3. CT
4. Ultrasound.

1. X-Ray:

“X” stands for “unknown”. X-ray imaging is also known as

- radiograph
- Röntgen imaging.

- Calcium in bones absorbs X-rays the most
- Fat and other soft tissues absorb less, and look gray
- Air absorbs the least, so lungs look black on a radiograph

2. CT(Computer Tomography):

Computed tomography (CT) is an integral component of the general radiography department. Unlike conventional radiography, in CT the patient lies on a couch that moves through into the imaging gantry housing the x-ray tube and an array of specially designed "detectors". Depending upon the system the gantry rotates for either one revolution around the patient or continuously in order for the detector array to record the intensity of the remnant x-ray beam. These recordings are then computer processed to produce images never before thought possible.

MRI(Magnetic Resonance Imaging):

This is an advanced and specialized field of radiography and medical imaging. The equipment used is very precise, sensitive and at the forefront of clinical technology. MRI is

not only used in the clinical setting, it is increasingly playing a role in many research investigations. Magnetic Resonance Imaging (MRI) utilizes the principle of Nuclear Magnetic Resonance (NMR) as a foundation to produce highly detailed images of the human body. The patient is firstly placed within a magnetic field created by a powerful magnet.

Image File Formats:

Image file formats are standardized means of organizing and storing digital images. Image files are composed of digital data in one of these formats that can be for use on a computer display or printer. An image file format may store data in uncompressed, compressed, or vector formats.

Image file sizes:

In raster images, **Image file size** is positively correlated to the number of pixels in an image and the color depth, or bits per pixel, of the image. Images can be compressed in various ways, however.

Major graphic file formats

The two main families of graphics Raster and Vector.

Raster formats

1. JPEG/JFIF

JPEG (Joint Photographic Experts Group) is a compression method; JPEG-compressed images are usually stored in the JFIF (JPEG File Interchange Format) file format. JPEG compression is (in most cases) lossy compression. The JPEG/JFIF filename extension is **JPG** or **JPEG**.

2. JPEG 2000

JPEG 2000 is a compression standard enabling both lossless and lossy storage. The compression methods used are different from the ones in standard JFIF/JPEG; they improve quality and compression ratios, but also require more computational power to process. JPEG 2000 also adds features that are missing in JPEG.

3. TIFF

The **TIFF** (Tagged Image File Format) format is a flexible format that normally saves 8 bits or 16 bits per color (red, green, blue) for 24-bit and 48-bit totals, respectively, usually using either the **TIFF** or **TIF** filename extension. TIFFs can be lossy and lossless;

4. GIF

GIF (Graphics Interchange Format) is limited to an 8-bit palette, or 256 colors. This makes the GIF format suitable for storing graphics with relatively few colors such as simple diagrams, shapes, logos and cartoon style images.

5. BMP

The **BMP file format** (Windows bitmap) handles graphics files within the Microsoft Windows OS. Typically, BMP files are uncompressed, hence they are large; the advantage is their simplicity and wide acceptance in Windows programs.

6. PNG

The **PNG** (Portable Network Graphics) file format was created as the free, open-source successor to GIF. The PNG file format supports 8 bit paletted images (with optional transparency for all palette colors) and 24 bit truecolor (16 million colors).

3. Explain the principle and working of Vidicon digital camera with neat diagram (Image sensing and Acquisition)

A digital camera is used to convert the optical information into a corresponding electrical signal, the amplitude which varies in accordance with the variations of brightness.

Photoelectric Effects

The two photoelectric effects used for converting variations of light intensity into electrical variations are (i) photoemission and (ii) photoconductivity.

Photo emission: Certain metals emit electrons when light falls on their surface. These emitted electrons are called photoelectrons and the emitting surface a photocathode. Light consists of small bundles of energy called photons. When light is made incident on a photocathode, the photons give away their energy to the outer valence electrons to allow them to overcome the potential-energy barrier at the surface. The number of electrons which can overcome the potential barrier and get emitted depends on the light intensity. Alkali metals are used as photocathode because they have very low work-function. Cesium-silver or bismuth-silver-cesium oxides are preferred as photo emissive surfaces because they are sensitive to incandescent light and have spectral response very close to the human eye.

Photo conductivity:-The second method of producing an electrical image is by photoconduction, where the conductivity or resistivity of the photosensitive surface varies in proportion to the intensity of light focused on it. In general the semiconductor metals including selenium, tellurium and lead with their oxides have this property known as photoconductivity. The variations in resistance at each point across the surface of the material are utilized to develop a varying signal by scanning it uniformly with an electron beam

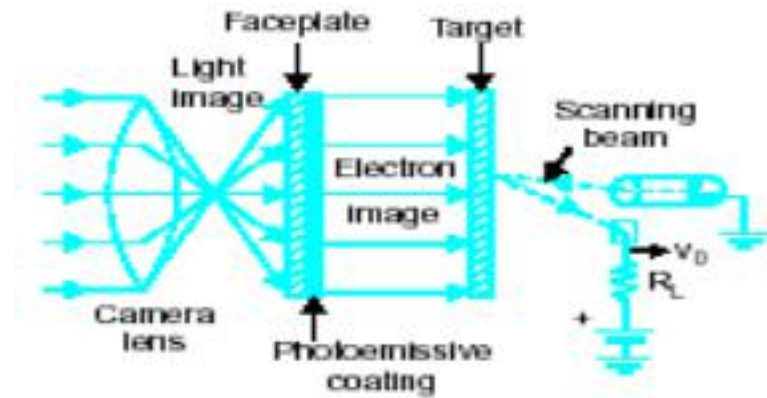


Fig-Photo emission

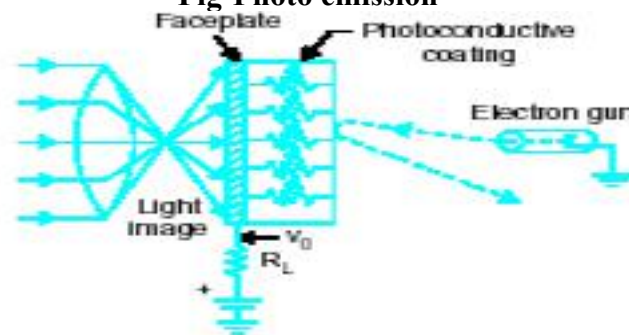


Fig-Photo conductivity

Types of camera tubes: 1) orthicon 2) vidicon 3) Plumbicon.

Vidicon digital camera:-

The principle of operation is photoconductivity. The resistance of the target decreases when exposed to light. The target consists of a thin photo conductive layer of either selenium or antimony compounds. This is deposited on a transparent conducting film coated on the inner surface of the face plate. This coating is called plate. Image side of photo layer is connected with the signal electrode and is connected to DC supply through a load resistance R. The beam that emerges from the electron gun is focused on the surface of the photo conductive layer by the combined action of magnetic field of an external coil and electrostatic field of grid 3. Grid 4 provides decelerating action so that the electrons hit the photo conductive layer with very less speed to prevent secondary emission. Deflection of the beam is done by the deflection coils placed around the tube.

Three sections of **image orthicon** are

1. Image section 2. Scanning section 3. electron gun multiplier.

1. Image section:- Inside of glass face plate at the front is coated with a silver antimony coating added with cesium, which acts as photocathode. Light from the scene being televised are focused on to photocathode surface and electrons are released from each point on photocathode in proportion to incident light intensity coming from the image. The distribution of electron generation depends on the brightness levels at different points of the picture. The electron image of the scene gets formed on the target side of the photo coating and extends towards the target. Target plate is very thin sheet of glass and can store the received charges. Target plate is 400V more positive wrt photocathode, and this causes acceleration of emitted electrons towards it.

The electron in motion has a tendency to repel each other, which may cause distortion in picture. To avoid this magnetic field is applied using a long focus coil. The magnetic coil focuses the emitted electrons on the target to get well defined electron image. The image side of the target has a small deposit of cesium and has high secondary emission. When electron image with high velocity hits the target secondary emission results. These secondary electrons are collected by a wire mesh screen, which is in front of the target on the image side and is maintained at a higher potential wrt target. Secondary emission leaves behind a positive charge distribution on the target plate.

The target plate is made up of thin glass, this retards the positive charge degradation and hence positive charge on the target plate build up. Due to high secondary emission the intensity of positive charge distribution is four to five times more, compared to the charge liberated by the photocathode. Increase in charge density at target compared to charge liberated at photocathode is known as image multiplication and increases the sensitivity of image orthicon. Target plate should have high resistivity laterally for storage action and low resistivity along its thickness for positive charge to conduct to other side which is scanned. This is the reason why the target plate is thin. Whatever charge distribution builds up on one side of target plate appears on the other side which is scanned.

2. Scanning section: Electron gun produces the beam focused using magnetic fields applied to grid 2, 3, 4. Alignment coil provides magnetic field that can be varied to adjust the position of the scanning beam to correct location. Deflection is by scanning fields of horizontal & vertical deflection coils mounted at yoke. These coils are connected to deflection sweep circuits. Grid 4 voltage decelerates the electron beam this prevents secondary emission from target. The electron beam strikes the target plate and the electrons from beam gets divided into two sets. 1) certain electrons neutralize the positive charged areas of the target. 2) certain electrons turn back and is called as signal current. Resultant beam that turns back are maximum for black areas and minimum for bright areas. The

electrons hitting the screen may cause stray electron field & may grid tangentially which may cause defocusing and loss resolution. The beam should strike at right angles with screen.

3. Electron gun multiplier. When the surface of a metal is bombarded by incident electrons having high velocities secondary emissions takes place. Camera tubes make use of secondary emissions to amplify small amount of photoelectric current that is later employed to develop video signal. The electron multiplier is a collection of anode cathode electrodes called dynodes. The electrons are accelerated by the positive potential of the dynode. The number of electrons available is multiplied each time the secondary electron strikes the emitting surface of the next positive dynode. The current amplification is noise free as it does not have any active or passive components.

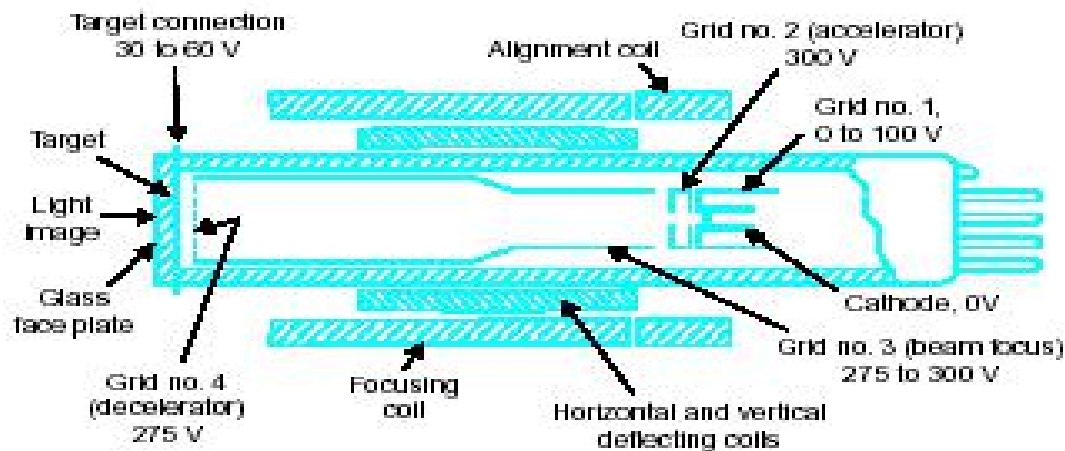


Fig-Digital camera(VIDICON)

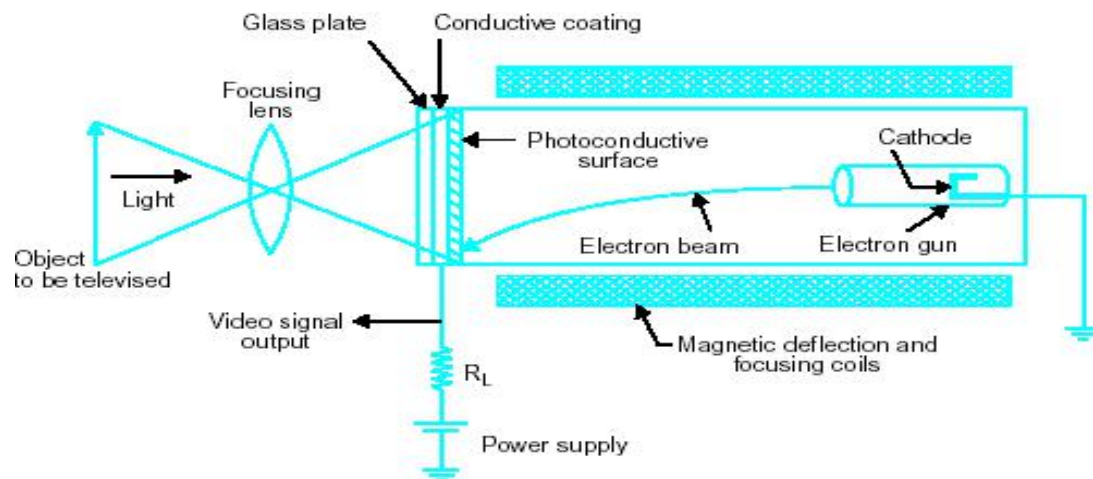


Fig-Digital camera (VIDICON)

Light from the scene is focused on a photosensitive surface known as the image plate, and the optical image thus formed with a lens system represents light intensity variations of the scene. The photoelectric properties of the image plate then convert different light intensities into corresponding electrical variations. In addition to this photoelectric conversion whereby the optical information is transduced to electrical charge distribution on the photosensitive image plate, it is necessary to pick-up this information as fast as possible

4. Explain in Detail the structure of the Human eye and also explain the image formation in the eye. (OR) Explain the Elements of Visual Perception.

Visual Perception: Means how an image is perceived by a human observer

The Shape of the human eye is nearly a sphere with an Average diameter = 20mm. It has 3 membranes:

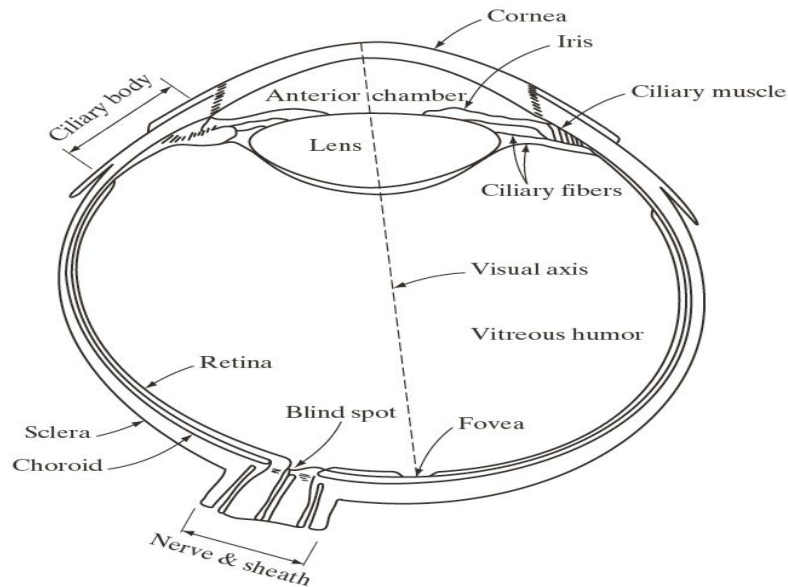
1. Cornea and Sclera – outer cover
2. Choroid
3. Retina -enclose the eye

Cornea: tough, transparent tissue covers the anterior surface of the eye.

Sclera: Opaque membrane, encloses the remainder of the optic globe

Choroid: Lies below the sclera .It contains a lot of blood vessels providing nutrition to the eye. The Choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe. At the anterior part, the choroid is divided into Ciliary body and the Iris diaphragm.

Structure of the Human Eye:



The Iris contracts or expands according to the amount of light that enters into the eye. The central opening of the iris is called pupil, which varies from 2 mm to 8 mm in diameter.

Lens is made of concentric layers of fibrous cells and is suspended by fibers that attach to the Ciliary body. It contains 60-70% of water, 6 % fat and more protein. Lens is colored by yellow pigment which increases with eye. Excessive clouding of lens can lead to poor or loss of vision, which is referred as **Cataracts**.

Retina: Innermost membrane of the eye which lines inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina. It consists of two photo receptors.

Receptors are divided into 2 classes:

1. Cones
2. Rods

1. Cones: These are 6-7 million in number, and are shorter and thicker. These are located primarily in the central portion of the retina called the **fovea**. They are highly sensitive to color.

Each is connected to its own nerve end and thus the human can resolve fine details.

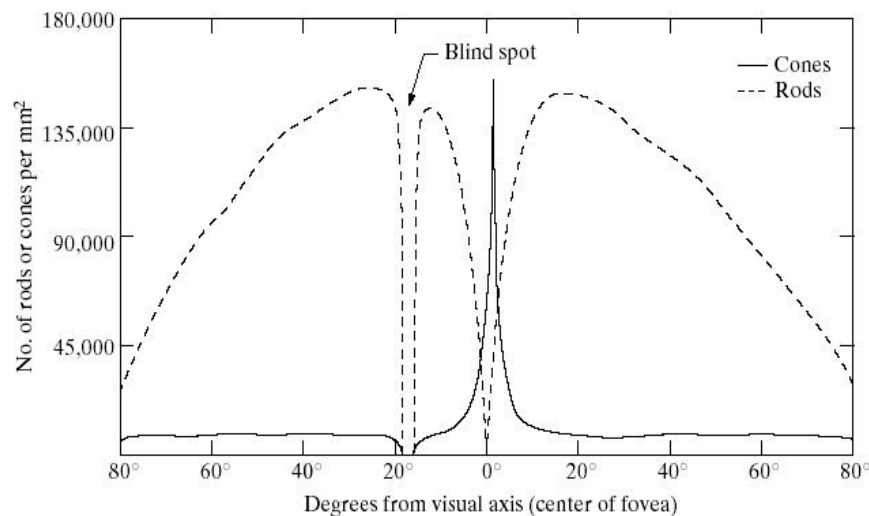
Cone vision is called **photopic or bright-light vision**.

2. Rods: These are 75-150 million in number and are relatively long and thin.

These are distributed over the retina surface. Several rods are connected to a single nerve end reduce the amount of detail discernible. Serve to give a general, overall picture of the field of view.

These are Sensitive to low levels of illumination. Rod vision is called **scotopic or dim-light vision**.

The following shows the distribution of rods and cones in the retina



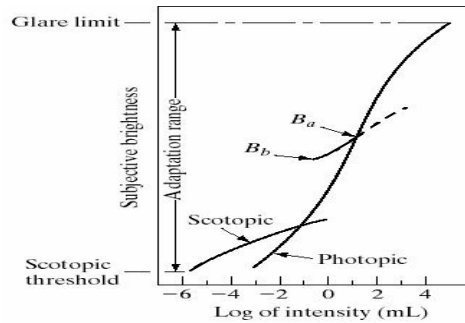
Blind spot: the absence of receptors in the retina is called as Blind spot. From the diagram,

1. Receptor density is measured in degrees from the fovea.
2. Cones are most dense in the center of the retina (in the area of the fovea).
3. Rods increase in density from the center out to approx. 20° off axis and then decrease in density out to the extreme periphery of the retina

5. Illustrate the concept of Brightness adaptation.(or)write short notes on 1. Mach band effect 2. Weber's Ratio 3. Simultaneous Contrast (May 2007)

Subjective Brightness: is defined as the intensity as preserved by the human visual system (HVS).

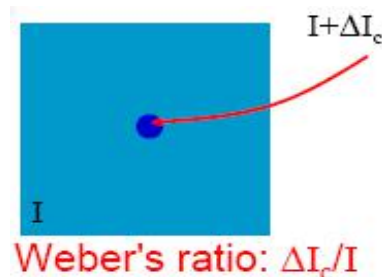
Brightness Adaptation means the HVS can operate only from scotopic to glare limit. It cannot operate over the range simultaneously.



From the diagram it is seen that the brightness is a logarithmic function of light intensity. B_a is a brightness adaptation level. Below B_a , the shades of colors cannot be discriminated.

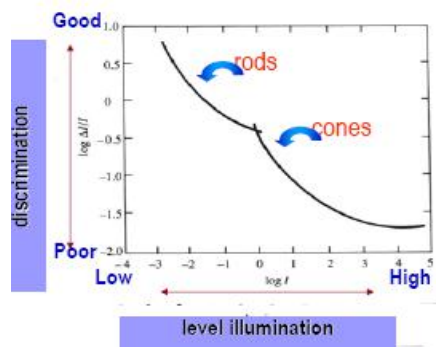
Contrast sensitivity: The ability of the eye to discrimination b/w changes in brightness at any specific adaptation level is of considerable interest

Weber's Ratio: The ratio of increment of illumination to background of illumination is called as Weber ratio.(ie) $\Delta I_c/I$



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

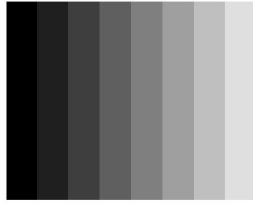
If the ratio ($\Delta I_c/I$) is large, then large percentage of change in intensity is needed (ie) poor brightness adaptation.



At low levels of illumination, when the Weber ratio is large, brightness discrimination is poor. Here vision is carried out by the Rods.

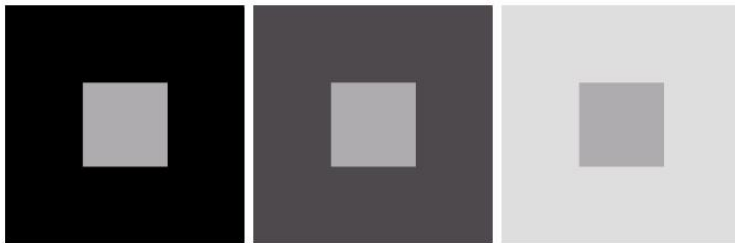
At high levels of illumination, when the Weber ratio is small, brightness discrimination is improved and better. Here vision is carried out by the Cones.

Mach Band: Mach bands are an optical illusion consisting of an image of two wide bands, one light and one dark, separated by a narrow strip with a light – to – dark gradient.



Thus mach band effect means the intensity of the stripes is constant. Therefore it preserves the brightness pattern near the boundaries and these bands are called as mach band.

Simultaneous Contrast: Two colors side by side interact with one another and change our perception accordingly. The effect of this interaction is called as Simultaneous Contrast



Here all the inner squares have the same intensity, but they appear progressively brighter as the background becomes lighter

5. Write short notes on (i) luminance,(ii) hue or tint, and (iii) saturation,iv) Brightness v) Contrast

Any color has three characteristics to specify its visual information. These are (i) luminance,(ii) hue or tint, and (iii) saturation. These are defined as follows:

(i) Luminance

This is the amount of light intensity as perceived by the eye regardless of the color. In black and white pictures, better lighted parts have more luminance than the dark areas. Different colours also have shades of luminance in the sense that though equally illuminated appear more or less bright. Thus on a monochrome TV screen, dark red colour will appear as black, yellow as white and a light blue colour as grey.

(ii) Hue

This is the predominant spectral colour of the received light. Thus the colour of any object is distinguished by its hue or tint. The green leaves have green hue and red tomatoes have red hue. Different hues result from different wavelengths of spectral radiation and are perceived as such by the sets of cones in the retina.

(iii) Saturation

This is the spectral purity of the colour light. Since single hue colours occur rarely alone, this indicates the amounts of other colours present. Thus saturation may be taken as an indication of how little the colour is diluted by white. A fully saturated colour has no white. vivid green is fully saturated and when diluted by white it becomes light green. The hue and saturation of a colour put together is known as chrominance. Note that it does not contain the brightness information. Chrominance is also called chroma.

(iv) Brightness:

Brightness is the overall or average intensity of illumination and it determines background light level in the reproduced picture. The brightness control can be varied to get average illumination of the scene.

(ii) Contrast:

Contrast is the difference in light intensity between black and white parts of the picture and above the average brightness level. Too much contrast makes the picture hard white little contrast gives the impression of a washed out picture.

6. Explain the Color image fundamentals (models) with neat diagram

According to the theory of the human eye, all colors are seen as variable combinations of the three so-called primary colors red (R), green (G), and blue (B). The following specific wavelength values to the primary colors:

Blue (B) = 435.8 nm

Green (G) = 546.1 nm

Red (R) = 700.0 nm

The primary colors can be added to produce the secondary colors

Magenta (red + blue), **cyan** (green + blue), and **yellow** (red + green), see Figure

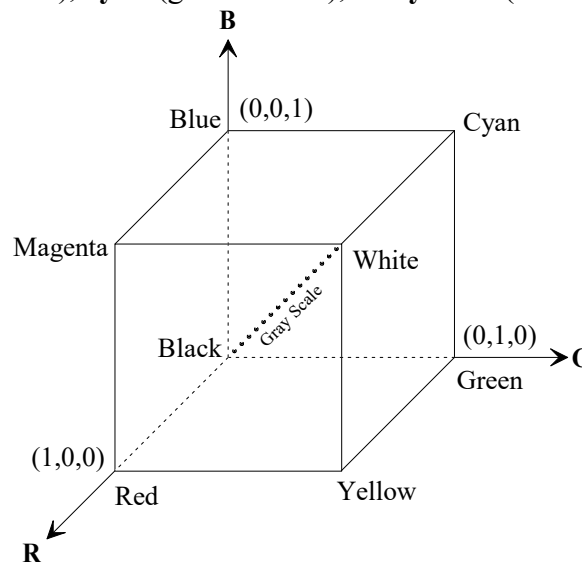
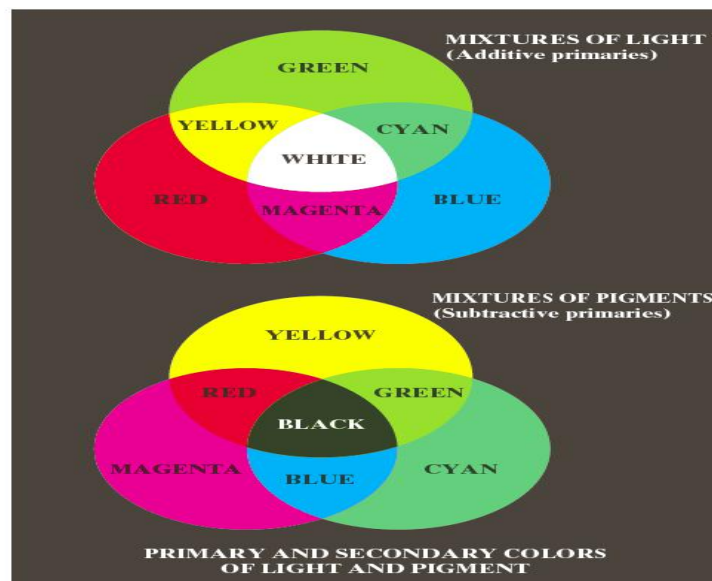


Figure -RGB color cube.



Mixing all the three primary colors results in white. Color television reception is based on this three color system with the additive nature of light.

There are several useful color models: RGB, CMY, YUV, YIQ, and HSI.

1. RGB color model

The colors of the RGB model can be described as a triple (R, G, B) , so that R, G, B . The RGB color space can be considered as a three-dimensional unit cube, in which each axis represents one of the primary colors, see Figure. Colors are points inside the cube defined by its coordinates. The primary colors thus are red= $(1,0,0)$, green= $(0,1,0)$, and blue= $(0,0,1)$. The secondary colors of RGB are cyan= $(0,1,1)$, magenta= $(1,0,1)$ and yellow= $(1,1,0)$.

The nature of the RGB color system is additive in the sense how adding colors makes the image brighter. **Black is at the origin, and white is at the corner** where $R=G=B=1$. The gray scale extends from black to white along the line joining these two points. Thus a shade of gray can be described by (x,x,x) starting from black= $(0,0,0)$ to white= $(1,1,1)$.

The colors are often normalized as given in equation . This normalization guarantees that $r+g+b=1$.

$$r = \frac{R}{(R + B + G)}$$

$$g = \frac{G}{(R + B + G)}$$

$$b = \frac{B}{(R + B + G)}$$

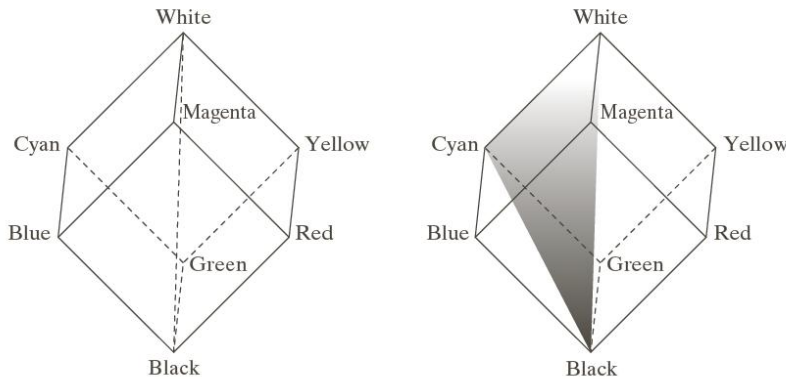
2. CMY color model

The CMY color model is closely related to the RGB model. Its primary colors are C (cyan), M (magenta), and Y (yellow). I.e. the secondary colors of RGB are the primary colors of CMY, and vice versa. The RGB to CMY conversion can be performed by

$$C = 1 - R$$

$$M = 1 - G$$

$$Y = 1 - B$$



The scale of C, M and Y also equals to unit: C, M, Y . The CMY color system is used in offset printing in a subtractive way, contrary to the additive nature of RGB. A pixel of color Cyan, for example, reflects all the RGB other colors but red. A pixel with the color of

magenta, on the other hand, reflects all other RGB colors but green. Now, if we mix cyan and magenta, we get blue, rather than white like in the additive color system.

ADDITIVE COLOUR MIXING:

In order to generate suitable colour signals it is necessary to know definite ratios in which red, green and blue combine form new colours. Since R,G and B can be mixed to create any colour including white, these are called primary colours.

$$R + G + B = \text{White colour}$$

$$R - G - B = \text{Black colour}$$

The primary colours R, G and B combine to form new colours

$$R + G = Y$$

$$R + B = \text{Magenta (purplish)}$$

$$G + B = \text{Cyan (greenish blue)}$$

GRASSMAN'S LAW:

Our eye is not able to distinguish each of the colours that mix to form a new colour but instead perceives only the resultant colour. Based on sensitivity of human eye to various colours, reference white for colour TV has been chosen to be a mixture of colour light in the ratio

$$100\% W = 30\%R + 59\%G + 11\%B$$

Similarly yellow can be produced by mixing 30% of red and 59% of green, magenta by mixing 30% of red and 11% of blue and cyan by mixing 59% of green to 11% of blue. Base on this it is possible to obtain white light by mixing 89% of yellow with 11% of blue or 70% of cyan with 30% of red.

Thus the Eye perceives new colours based on the algebraic sum of red, green and blue light fluxes. This forms the basic of colour signal generation and is known as GRASSMAN'S LAW.

3. YUV color model

The basic idea in the YUV color model is to separate the color information apart from the brightness information. The components of YUV are:

$$Y = 0.3 \cdot R + 0.6 \cdot G + 0.1 \cdot B$$

$$U = B - Y$$

$$V = R - Y$$

Y represents the *luminance* of the image, while U, V consists of the color information, i.e. *chrominance*. The luminance component can be considered as a gray-scale version of the RGB image, The advantages of YUV compared to RGB are:

1. The brightness information is separated from the color information
2. The correlations between the color components are reduced.
3. Most of the information is collected to the Y component, while the information content in the U and V is less. where the contrast of the Y component is much greater than that of the U and V components.

The latter two properties are beneficial in image compression. This is because the correlations between the different color components are reduced, thus each of the components can be compressed separately. Besides that, more bits can be allocated to the Y component than to U and V . The YUV color system is adopted in the JPEG image compression standard.

3. YIQ color model

YIQ is a slightly different version of YUV. It is mainly used in North American television systems. Here Y is the luminance component, just as in YUV. I and Q correspond to U and V of YUV color systems. The RGB to YIQ conversion can be calculated by

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

$$I = 0.596 \cdot R - 0.275 \cdot G - 0.321 \cdot B$$

$$Q = 0.212 \cdot R - 0.523 \cdot G + 0.311 \cdot B$$

The YIQ color model can also be described corresponding to YUV:

$$Y = 0.3 \cdot R + 0.6 \cdot G + 0.1 \cdot B$$

$$I = 0.74 \cdot V - 0.27 \cdot U$$

$$Q = 0.48 \cdot V + 0.41 \cdot U$$

4. HSI color model

The HSI model consists of *hue* (H), *saturation* (S), and *intensity* (I).

Intensity corresponds to the luminance component (Y) of the YUV and YIQ models.

Hue is an attribute associated with the dominant wavelength in a mixture of light waves, i.e. the dominant color as perceived by an observer.

Saturation refers to relative purity of the amount of white light mixed with hue. The advantages of HSI are:

The intensity is separated from the color information (the same holds for the YUV and YIQ models though).

The hue and saturation components are intimately related to the way in which human beings perceive color.

The RGB to HSI conversion can be summarized as follows:

$$H = \frac{1}{360^\circ} \cdot \cos^{-1} \left[\frac{\frac{1}{2} \cdot [(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right], \text{ if } B \leq G$$
$$H = 1 - \frac{1}{360^\circ} \cdot \cos^{-1} \left[\frac{\frac{1}{2} \cdot [(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right], \text{ otherwise}$$
$$S = 1 - \frac{3}{(R + G + B)} \cdot \min(R, G, B)$$
$$I = \frac{1}{3} \cdot (R + G + B)$$

Table -Summary of image types.

Image type	Typical bpp	No. of colors	Common file formats
Binary image	1	2	JBIG, PCX, GIF, TIFF
Gray-scale	8	256	JPEG, GIF, PNG, TIFF
Color image	24	$16.6 \cdot 10^6$	JPEG, PNG, TIFF
Color palette image	8	256	GIF, PNG
Video image	24	$16.6 \cdot 10^6$	MPEG

7. Discuss the role of sampling and quantization in the context of image encoding applications.

Image Sampling and Quantization:

To generate digital images from sensed data. The output of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes; sampling and quantization.

Basic Concepts in Sampling and Quantization:

The basic idea behind sampling and quantization is illustrated in figure, figure (a) shows continuous image, $f(x, y)$, that we want to convert to digital form. An image may be continuous with respect to the x-and y-coordinates, and also in amplitude. To convert it to digital form, we have to sample the function in both coordinates and in amplitude, Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization.

The one dimensional function shown in figure (b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in figure (a). The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in figure (c). The location of each sample is given by a vertical tick mark in the bottom part of the figure. The samples are shown as small white squares superimposed on the function. The set of these discrete locations gives the sampled function. However, the values of the samples still span (vertically) a continuous range of gray-level values. In order to form a digital function, the gray-level values also must be converted (quantized) into discrete quantities. The right side of figure (c) shows the gray-level scale divided into eight discrete levels, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight gray levels. The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown in figure (d). Starting at the top of the image and carrying out this procedure line by line produces a two dimensional digital image.

Sampling in the manner just described assumes that we have a continuous image in both coordinate directions as well as in amplitude. In practice, the method of sampling is determined by the sensor arrangement used to generate the image. When an image is generated by a single sensing element combined with mechanical motion, as in figure, the output of the sensor is quantized in the manner described above. However, sampling is accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data. Mechanical motion can be made very exact so, in principle, there is almost no limit as to how fine we can sample and image. However, practical limits are established by imperfections in the optics used to focus on the sensor an illumination spot that is inconsistent with the fine resolution achievable with mechanical displacements.

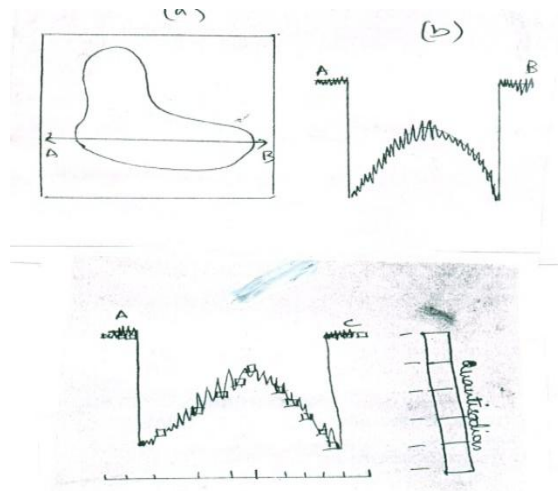


Figure: Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

When a sensing strip is used for image acquisition, the number of sensors in the strip establishes the sampling limitations in one image direction. Mechanical motion in the other direction can be controlled more accurately, but it makes little sense to try to achieve sampling density in one direction that exceeds the sampling limits established by the number of sensors in the other. Quantization of the sensor outputs completes the process of generating a digital image.

Quantization

This involves representing the sampled data by a finite number of levels based on some criteria such as minimization of the quantizer distortion, which must be meaningful.

Quantizer design includes input (decision) levels and output (reconstruction) levels as well as number of levels. The decision can be enhanced by psychovisual or psychoacoustic perception.

Quantizers can be classified as memory less (assumes each sample is quantized independently or with memory (takes into account previous sample))

Alternative classification of quantisers is based on uniform or non- uniform quantization. They are defined as follows.

Uniform quantizers

They are completely defined by (1) the number of levels it has (2) its step size and whether it is midriser or midtreader. We will consider only symmetric quantizers i.e. the input and output levels in the 3rd quadrant are negative of those in 1st quadrant.

Non-uniform quantizers

The Step sizes are not constant. Hence non-uniform quantization is specified by input and output levels in the 1st and 3rd quadrants.

8. Explain 2D Sampling theory

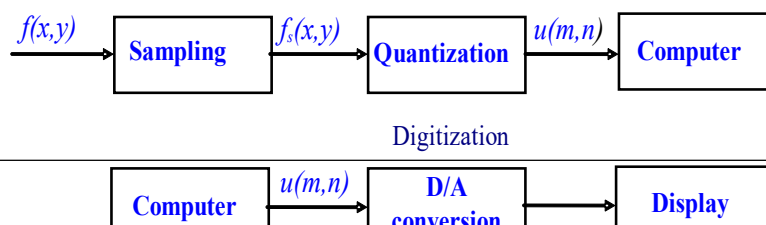
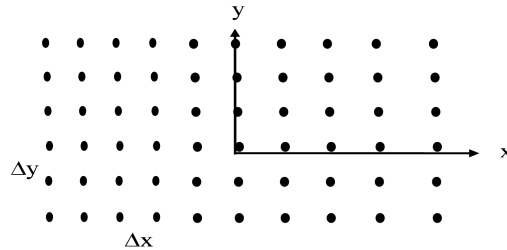


Fig 1 Image sampling and quantization / Analog image display



The common sampling grid is uniformly spaced, rectangular grid. Image sampling = read from the original, spatially continuous, brightness function $f(x,y)$, **only** in the black dots positions (\Leftrightarrow **only** where the grid allows):

$$\Rightarrow f_s(x, y) = \begin{cases} f(x, y), & x = m \Delta x, y = n \Delta y \\ 0, & \text{otherwise} \end{cases},$$

$$m, n \in \mathbf{Z}.$$

The sampling intervals $\Delta x, \Delta y$ needed to have no loss of information depend on the spatial frequency content of the image.

$$f_s(x, y) = f(x, y) \times \text{comb}(x, y) - 1$$

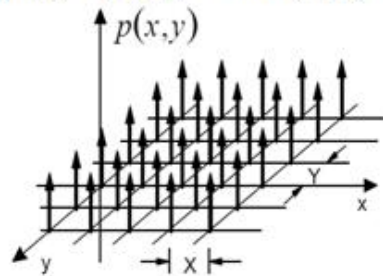


Fig-2.2-D rectangular sampling grid

$$\text{comb}(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \delta(x - m\Delta x, y - n\Delta y) - 2$$

$$f_s(x, y) = f(x, y) \times \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \delta(x - m\Delta x, y - n\Delta y) - 3$$

$$y(n) = x(n) * h(n) - 4, Y(W) = X(W) * H(W) - 5$$

Sampling conditions for no information loss – derived by examining the **spectrum of the image** \Leftrightarrow by performing the Fourier analysis:

$$f(x, y) \xrightarrow{FT} F(\epsilon 1, \epsilon 2), \text{comb}(x, y) \xrightarrow{FT} \text{COMB}(\epsilon 1, \epsilon 2) - 6$$

Where $F(\epsilon 1, \epsilon 2)$ = the Fourier transform of the image:

Therefore the Fourier transform of f_s is:

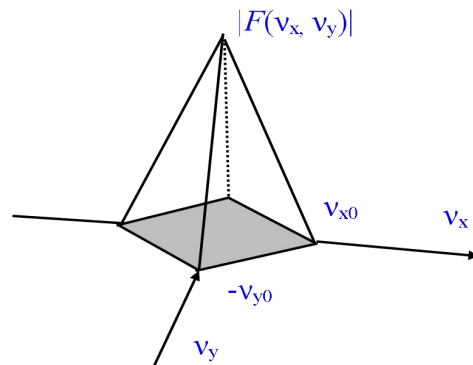
$$Fs(\epsilon 1, \epsilon 2) = F(\epsilon 1, \epsilon 2) \times \text{COMB}(\epsilon 1, \epsilon 2) - 7$$

The **sampling grid** function is periodical with period $(\Delta x, \Delta y) \Rightarrow$ can be expressed by its Fourier series expansion:

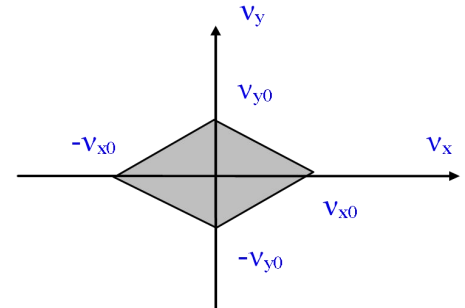
$$\text{COMB}(\epsilon 1, \epsilon 2) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \delta(\epsilon 1 - k \epsilon x_s, \epsilon 2 - l \epsilon y_s) * \epsilon x_s * \epsilon y_s - 8$$

$$\epsilon x_s = \frac{1}{\Delta x}, \epsilon y_s = \frac{1}{\Delta y} - 9$$

$$Fs(\epsilon 1, \epsilon 2) = F(\epsilon 1, \epsilon 2) \times \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \delta(\epsilon 1 - k \epsilon x_s, \epsilon 2 - l \epsilon y_s) * \epsilon x_s * \epsilon y_s - 10$$

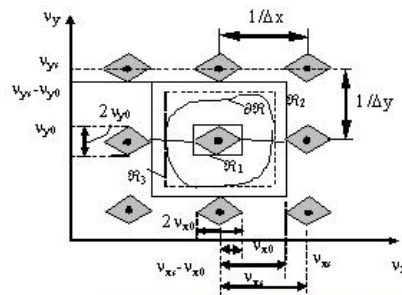


The Fourier transform of the limited spectrum image



The spectral support region

The spectrum of a limited bandwidth image and its spectral support



The sampled image spectrum

$$F_s(\xi_1, \xi_2) = F(\xi_1 - k \xi_s, \xi_2 - l \xi_s) \times \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \delta(\xi_1, \xi_2) * \xi_s * \xi_s - 11$$

$$F_s(\xi_1, \xi_2) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} F(\xi_1 - k \xi_s, \xi_2 - l \xi_s) * (\xi_s * \xi_s) - 12$$

The spectrum of the sampled image = the collection of an infinite number of scaled spectral replicas of the spectrum of the original image, centered at multiples of spatial frequencies $1/\Delta x$, $1/\Delta y$.

Image reconstruction from its samples

$$F_s(\xi_1, \xi_2) = \frac{F(\xi_1, \xi_2)}{H(\xi_1, \xi_2)} - 13$$

$$F(\xi_1, \xi_2) = F_s(\xi_1, \xi_2) * H(\xi_1, \xi_2) - 14$$

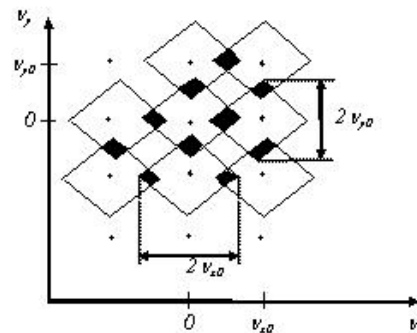
Let us assume that the filtering region R is rectangular, at the middle distance between two spectral replicas

$$H(\xi_1, \xi_2) = \frac{1}{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \xi_s * \xi_s} - 15$$

$$\xi_s \geq 2 \xi_0, \xi_s \geq 2 \xi_0 - 16$$

$$F(\xi_1, \xi_2) = 0 \text{ When } \xi_1 > \xi_0, \xi_2 > \xi_0 - 17$$

$$\tilde{f}(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f(x - m\Delta x, y - n\Delta y) \times \sin(x \xi_s - m) \sin(y \xi_s - n) - 18$$



Aliasing - fold-over frequencies

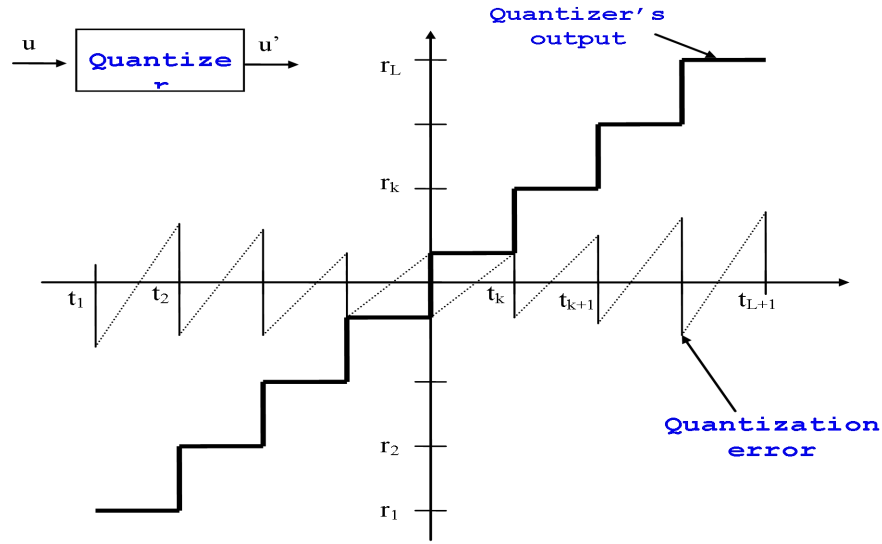
Since the sinc function has infinite extent \Rightarrow it is impossible to implement in practice the ideal LPF it is impossible to reconstruct in practice an image from its samples without error if we sample it at the Nyquist rates

9. Explain Uniform Quantization and Non-Uniform Quantization

1. Non-Uniform Quantization:

Max Lloyd Quantizer

Given the range of input u as from a_L to a_u , and the number of output levels as L , we design the quantizer such that MSQE is minimum.



Let t_k =decision level (i / p level) and r_k =reconst level (o / p level)

If $t_k \leq u \leq t_{k+1}$ then $u' = Q(u) = r_k$ and $e = u - u'$

$$MSQE = \varepsilon = E[(u - u')^2] = \int_{t_k}^{t_{k+1}} (u - r_k)^2 p(u) du$$

where $p(u)$ = Probability density function of random variable u

The error ε can also be written as

$$\varepsilon = \sum_{k=1}^L \int_{t_k}^{t_{k+1}} (u - r_k)^2 p(u) du$$

$\therefore r_k$ and t_k are variables. To minimize ε we set $\frac{\partial \varepsilon}{\partial r_k} = \frac{\partial \varepsilon}{\partial t_k} = 0$

Find t_k :

$$u = t_k, \quad u' = r_k$$

$$\varepsilon = \sum_{k=1}^L \int_{t_{k-1}}^{t_k} (u - r_{k-1})^2 p(u) du + \int_{t_k}^{t_{k+1}} (u - r_k)^2 p(u) du$$

$$\frac{\partial \varepsilon}{\partial t_k} = \frac{\partial}{\partial t_k} \left[\int_{t_{k-1}}^{t_k} (t_k - r_{k-1})^2 p(t_k) dt_k + \int_{t_k}^{t_{k+1}} (t_k - r_k)^2 p(t_k) dt_k \right]$$

$$\Rightarrow 2(t_k - r_{k-1})^2 p(t_k) + (t_k - r_k)^2 p(t_k) = 0$$

$$\Rightarrow (t_k - r_{k-1}) = \pm (t_k - r_k)$$

Since $(t_k - r_{k-1}) > 0$ and $(t_k - r_k) < 0$ solution that is valid is

$$(t_k - r_{k-1}) = -(t_k - r_k) \text{ or } t_k = (r_k + r_{k-1}) / 2$$

i.e. input level is average of two adjacent output levels.

Find r_k

$$\frac{d\varepsilon}{dr_k} = \frac{\partial}{\partial r_k} \left[\int_{t_k}^{t_{k+1}} (u - r_k)^2 p(u) du \right] = 0 \text{ --- 6}$$

$$\frac{\partial \varepsilon}{\partial r_k} = -2 \int_{t_k}^{t_{k+1}} (u - r_k) p(u) du = 0 \text{ --- 7}$$

Hence

$$r_k = \left[\int_{t_k}^{t_{k+1}} u p(u) du \right] / \left[\int_{t_k}^{t_{k+1}} p(u) du \right] \text{ --- 8}$$

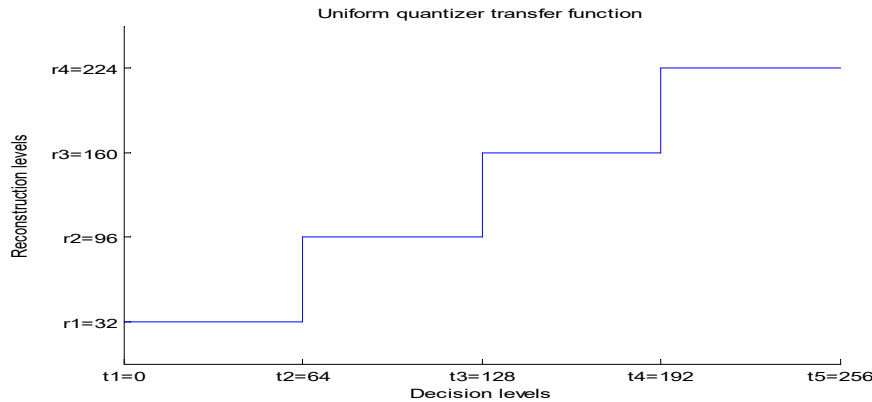
$$\therefore r_k = \frac{t_{k+1} + t_k}{2} \text{ --- 9}$$

Output level is the centroid of adjacent input levels. This solution is not closed form. To find input level t_k , one has to find r_k and vice versa. However by iterative techniques both r_k and t_k can be found using Newton's method.

i.e. each reconstruction level lies mid-way between two adjacent decision levels.

2. Uniform Quantizer:

When $p(u) = \frac{1}{a_u - a_L} = \frac{1}{A}$, $p(e) = 1/q$, as shown



Since quantization error is distributed uniformly over each step size (q) with zero mean i.e. over $(-q/2, q/2)$, the variance is,

$$\sigma_e^2 = \int_{-q/2}^{q/2} e^2 p(e) de \text{ --- 1}$$

$$\sigma_e^2 = \frac{1}{q} \int_{-q/2}^{q/2} e^2 de = q^2 / 12 \text{ --- 2}$$

Let range of u be A and σ_u^2 the variance;

$$\sigma_u^2 = \int_{-A/2}^{A/2} u^2 p(u) du \text{ --- 3}$$

$$\text{Then } \sigma_u^2 = \frac{1}{A} \int_{-A/2}^{A/2} u^2 du = A^2 / 12 \text{ --- 4}$$

A b -bit quantizer can represent L output levels as $2^b = L \therefore$ quantization step size

$$q = \frac{A}{2^b}.$$

$$\text{Hence } \sigma_e^2 = \frac{1}{12} A^2 / 2^{2b}; \dots \dots 5$$

$$\text{SNR for uniform quantizer in dB} = 10 \log(\sigma_u^2 / \sigma_e^2) = 2^{2b} \approx 6\text{dB} \dots \dots 6$$

Properties of optimum Mean square quantizers:.

1. The quantizer output is an unbiased estimate of input i.e. $E[u'] = E[u]$
 \Rightarrow SNR for uniform quantizer increases by 6 dB /bit
2. The quantizer error is orthogonal to the quantizer output i.e.,
 $E[(u - u')u'] = 0 \Rightarrow$ quantization noise is uncorrelated with quantized output.
3. The variance of quantizer o/p is reduced by the factor $[1 - f(B)]$ where $f(B)$ denotes the mean square distortion of the B-bit quantizer for unity variance inputs.

$$\text{i.e. } \sigma_u'^2 = (1 - f(B))\sigma_u^2$$

4. It is sufficient to design mean square quantizers for zero mean and unity variance distributions.

10. Explain in detail the ways to represent the digital image.

Assume that an image $f(x,y)$ is sampled so that the resulting digital image has M rows and N columns. The values of the coordinates (x,y) now become discrete quantities. For notational clarity and convenience, we shall use integer values for these discrete coordinates. Thus, the values of the coordinates at the origin are (x,y) = (0,0). The next coordinate values along the first row of the image are represented as (x,y) = (0,1). Figure shows the coordinate convention used.

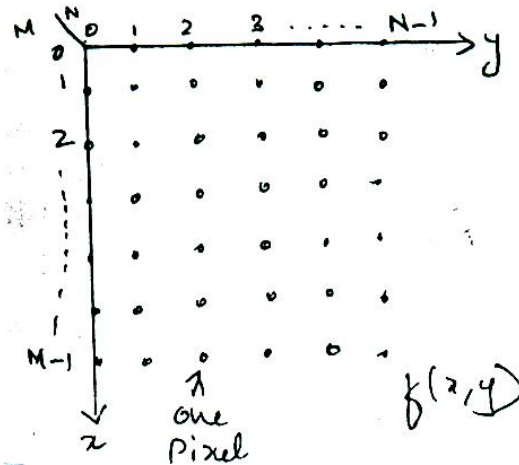


Figure: Coordinate convention used in this book to represent digital images.

The notation introduced in the preceding paragraph allows us to write the complete M X N digital image in the following compact matrix from:

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

The right side of this equation is by definition a digital image. Each element of this matrix array is called an image element, picture element, pixel. or pel. The terms image

and pixel will be used throughout the rest of our discussions to denote a digital image and its elements.

In some discussions, it is advantageous to use a more traditional matrix notation to denote a digital image and its elements:

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

Clearly, $a_{ij} = f(x = i, y = j) = f(i,j)$, so Equations are identical matrices.

This digitization process requires decisions about values for M, N . and for the number, L , of discrete gray levels allowed for each pixel. There are no requirements on M and N , other than that they have to be positive integers. However, due to processing, storage, and sampling hardware considerations, the number of gray levels typically is an integer power of 2:

$$L = 2^k$$

We assume that the discrete levels are equally spaced and that they are integers in the interval $[0, L - 1]$.

The number, b , of bits required to store a digitized image is

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

When $M = N$, this equation becomes $b = N^2 k$.

Number of storage bits for various values of N and k .

Nk	1($L=2$)	2($L=4$)	3($L=8$)	4($L=16$)	5($L=32$)	6($L=64$)	7($L=128$)	8($L=256$)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

11. Write short notes on: (a) Neighbors of Pixels (b) Distance Measures c.

Connectivity d. Adjacency

1. Neighbors of a pixels:

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

This set of pixels, called the *4-neighbors* of p , is denoted by $N_4(p)$. Each pixel is a unit distance from (x, y) , and some of the neighbors of p lie outside the digital image if (x, y) is on the border of the image.

The four *diagonal* neighbors 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-neighbors, are called the *8-neighbors* of p , denoted by $N_8(p)$. As before, some of the points in $N_D(p)$ and $N_8(p)$ fall outside the image if (x, y) is on the border of the image.

8-neighbors of p : $N_8(p)$: All the 4- and diagonal neighbors

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

2. Distance Measures;

For pixels p , q , and z , with coordinates (x, y) , (s, t) , and (u, w) , respectively, D is a *distance function or metric* if

$$(a) D(p, q) \geq 0 \text{ (} D(p, q) = 0 \text{ iff } p = q),$$

$$(b) D(p, q) = D(q, p), \text{ and}$$

$$(c) D(p, z) \leq D(p, q) + D(q, z).$$

The *Euclidean distance* between p and q is defined as

$$D(p, q) = [(x - s)^2 + (y - t)^2]^{1/2}. \quad \dots\dots\dots 1$$

For this distance measure the pixels having a distance less than or equal to some value r from (x, y) are the points contained in a disk of radius r centered at (x, y) .

The D_4 distance (also called *city-block distance*) between p and q is defined as

$$D_4(p, q) = |x - s| + |y - t|. \quad \dots\dots\dots 2$$

In this case, the pixels having a D_4 distance from (x, y) less than or equal to some value r form a diamond centered at (x, y) . For example, the pixels with D_4 distance ≤ 2 from (x, y) (the center point) form the following contours of constant distance:

$$\begin{array}{ccccc} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{array}$$

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

The D_8 distance (also called *chessboard distance*) between p and q is defined as

$$D_8(p, q) = \max(|x - s|, |y - t|). \quad \dots\dots\dots 3$$

In this case, the pixels with D_8 distance from (x, y) less than or equal to some value r form a square centered at (x, y) . For example, the pixels with D_8 distance ≤ 2 from (x, y) (the center point) form the following contours of constant distance.

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

The pixels with $D_8 = 1$ are the 8-neighbors of (x, y) .

3.Connectivity

Connectivity between pixels is a fundamental concept that simplifies the definition of numerous digital image concepts, such as regions and boundaries. To establish if two pixels are connected, it must be determined if they are neighbors and if their gray levels satisfy a specified criterion of similarity (say, if their gray levels are equal). For instance, in a binary image with values 0 and 1, two pixels may be 4-neighbors, but they are said to be connected only if they have the same value.

4.Adjacency: Consider V as a set of gray-level values, and p and q as two pixels.

4-adjacency: two pixels p and q from V are 4-adjacent if q is in $N_4(p)$.

8-adjacency: two pixels p and q from V are 8-adjacent if q is in $N_8(p)$.

m-adjacency: two pixels p and q from V are m-adjacent if q is in $N_4(p)$, or in $N_D(p)$ and $N_4(p) \cap N_4(q)$ has no pixels whose values are from V .

Pixel Arrangement	8-Adjacent Pixels	m-Adjacent Pixels
0 1 1	0 1 1	0 1 1
0 1 0	0 1 0	0 1 0
0 0 1	0 0 1	0 0 1

$V\{1\}$: The set V consists of pixels with value 1.

Two image subsets S_1, S_2 are adjacent if at least one pixel from S_1 and one pixel from S_2 are adjacent.

12. Write short notes on two dimensional Orthogonal and Unitary transforms

Two dimensional unitary transforms play an important role in image processing. The term image transform refers to a class of unitary matrices used for representation of images.

In analogy, with I-D signals that can be represented by an orthogonal series of basis functions, we can similarly represent an image in terms of a discrete set of basis arrays called “**basis images**”. These are generated by unitary matrices.

Alternatively an $(N \times N)$ image can be represented as $(N^2 \times 1)$ vector. An image transform provides a set of coordinates or basis vectors for the vector space.

I-D-Transforms:

For a one dimensional sequence $\{ u(n), \quad n = 0, 1, \dots, N-1 \}$ representing a vector \vec{u} of size N , a unitary transform is: $\vec{v} = \underline{A} \vec{u}$

$$\Rightarrow v(k) = \sum_{n=0}^{N-1} a(k,n)u(n), \text{ for } 0 \leq k \leq N-1 \quad (1)$$

where $\underline{A}^{-1} = \underline{A}^{*T}$ (unitary)

This implies, $\vec{u} = \underline{A}^{*T} \vec{v}$

$$\text{or, } u(n) = \sum_{k=0}^{N-1} v(k)a^*(k,n), \text{ for } 0 \leq n \leq N-1 \quad (2) \quad \text{Equation}$$

(2) can be viewed as a series representation of sequence $u(n)$. The columns of

\underline{A}^{*T} i.e the vectors $\vec{a}_k^* = \{a^*(k,n), 0 \leq n \leq N-1\}^T$ are called the “basis vectors” of

\underline{A} .

The series coefficients $v(k)$ give a representation of original sequence $u(n)$ and are useful in compression, filtering, feature extraction and other analysis.

Two dimensional Orthogonal and Unitary transforms:

As applied to image processing, a general orthogonal series expansion for an $N \times N$ image is a pair of transformations of the form :

$$v(k,l) = \sum_{m,n=0}^{N-1} u(m,n)a_{k,l}(m,n), \quad 0 \leq k,l \leq N-1 \quad (3)$$

$$u(m,n) = \sum_{k,l=0}^{N-1} v(k,l)a_{k,l}^*(m,n), \quad 0 \leq m,n \leq N-1 \quad (4)$$

where $\{a_{k,l}(m,n)\}$ is called an “image transform.”

It is a set of complete orthogonal discrete basis functions satisfying the properties:-

$$\begin{aligned} 1) \text{ Orthonormality: } & \sum_{m,n=0}^{N-1} a_{k,l}(m,n)a_{k',l'}^*(m,n) \\ & = \delta(k-k', l-l') \end{aligned}$$

$$\begin{aligned} 2) \text{ Completeness : } & \sum_{k,l=0}^{N-1} a_{k,l}(m,n)a_{k,l}^*(m',n') \\ & = \delta(m-m', n-n') \end{aligned}$$

The elements $v(k,l)$ are transform coefficients and $\underline{V} \triangleq \{v(k,l)\}$ is the transformed image.

The orthonormality property assures that any truncated series expansion of the form

$$U_{P,Q}(m,n) \triangleq \sum_{k=0}^{P-1} \sum_{l=0}^{Q-1} v(k,l)a_{k,l}^*(m,n), \text{ for } P \leq N,$$

$$Q \leq N$$

will minimize the sum of squares error

$$\sigma_e^2 = \sum_{m,n=0}^{N-1} [u(m,n) - U_{P,Q}(m,n)]^2$$

where coefficients $v(k,l)$ are given by (3).

The completeness property assures that this error will be zero for $P = Q = N$

Separable Unitary Transforms:

The number of multiplications and additions required to compute transform coefficients $v(k,l)$ in equation(3) is $O(N^4)$. This is too large for practical size images.

If the transform is restricted to be separable,

$$\begin{aligned} \text{i.e } a_{k,l}(m,n) &= a_k(m)b_l(n) \\ &\triangleq a(k,m)b(l,n) \end{aligned}$$

where $\{a_k(m), k = 0(1)n-1\}$,

and $\{b_l(n), l = 0(1)N-1\}$ are 1D complete orthogonal sets of basis vectors.

On imposition of completeness and orthonormality properties we can show that $\underline{A} \triangleq \{a(k,m)\}$,

and $\underline{B} \triangleq \{b(l,n)\}$ are unitary matrices.

$$\text{i.e } \underline{A}\underline{A}^{*T} = \underline{I} = \underline{A}^T \underline{A}^* \text{ and } \underline{B}\underline{B}^{*T} = \underline{I} = \underline{B}^T \underline{B}^*$$

Often one chooses \underline{B} same as \underline{A}

$$\begin{aligned} \therefore v(k,l) &= \sum_{m,n=0}^{N-1} a(k,m)u(m,n)a(l,n) \\ \Leftrightarrow \underline{V} &= \underline{A} \underline{U} \underline{A}^T \end{aligned} \quad (5)$$

$$\begin{aligned} \text{And } u(m,n) &= \sum_{k,l=0}^{N-1} a^*(k,m)v(k,l)a^*(l,n) \\ \Leftrightarrow \underline{U} &= \underline{A}^{*T} \underline{V} \underline{A}^* \end{aligned} \quad (6)$$

Eqn (5) can be written as $\underline{V}^T = \underline{A}(\underline{A}\underline{U})^T$

\Rightarrow Eqn (5) can be performed by first transforming each column of \underline{U} and then transforming each row of the result to obtain rows of \underline{V} .

Basis Images : Let \bar{a}_k^* denote k^{th} column of \underline{A}^{*T} .

Let us define the matrices $\underline{A}_{k,l}^* = \bar{a}_k^* \bar{a}_l^{*T}$ and matrix inner product of two $N \times N$ matrices \underline{F} and \underline{G} as

$$\langle \underline{F}, \underline{G} \rangle = \sum_{m,n=0}^{N-1} f(m,n)g^*(m,n)$$

Then equ (6) and (5) give a series representation.

$$\underline{U} = \sum_{k,l=0}^{N-1} v(k,l) \underline{A}_{k,l}^*$$

$$\text{and } v(k,l) = \langle \underline{u}, \underline{A}_{k,l}^* \rangle$$

Any image \underline{U} can be expressed as linear combination of N^2 matrices. $\underline{A}_{k,l}^*$ called “basis images”.

Therefore any $N \times N$ image can be expanded in a series using a complete set of N^2 basis images.

Example: Let $\underline{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$; $\underline{U} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Transformed image $\underline{V} = \underline{A} \underline{U} \underline{A}^T = \begin{pmatrix} 5 & -1 \\ -2 & 0 \end{pmatrix}$

And Basis images are found as outer product of columns of \underline{A}^{*T} i.e

$$\underline{A}_{0,0}^* = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1 \quad 1)$$

$$\underline{A}_{0,1}^* = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1 \quad -1)$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} = \underline{A}_{1,0}^{*T}$$

$$\underline{A}_{1,1}^* = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1 \quad -1)$$

The inverse transformation

$$\begin{aligned} \underline{A}^{*T} \underline{V} \underline{A}^* &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 & -1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \underline{U} \end{aligned}$$

13. Explain the 2-D DFT transformations

The two dimensional Discrete Fourier Transform (2-D DFT)

If $f(x, y)$ is an $M \times N$ array, such as that obtained by sampling a continuous function of two dimensions at dimensions M and N on a rectangular grid, then its two dimensional Discrete Fourier transform (DFT) is the array given by

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

$$u = 0, \dots, M-1, \quad v = 0, \dots, N-1$$

and the inverse DFT (IDFT) is

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

When images are sampled in a square array, $M = N$ and

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

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

It is straightforward to prove that the two dimensional Discrete Fourier Transform is separable, symmetric and unitary.

14. Explain in detail the discrete Cosine Transform (DCT) and also explain its properties.

The $N \times N$ cosine transform matrix $C = \{c(K, n)\}$, also called the discrete cosine transform (DCT), is defined as

$$C(k, n) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0, 0 \leq n \leq N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N}, & 1 \leq k \leq N-1, 0 \leq n \leq N-1 \end{cases}$$

The one dimensional DCT of a sequence $\{u(n), 0 \leq n \leq N-1\}$ is defined as

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos \left[\frac{\pi(2n+1)k}{2N} \right], \quad 0 \leq k \leq N-1$$

Where

$$\alpha(0) = \sqrt{\frac{1}{N}}, \quad \alpha(k) = \sqrt{\frac{2}{N}} \text{ for } 1 \leq k \leq N-1$$

The inverse transformation is given by

$$U(n) = \sum_{k=0}^{N-1} \alpha(k) v(k) \cos \left[\frac{\pi(2n+1)k}{2N} \right], \quad 0 \leq n \leq N-1$$

The two-dimensional cosine transform pair is obtained by substituting $A = A^* = C$ in (5.11) and (5.12). The basis images of the 8×8 two-dimensional cosine transform are shown in figures shows examples of the cosine transform of different images.

Properties of the Cosine Transform

1. The cosine transform is real and orthogonal, that is, $C = C^* \Rightarrow C^{-1} = C^T$
2. The cosine transform is not the real part of the unitary DFT. This can be seen by inspection of C and the DFT matrix F . (Also see Problem) However, the cosine transform of a sequence is related to the DFT of its symmetric extension (see Problem)
3. The cosine transform is a fast transform. The cosine transform of a vector of N elements can be calculated in $O(N \log_2 N)$ operations via an N -point FFT [19]. To show this we define a new sequence $\bar{u}(n)$ by reordering the even and odd elements of $u(n)$ as

$$\begin{cases} \bar{u}(n) = u(2n) \\ \bar{u}(N-n-1) = u(2n+1) \end{cases}, \quad 0 \leq n \leq \left(\frac{N}{2}\right) - 1$$

Now, we split the summation term in into even and odd terms and use to obtain.

$$v(k) = \alpha(k) \left\{ \sum_{n=0}^{(N/2)-1} u(2n) \cos \left[\frac{\pi(4n+1)k}{2N} \right] + \sum_{n=0}^{(N/2)-1} u(2n+1) \cos \left[\frac{\pi(4n+3)k}{2N} \right] \right\}$$

$$\alpha(k) \left\{ \sum_{n=0}^{(N/2)-1} u(n) \cos \left[\frac{\pi(4n+1)k}{2N} \right] + \sum_{n=0}^{(N/2)-1} u(N-n-1) \cos \left[\frac{\pi(4n+3)k}{2N} \right] \right\}$$

Changing the index of summation in the second term to $n' = N - n - 1$ and combining terms, we obtain

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos \left[\frac{\pi(4n+1)k}{2N} \right]$$

$$\text{Re} = \left[\alpha(k) e^{-j\pi k/2N} \sum_{n=0}^{N-1} u(n) e^{-j2\pi kn/N} \right] = \text{Re} \left[\alpha(k) W_{2N}^{k/2} \text{DFT}\{u(n)\} N \right]$$

Which proves the previously stated result For inverse cosine transform we write for even data points as

$$u(2n) = \hat{u}(2n) \quad \text{Re} \left[\sum_{k=0}^{N-1} \left[\alpha(k) v(k) e^{j\pi/2N} \right] e^{j2\pi nk/N} \right],$$

$$0 \leq n \leq \left(\frac{N}{2} \right) - 1$$

The odd data points are obtained by nothing that

$$U(2n+1) = \hat{u}[2(n-1-n)], \quad 0 \leq n \leq \left(\frac{N}{2} \right) - 1$$

Therefore, if we calculate the N -point inverse FFT of the sequence $\alpha(k) v(k) \exp(j\pi k/2N)$, we can also obtain the inverse DCT in $O(N \log N)$ operations. Direct algorithms that do not require FFT as an intermediate step, so that complex arithmetic is avoided, are also possible [18]. The computational complexity of the direct as well as the FFT based methods is about the same.

4. The cosine transform has excellent energy compaction for highly correlated data. This is due to the following properties.

5. The basis vectors of the cosine transform (that is, rows of C) are the eigenvectors of the symmetric tridiagonal matrix Q_c , defined as

$$Q_c = \begin{pmatrix} 1-\alpha & -\alpha & 0 \\ -\alpha & 1 & 1-\alpha \\ 0 & \alpha & 1-\alpha \end{pmatrix}$$

6. The $N \times N$ cosine transform is very close to the KL transform of a first-order stationary Markov sequence of length N whose covariance matrix is given by when the correlation parameter ρ is close to 1. The reason is that R^{-1} is a symmetric tridiagonal matrix, which for a scalar $\beta^2 = (1-\rho^2)/(1+\rho^2)$ and $\alpha = \rho/(1+\rho^2)$ satisfies the relation

$$\beta^2 R^{-1} = \begin{pmatrix} 1-\rho\alpha & -\alpha & 0 \\ -\alpha & 1 & 1-\alpha \\ 0 & \alpha & 1-\rho\alpha \end{pmatrix}$$

This gives the approximation $\beta^2 R^{-1} \cong Q_c$ for $\rho \cong 1$

Hence the eigenvectors of R and the eigenvectors of Q_c , that is, the cosine transform, will be quite close. These aspects are considered in greater depth in Section on sinusoidal transforms.

This property of the cosine transform together with the fact that it is a fast transform has made it a useful substitute for the KL transform of highly correlated first order Markov sequences.

UNIT II IMAGE ENHANCEMENT

Spatial Domain: Gray level transformations – Histogram processing – Basics of Spatial Filtering– Smoothing and Sharpening Spatial Filtering, Frequency Domain: Introduction to Fourier Transform– Smoothing and Sharpening frequency domain filters – Ideal, Butterworth and Gaussian filters, Homomorphic filtering, Color image enhancement.

PART A

1. Define Image Enhancement.

Image enhancement is a process an image so that the result is more suitable than the original image for specific application. Image enhancement is useful in feature extraction, image analysis and visual information display.

2. What are the two categories of Image Enhancement?

Image Enhancement approaches falls into two broad categories: they are

1. Spatial domain approaches
2. Frequency domain approaches
 - i) Spatial domain refers to image plane itself & approaches in this category are based on direct manipulation of picture image.
 - ii) Frequency domain methods based on modifying the image by Fourier transform.

3. Write the transfer function for Butterworth filter. (May/June 2009)

The transfer function for a Butterworth Low pass filter is,

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

The transfer function for a Butterworth High pass filter is,

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

Where, n = filter order, D_o = cutoff frequency and $D(u, v) = (u^2 + v^2)^{1/2}$ = distance from point (u, v) to the origin

4. Specify the objective of image enhancement technique. (Or) Why the image enhancement is needed in image processing technique. (May / June 2009)

1. 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. Image Enhancement process consists of a collection of technique that seek to improve the visual appearance of an image. Thus the basic aim is to make the image look better.
3. Image enhancement refers to accentuation or sharpening of image features such as edges, boundaries or contrast to make a graphic display more useful for display and analysis.
4. Image enhancement includes gray level and contrast manipulation, noise reduction, edge crispening & sharpening, filtering, interpolation & magnification, pseudo coloring, so on...

5. Mention the various image enhancement techniques.

a. Point operations or basic gray level transformations

- a) Contrast stretching
- b) Noise clipping
- c) Window slicing
- d) Histogram modeling

b. Spatial domain methods

- a) Noise smoothing
- b) Median filtering
- c) LP,HP,BP filtering
- d) Zooming
- e) Unsharp masking

c. Frequency domain methods

- a) Linear filtering
- b) Homomorphic filtering
- c) Root filtering

d. Pseudo coloring

- a) False coloring
- b) Pseudo coloring

6. What is contrast stretching? (May/June 2009)

1. Contrast stretching reduces an image of higher contrast than the original by darkening the levels below m and brightening the levels above m in the image.

2. Contrast stretching is used to increase the dynamic range of the gray levels in the image being processed.

Low contrast images occur often due to poor / non-uniform lighting conditioning or due to non-linearity or small dynamic range of the image sensor.

7. What is meant by histogram of a digital image? (May / June 2006), (May / June 2009), (May / June 2007)

1. Histogram of an image represents the relative frequency of occurrence of the various gray levels in the image.

2. Histogram modeling techniques modify an image so that its histogram has a desired shape. This is useful in stretching the low contrast levels of images with narrow histograms.

3. Thus histogram is used to either increase or decrease, the intensity level of the pixels by any kind of transformation.

The histogram of a digital image with gray levels in the image $[0,L-1]$ is a discrete function given as,

$$h(r_k) = n_k, \text{ where } r_k \text{ is the } k\text{th grey level, } k = 0,1,\dots,L-1$$

$$n_k \text{ is the number of pixels in the image with grey level } r_k$$

8. What is Median Filtering? (May / June 2009) (May / June 2007)

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

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

2. Median filters are useful for removing isolated lines or points (pixels) while preserving spatial resolutions. They perform very well on images containing binary (**salt and pepper**) noise but perform poorly when the noise is Gaussian.

9. What do you mean by Point processing?

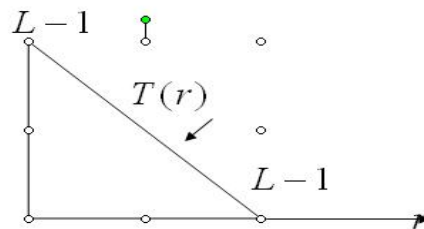
When enhancement is done with respect to a single pixel, the process is called point processing. i.e., enhancement at any point in an image depends only on the gray level at that point.

10. What is Image Negatives?

A negative image can be obtained by reversing the intensity levels of an image, according to the transformation,

$$s = L - 1 - r$$

, when the gray level range is $[0, L-1]$



1. Digital negatives are useful in the display of medical images and to produce negative prints of image.

2. It is also used for enhancing white or gray details embedded in dark region of an image.

11. Write the steps involved in frequency domain filtering.

1. Multiply the input image by $(-1)^{x+y}$ 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 $(-1)^{x+y}$

12. What is meant by Laplacian filter?

It is a linear operator and is the simplest isotropic derivative operator, which is defined as,

$$\nabla^2 f = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \quad \text{-----(1)}$$

Where,

$$\frac{\partial^2 f(x,y)}{\partial x^2} \cong [f(x+1,y) + f(x-1,y) - 2f(x,y)] \quad \text{-----(2)}$$

$$\frac{\partial^2 f(x,y)}{\partial y^2} \cong [f(x,y+1) + f(x,y-1) - 2f(x,y)] \quad \text{-----(3)}$$

Sub (2) & (3) in (1), we get,

$$\nabla f^2 \cong [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)]$$

The above digital laplacian equation is implemented using filter mask,

0	1	0
1	-4	1
0	1	0

13. 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.	$R = w(-1,-1) f(x-1,y-1) + w(-1,0) f(x-1,y) + \dots + w(0,0) f(x,y) + \dots + w(1,0) f(x+1,y) + w(1,1) f(x+1,y+1).$	$R = w_1z_1 + w_2z_2 + \dots + w_9z_9$ $= \sum_{i=1}^9 w_i z_i$

14. 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.

15. What is meant by masking?

Mask is the small 2-D array in which the value of mask co-efficient determines the nature of process. The enhancement technique based on this type of approach is referred to as mask processing. Image sharpening is often referred to as mask processing or filtering.

16. Give the formula for log transformation.

Log transformations are used to compress the dynamic range of the image data, using the equation,

$$s = c \log_{10} (1 + |r|)$$

Where c is constant & $u \geq 0$

17. Write the application of sharpening filters?

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

18. Define clipping.

A special case of contrast stretching where the slope values $\alpha = r=0$ is called slipping. This is useful for noise reduction.

19. Name the different types of derivative filters?

1. Laplacian operators
2. Prewitt operators
3. Roberts cross gradient operators

20. Define Spatial Domain Process.

The term spatial domain refers to the image plane itself and approaches in this category are based on direct manipulation of pixels in an image. Spatial domain processes is denoted by the expression.

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

Where $f(x,y) \rightarrow$ input image $G(x,y) \rightarrow$ processed image
 $T \rightarrow$ operator of f , defined over some neighborhood of (x,y)

21. Define Frequency domain process

The term frequency domain refers to the techniques that are based on modifying the Fourier transform of an image.

22. What is meant by gray-level transformation function?

The spatial domain process is denoted by the expression.

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

From this expression the simplest form of T is when the neighborhood is of single pixel and 'g' depends only on the value of f at (x,y) and T becomes a gray level transformation function which is given by

$r \rightarrow$ grey level of $f(x,y)$
 $S \rightarrow$ grey level of $g(x,y)$

$$S = T(r)$$

23. Define contrast stretching.

The contrast stretching is a process of to increase the dynamic range of the gray levels in the image being processed.

24. Define Thresholding.

Thresholding is a special case of clipping where $a=b$ and output becomes binary. For examples, a seemingly binary image, such as printed page, does not give binary output when scanned because of sensor noise and background illumination variations. Thresholding is used to make such an image binary.

25. What are the types of functions of gray level Transformations for Image Enhancement?

Three basic types of functions used frequently for image enhancement are

1. Linear (negative and identity transformations)
2. Logarithmic Transformations
3. Power law Transformations.

26. Give the expression of Log Transformations?

The general form of the log Transformation is given

$$S = C \log (1+r).$$

Where C is a construct and it is assumed that $r \geq 0$. this transformation enhances the small magnitude pixels compared to those pixels with large magnitudes.

27. Write the expression for power-Law Transformations.

Power-law transformations have the basic form

$$S=Cr^\gamma$$

Where c and r are positive constants.

28. What is meant by gamma and gamma correction?

Power law transformation is given by

$$S=Cr^\gamma$$

The exponent γ in the power-law equation is referred to as gamma. The process used to correct the power law response phenomena is called gamma corrections.

29. What is the advantage & disadvantages of piece-wise Linear Transformations?

The principal advantage of piecewise Linear functions over the other types of functions is that piece wise functions can be arbitrarily complex.

The main disadvantage of piece wise functions is that their specification requires more user input.

30. What is meant by gray-level slicing?

Gray-level slicing is used in areas where highlighting a specific range of gray level in a image. There are 2 basic approaches gray level slicing.

1. Highlights range $\{A,B\}$ of gray levels and reduces all other to a constant level.
2. Second approach is that highlights range $[A,B]$ but preserves all other levels.

31. What is meant by Bit-plane slicing?

Instead of high lighting grey-level ranges, highlighting the contributes made to total image appearance by specific bits is called Bit-plane slicing.

use of bit-plane slicing: This transformation is useful in determining the visually significant bits.

32. What is histogram and write its significance?

The histogram of an image represents the relative frequency of occurrence of the various gray levels in the image. They are the basis for numerous spatial domain processing techniques. Histogram manipulation can be used effectively for image enhancement.

33. Define histogram Equalization.

Histogram Equalization refers to the transformation (or) mapping for which the processed (output) image is obtained by mapping each pixel into the input image with the corresponding pixel in the output image.

34. Define Histogram matching (or) histogram specification.

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

35. What is meant by Image subtraction?

The difference between two images $f(x,y)$ and $h(x,y)$ expressed as

$$Y(x,y)=f(x,y)-h(x,y)$$

is obtained by computing the difference all pairs off corresponding pixels from f and h. the key usefulness of subtraction is the enhancement of differences between images.

36. Define mask & what is meant by spatial mask?

Same neighborhood operations work with the values of eh image pixels in the neighborhood and the corresponding values of a sub image that has the same dimension as

the neighborhood. The sub image is called filter (or) mask (or) kernel (or) template (or) window.

When the image is convolved with a finite impulses response filter it is called as spatial mask.

37. Define spatial filtering.

Filtering operations that are performed directly on the pixels of an image is called as spatial filtering.

38. What are smoothing filters (spatial)?

Smoothing filters are used for blurring and for noise reduction. Blurring is used in preprocessing steps.

The output of a smoothing linear spatial filter is simply the average of pixels contained in the neighborhood of the filter mask.

39. Why smoothing Linear filters also called as averaging filters?

The output of a smoothing, linear spatial filter is simply the average of the pixels contained in the neighborhood of the filter mask. For this reason these filters are called as averaging filters. They are also called as low pass filters.

40. What are sharpening filters? Give an example?

The principal objective of sharpening is to highlight fine detail in an image (or) to enhance detail that has been blurred, either in error or as natural effect of a particular method of image acquisition. Uses of image sharpening applications such as electronic printing, medical imaging and autonomous guidance in military systems

41. Define unsharp masking.

The process of subtracting a blurred version of an image from the image itself is called as unsharp masking which is expressed as

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

42. What is meant by arithmetic filer?

Compute the average value of the corrupted image $g(x, y)$ in the area defined

The value of the restored image at any point (x, y)

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

Note: Using a convolution mask in which all coefficients have value $1/mn$. Noise is reduced as a result of blurring.

43. What is meant by geometric mean filer?

Geometric mean filter is given by the expression

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

44. What is meant by harmonic mean filer?

The harmonic mean filter operation is given by the expression

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

45. What is meant by contra harmonic mean filter?

The contra harmonic mean filter operation is given by the expression

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

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

PART – B

IMAGE ENHANCEMENT – Introduction

Image Enhancement process consists of a collection of technique that seek to improve the visual appearance of an image. Thus the basic aim is to make the image look better.

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.

Image enhancement refers to accentuation or sharpening of image features such as edges, boundaries or contrast to make a graphic display more useful for display and analysis.

Image enhancement includes gray level and contrast manipulation, noise reduction, edge crispening & sharpening, filtering, interpolation & magnification, pseudo coloring, so on...

The 2 categories of image enhancement.

- i) Spatial domain refers to image plane itself & approaches in this category are based on direct manipulation of picture image.
- ii) Frequency domain methods based on modifying the image by Fourier transform.

a.Spatial Domain Filters

1. Spatial (time) domain techniques are techniques that operate directly on pixels whereas Frequency domain techniques are based on modifying the Fourier transform of an image.

2. The use of a spatial mask for image processing is called spatial filtering.

3. Mechanics of spatial filtering:

The process of spatial filtering consists of moving the filter mask from point to point in an image. At each point(x,y), the response of the filter at that point is calculated using a predefined relationship.

4. Types of spatial filtering:

- i) Linear filter
 - a) Lowpass filters
 - b) High pass filters
 - c) Bandpass filters
- ii) Non linear filter
 - a) Median filter
 - b) Max filter
 - c) Min filter

b.Frequency Domain Filters

1. Frequency domain techniques are based on modifying the Fourier transform of an image whereas Spatial (time) domain techniques are techniques that operate directly on pixels.
2. Edges and sharp transitions (e.g., noise) in an image contribute significantly to high-frequency content of FT.
3. Low frequency contents in the FT correspond to the general gray level appearance of an image over smooth areas.
4. High frequency contents in the FT correspond to the finer details of an image such as edges, sharp transition and noise.
5. Blurring (smoothing) is achieved by attenuating range of high frequency components of FT.
6. Filtering in Frequency Domain with $H(u,v)$ is equivalent to filtering in Spatial Domain with $f(x,y)$.
 - Low Pass Filter: attenuates high frequencies and allows low frequencies. Thus a LP Filtered image will have a smoothed image with less sharp details because HF components are attenuated
 - High Pass Filter: attenuates low frequencies and allows high frequencies. Thus a HP Filtered image will have a

There are three types of frequency domain filters, namely,

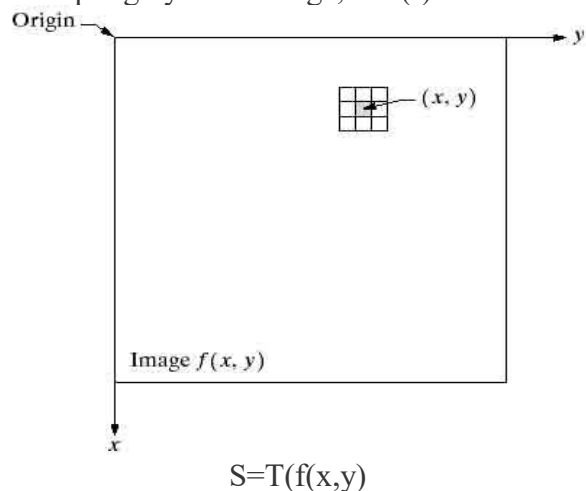
- i) Smoothing filters
 - ii) Sharpening filters
 - iii) Homomorphic filters
1. **Explain Enhancement using point operations. (Nov / Dec 2005) (May / June 2006) (May / June 2009)(or)Write short notes on Contrast stretching and Grey level slicing. (May / June 2009)**

Point operations:

Point operations are zero memory operations where a given gray level $r \in [0,L]$ is mapped into a gray level $s \in [0,L]$ according to the transformation,

$s = T(r)$, where s = transformed gray level image

r = input gray level image, $T(r)$ = transformation



The center of the sub image is moved from pixel to pixel starting, say, at the top left

corner. The operator T is applied at each location (x, y) to yield the output, g , at that location. The process utilizes only the pixels in the area of the image spanned by the neighborhood.

Where $f(x, y)$ is the input image, $g(x, y)$ is the processed image, and T is an operator on f , defined over some neighborhood of (x, y) . In addition, T can operate on a *set* of input images, such as performing the pixel-by-pixel sum of K images for noise reduction

Since enhancement is done with **respect to a single pixel**, the process is called **point processing**.ie., enhancement at any point in an image depends only on the gray level at that point.

Applications of point processing:

- To change the contrast or dynamic range of an image
- To correct non linear response of a sensor used for image enhancement

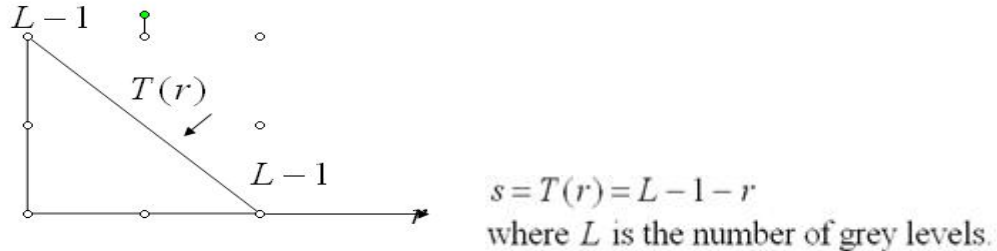
1. Basic Gray Level Transformations

The three basic types of functions used frequently for image enhancement

(a) Image Negatives: or Digital Negatives:

A negative image can be obtained by **reversing the intensity levels** of an image, according to the transformation,

$$s = L - 1 - r, \text{ when the gray level range is } [0, L-1]$$



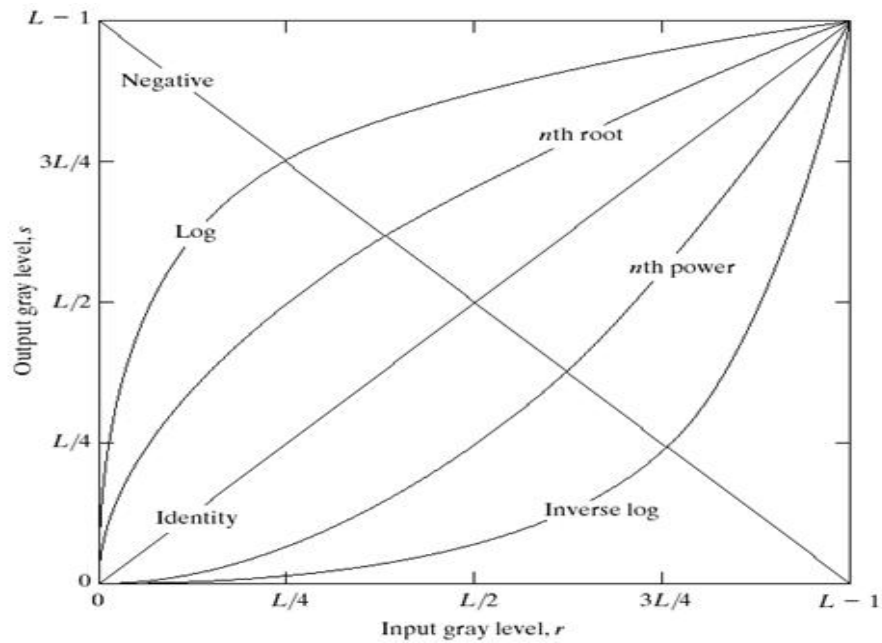
Digital negatives are useful in the display of medical images and to produce negative prints of image. It is also used for enhancing white or gray details embedded in dark region of an image.

(b) Log Transformations:

The dynamic range of gray levels in an image can be very large and it can be compressed. Hence, log transformations are **used to compress the dynamic range of the image data**, using the equation,

$$s = c \log_{10} (1 + |r|)$$

Where c is constant & $r \geq 0$



From the above diagram, it can be seen that this transformation maps narrow range of low gray-level values in the input image into a wider range of output levels. The opposite is true which is used to expand the values of dark pixels in an image while compressing the higher-level values. This done by the inverse log transformation

(c) Power Law Transformations:

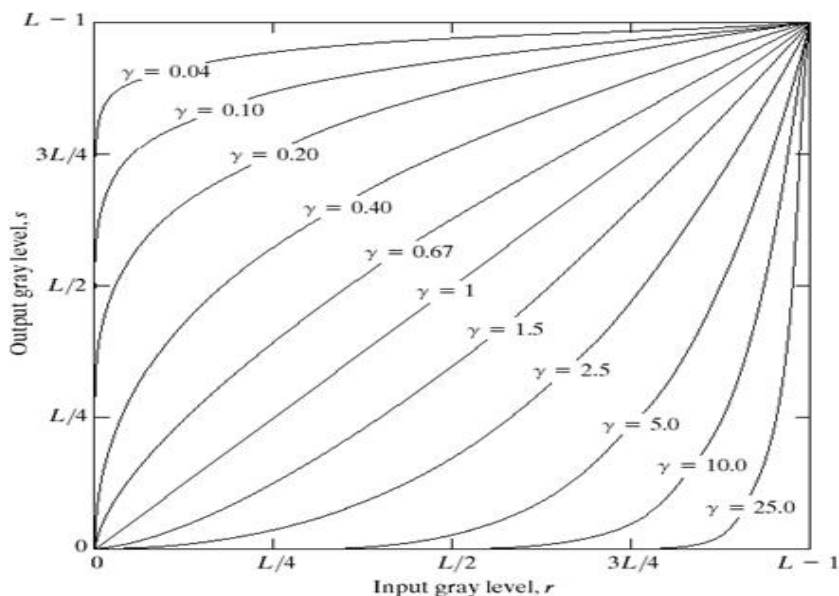
The power law transformation equation is given as,

$$s = c r^\gamma$$

Where c and γ are the constants

The above equation can also be written as,

$$s = c (r+\epsilon)^\gamma$$



This transformation maps a narrow range of dark input values into a wider range of output values, for fractional values of γ .

For $c = \gamma = 1$, identity transformation is obtained.

For $\gamma > 1$, the transformation is opposite for that obtained for $\gamma < 1$

The process which is used to convert this power law response phenomenon is called **gamma correction**.

$\gamma > 0$ Compresses dark values and Expands bright values

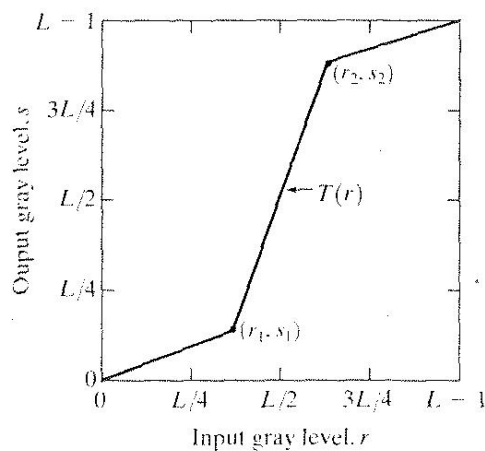
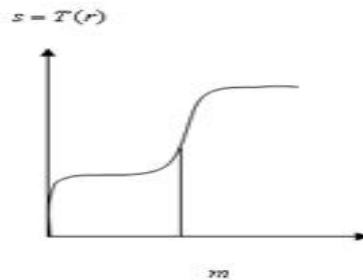
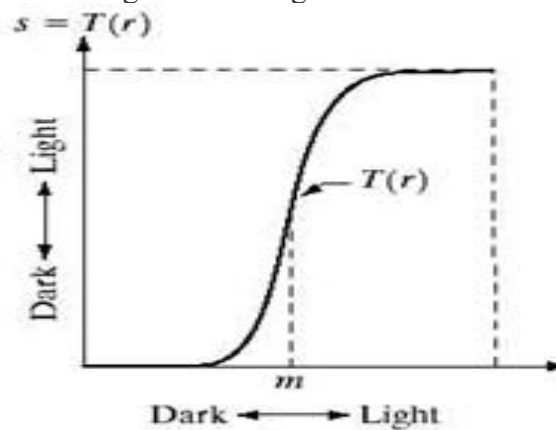
$\gamma < 0$ Expands dark values and Compresses bright values

2. Piecewise-linear transformations:

(a) Contrast Stretching:

Contrast stretching is used to **increase the dynamic range of the gray levels** in the image being processed.

Low contrast images occur often due to poor / non-uniform lighting conditioning or due to non-linearity or small dynamic range of the image sensor.



Produce an image of higher contrast than the original by darkening the levels below m and brightening the levels above m in the original image. In this technique, known as *contrast stretching*.

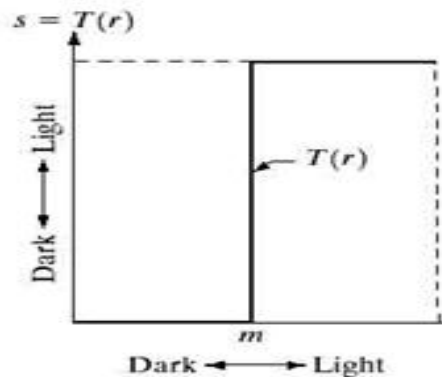
(b) Clipping:

A special case of Contrast stretching, where $\alpha = \gamma = 0$ is called clipping. This transformation **is used for noise reduction**, when the input signal is said to lie in the range $[a,b]$.

(c) Thresholding:

The values of r below m are compressed by the transformation function into a narrow range of s , toward black. The opposite effect takes place for values of r above m . $T(r)$ produces a two-level (binary) image. A mapping of this form is called a *thresholding* function.

A special case of Clipping, where $a = b = t$ is called thresholding and the output becomes binary. Thus it is **used to convert gray level image into binary**.



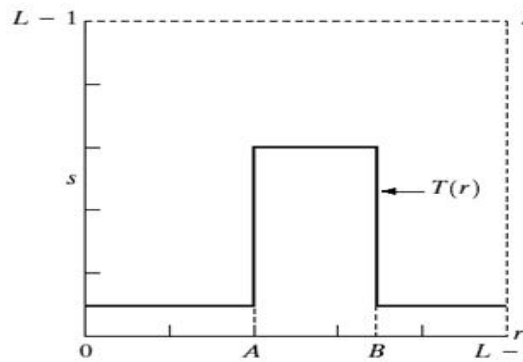
(d) Gray level slicing or Intensity level slicing or window slicing:

This transformation is **used to segment certain gray level regions from the rest of the image**. This technique is useful when different features of an image are contained in different gray levels.

Example: Enhancing features such as masses of water in satellite images and enhancing flaws in X-ray images.

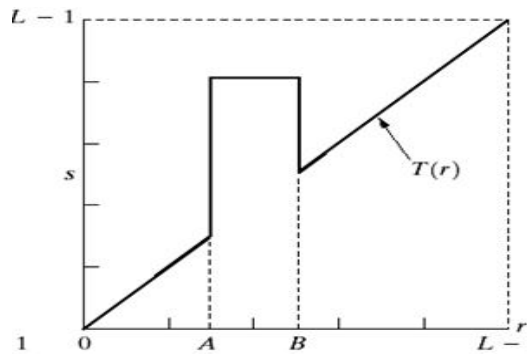
There are two ways of approaching this transformation

- To display a high value for all gray levels which is in the range of interest and a low value for all other gray levels. It fully illuminates the pixels lying in the interval $[A,B]$ and removes the background.
i.e., without background,



$$s = \begin{cases} L, & A \leq r \leq B \\ 0, & \text{else} \end{cases}$$

- b) To brighten the desired range of gray levels while preserving the background of the image.
i.e., with background,

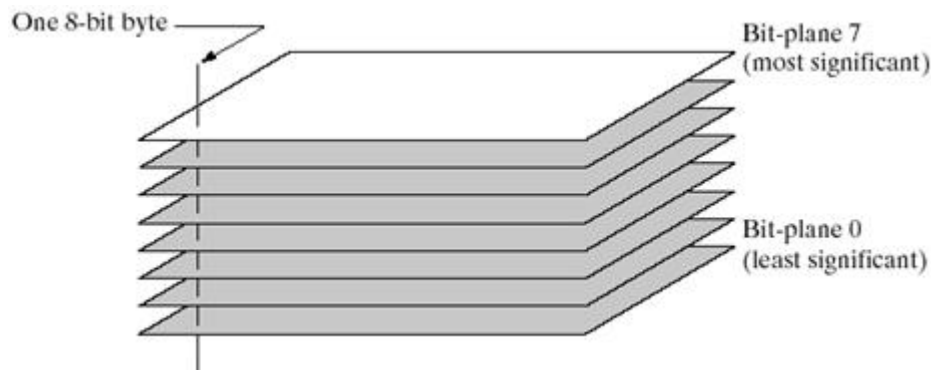


$$s = \begin{cases} L, & A \leq r \leq B \\ r, & \text{else} \end{cases}$$

(e) Bit plane slicing or Bit extraction:

This transformation is **used to determine the number of visually significant bits in an image.** It extracts the information of a single bit plane.

Consider each pixel in an image to be represented by 8 bits and imagine that the image is made up of eight 1-bit plane, ranging from bit plane 0 for LSB to bit plane 7 for MSB, as shown below



In terms of 8 bit bytes, plane0 contains all the lowest order bits in the bytes comprising the pixels in the image and plane7 contains all the high order bits.

Thus, the high-order bits contain the majority of the visually significant data while other bits do not convey an information about the image structure

Bit Extraction:

If each pixel of an image is uniformly quantized to B bits, then the nth MSB can be extracted and displayed as follows,

Let, $r = k_1 2^{B-1} + k_2 2^{B-2} + \dots + k_{B-1} 2 + k_B$

Then the output is given

$$s = \begin{cases} L, & \text{if } k_n = 1 \\ 0, & \text{else} \end{cases}$$

Where, $k_n = [i_n - 2 i_{n-1}]$ $i_n = \text{Int} [u / 2^{B-n}]$

$\text{Int}[x]$ = Integer part of x , $n = 1, 2, \dots, B$, B = no. of bits used to represent r , as an integer

Bit Removal:

For MSB removal $s = 2r \text{ modulo } (L+1)$, $0 \leq r \leq L$

For LSB removal, $s = 2 \text{ Int} [r/2]$

2. Explain briefly about histogram modeling

HISTOGRAM

Histogram of an image **represents the relative frequency of occurrence of the various gray levels in the image.**

Histogram modeling techniques modify an image so that its histogram has a desired shape. This is useful in stretching the low contrast levels of images with narrow histograms. Thus histogram is **used to either increase or decrease**, the intensity level of the pixels by any kind of transformation.

The histogram of a digital image with gray levels in the image $[0, L-1]$ is a discrete function given as,

$$h(r_k) = n_k, \text{ where } r_k \text{ is the } k\text{th grey level, } k = 0, 1, \dots, L-1$$

n_k is the number of pixels in the image with grey level r_k

Normalized histogram is given as,

$$p(r_k) = \frac{n_k}{n}$$

Where, n is the total number of pixels in the image and $n \geq n_k$

The **Normalized histogram** is the histogram **divided by the total number of pixels in the source image.**

1. The sum of all values in the normalized histogram is 1.

2. The value given by the normalized histogram for a certain gray value can be read as the probability of randomly picking a pixel having that gray value

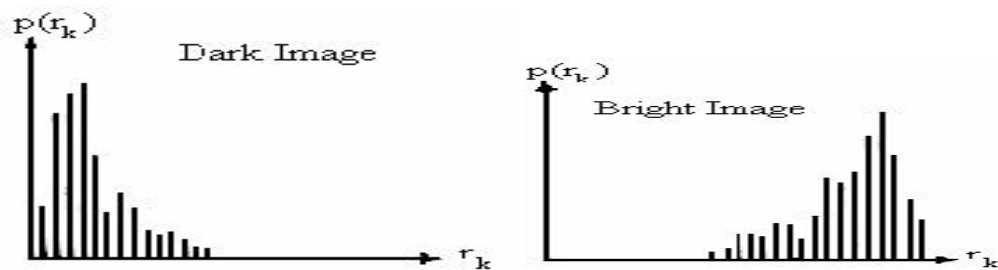
3. The horizontal axis of each histogram plot corresponds to grey level values r_k in the image and the vertical axis corresponds to $h(r_k) = n_k$ or $p(r_k) = \frac{n_k}{n}$, if the values are

normalized.

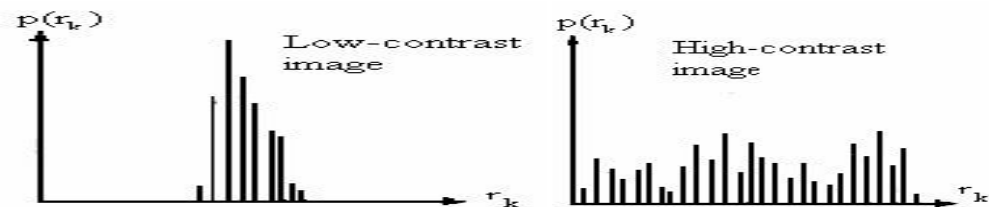
Histogram of **four basic gray level characteristics:**

- Dark image: (under exposed image): Here components of the histogram are concentrated on the low or dark side of the gray scale.
- Bright image: (over exposed image): Here components of the histogram are

concentrated on the high or bright side of the gray scale.



- Low-contrast image: Here components of the histogram at the middle of the gray scale and are narrow.
- High-contrast image: Here components of the histogram cover a broad range of the gray scale and they are almost uniformly distributed.



3. Write short notes on Histogram equalization and Histogram modification. (Nov / Dec 2005) (May / June 2009) (May / June 2007)

Histogram equalization: A technique which is used to obtain uniform histogram is known as histogram equalization or histogram linearization.

Let r represent the grey levels in the image to be enhanced. Assume r to be normalized in the interval $[0, 1]$, with $r = 0$ representing black and $r = 1$ representing white.

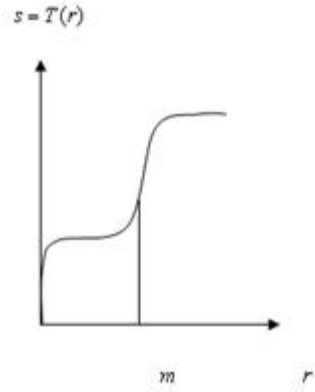
For any value r in the interval $[0, 1]$, the image transformation is given as,

$$S = T(r), 0 \leq r \leq 1$$

This transformation produces a level s for every pixel value r in the original image. The transformation function $T(r)$ satisfies the following conditions,

- $T(r)$ is single-valued and monotonically increases in the interval $0 \leq r \leq 1$ (ie., preserves the order from black to white in the output image) and
- $0 \leq T(r) \leq 1$ for $0 \leq r \leq 1$ (ie., guarantees that the output gray levels will be in the same range as the input levels)

The following shows the transformation function satisfying the above two conditions.



The Inverse transformation is given as,

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

If the gray levels in an image can be viewed as random variables in the interval $[0,1]$, then $P_r(r)$ and $P_s(s)$ denote the pdfs of the random variables r and s .

If $P_r(r)$ and $T(r)$ are known and $T^{-1}(s)$ satisfies the condition (a) , then

$$P_s(s) = \left[P_r(r) \left| \frac{dr}{ds} \right| \right]_{r = T^{-1}(s)} \quad \text{-----(1)}$$

Thus , the pdf of the transformed variables is determined by the gray level pdf of the input image and by the chosen transformation function.

For Continuous gray level values:

Consider the transformation,

$$s = T(r) = \int_0^r p_r(w)dw \quad \text{-----(2)}$$

where w is a dummy variable of integration and

$\int_0^r p_r(w)dw =$ cumulative distribution function [CDF] of random variable ' r ' con

dition (a) and (b) are satisfied by this transformation as CDF increases monotonically from 0 to 1 as a function of r .

When $T(r)$ is given, we can find $P_s(s)$ using equation (1),

We know that,

$$s = T(r)$$

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d}{dr} \int_0^r p_r(w)dw = p_r(r)$$

$$\frac{dr}{ds} = \frac{1}{\Pr(r)} \text{-----(3)}$$

Substituting (3) in (1), we get,

$$\begin{aligned} p_s(s) &= \left[p_r(r) \left| \frac{dr}{ds} \right| \right], \quad 0 \leq s \leq 1 \\ &= \left[p_r(r) \left| \frac{1}{p_r(r)} \right| \right] \\ &= 1, \quad 0 \leq s \leq 1 \end{aligned}$$

This indicates that $P_s(s)$ is a uniform pdf.

For Discrete gray level values:

The probability of occurrence of gray level r_k in an image is given as,

$$p_r(r_k) = \frac{n_k}{n}, k = 0, 1, \dots, L-1$$

Where, n = total no. of pixels in the image

n_k = No. of pixels that have gray level r_k

L = total no. of possible gray levels in the image

The discrete form of the transformation function is ,

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, \quad 0 \leq r_k \leq 1, \quad k = 0, 1, \dots, L-1$$

Thus the processed output image is obtained by mapping each pixel with level r_k in the input image into a corresponding pixel with level s_k in the output image.

- ♦ The inverse transformation is given as ,

$$r_k = T^{-1}(s_k), k = 0, 1, 2, \dots, L-1$$

Histogram Modification: A technique which is used to modify the histogram is known as histogram modification.

Here the input gray level r is first transformed non-linearly by $T(r)$ and the output is uniformly quantized.

In histogram equalization , the function $T(r) = \sum_{j=0}^k p_r(r_j), \quad k = 0, 1, \dots, L-1$

The above function performs a compression of the input variable.

Other choices of $T(r)$ that have similar behavior are,

$$T(r) = \log(1+r), \quad r \geq 0$$

$$T(r) = r^{1/n}, \quad r \geq 0, \quad n = 2, 3, \dots$$

4. Write short notes on histogram specification. (May / June 2009)

Histogram Specification:

It is defined as a method which is **used to generate a processed image that has a specified histogram**. Thus, it is used to specify the shape of the histogram, for the processed image & the shape is defined by the user. Also called **histogram matching**

For Continuous case:

Let r be an input image, Z be an output image

$p_r(r)$ is the original probability density function

$p_z(z)$ is the desired probability density function

Let S be a random variable and if the histogram equalization is first applied on the original image r , then

$$s = T(r) = \int_0^r p_r(w)dw$$

Suppose that the desired image z is available and histogram equalization is given as,

$$s = G(z) = \int_0^z p_z(z)dz$$

Where G is the transformation of $z = s$.

From equation (1) and (2),

$$G(z) = s = T(r)$$

The inverse process $z = G^{-1}(s) = G^{-1}[T(r)]$

Therefore, the process of histogram specification can be summarised in the following steps.

- i) Obtain the transformation function $T(r)$ using equation (1)
- ii) Obtain the transformation function $G(z)$ using equation (2)
- iii) Obtain the Inverse transformation function G^{-1}
- iv) Obtain the output image by applying using equation (3) to all pixels in the input image

This procedure yields an image whose gray levels z have the specified pdf $p_z(z)$

For Discrete gray level values:

The probability of occurrence of gray level r_k in an image is given as,

$$p_r(r_k) = \frac{n_k}{n}, k = 0, 1, \dots, L-1$$

Where, n = total no. of pixels in the image

n_k = No. of pixels that have gray level r_k

L = total no. of possible gray levels in the image

The discrete form of the transformation function is ,

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, 0 \leq r_k \leq 1, k = 0, 1, \dots, L-1$$

Also, the discrete form of the desired transformation function is ,

$$s_k = G(z_k) = \sum_{i=0}^k p_z(z_i),$$

From (1) and (2), the inverse transformation is given as ,

$$G(z_k) = s_k = T(r_k)$$

$$z_k = G^{-1}(s_k) = G^{-1}[T(r_k)]$$

5. Write short notes on the various Noise models (Noise distribution) in detail. (Nov / Dec 2005)

The principal noise sources in the digital image arise during image acquisition or digitization or transmission. The performance of imaging sensors is affected by a variety of factors such as environmental conditions during image acquisition and by the quality of the sensing elements.

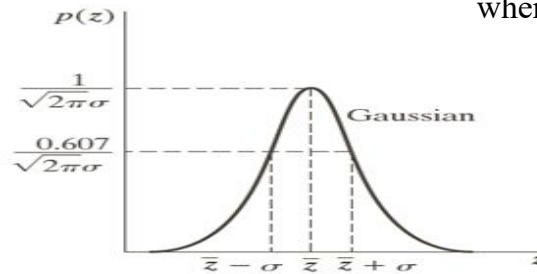
Assuming the noise is independent of spatial coordinates and it is uncorrelated with respect to the image itself.

The various Noise probability density functions are as follows,

1. Gaussian Noise: It occurs due to electronic circuit noise, sensor noise due to poor illumination and/or high temperature

The PDF of Gaussian random variable, z , is given by

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



where, z represents intensity

\bar{z} is the mean (average) value of z

σ is the standard deviation

The PDF of Gaussian random variable, z , is given by

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

70% of its values will be in the range

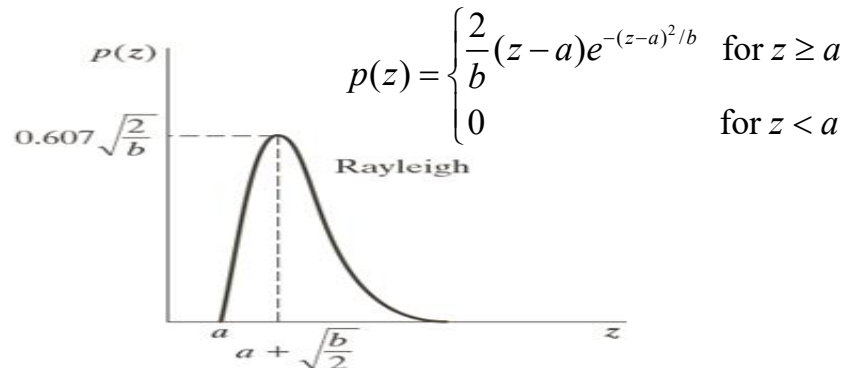
95% of its values will be in the range

$$\begin{aligned} &[(\mu - \sigma), (\mu + \sigma)] \\ &[(\mu - 2\sigma), (\mu + 2\sigma)] \end{aligned}$$

2. Rayleigh Noise

It occurs due to Range imaging and is used for approximating skewed histograms.

The PDF of Rayleigh noise is given by



The mean and variance of this density are given by

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

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

3. Erlang (Gamma) Noise

It occurs due to laser imaging

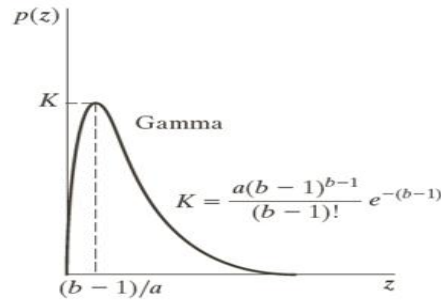
The PDF of Erlang noise is given by

The mean and variance of this density are given by

$$\bar{z} = b / a$$

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

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



Where $a > 0$, b is a positive integer and “!” indicates factorial.

4. Exponential Noise

It occurs due to laser imaging

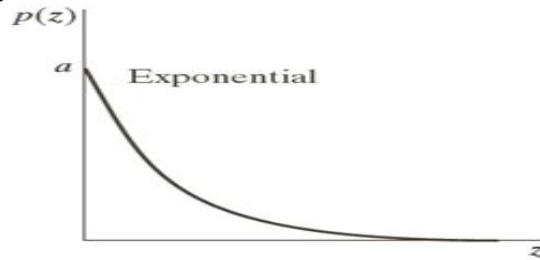
The PDF of exponential noise is given by

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

The mean and variance of this density are given by

$$\bar{z} = 1 / a$$

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



5. Uniform Noise

It is a Least descriptive and basis for numerous random number generators

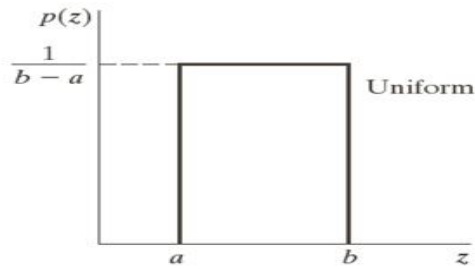
The PDF of uniform noise is given by

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

The mean and variance of this density are given by

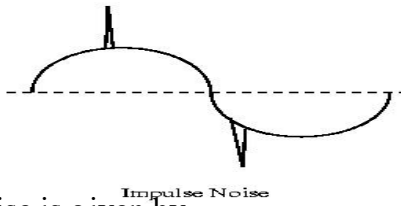
$$\bar{z} = (a + b) / 2$$

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



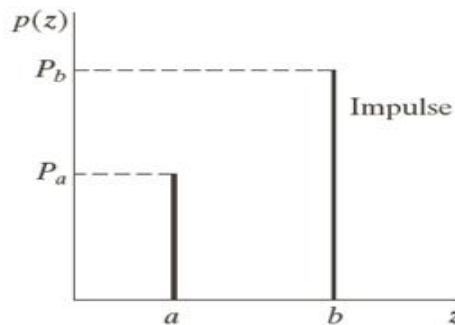
6. Impulse (Salt-and-Pepper) Noise

It occurs due to quick transients such as faulty switching



The PDF of (bipolar) impulse noise is given by

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



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

If either P_a or P_b is zero, then the impulse noise is called Unipolar

When $a = \text{minimum gray level value} = 0$, denotes negative impulse appearing black (pepper) points in an image and when $b = \text{maximum gray level value} = 255$, denotes positive impulse appearing white (salt) noise, hence it is also called **as salt and pepper noise or shot and spike noise**

7. Periodic Noise

Periodic noise in an image arises typically from electrical or electromechanical interference during image acquisition. It is a type of spatially dependent noise

Periodic noise can be reduced significantly via frequency domain filtering

6. Explain the detail of Spatial Filters.

Some neighborhood operations work with the values of the image pixels in the neighborhood and the corresponding values of a sub image that has the same dimensions as the neighborhood. **The sub image is called a filter, mask, kernel, template, or window** The values in a filter sub image are referred to as coefficients, rather than pixels.

The concept of filtering has its roots in the use of the Fourier transform for signal processing in the so-called frequency domain.

The process consists simply of moving the filter mask from point to point in an image. At each point (x,y), the response of the filter at that point is calculated using a predefined relationship.

Linear Spatial Filter:

Linear filter means that the transfer function and the impulse or point spread function of a linear system are inverse Fourier transforms of each other.

For a linear spatial filtering, 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.

Consider a 3 x 3 mask, the response R of linear spatial filtering with the filter mask at a point (x,y) in the image is,

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

Thus the coefficients w(0,0) coincides with image value f(x,y), indicating that the mask is centered at (x,y), when the computation of the sum of products takes place.

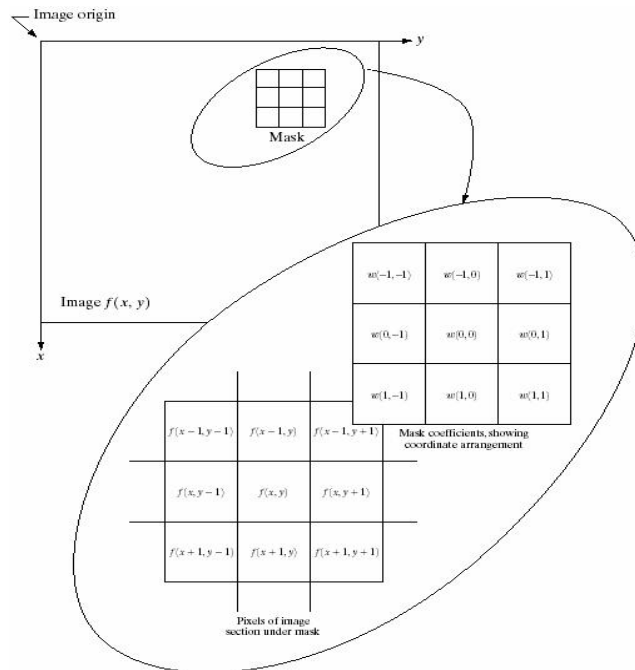
Hence the filter mask size is m x n, where m = 2a+1 and n = 2b+1 and a,b are non-negative integers. In general, the linear filtering of an image f of size M x N with a filter mask of size m x n is given as,

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t), \text{ where } \dots a = \frac{(m-1)}{2}, b = \frac{(n-1)}{2}$$

Where,

$$x = 0, 1, 2, \dots, M-1$$

$$y = 0, 1, 2, \dots, N-1$$



From the general expression, it is seen that the linear spatial filtering is just “convolving a mask with an image”. Also the filter mask is called as Convolution mask or Convolution kernel.

The following shows the 3 x 3 spatial filter mask,

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

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

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

In general

$$R = w_1 z_1 + w_2 z_2 + \dots + w_{mn} z_{mn}$$

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

Where, w 's are mask coefficients, z 's are values of the image gray levels corresponding to the mask coefficients and mn is the total number of coefficients in the mask.

Non-Linear Spatial filter:

This filter is based directly on the values of the pixels in the neighborhood and they don't use the coefficients. These are used to reduce noise.

7.Explain the various smoothing filters in the spatial domain.

Smoothing Spatial Filters or Low pass spatial filtering:

1.Smoothing filters are used for blurring and for noise reduction

2.Blurring is used in preprocessing steps, such as removal of small details from an image prior to object extraction and bridging of small gaps in lines or curves.It includes both linear and non-linear filters

1.Neighborhood Averaging:

1.The response of a smoothing linear spatial filter is simply the average of the pixels in the neighborhood of the filter mask,i.e., each pixel in an image is replaced by the average of the gray levels in the neighborhood defined by the filter mask. Hence this filter is called Neighborhood Averaging or low pass filtering.

2.Thus, spatial averaging is used for noise smoothening, low pass filtering and sub sampling of images.

3.This process reduces sharp transition in gray level

4.Thus the averaging filters are used to reduce irrelevant detail in an image.

Consider a 3 x 3 smoothing or averaging filter

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

The response will be the sum of gray levels of nine pixels which could cause R to be out of the valid gray level range.

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

Thus the solution is to scale the sum by dividing R by 9.

$$\begin{aligned} R &= w_1 z_1 + w_2 z_2 + + w_9 z_9 \\ &= \sum_{i=1}^9 w_i z_i \\ &= \frac{1}{9} \sum_{i=1}^9 z_i, \text{ when } .. w_i = 1 \end{aligned}$$

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

This is nothing but the average of the gray levels of the pixels in the 3 x 3 neighborhood defined by the pixels.

Other spatial low pass filters of various sizes are ,

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

Thus a m x n mask will have a normalizing constant equal to 1/mn. **A spatial average filter in which all the coefficients are equal is called as a Box Filter.**

2. Weighted Average Filter: For filtering a M x N image with a weighted averaging filter of size m x m is given by,

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

Example:

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

Here the pixel at the centre of the mask is multiplied by a higher value than any other, thus giving this pixel more importance.

The diagonals are weighed less than the immediate neighbors of the center pixel. **This reduces the blurring.**

3.Uniform filtering

The most popular masks for low pass filtering are masks with all their coefficients positive and equal to each other as for example the mask shown below. Moreover, they sum up to 1 in order to maintain the mean of the image.

$$\frac{1}{9} \times$$

1	1	1
1	1	1
1	1	1

4.Gaussian filtering

The two dimensional Gaussian mask has values that attempts to approximate the continuous function

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}}$$

. The following shows a suitable integer-valued convolution kernel that approximates a Gaussian with a σ of 1.0.

$$\frac{1}{273} \times$$

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

5.Median filtering(Non linear filter)

Median filtering is the operation that replaces each pixel by the median of the grey level in the neighborhood of that pixel.

Process is replaces the value of a pixel by the median of the gray levels in region S_{xy} of that pixel:

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

$$V(m,n) = \text{median}\{y(m-k,n-l), (k,l) \in w\}$$

Where w is suitably chosen window. The window size N_w is chosen to be odd.

If N_w is even, then the median is taken as the average of the two values in the middle.

Typical window size s are $3 \times 3, 5 \times 5, 7 \times 7$ or 5 point window

Properties:

1. Unlike average filtering, **median filtering does not blur** too much image details.

Example: Consider the example of filtering the sequence below using a

3-pt median filter: 16 14 15 12 2 13 15 52 51 50 49

The output of the median filter is: 15 14 12 12 13 15 51 51 50

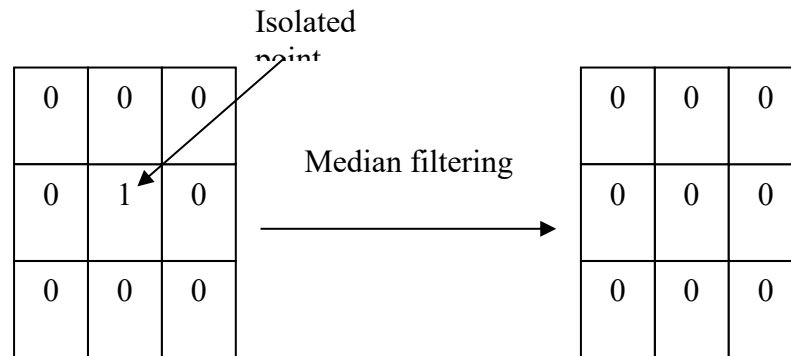
2. **Median filters are non linear filters** because for two sequences $x(n)$ and $y(n)$

$$\text{median}\{x(n) + y(n)\} \neq \text{median}\{x(n)\} + \text{median}\{y(n)\}$$

3. **Median filters are useful for removing isolated lines or points** (pixels) while preserving spatial resolutions.

4. **It works very well on images containing binary (salt and pepper) noise** but performs poorly when the noise is Gaussian.

5. Their performance is also **poor when the number of noise pixels in the window is greater than or half the number of pixels** in the window

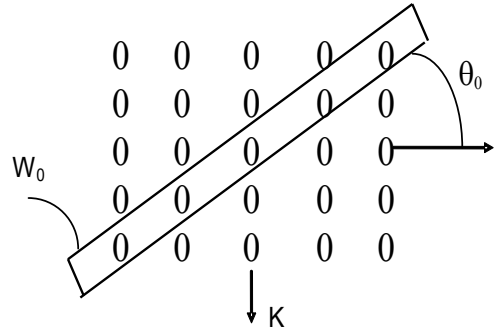


6. Directional smoothing

To protect the edges from blurring while smoothing, a directional averaging filter can be useful. Spatial averages $g(x, y; \theta)$ are calculated in several selected directions (for example could be horizontal, vertical, main diagonals)

$$g(x, y; \theta) = \frac{1}{N_\theta} \sum_{(k, l) \in W_\theta} f(x - k, y - l)$$

and a direction θ^* is found such that $|f(x, y) - g(x, y; \theta^*)|$ is minimum. (Note that W_θ is the neighbourhood along the direction θ and N_θ is the number of pixels within this neighbourhood). Then by replacing $g(x, y; \theta)$ with $g(x, y; \theta^*)$ we get the desired result.



7.Max Filter & Min Filter:

The 100th percentile filter is called **Maximum filter**, which is used to **find the brightest points** in an image.

$$R = \max \{z_k, k = 1, 2, 3, \dots, 9\}, \text{ for } 3 \times 3 \text{ matrix}$$

The 0th percentile filter is called **Minimum filter**, which is used to **find the darkest points** in an image.

$$R = \min \{z_k, k = 1, 2, 3, \dots, 9\} \text{ for } 3 \times 3 \text{ matrix}$$

(i.e) Using the 100th percentile results in the so-called max filter, given by

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

This filter is useful for finding the brightest points in an image. Since pepper noise has very low values, it is reduced by this filter as a result of the max selection processing the sub image area S_{xy} .

The 0th percentile filter is min filter:

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

This filter is useful for finding the darkest points in an image. Also, it reduces salt noise as a result of the min operation

8.Mean Filters

This is the simply methods to reduce noise in spatial domain.

1.Arithmetic mean filter

2.Geometric mean filter

3.Harmonic mean filter

4.Contraharmonic mean filter

Let S_{xy} represent the set of coordinates in a rectangular sub image window of size $m \times n$, centered at point (x, y) .

1.Arithmetic mean filter

Compute the average value of the corrupted image $g(x, y)$ in the area defined by $S_{x, y}$. The value of the restored image at any point (x, y)

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

Note: Using a convolution mask in which all coefficients have value $1/mn$. Noise is reduced as a result of blurring.

2. Geometric mean filter

Geometric mean filter is given by the expression

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

3. Harmonic mean filter

The harmonic mean filter operation is given by the expression

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

4. Contraharmonic mean filter

The contra harmonic mean filter operation is given by the expression

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

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

8. Explain the various sharpening filters in spatial domain. (May / June 2009)

sharpening filters in spatial domain:

This is used to highlight fine details in an image, to enhance detail that has been blurred and to sharpen the edges of objects

Smoothing of an image is similar to integration process,

Sharpening = 1 / smoothing

Since it is inversely proportional, sharpening is done by differentiation process.

Derivative Filters:

1. The derivatives of a digital function are defined in terms of differences.

2. It is used to highlight fine details of an image.

Properties of Derivative:

1. For contrast gray level, derivative should be equal to zero

2. It must be non-zero, at the onset of a gray level step or ramp

3. It must be non zero along a ramp [constantly changes]

1st order derivative:

The first order derivative of one dimensional function $f(x)$ is given as,

$$\frac{\partial f}{\partial x} \cong [f(x+1) - f(x)]$$

IInd order derivative:

The second order derivative of one dimensional function $f(x)$ is given as,

$$\frac{\partial^2 f}{\partial x^2} \cong [f(x+1) + f(x-1) - 2f(x)]$$

Conclusion:

- I. First order derivatives produce thicker edges in an image.
- II. Second order derivatives have a stronger response to fine details , such as thin lines and isolated points
- III. First order derivatives have a stronger response to a gray level step
- IV. Second order derivatives produces a double response at step changes in gray level

Laplacian Operator:

It is a linear operator and is the simplest isotropic derivative operator, which is defined as,

$$\text{Laplacian} \quad \nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial^2 x} + \frac{\partial^2 f(x, y)}{\partial^2 y} \quad \text{-----(1)}$$

Where,

$$\frac{\partial^2 f(x, y)}{\partial^2 y} \approx f(x, y+1) + f(x, y-1) - 2f(x, y) \quad \text{-----(2)-}$$

$$\frac{\partial^2 f(x, y)}{\partial^2 x} \approx f(x+1, y) + f(x-1, y) - 2f(x, y) \quad \text{-----(3)}$$

Sub (2) & (3) in (1), we get,

$$\nabla f^2 \cong [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)] \quad \text{-----(4)}$$

The above digital laplacian equation is implemented using filter mask,

0	1	0
1	-4	1
0	1	0

The digital laplacian equation including the diagonal terms is given as,

$$\nabla f^2 \cong \left[\begin{aligned} &f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) + f(x+1, y+1) + f(x+1, y-1) \\ &+ f(x-1, y-1) + f(x-1, y+1) - 8f(x, y) \end{aligned} \right]$$

the corresponding filter mask is given as,

1	1	1
1	-8	1
1	1	1

Similarly , two other implementations of the laplacian are,

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

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

The sharpened image using the laplacian operator is given as,

$$g(x, y) \cong \begin{aligned} &f(x, y) - \nabla^2 f(x, y), \\ &f(x, y) + \nabla^2 f(x, y), \end{aligned}$$

Substitute equation (4) in (6),

$$g(x, y) \cong f(x, y) - [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)]$$

$$g(x, y) \cong 5f(x, y) - [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)]$$

The above digital laplacian equation is implemented using filter mask,

0	-1	0
-1	5	-1
0	-1	0

This is called as composite laplacian mask.

Substitute equation (5) in (6),

$$g(x, y) \cong f(x, y) - \left[\begin{aligned} &f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) + f(x+1, y+1) + f(x+1, y-1) \\ &+ f(x-1, y-1) + f(x-1, y+1) - 8f(x, y) \end{aligned} \right]$$

$$g(x, y) \cong 9f(x, y) - \left[\begin{array}{l} f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) + f(x+1, y+1) + f(x+1, y-1) \\ + f(x-1, y-1) + f(x-1, y+1) \end{array} \right]$$

The above digital laplacian equation is implemented using filter mask,

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

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

This is called the second composite mask.

Gradient operator:

For a function $f(x, y)$, the gradient f at co-ordinate (x, y) is defined as the 2-dimensional column vector

$$\Delta f = \begin{bmatrix} G_x \\ G_y \end{bmatrix}$$

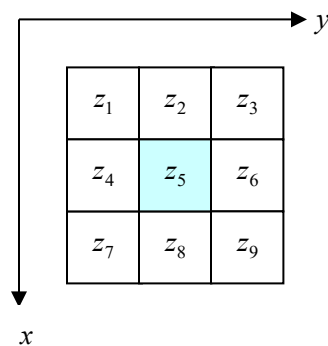
$$= \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix}$$

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

$$\cong |G_x| + |G_y|$$

1. Roberts operator

consider a 3×3 mask is the following.



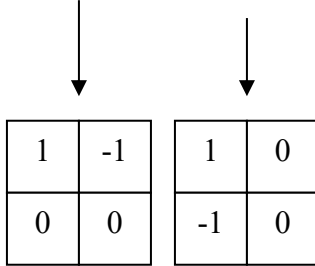
It can be approximated at point z_5 in a number of ways. The simplest is to use the difference $(z_8 - z_5)$ in the x direction and $(z_6 - z_5)$ in the y direction. This approximation is known as the **Roberts** operator, and is expressed mathematically as follows.

$$Gx = \left| \frac{\partial f(x, y)}{\partial x} \right| = f(x+1) - f(x) = |z_6 - z_5|$$

$$Gy = \left| \frac{\partial f(x, y)}{\partial y} \right| = f(y+1) - f(y) = |z_8 - z_5|$$

$$\nabla f \cong |Gx| + |Gy|$$

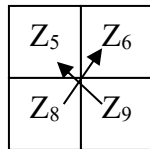
$$\cong |z_6 - z_5| + |z_8 - z_5|$$



Roberts operator

Another approach is to use **Roberts cross differences**

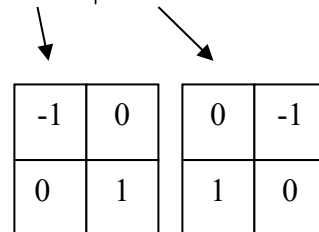
Take



This approximation is known as the **Roberts** cross gradient operator, and is expressed mathematically as follows

$$\nabla f \cong |z_9 - z_5| + |z_8 - z_6|$$

The above Equation can be implemented using the following mask

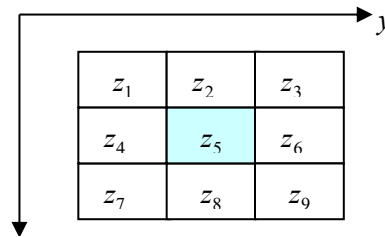


Roberts operator

The original image is convolved with both masks separately and the absolute values of the two outputs of the convolutions are added.

2. Prewitt operator

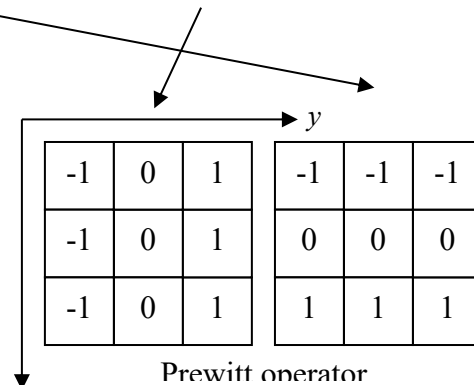
consider a 3×3 mask is the following.



$$Gx = \left| \frac{\partial f(x, y)}{\partial x} \right| = 3^{\text{rd}} \text{ row} - 1^{\text{st}} \text{ row} = |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)|$$

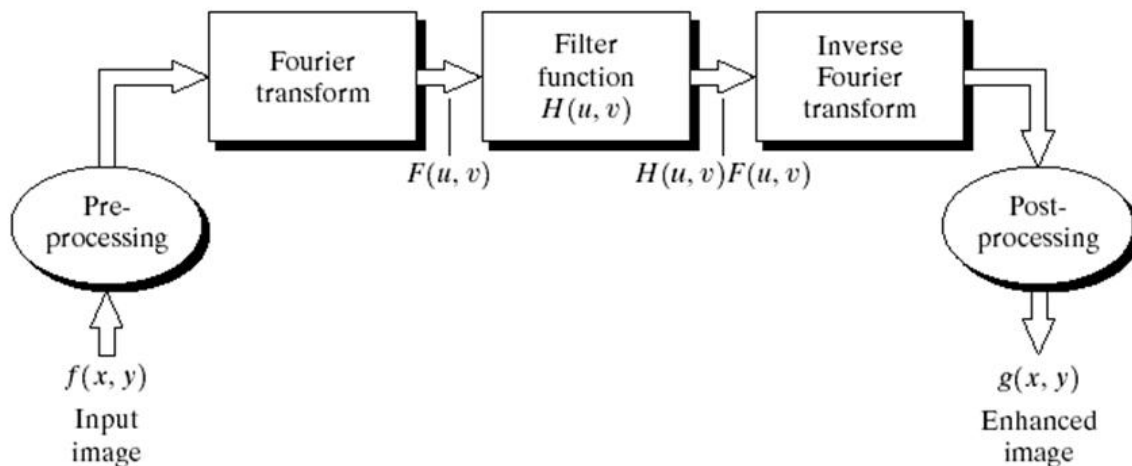
$$Gy = \left| \frac{\partial f(x, y)}{\partial y} \right| = 3^{\text{rd}} \text{ col} - 1^{\text{st}} \text{ col} = |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$

$$\nabla f \cong |Gx| + |Gy| = |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| + |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$



This approximation is known as the **Prewitt operator**.

9. Explain the basic steps for filtering (image enhancement) in Frequency domain.



1. Multiply the input image by $(-1)^{x+y}$ to center the transform, image dimensions $M \times N$
2. Compute $F(u, v)$ DFT of the given image, DC at $M/2, N/2$.
3. Multiply $F(u, v)$ by a filter function $H(u, v)$ to get the FT of the output image,

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

Here each component of H is multiplied with both the real and imaginary parts of the corresponding component in F , such filters are called **zero-phase shift filters**. DC for $H(u, v)$ at $M/2, N/2$.

4. Compute the inverse DFT of result in (c), so that the

$$\text{filtered image} = F^{-1}[G(u,v)]$$

5. Take real part of the result in (d)

6. The final enhanced image is obtained by Multiplying result in (e) by $(-1)^{x+y}$

10. Discuss in detail the significance of Homomorphic filtering in image enhancement. (May / June 2009) (May / June 2007)

Definition: The filter which controls **both high frequency and low frequency** components are called Homomorphic filtering.

Homomorphic filtering is a generalized technique for signal and image processing, involving a nonlinear mapping to a different domain in which linear filter techniques are applied, followed by mapping back to the original domain.

Features & Application:

1. Homomorphic filter is used **for image enhancement**.
2. It simultaneously **normalizes the brightness** across an image and **increases contrast**.
3. It is also used to remove **multiplicative noise**.

Images normally consist of light reflected from objects. The basic nature of the image $f(x,y)$ may be characterized by two components:

(1) The amount of source light incident on the scene being viewed, &

(2) The amount of light reflected by the objects in the scene.

These portions of light are called the ***illumination*** and ***reflectance*** components, and are denoted $i(x,y)$ and $r(x,y)$ respectively. The functions i and r combine multiplicatively to give the image function F :

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

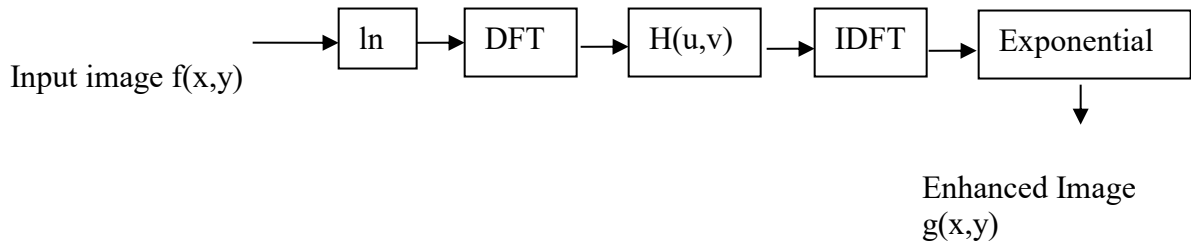
where $0 < i(x,y) < 1$ -----indicates perfect black body -----indicates perfect absorption , and $0 < r(x,y) < 1$ -----indicates perfect white body ----indicates perfect reflection.

Since i and r combine multiplicatively, they can be added by taking log of the image intensity, so that they can be separated in the frequency domain.

Illumination variations can be thought as a multiplicative noise and can be reduced by filtering in the log domain. To make the illuminations of an image more even, the HF components are increased and the LF components are filtered, because the HF components are assumed to represent the reflectance in the scene whereas the LF components are assumed to represent the illumination in the scene. i.e., High pass filter is used to suppress LF's and amplify HF's in the log intensity domain.

Illumination component tends to vary slowly across the image and the reflectance tends to vary rapidly. Therefore, by applying a frequency domain filter the intensity variation across the image can be reduced while highlighting detail. This approach is shown below.

BLOCK DIAGRAM:



Analysis:

$$\text{WKT, } f(x,y) = i(x,y)r(x,y) \text{ -----(1)}$$

Taking natural log on both sides of the above equation, we get,

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

$$\text{Let } z = \ln[f(x,y)] = \ln[i(x,y)] + \ln[r(x,y)]$$

Taking DFT on both sides of the above equation, we get,

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

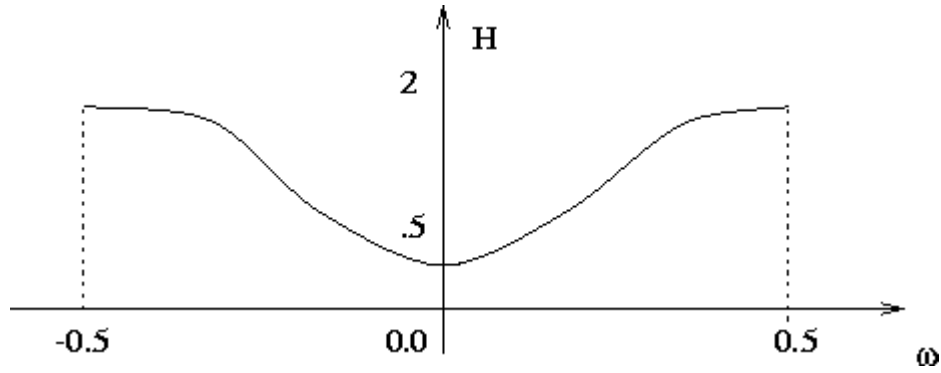
$$\begin{aligned} \text{DFT}[z(x,y)] &= \text{DFT}\{\ln[f(x,y)]\} = \text{DFT}\{\ln[i(x,y)] + \ln[r(x,y)]\} \\ &= \text{DFT}\{\ln[i(x,y)]\} + \text{DFT}\{\ln[r(x,y)]\} \text{-----(2)} \end{aligned}$$

Since $\text{DFT}[f(x,y)] = F(u,v)$, equation (2) becomes,

$$Z(u,v) = F_i(u,v) + F_r(u,v) \text{ -----(3)}$$

The function Z represents the Fourier transform of the *sum* of two images: a low frequency illumination image and a high frequency reflectance image.

Figure : Transfer function for homomorphic filtering.



If we now apply a filter with a transfer function that suppresses low frequency components and enhances high frequency components, then we can suppress the illumination component and enhance the reflectance component.

Thus, the Fourier transform of the output image is obtained by multiplying the DFT of the input image with the filter function $H(u,v)$.

$$\text{i.e., } S(u,v) = H(u,v) Z(u,v) \text{-----(4)}$$

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

Substitute equation (3) in (4), we get,

$$S(u,v) = H(u,v) [F_i(u,v) + F_r(u,v)] = H(u,v) F_i(u,v) + H(u,v) F_r(u,v) \text{--(5)}$$

Applying IDFT to equation (6), we get,

$$\begin{aligned} T^{-1}[S(u,v)] &= T^{-1} [H(u,v) F_i(u,v) + H(u,v) F_r(u,v)] \\ &= T^{-1}[H(u,v) F_i(u,v)] + T^{-1}[H(u,v) F_r(u,v)] \end{aligned}$$

$$\text{➤ } s(x,y) = i'(x,y) + r'(x,y) \text{-----(6)}$$

The Enhanced image is obtained by taking exponential of the IDFT $s(x,y)$,

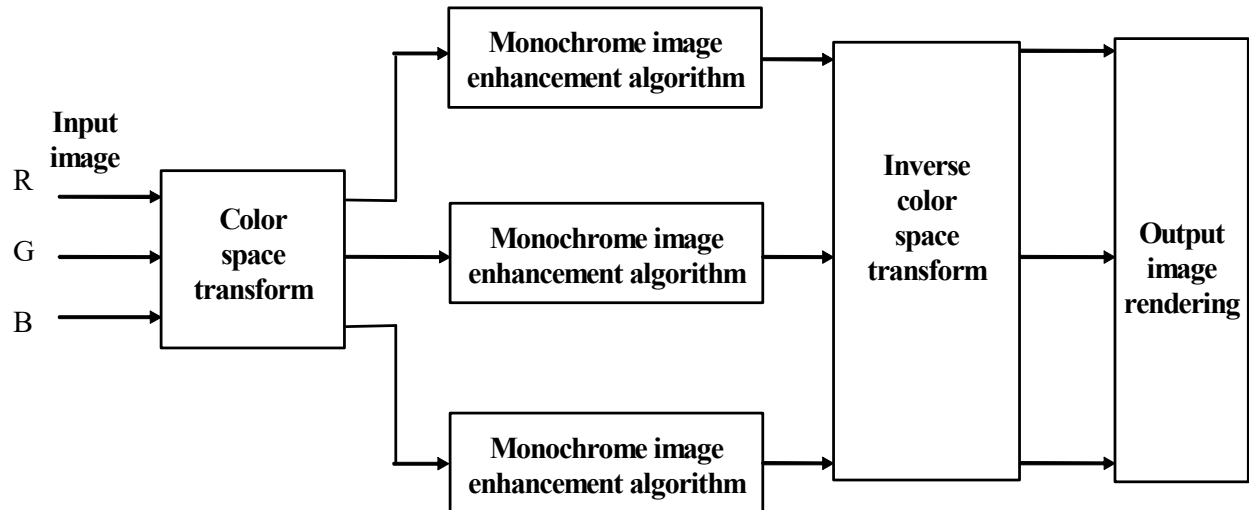
$$\text{i.e., } g(x,y) = e^{s(x,y)} = e^{i'(x,y)} e^{r'(x,y)} = i_o(x,y) r_o(x,y)$$

where, $i_o(x,y) = e^{i'(x,y)}$, $r_o(x,y) = e^{r'(x,y)}$ are the illumination and reflection components of the enhanced output image.

11. Write short notes on Color image enhancement

Color image enhancement:

An image enhancement technique that consists of changing the colors of an image or assigning colors to various parts of an image.



1. Color code regions of an image based on frequency content
2. The Fourier transform of an image is modified independently by three filters to produce three images used as Fourier transform of the R, G, B components of a color image
3. Additional processing can be any image enhancement algorithm like histogram equalization
4. Take inverse color transformation to obtain R' , G' , B' components

UNIT III -IMAGE RESTORATION

Image Restoration - degradation model, unconstrained restoration - Lagrange multiplier and constrained restoration, Inverse filtering-removal of blur caused by uniform linear motion, Wiener filtering, Geometric transformations-spatial transformations.

PART – A

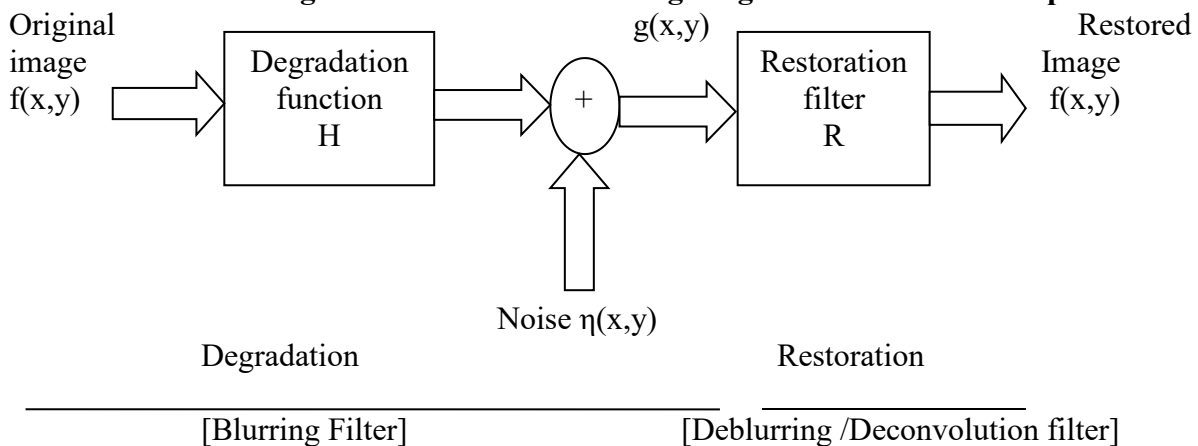
1. Define Image Restoration.

Image restoration refers to removal or minimization of known degradations in an image i.e. Restoration attempts to reconstruct or recover an image that has been degraded.

2. Compare Image Enhancement with image restoration?

Image Enhancement	Image Restoration
Improve visual appearance or increase contrast i.e. highlight certain features in an image or bring out details parts	Image restoration refers to removal or minimization of known degradations in an image i.e. Restoration attempts to reconstruct or recover an image that has been degraded.

3. Show the block diagram of a model of the image degradation /restoration process.



4. Classify Image restoration techniques.

Image restorations are classified as following:

Image Restoration		
Restoration models	Linear Filtering	Other Method
<ul style="list-style-type: none">Image formation modelsDetector / RecorderNoise models	<ul style="list-style-type: none">Inverse FilterWeiner FilterFIR filterLeast square, SVDRecursive (Kalman) filter	<ul style="list-style-type: none">Blind deconvolution

5. Define blind deconvolution.

The process of restoring an image by using a degradation function that has been estimated in some way sometimes is called blind deconvolution.

6. Define inverse filtering.

Inverse filtering is the process of recovering the input of a system from its output. Inverse filters are useful for precorrecting an input signal in anticipation of the degradations caused by the system, such as correcting the nonlinearity of a display.

7. Define Pseudo inverse filter.

The Pseudo-inverse filter is a stabilized version of the inverse filter. For a linear shift invariant system with frequency response $H(\omega_1, \omega_2)$, the pseudo inverse filter.

$$H^-(\omega_1, \omega_2) = \begin{cases} \frac{1}{H(\omega_1, \omega_2)} & , H \neq 0 \\ 0 & H = 0 \end{cases}$$

8. Give the limitations for inverse filtering and pseudo filtering.

The main limitation of inverse and pseudo-inverse filtering is that these filter remain very sensitive to noise.

9. What is meant by Weiner filter?

Weiner filter is a method of restoring images in the presence of blur as well as noise.

10. What is the expression for the error measure for Weiner filter?

The error measure is given by

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

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

11. What is the main objective of Weiner filtering?

The main objective of Weiner filtering is to find an estimate \hat{f} of the uncorrupted image f such that the mean square error between them is minimized.

12. Give the expression for Weiner filtering and give its other names for Weiner filter.

The expression for Weiner filtering is given by

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

The filter which consists of terms inside the brackets also is commonly referred to as minimum mean square error filter (or) least square error filter.

13. What is meant geometric transformation?

Geometric transformations modify the spatial relationships between pixels in an image. Geometric transformation often are called *rubber-sheet transformations*

14. What are the two basic operations of geometric transformation?

The two basic operations of geometric transformation are,

- i) A special transformation which defines the rearrangement of pixels on the image plane.
- ii) Gray level interpolation, which deals with the assignment of gray levels to pixels in the spatially transformed image.

15. What is meant spatial transformation?

An image “f” with pixel co=ordinates (x,y) undergoes geometric distortion to produce an image “g” with co-ordinates (x', y'). this transformation may be expressed as

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

and

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

where r(x,y) & s(x,y) are the spatial transformations.

16. Write an equation for Fredholm Integral of first kind and also define PSF

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

But the impulse response of H is given as, $h(x, \alpha, y, \beta) = H[\delta(x - \alpha, y - \beta)]$

Where $h(x, \alpha, y, \beta)$ is called Ponit spread function (PSF).

Therefore equation becomes,

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

This equation is called as Superposition or **Fredholm Integral of first kind**.

17. Write the draw backs of inverse filtering

1. Doesn't perform well when used on noisy images
2. In the presence of noise, when $H(u,v)$ is small, then the noise $\eta(u,v)$ dominates.

18. What is meant by constrained least squares restoration?. (May / June 09).

In Wiener filter, the power spectra of the undegraded image and noise must be known. Constrained least squares filtering just require the mean and variance of the noise.

19. Define Lagrange Multipliers

Lagrange multiplier is a method to find maximum or minimum value of a function subject to some constrains (side conditions) on the domain of the function.

Examples:

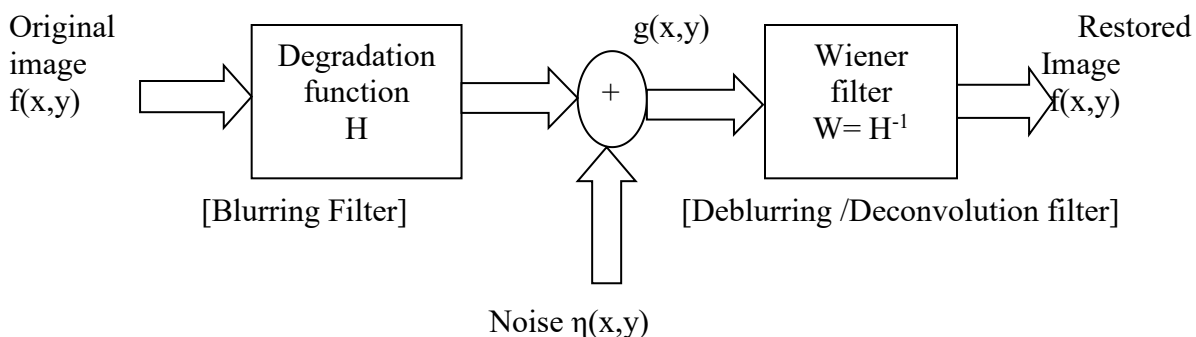
Maximize or minimize $f(x,y)$ subject to **condition** $g(x,y) = c$.

Maximize or minimize $f(x,y,z)$ subject to **condition** $g(x,y,z) = c$.

Maximize or minimize $f(x,y,z)$ subject to **conditions** $g(x,y,z) = c_1$, and $h(x,y,z) = c_2$.

20. Show the block diagram of a Wiener Model with Noise

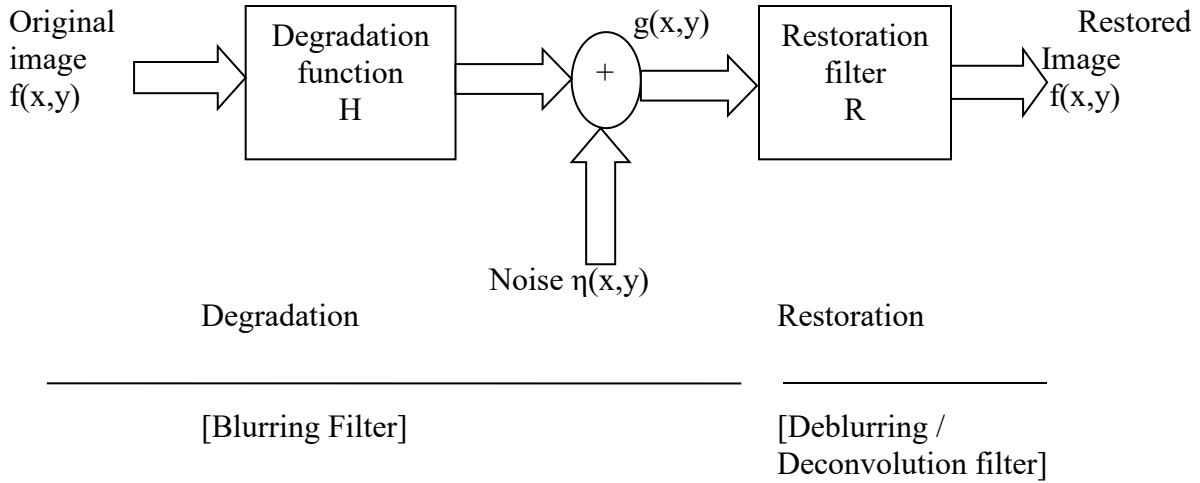
Wiener Model with Noise:



PART B

1. Explain the image degradation model and its properties. (May / June 2009) (or) Explain the mathematical model of Image degradation process

A simplified version for the image restoration / degradation process model is



The image $f(x,y)$ is degraded with a function H and also a noise signal $\eta(x,y)$ is added. The restoration filter removes the degradation from the degraded image $g(x,y)$ and restores the original image $f(x,y)$.

The Degradation model is given as,

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

Taking Fourier transform,

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

The Restoration model is given as,

$$\hat{f}(x, y) = r(x, y) * g(x, y)$$

Taking Fourier transform,

$$\hat{F}(u, v) = R(u, v).G(u, v) = H^{-1}(u, v).G(u, v) = \frac{G(u, v)}{H(u, v)}$$

where $g(x, y)$ is the degraded image, $f(x, y)$ is the original image

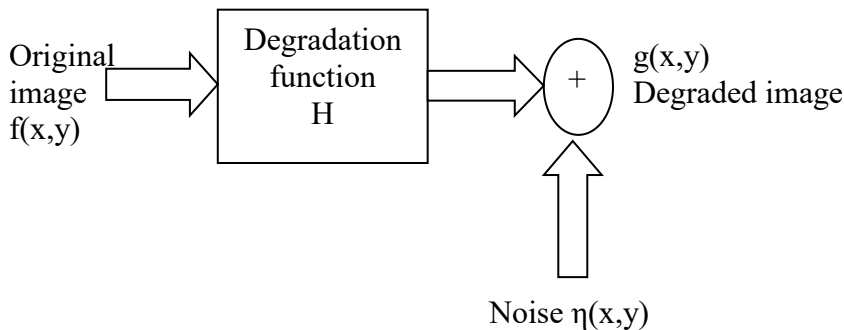
H is an operator that represents the degradation process

$\eta(x, y)$ is the external noise which is assumed to be image-independent

$\hat{f}(x, y)$ is the restored image. R is an operator that represents the restoration process

Properties of the Degradation model:

A simplified version for the image degradation model is



Consider the general degradation model

$$g(x, y) = H[f(x, y)] + \eta(x, y) \text{ -----(1)}$$

a) Property of Linearity:

If the external noise $\eta(x, y)$ is assumed to be zero, then equation (1) becomes,

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

H is said to be **linear** if

$$H[af_1(x, y) + bf_2(x, y)] = aH[f_1(x, y)] + bH[f_2(x, y)] \text{ -----(2)}$$

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

The linear operator possesses both additivity and homogeneity property.

b) Property of Additivity:

If $a = b = 1$, then equation (2) becomes,

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

Is called the property of additivity. It states that, if H is a linear operator, then the response to a sum of two inputs is equal to the sum of two responses.

c) Property of Homogeneity:

If $f_2(x, y) = 0$, then equation (2) becomes,

$H[af_1(x, y)] = aH[f_1(x, y)]$ is called the property of homogeneity. It states that the response to a constant multiple of any input is equal to the response of that input multiplied by the same constant.

d) Property of Position or space invariant:

An operator having the input – output relationship as $g(x, y) = H[f(x, y)]$ is said to be **position** or **space invariant** if

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

For any $f(x, y)$ and any α and β . It states that the response at any point in the image depends only on the value of the input at that point and not on its position.

2. Explain the degradation model for continuous function and for discrete function.
(May / June 2009)

a. degradation model for continuous function

An image in terms of a continuous impulse function is given as,

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta \text{ -----(1)}$$

Consider the general degradation model

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

If the external noise $\eta(x, y)$ is assumed to be zero, then equation becomes,

$$g(x, y) = H[f(x, y)] \text{ -----(2)}$$

Substituting (1) in (2), we get,

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

Since H is a linear operator, extending the additive property to integrals, we get,

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

Since $f(\alpha, \beta)$ is independent of x & y , using homogeneity property, we get,

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta) H[\delta(x - \alpha, y - \beta)] d\alpha d\beta \text{ -----(3)}$$

But the impulse response of H is given as, $h(x, \alpha, y, \beta) = H[\delta(x - \alpha, y - \beta)]$

Where $h(x, \alpha, y, \beta)$ is called Ponit spread function (PSF).

Therefore equation (3) becomes,

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta) h(x, \alpha, y, \beta) d\alpha d\beta \text{ -----(4)}$$

This equation is called as Superposition or **Fredholm Integral of first kind**.

It states that if the response of H to an impulse is known, then the response to any input $f(\alpha, \beta)$ can be calculated using the above equation.

Thus, a linear system H is completely characterized by its impulse response.

If H is *position* or *space invariant*, that is,

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

Then equation (4) becomes,

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

This is called the Convolution integral, which states that if the impulse of a linear system is known, then the response g can be calculated for any input function.

In the presence of additive noise, equation (4) becomes,

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

This is the linear degradation model expression for a continuous image function.

If H is position Invariant then,

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

b. degradation model for discrete function

Let two functions $f(x)$ and $h(x)$ be sampled uniformly to form arrays of dimension A and B .

$$\begin{aligned} x &= 0, 1, \dots, A-1, \text{ for } f(x) \\ &= 0, 1, \dots, B-1, \text{ for } h(x) \end{aligned}$$

Assuming sampled functions are periodic with period M , where $M \geq A + B - 1$ is chosen to avoid overlapping in the individual periods.

Let $f_e(x)$ and $h_e(x)$ be the extended functions, then for a time invariant degradation process the discrete convolution formulation is given as,

$$g_e(x) = \sum_{m=0}^{M-1} f_e(m) h_e(x - m)$$

In matrix form

$$\mathbf{g} = \mathbf{H}\mathbf{f},$$

where \mathbf{g} and \mathbf{h} are M dimensional column vectors

$$\mathbf{f} = \begin{bmatrix} f_e(0) \\ f_e(1) \\ \vdots \\ f_e(M-1) \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_e(0) \\ g_e(1) \\ \vdots \\ g_e(M-1) \end{bmatrix}, \text{ and } \mathbf{H} \text{ is a } M \times M \text{ matrix given as,}$$

$$\mathbf{H}_{(M \times M)} = \begin{bmatrix} h_e(0) & h_e(-1) & \dots & h_e(-M+1) \\ h_e(1) & h_e(0) & \dots & h_e(-M+2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(M-1) & h_e(M-2) & \dots & h_e(0) \end{bmatrix}$$

Assuming $h_e(x)$ is periodic, then $h_e(x) = h_e(M + x)$ and hence,

$$\mathbf{H}_{(M \times M)} = \begin{bmatrix} h_e(0) & h_e(M-1) & \dots & h_e(1) \\ h_e(1) & h_e(0) & \dots & h_e(2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(M-1) & h_e(M-2) & \dots & h_e(0) \end{bmatrix}$$

Considering the extension function $f_e(x)$ and $h_e(x)$ to be periodic in 2D with periods M and N , then, for a space invariant degradation process we obtain

$$g_e(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m, n) h_e(x - m, y - n) + n_e(x, y)$$

In matrix form,

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\eta}$$

where \mathbf{f} , \mathbf{g} and $\boldsymbol{\eta}$ are MN – dimensional column vectors and \mathbf{H} is $MN \times MN$ dimension. It consists of M^2 partition, each partition having dimension of size $M \times N$ and ordered accordingly and is given as,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{M-1} & \dots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \dots & \mathbf{H}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \dots & \mathbf{H}_0 \end{bmatrix}$$

Each partition H_j is constructed from the j th row of the extended function $h_e(x, y)$ is,

$$\mathbf{H}_j = \begin{bmatrix} h_e(j,0) & h_e(j,N-1) & \dots & h_e(j,1) \\ h_e(j,1) & h_e(j,0) & \dots & h_e(j,2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(j,N-1) & h_e(j,N-2) & \dots & h_e(j,0) \end{bmatrix}$$

3. Write notes on inverse filtering as applied to image restoration. (May / June 2009)
(or) Write short notes on Inverse filter (Nov / Dec 2005)

Inverse filter:

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

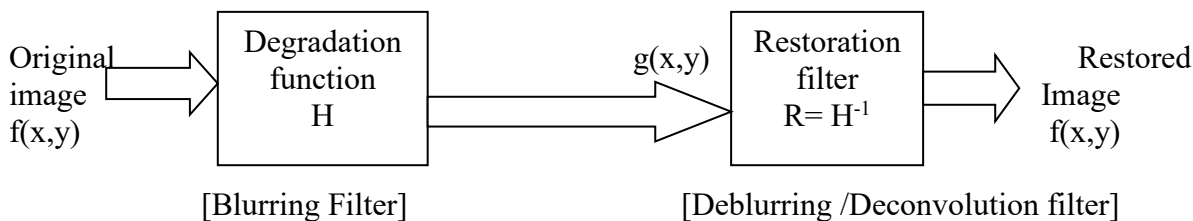
$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}$$

Problem statement:

Given a known image function $f(x, y)$ and a blurring function $h(x, y)$, we need to recover $f(x, y)$ from the convolution $g(x, y) = f(x, y) * h(x, y)$

Solution by inverse filtering:

a. Without Noise:



From the model,

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

$$\text{i.e., } G(u, v) = F(u, v)H(u, v)$$

$$F(u, v) = \frac{G(u, v)}{H(u, v)} \text{-----(1)}$$

Also, from the model,

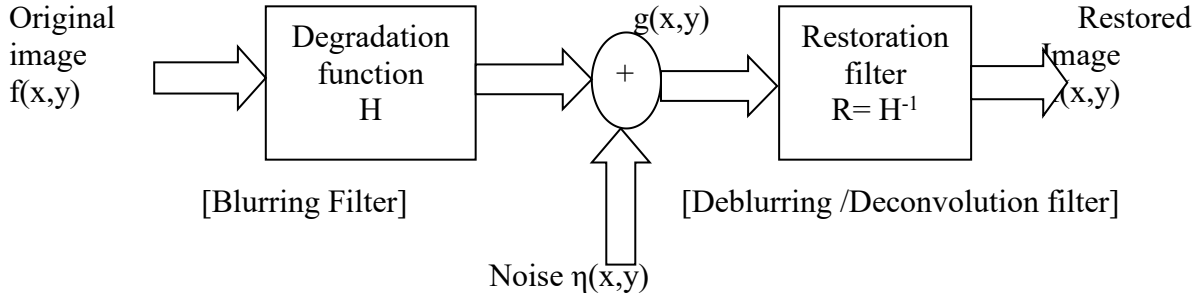
$$\hat{F}(u, v) = R(u, v)G(u, v) = H^{-1}(u, v)G(u, v) = \frac{G(u, v)}{H(u, v)} \text{-----(2)}$$

From (1) & (2), we get,

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

Hence exact recovery of the original image is possible.

b. With Noise:



The restored image is given as,

$$\hat{F}(u, v) = R(u, v)G(u, v) = H^{-1}(u, v)G(u, v) = \frac{G(u, v)}{H(u, v)} \text{-----}(3)$$

But, $G(u, v) = F(u, v) \cdot H(u, v) + \eta(u, v)$ ------(4)

Substitute equation (4) in (3), we get,

$$\begin{aligned} \hat{F}(u, v) &= \frac{G(u, v)}{H(u, v)} = \frac{F(u, v)H(u, v) + \eta(u, v)}{H(u, v)} \\ &= \frac{F(u, v)H(u, v)}{H(u, v)} + \frac{\eta(u, v)}{H(u, v)} \end{aligned}$$

$$\hat{F}(u, v) = F(u, v) + \frac{\eta(u, v)}{H(u, v)} \text{-----}(5)$$

Here, along with the original image, noise is also added. Hence the restored image consists of noise.

Drawbacks:

1. Doesnot perform well when used on noisy images

2. In the absence of noise, If $H(u, v) \neq 0$, then $R(u, v) = H^{-1}(u, v) = \frac{1}{H(u, v)} = \frac{1}{0} = \text{Infinite}$,

from equation (2)

3. In the presence of noise, when $H(u, v)$ is small, then the noise $\eta(u, v)$ dominates (from equation (5))

Pseudo inverse filtering: The solution is to carry out the restoration process in a limited neighborhood about the origin where $H(u, v)$ is not very small. This procedure is called

Pseudo inverse filtering. In this case,

$$\hat{F}(u,v) = \begin{cases} \frac{1}{|H(u,v)|^2} & |H(u,v)| > \delta \\ 0 & |H(u,v)| \leq \delta \end{cases}$$

The threshold δ is defined by the user. In general, the noise may possess large components at high frequencies (u,v) , while $H(u,v)$ and $G(u,v)$ normally will be dominated by low frequency components.

4. Write short notes on the Wiener filter characteristics. (Nov / Dec 2005) (or)
Derive the Wiener filter equation. Under what conditions does it become the pseudo inverse filter?

OBJECTIVE:

To find an estimate of the uncorrupted image f such that the mean square error between them is corrupted. This error measure is given as,

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

ASSUMPTION:

1. Noise and Image are uncorrelated
2. Then either noise or image has a zero mean
3. Gray levels in the estimate are a linear function of the levels in the degraded image

Let, $H(u,v)$ = Degradation function

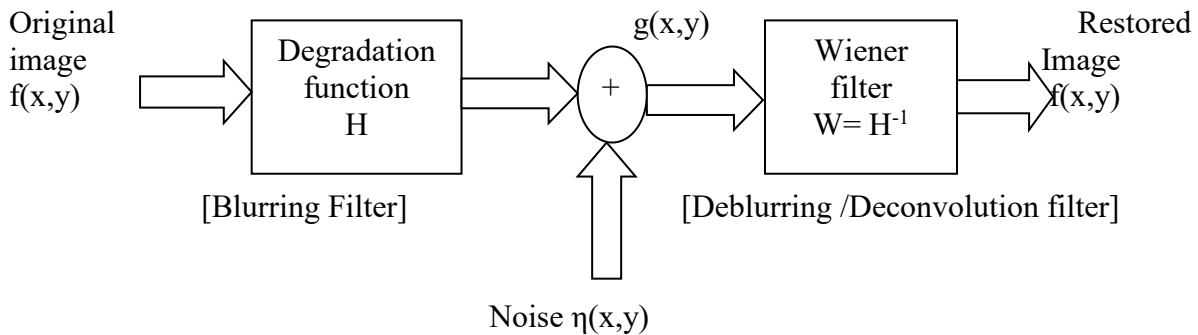
$H^*(u,v)$ = Complex conjugate of $H(u,v)$

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

$S_\eta(u,v) = |N(u,v)|^2$ = Power spectrum of the noise

$S_f(u,v) = |F(u,v)|^2$ = Power spectrum of the undegraded image

Wiener Model with Noise:



The restored image is given as,

$$\hat{F}(u,v) = W(u,v)G(u,v) = H^{-1}(u,v)G(u,v) = \frac{G(u,v)}{H(u,v)} \text{-----(1)}$$

But,

$$W(u,v) = \frac{H^*(u,v).S_f(u,v)}{|H(u,v)|^2 S_f(u,v) + S_\eta(u,v)} \text{-----(2)}$$

Where $W(u,v)$ is the Wiener filter
 Substitute equation (2) in (1), we get,

$$\text{We get, } \hat{F}(u, v) = \left[\frac{H^*(u, v) S_f(u, v)}{|H(u, v)|^2 S_f(u, v) + S_\eta(u, v)} \right] \cdot G(u, v) \text{ -----(3)}$$

But,

$$|H(u, v)|^2 = H^*(u, v) \cdot H(u, v) \text{ -----(4)}$$

To find an optimal Linear filter with Mean Square error as minimum,

Substitute equation (4) in (3), we get,

$$\begin{aligned} \hat{F}(u, v) &= \left[\frac{H^*(u, v) S_f(u, v)}{H^*(u, v) \cdot H(u, v) S_f(u, v) + S_\eta(u, v)} \right] \cdot G(u, v) \\ &= \left[\frac{H^*(u, v) S_f(u, v)}{S_f(u, v) [H^*(u, v) \cdot H(u, v) + \frac{S_\eta(u, v)}{S_f(u, v)}]} \right] \cdot G(u, v) \\ &= \left[\frac{H^*(u, v)}{H^*(u, v) \cdot H(u, v) + \frac{S_\eta(u, v)}{S_f(u, v)}} \right] \cdot G(u, v) \\ &= \left[\frac{1}{H(u, v) + \frac{S_\eta(u, v)}{S_f(u, v)} \cdot H^*(u, v)} \right] \cdot G(u, v) \end{aligned}$$

Thus when the filter consists of the terms inside the bracket, it is called as the minimum mean square error filter or the least square error filter.

Case 1: If there is no noise, $S_\eta(u, v) = 0$, then the Wiener filter reduces to pseudo Inverse filter.

From equation (2),

$$W(u, v) = \frac{H^*(u, v) \cdot S_f(u, v)}{|H(u, v)|^2 S_f(u, v) + 0} = \frac{H^*(u, v)}{H^*(u, v) H(u, v)} = \begin{cases} \frac{1}{H(u, v)} \\ 0 \end{cases}$$

,when $H(u,v) \neq 0$

, when $H(u,v)=0$

Case 2: If there is no blur, $H(u,v) = 1$, then from equation (2),

$$\begin{aligned}
 W(u,v) &= \frac{H^*(u,v) \cdot S_f(u,v)}{|H(u,v)|^2 S_f(u,v) + S_n(u,v)} \\
 &= \left[\frac{H^*(u,v)}{H^*(u,v) \cdot H(u,v) + \frac{S_n(u,v)}{S_f(u,v)}} \right] \\
 &= \left[\frac{1}{1 + \frac{S_n(u,v)}{S_f(u,v)}} \right] \\
 &= \left[\frac{1}{1 + K} \right]
 \end{aligned}$$

Where

$$K = \frac{S_n(u,v)}{S_f(u,v)}$$

Thus the Wiener denoising Filter is given as,

$$W(u,v) = \left[\frac{1}{1 + K} \right] = \left[\frac{S_f(u,v)}{S_f(u,v) + S_n(u,v)} \right]$$

and

$$W(u,v) = \frac{S_f(u,v)}{S_f(u,v) + S_n(u,v)} = \frac{(SNR)}{(SNR) + 1}$$

$$(i) \quad (SNR) \gg 1 \Rightarrow W(u,v) \cong 1$$

$$(ii) \quad (SNR) \ll 1 \Rightarrow W(u,v) \cong (SNR)$$

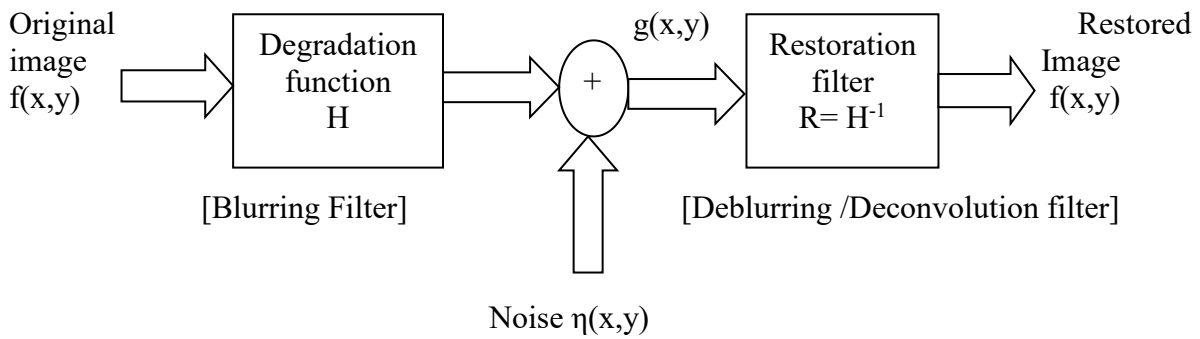
(SNR) is high in low spatial frequencies and low in high spatial frequencies so $W(u,v)$ can be implemented with a low pass (smoothing) filter.

Wiener filter is more robust to noise and preserves high frequency details SNR is low, Estimate is high

5. Explain in detail the constrained least squares restoration. (May / June 09).

In Wiener filter, the power spectra of the undegraded image and noise must be known. Although a constant estimate is sometimes useful, it is not always suitable. Constrained least squares filtering just require the mean and variance of the noise.

With Noise:



The blurred image is given as,

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

In matrix form , the above equation is expressed as,

$$g = Hf + \eta$$

$$\boxed{\eta = g - Hf} \text{---(1)}$$

Problem statement:

To find the minimum of a criterion function C defined as,

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

Subject to the constraint

$$\left\| g - H \hat{f} \right\|^2 = \|\eta\|^2 \text{-----(2)}$$

Where $\|w\|^2 = w^T w$ is the Euclidean vector norm,

\hat{f} is the estimate of the undegraded image, and

$\nabla^2 f$ is the laplacian operator

Solution to the Problem:

The frequency domain solution to this optimization problem is given by,

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma |P(u,v)|^2} \right] \text{-----(3)}$$

Where γ is a parameter that must be adjusted so that the following constraint is satisfied,

$$\|\eta\|^2 = \|g - H \hat{f}\|^2, \text{-----(4)}$$

And $P(u,v)$ is the Fourier transform of the function,

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

Here the noise is replaced by $P(u,v)$...i.e., Noise is approximated to Laplace such that mean = 0. If mean is zero, the amount of DC component is zeroi.e, overall brightness of image is zero. The value of γ is selected by iteration such that $\hat{F}(u,v)$ is a good approximation of $F(u,v)$.

One Approach to find the value of γ ,

A residual vector r is defined as, $r = g - H \hat{f}$,-----(5)

Since $\hat{F}(u,v)$ is a function of γ , then γ is also a function of r ,

$$\text{i.e., } \|r\|^2 = \Phi(\gamma),$$

from equation (3) and (5), if $\|r\|^2 = \|\eta\|^2$, then $\hat{f} = f$ is a good estimation

Therefore, $\|r\|^2 = \|\eta\|^2 \pm a$,

Steps:

- I. Specify an initial value of “ γ ”
- II. Compute $\|r\|^2$, which is given as a function of γ ,

$$\text{i.e., } \|r\|^2 = \Phi(\gamma),$$

- III. Check whether $\|r\|^2 = \|\eta\|^2 \pm a$, is satisfied, else increase γ if $\|r\|^2 < \|\eta\|^2 - a$, or decrease γ if $\|r\|^2 > \|\eta\|^2 + a$,

- IV. Use the new value of γ , to recompute the optimum estimate $\hat{F}(u, v)$

Another approach to find the value of γ is Newton – raphson algorithm:

By knowing the mean and variance of noise, we can estimate the restored image. We know that the residual vector is defined as,

$$r = g - H f^{\wedge}$$

$$\text{Where } \|r\|^2 = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} r^2(x, y),$$

The variance of the noise over the entire image is given by the sample average method,

$$\sigma_{\eta}^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\eta(x, y) - m_{\eta}]^2,$$

$$\text{Where sample mean is } m_{\eta} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\eta(x, y)],$$

From the above expressions, $\|\eta\|^2 = MN[\sigma_{\eta}^2 - m_{\eta}]$

Thus the value for the constraint is determined in terms of noise mean and variance.

6. Discuss about the Lagrange Multipliers.

Lagrange Multipliers

Lagrange multiplier is a method to find maximum or minimum value of a function subject to some constraints (side conditions) on the domain of the function.

Examples:

Maximize or minimize $f(x, y)$ subject to **condition** $g(x, y) = c$.

Maximize or minimize $f(x, y, z)$ subject to **condition** $g(x, y, z) = c$.

Maximize or minimize $f(x, y, z)$ subject to **conditions** $g(x, y, z) = c_1$, and $h(x, y, z) = c_2$.

Lagrange Theorem:

Let f and g have continuous first partial derivatives such that f has a local maximum or local minimum at (x_0, y_0) on the smooth constraint curve $g(x, y) = c$. If $\nabla g(x_0, y_0) = 0$, then there is a real number λ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

λ is Lagrange multiplier.

Lagrange Method:

To find extrema of $f(x, y)$ subject to constraint $g(x, y) = 0$:

I. Solve the equations

$\nabla f(x, y) = \lambda \nabla g(x, y)$ and $g(x, y) = 0$, by solving

$$f_x(x, y) = \lambda g_x(x, y)$$

$$f_y(x, y) = \lambda g_y(x, y)$$

$$g(x, y) = 0.$$

Find all solutions (x_1, y_1, λ_1) , (x_2, y_2, λ_2) , (x_3, y_3, λ_3) , . . .

Discard the Lagrange multipliers $\lambda_1, \lambda_2, \lambda_3, \dots$

II. Evaluate $f(x, y)$ at each solution point obtained above.

The largest value yields the maximum and the smallest value yields the minimum subject to constraint $g(x, y) = 0$.

The Lagrange Multiplier Method

- i. To find the only possible solutions of the problem
 Maximize or minimize a function $f(x, y)$ subject to $g(x, y) = c$

- ii. 1. Write down the Lagrangian function

$$L(x, y) = f(x, y) - \lambda(g(x, y) - c)$$

where λ is a constant.

2. Differentiate L with respect to x and y , and equate partials to 0.
 3. We shall get two equations together with the constraint $g(x, y) = c$

$$L_x(x, y) = f_x(x, y) - \lambda g_x(x, y)$$

$$L_y(x, y) = f_y(x, y) - \lambda g_y(x, y)$$

4. Solve them simultaneously for x, y and λ

The conditions in 3 are called **first-order conditions** for the problem of maximizing or minimizing a function $f(x, y)$ subject to $g(x, y) = c$

Example: Solve the problem: Maximize $f(x, y) = x^2 + y^2$ subject to

$$g(x, y) = x^2 + xy + y^2 = 3$$

The Lagrangian is $L(x, y) = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$

$$L_x(x, y) = 2x - \lambda(2x + y) = 0$$

$$L_y(x, y) = 2y - \lambda(x + 2y) = 0$$

$$x^2 + xy + y^2 - 3 = 0$$

$$\lambda = \frac{2x}{2x + y} \text{ if } 2x + y \neq 0 \quad \text{and} \quad \lambda = \frac{2y}{y + 2x} \text{ if } y + 2x \neq 0$$

Combining them together we get $\frac{2x}{2x + y} = \frac{2y}{y + 2x}$ and $x^2 = y^2$ or $x = y$ or $x = -y$

Suppose $y = x$ From 3 we get $x^2 = 1$, so $x = 1$ or $x = -1$ and it gives two solutions

$$(x, y) = (1, 1) \text{ and } (-1, -1) \text{ with } \lambda = \frac{2}{3}$$

Suppose $y = -x$ From 3 we get $x^2 = 3$, so $x = \sqrt{3}$ or $x = -\sqrt{3}$ and it gives two solutions

$$(x, y) = (\sqrt{3}, -\sqrt{3}) \text{ and } (-\sqrt{3}, \sqrt{3}) \text{ with } \lambda = 2$$

Two more cases $2x + y = 0$ and $2y + x = 0$ yield no solutions since they result in $x = 0$ or $y = 0$ which contradicts the constraint $x^2 + xy + y^2 - 3 = 0$.

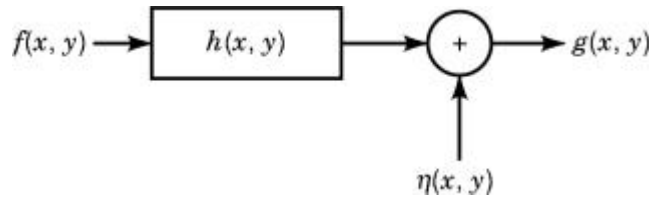
Two solutions $(x, y) = (1, 1)$ and $(-1, -1)$ lead to $f(1, 1) = (-1, -1) = 2$ while the solutions

$$(x, y) = (\sqrt{3}, -\sqrt{3}) = (-\sqrt{3}, \sqrt{3}) \text{ lead to } f(\sqrt{3}, -\sqrt{3}) = f(-\sqrt{3}, \sqrt{3}) = 6$$

The first two solutions solve minimization problem while the last two solve maximization problem.

7.Explain unconstrained restoration and constrained restoration

The original image signal $f(x, y)$ is subjected to the linear degrading function $h(x, y)$. An arbitrarily noise signal $\eta(x, y)$ is then added to create the degraded image signal $g(x, y)$.



Linear Algebraic Restoration Techniques

These techniques utilize matrix algebra and discrete mathematics for solving the problem of image restoration. To extract the discrete model for the restoration system, we make the following assumptions:

1. The digitized original image $f(x, y)$ and the restored image $f_r(x, y)$ are stored in the $M^2 \times 1$ column vectors \mathbf{f} and \mathbf{f}_r , respectively.
2. The digitized degrading function $h(x, y)$ is stored in a square matrix \mathbf{H} ($M^2 \times M^2$).
3. The degraded image $g(x, y)$ and noise $\eta(x, y)$ is stored in the $M^2 \times 1$ column vectors \mathbf{g} and $\mathbf{\eta}$, respectively.

We can express the observed (degraded) image vector in the compact form:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} \quad (6)$$

a. Unconstrained Reconstruction

If we know very little about the noise η , then we try to find an estimate image \mathbf{f}_r , such that $\mathbf{H}\mathbf{f}_r$ approximates \mathbf{g} in a least-squares manner. This can be accomplished by minimizing the norm of the noise η . Squaring the norm of both sides of Eq. (6) after substituting for \mathbf{f} by the estimate vector \mathbf{f}_r and moving $\mathbf{H}\mathbf{f}_r$ to the other side of the equation, we get

$$\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\mathbf{f}_r\|^2 = (\mathbf{g} - \mathbf{H}\mathbf{f}_r)(\mathbf{g} - \mathbf{H}\mathbf{f}_r)' \quad (7)$$

Where $\|\mathbf{a}\|^2$ is the square of the norm of vector \mathbf{a} and is given by $\|\mathbf{a}\|^2 = \mathbf{a}\mathbf{a}'$, where \mathbf{a}' is the transpose of the vector \mathbf{a} .

Consider the error function \mathbf{E} , where

$$\mathbf{E}(\mathbf{f}_r) = \|\mathbf{g} - \mathbf{H}\mathbf{f}_r\|^2 = (\mathbf{g} - \mathbf{H}\mathbf{f}_r)(\mathbf{g} - \mathbf{H}\mathbf{f}_r)' \quad (8)$$

Then our goal is to minimize E with respect to \mathbf{f}_r . This can be accomplished by taking the derivative of E with respect to \mathbf{f}_r , and equating the result to zero, that is

$$\frac{\partial E(\mathbf{f}_r)}{\partial \mathbf{f}_r} = -2\mathbf{H}'(\mathbf{g} - \mathbf{H}\mathbf{f}_r) = 0 \quad (9)$$

Assuming that \mathbf{H}^{-1} exists, the solution for this equation is given as

$$\mathbf{f}_r = \mathbf{H}^{-1}\mathbf{g} \quad (10)$$

b. Constrained Reconstruction

To account for the noise term in Eq. (6), we introduce the square matrix \mathbf{Q} ($M^2 \times M^2$) to represent some linear operator on \mathbf{f} . By selecting different \mathbf{Q} 's, we are able to set the goal of restoration as desired. Equation (8) is now modified to

$$E(\mathbf{f}_r) = \|\mathbf{Q}\mathbf{f}_r\|^2 + \lambda(\|\mathbf{g} - \mathbf{H}\mathbf{f}_r\|^2 - \|\mathbf{n}\|^2) \quad (11)$$

Where λ constant is called the LaGrange multiplier. Again we try to minimize the error function E , by taking its derivative with respect to \mathbf{f}_r and equating the result to zero, that is

$$\frac{\partial E(\mathbf{f}_r)}{\partial \mathbf{f}_r} = 2\mathbf{Q}'\mathbf{Q}\mathbf{f}_r - 2\lambda\mathbf{H}'(\mathbf{g} - \mathbf{H}\mathbf{f}_r) = 0 \quad (12)$$

Solving for \mathbf{f}_r , we get

$$\mathbf{f}_r = (\mathbf{H}\mathbf{H} + \alpha\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{H}'\mathbf{g} \quad (13)$$

Where $\alpha = 1/\lambda$ is a constant that we adjust to satisfy the constraint.

9. Discuss about the concepts and uses of Geometric Transformations in the context of image restoration.

Geometric Transformations

Geometric transformations modify the spatial relationships between pixels in an image. Geometric transformation often are called *rubber-sheet transformations*, because they may be viewed as the process of “printing” an image on a sheet of rubber and then stretching this sheet according to some predefined set of rules.

In terms of digital image processing, a geometric transformation consists of two basic operations: (1) a *spatial transformation*, which defines the “rearrangement” of pixels on the image plane; and (2) *gray-level interpolation*, which deals with the assignment of gray levels to pixels in the spatially transformed image.

Spatial Transformations:

Suppose that an image f with pixel coordinates (x, y) undergoes geometric distortion to produce an image g with coordinates (x^1, y^1) . This transformation may be expressed as

$$x^1 = r(x, y) \quad \dots\dots\dots 1$$

and $y^1 = s(x, y) \quad \dots\dots\dots 2$

where $r(x, y)$ and $s(x, y)$ are the spatial transformations that produced the geometrically distorted image $g(x^1, y^1)$. For example, if $r(x, y) = x/2$ and $s(x, y) = y/2$, the “distortion” is simply a shrinking of the size of $f(x, y)$ by one half in both spatial directions. If $r(x, y)$ and $s(x, Y)$ were known analytically, recovering $f(x, y)$ from the distorted image $g(x^1, y^1)$ by applying the transformation in reverse might be possible theoretically. In practice, however, formulating a single set of analytical functions $r(x, y)$ and $s(x, y)$ that

describe the geometric distortion process over the entire image plane generally is not possible. The method used most frequently to overcome this difficulty is to formulate the spatial relocation of pixels by the use of *tiepoints*, which are a subset of pixels whose location in the input (distorted) and output (corrected) images is known precisely.

Figure shows quadrilateral regions in a distorted and corresponding corrected image. The vertices of the quadrilaterals are corresponding tiepoints.

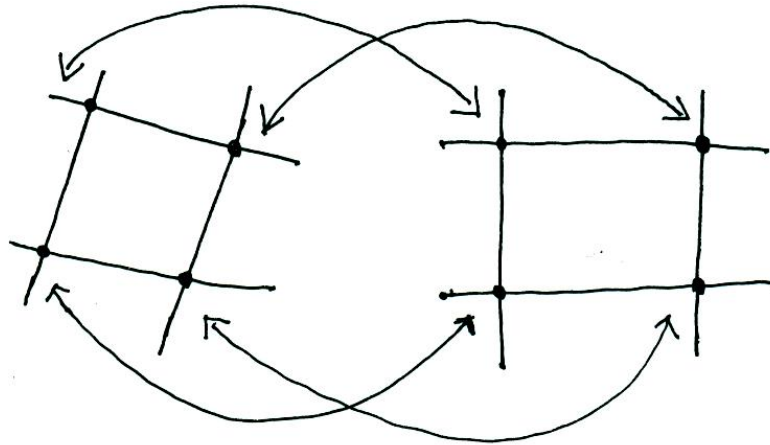


Figure: Corresponding tie points in two image segments.

Suppose that the geometric distortion process within the quadrilateral regions is modeled by a pair of bilinear equations so that

$$r(x, y) = c_1x + c_2y + c_3xy + c_4 \dots\dots 3$$

and $s(x, y) = c_5x + c_6y + c_7xy + c_8 \dots\dots\dots 4$

Then, from Equations (1) and (2),

$$x^1 = c_1x + c_2y + c_3xy + c_4 \dots\dots\dots 5$$

and $y^1 = c_5x + c_6y + c_7xy + c_8 \dots\dots\dots 6$

Since there are a total of eight known tiepoints, these equations can be solved for the eight coefficients c_i , $i = 1, 2, \dots\dots\dots 8$. The coefficients constitute the geometric distortion model used to transform *all* pixels within the quadrilateral region defined by the tiepoints used to obtain the coefficients. In general, enough tiepoints are needed to generate a set of quadrilaterals that cover the entire image, with each quadrilateral having its own set of coefficients.

Once we have the coefficients, the procedure used to generate the corrected (i.e., restored) image is not difficult. If we want to find the value of the undistorted image at any point (x_0, y_0) , we simply need to know where in the distorted image $f(x_0, y_0)$ was mapped. This we find out by substituting (x_0, y_0) into Equations (5) and (6) to obtain the geometrically distorted coordinates (x_0^1, y_0^1) . The value of the point in the undistorted image that was mapped to (x_0^1, y_0^1) is $g(x_0^1, y_0^1)$. So we obtain the restored image value simply by letting $\hat{f}(x_0, y_0) = g(x_0^1, y_0^1)$. For example, to generate $\hat{f}(0, 0)$, we substitute $(x, y) = (0, 0)$

into Equations (5) and (6) to obtain a pair of coordinates (x^1, y^1) from those equations. Then we let $\hat{f}(0,0) = g(x^1, y^1)$, where x^1 and y^1 are the coordinate values just obtained. Next, we substitute $(x, y) = (0, 1)$ into Equations (5) and (6), obtain another pair of values (x^1, y^1) , and let $\hat{f}(0, 1) = g(x^1, y^1)$ for those coordinate values. The procedure continues pixel by pixel and row by row until an array whose size does not exceed the size of image g is obtained. A column (rather than a row) scan would yield identical results. Also, a bookkeeping procedure is needed to keep track of which quadrilaterals apply at a given pixel location in order to use the proper coefficients.

Tie points are established by a number of different techniques, depending on the application. For instance, some image generation systems having physical artifacts (such as metallic points) embedded on the imaging sensor itself. These produce a *known* set of points (called *reseau marks*) directly on the image as it is acquired. If the image is distorted later by some other process (such as an image display or image reconstruction process), then the image can be geometrically corrected using the technique just described.

Gray-Level Interpolation

The method discussed in the preceding section steps through integer values of the coordinates (x, y) to yield the restored image. However, depending on the values of the coefficients C_i , Equations (5) and (6) can yield no integer values

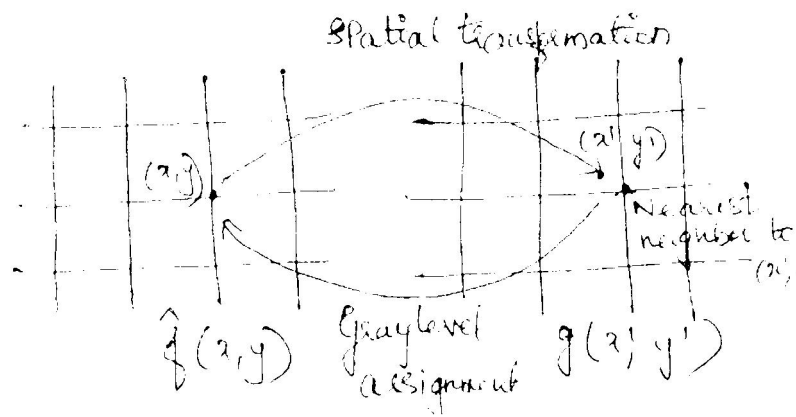


Figure: Gray-level interpolation based on the nearest neighbor concept.

for x^1 and y^1 . Because the distorted image g is digital, its pixel values are defined only at integer coordinates. Thus using non-integer values for x^1 and y^1 causes a mapping into locations of g for which no gray levels are defined. Inferring what the gray-level values at those locations should be, based only on the pixel values at integer coordinate locations, then becomes necessary. The technique used to accomplish this is called *gray-level interpolation*.

The simplest scheme for gray-level interpolation is based on a nearest neighbor approach. This method, also called *zero-order interpolation*, is illustrated in Figure. This figure shows the mapping of integer (x, y) coordinates into fractional coordinates (x^1, y^1) by means of Equations (5) and (6); the selection of the closest integer coordinate neighbor to (x^1, y^1) ; the assignment of the gray level of this nearest neighbor to the pixel rated at (x, y) .

Although nearest neighbour interpolation is simple to implement, this method often has the drawback of producing undesirable artifacts, such as distortion of straight edges in images of high resolution. For general-purpose image processing a bilinear interpolation

approach that uses the gray levels of the four nearest neighbors usually is adequate. This approach is straightforward. Because the gray level of each of the four integral nearest neighbors of a nonintegral pair of coordinates (x^1, y^1) is known, the gray-level value at these coordinates, denoted $v(x^1, y^1)$, can be interpolated from the values of its neighbors by using the relationship

$$v(x^1, y^1) = ax + 1 + by^1 + cx^1 y^1 + d \dots\dots\dots 7$$

where the four coefficients are easily determined from the four equations in four unknowns that can be written using the four known neighbors of (x^1, y^1) . When these coefficients have been determined, $v(x^1, y^1)$ is computed and this value is assigned to the location in $f(x, y)$ that yields the spatial mapping into location (x^1, y^1) . It is easy to visualize this procedure with the aid of Figure. The exception is that, instead of using the gray-level value of the nearest neighbor to (x^1, y^1) , we actually interpolate a value at location (x^1, y^1) and use this value for the gray-level assignment at (x, y) .

UNIT IV IMAGE SEGMENTATION

Edge detection, Edge linking via Hough transform – Thresholding - Region based segmentation – Region growing – Region splitting and merging – Morphological processing- erosion and dilation, Segmentation by morphological watersheds – basic concepts – Dam construction – Watershed segmentation algorithm.

PART A

1. Define Image Segmentation (May / June 2009)

Segmentation is a process of subdividing an image in to its constituted regions or objects or parts. The level to which the subdivision is carried depends on the problem being solved, i.e., segmentation should be stopped when object of interest in an application have been isolated.

Image segmentation techniques are used to isolate the desired object from the scene so that measurements can be made on it subsequently.

Applications of segmentation.

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

2. What are the different image segmentation techniques?

Image segmentation is classified as follows:

- ❖ Template matching
- ❖ Thresholding
- ❖ Boundary detection
- ❖ Clustering
- ❖ Quad- trees
- ❖ Texture matching

3. What are the two properties that are followed in image segmentation?

Image segmentation algorithms generally are based on one of the basic properties of intensity values

1. discontinuity
2. similarity

4. Explain the property of discontinuity?

The principal approach of this property is to partition an image based on abrupt changes in intensity such as edges in an image.

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

Points, lines and edges.

6. What is the idea behind the similarity property?

The principal approach in the property is based on portioning an image into regions that are similar according to a set of predefined criteria. Thresholding, region growing and region splitting and merging are examples of this method.

7. What is edge?

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

8. What is edge detection? Explain. (May / June 2009)

Edge direction is fundamental importance in image analysis. Edges characterize object boundaries and are therefore useful for segmentation, registration and identification of objects in scenes. Edge points can be thought of as pixel locations of abrupt gray- level change.

Edge detection is used for detecting discontinuities in gray level. First and second order digital derivatives are implemented to detect the edges in an image.

9. Define zero crossing property of edge detection.

An imaginary straight line joining the extreme positive and negative values of the second derivative could cross zero near the midpoint of the edge. Zero crossing property is useful for locating the centers of thick edges.

10. What is meant by gradient operators?

First order derivatives in an image are computed using the gradient operators. The gradient of an image $f(x,y)$ at location (x,y) is defined as the vector.

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

The ∇f is called as the gradient. Magnitude of the vector is

$$\Delta f = \text{mag } f = [G_x^2 + G_y^2]^{1/2}$$

$$\alpha(x,y) = \tan^{-1}(G_y/G_x), \alpha(x,y) \text{ is the direction angle of vector } \Delta f$$

11. Define Laplacian.

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

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

12. Define Laplacian of a Gaussian function.

The laplacian of h (the second derivative of h with respect to r) is

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

This function is commonly referred to as the Laplacian of a Gaussian (LOG).

13. 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 straight line joining the extreme positive and negative values of the second derivative would cross zero near the midpoint of the edge.

14. 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.

15. 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.

16. 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.

17. How do you detect isolated points in an image? (May/June 2009)

To detect the isolated points due to noise or interference, the general mask is shown below,

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

18. What is the role of gradient operator and laplacian operator in segmentation?

The first and second derivative at any point in an image can be obtained by using the magnitude of the gradient operators at that point and laplacian operators

19. What is the role of first derivative and second derivative in segmentation?

The magnitude of the first derivative is used to identify the presence of an edge in the image.

The second derivative is positive for the part of the transition which is associated with the dark side of the edge. It is zero for pixel which lies exactly on the edge.

It is negative for the part of transition which is associated with the light side of the edge. The sign of the second derivative is used to find that whether the edge pixel lies on the light side or dark side.

20. What are the types of thresholding?

There are three different types of thresholds. They are

1. global thresholding

2. local thresholding
3. dynamic (or) Adaptive

21. 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**.

22. 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.

23. What is meant by region – Based segmentation?

The objective of segmentation is to partition an image into regions. These segmentation techniques are based on finding the regions directly.

24. Define region growing?

Region growing is a procedure that groups pixels or sub regions 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 neighboring pixels that have properties similar to the seed.

25. Specify the steps involved in splitting and merging?

The region splitting and merging is to subdivide an image initially into a set of arbitrary, disjointed regions and then merge and / or split the regions in attempts to satisfy the conditions.

Split into 4 disjoint quadrants any region R_i for which $P(R_i) = \text{FALSE}$. Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k) = \text{TRUE}$. Stop when no further merging or splitting is positive.

26. What is the main objective of watershed algorithm?

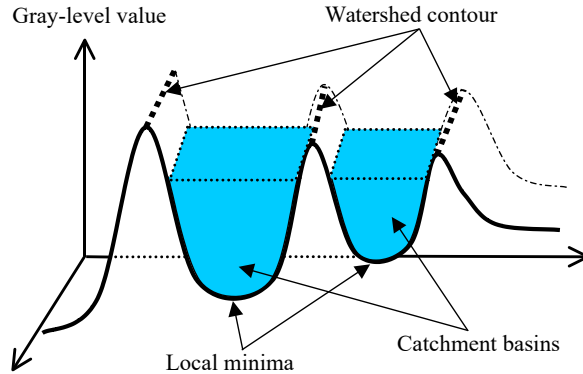
The concept of watersheds is based on visualizing an image in three dimensions: two spatial coordinates versus gray levels. In such a "topographic" interpretation, we consider three types of points:

- a. Points belonging to a regional minimum
- b. Catchment basin / watershed of a regional minimum

Points at which a drop of water will certainly fall to a single minimum

- c. Divide lines / Watershed lines

Points at which a drop of water will be equally likely to fall to more than one minimum.



The principal objective of segmentation algorithms based on these concepts is to find the watershed lines

27. What is the use of Dam?

A dam is built to prevent the merging. Dam construction is based on binary images; the simplest way to construct dams separating sets of binary points is to use morphological dilation.

28. What is meant by markers?

it used to control over segmentation is based on markers. Marker is a connected component belonging to an image. We have internal markers, associated with objects of interest and external markers associated with background.

Part-B

1. What is segmentation? (May / June 2009)

Segmentation is a process of subdividing an image in to its constitute regions or objects or parts. The level to which the subdivision is carried depends on the problem being solved, i.e., Segmentation should be stopped when object of interest in an application have been isolated.

Applications of segmentation.

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

Segmentation algorithms for monochrome images (gray-scale images) generally are based on one or two basic properties of gray-level values: **discontinuity and similarity**.

1. Discontinuity is to partition an image based on abrupt changes in gray level. Example:

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

2. Similarity is to partition an image into regions that are similar according to a set of predefined criteria.

Example: **thresholding, region growing, or region splitting and merging.**

Formally **image segmentation can be defined as** a process that divides the image R into *regions* R_1, R_2, \dots, R_n such that:

- a. The segmentation is complete: $\bigcup_{i=1}^n R_i = R$
- b. Each region is uniform.
- c. The regions do not overlap each other: $R_i \cap R_j = \emptyset, \forall i \neq j$

- d. The pixels of a region have a common property (same gray level): $P(R_i) = TRUE$
- e. Neighboring regions have different properties:
 $P(R_i) \neq P(R_j), \forall R_i, R_j$ are neighbors.

2. Explain the various detection of discontinuities in detail.(or) Explain Edge detection in detail

Detection of discontinuities

Discontinuities such as isolated points, thin lines and edges in the image can be detected by using similar masks as in the low- and high-pass filtering. The absolute value of the weighted sum given by equation (2.7) indicates how strongly that particular pixel corresponds to the property described by the mask; the greater the absolute value, the stronger the response.

1. Point detection

The mask for detecting *isolated points* is given in the Figure. A point can be defined to be isolated if the response by the masking exceeds a predefined threshold:

$$|f(x)| > T$$

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

Figure : Mask for detecting isolated points.

2. Line detection

- *line detection* is identical to point detection. I
- Instead of one mask, four different masks must be used to cover the four primary directions namely, horizontal, vertical and two diagonals.
- Thus lines are detected using the following masks

Horizontal	+45°	Vertical	-45°																																				
<table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	-1	-1	-1	2	2	2	-1	-1	-1	<table><tr><td>-1</td><td>-1</td><td>2</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>2</td><td>-1</td><td>-1</td></tr></table>	-1	-1	2	-1	2	-1	2	-1	-1	<table><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr></table>	-1	2	-1	-1	2	-1	-1	2	-1	<table><tr><td>2</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>2</td></tr></table>	2	-1	-1	-1	2	-1	-1	-1	2
-1	-1	-1																																					
2	2	2																																					
-1	-1	-1																																					
-1	-1	2																																					
-1	2	-1																																					
2	-1	-1																																					
-1	2	-1																																					
-1	2	-1																																					
-1	2	-1																																					
2	-1	-1																																					
-1	2	-1																																					
-1	-1	2																																					

Figure Masks for line detection.

- Horizontal mask will result with max response when a line passed through the middle row of the mask with a constant background.
- Thus a similar idea is used with other masks.
- The preferred direction of each mask is weighted with a larger coefficient (i.e.,2) than other possible directions.

Steps for Line detection:

- Apply every masks on the image
- let R_1, R_2, R_3, R_4 denotes the response of the horizontal, +45 degree, vertical and -45 degree masks, respectively.
- if, at a certain point in the image $|R_i| > |R_j|$,
- for all $j \neq i$, that point is said to be more likely associated with a line in the direction of mask i .
- Alternatively, to detect all lines in an image in the direction defined by a given mask, just run the mask through the image and threshold the absolute value of the result.
- The points that are left are the strongest responses, which, for lines one pixel thick, correspond closest to the direction defined by the mask.

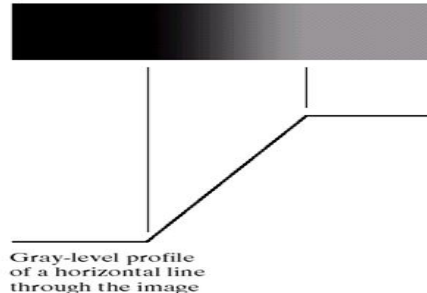
3.Edge detection

- The most important operation to detect discontinuities is *edge detection*.
- **Edge detection is used for detecting discontinuities in gray level. First and second order digital derivatives are implemented to detect the edges in an image.**
- Edge is defined as the boundary between two regions with relatively distinct gray-level properties.
- An edge is a set of connected pixels that lie on the boundary between two regions.
- Approaches for implementing edge detection
 - first-order derivative (Gradient operator)**
 - Second-order derivative (Laplacian operator)**
- Consider the following diagrams:

Model of an ideal digital edge

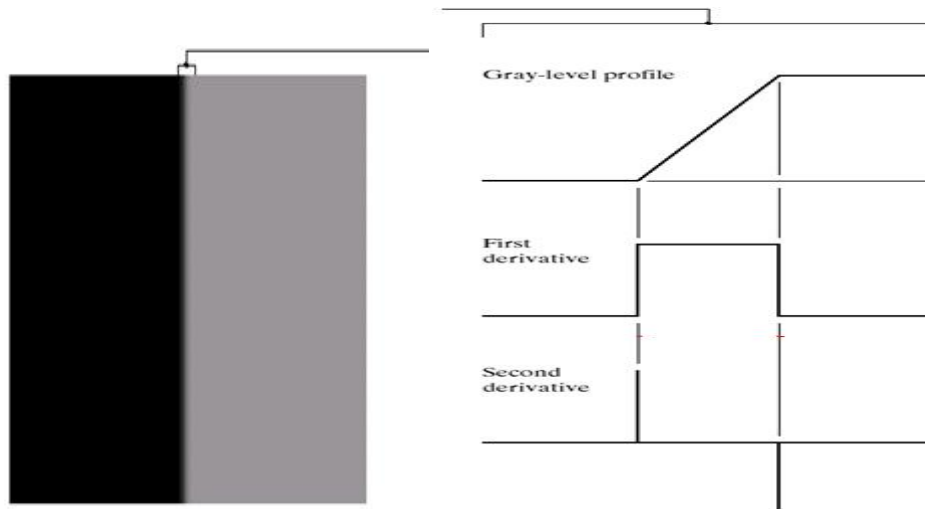


Model of a ramp digital edge



THICK EDGE:

- The slope of the ramp is inversely proportional to the degree of blurring in the edge.
- We no longer have a thin (one pixel thick) path.
- Instead, an edge point now is any point contained in the ramp, and an edge would then be a set of such points that are connected.
- The thickness is determined by the length of the ramp.
- The length is determined by the slope, which is in turn determined by the degree of blurring.
- **Blurred edges tend to be thick and sharp edges tend to be thin**



- The first derivative of the gray level profile is positive at the trailing edge, zero in the constant gray level area, and negative at the leading edge of the transition
- The magnitude of the first derivative is used to identify the presence of an edge in the image
- The second derivative is positive for the part of the transition which is associated with the dark side of the edge. It is zero for pixel which lies exactly on the edge.
- It is negative for the part of transition which is associated with the light side of the edge
- The sign of the second derivative is used to find that whether the edge pixel lies on the light side or dark side.
- The second derivative has a zero crossings at the midpoint of the transition in gray level. The first and second derivative at any point in an image can be obtained by using the magnitude of the gradient operators at that point and laplacian operators.

1. FIRST DERIVATIVE. This is done by calculating the *gradient* of the pixel relative to its neighborhood. A good approximation of the first derivative is given by the two *Sobel operators* as shown in the following Figure, with the advantage of a smoothing effect. Because derivatives enhance noise, the smoothing effect is a particularly attractive feature of the Sobel operators.

- First derivatives are implemented using the magnitude of the gradient.

i.Gradient operator:

For a function $f(x, y)$, the gradient ∇f at co-ordinate (x, y) is defined as the 2-dimensional column vector

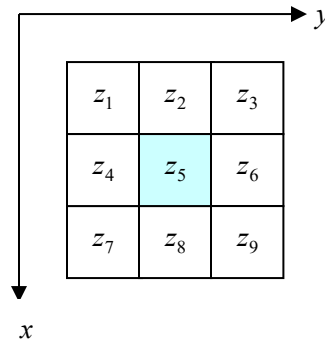
$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

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

- Let $\alpha(x, y)$ represent the direction angle of the vector ∇f at (x, y)
 $\alpha(x, y) = \tan^{-1}(G_y / G_x)$
- The direction of an edge at (x, y) is perpendicular to the direction of the gradient vector at that point
- The *direction* of the gradient is: $\alpha = \tan^{-1}\left(\frac{G_x}{G_y}\right)$

a) Roberts operator

consider a 3×3 mask is the following.



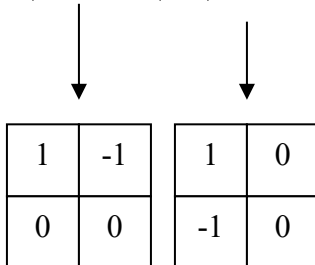
It can be approximated at point z_5 in a number of ways. The simplest is to use the difference $(z_8 - z_5)$ in the x direction and $(z_6 - z_5)$ in the y direction. This approximation is known as the **Roberts** operator, and is expressed mathematically as follows.

$$G_x = \left| \frac{\partial f(x, y)}{\partial x} \right| = f(x+1) - f(x) = |z_6 - z_5|$$

$$G_y = \left| \frac{\partial f(x, y)}{\partial y} \right| = f(y+1) - f(y) = |z_8 - z_5|$$

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

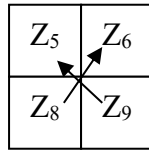
$$\cong |z_6 - z_5| + |z_8 - z_5|$$



Roberts operator

Another approach is to use **Roberts cross differences**

Take



This approximation is known as the **Roberts** cross gradient operator, and is expressed mathematically as follows

$$\nabla f \cong |z_9 - z_5| + |z_8 - z_6|$$

-1	0
0	1

0	-1
1	0

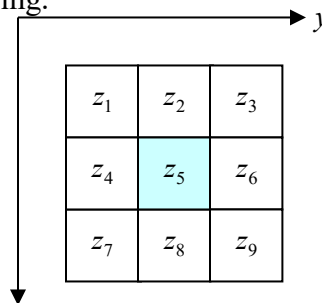
Roberts operator

The above Equation can be implemented using the following mask.

The original image is convolved with both masks separately and the absolute values of the two outputs of the convolutions are added.

b)Prewitt operator

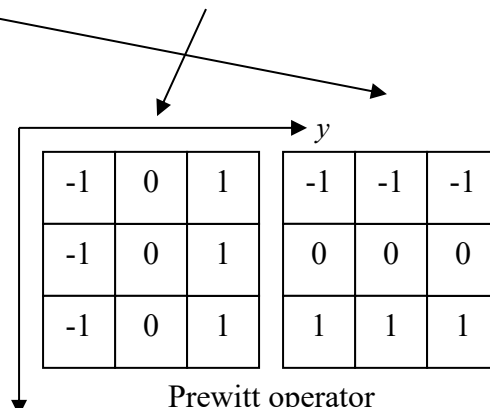
consider a 3×3 mask is the following.



$$Gx = \left| \frac{\partial f(x, y)}{\partial x} \right| = 3^{rd} \text{ row} - 1^{st} \text{ row} = |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)|$$

$$Gy = \left| \frac{\partial f(x, y)}{\partial y} \right| = 3^{rd} \text{ col} - 1^{st} \text{ col} = |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$

$$\nabla f \cong |Gx| + |Gy| = |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| + |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$



Prewitt operator

This approximation is known as the **Prewitt operator**.

The above Equation can be implemented by using the following masks.

c) Sobel operator.

- The gradients are calculated separately for horizontal and vertical directions:

$$G_x = (x_7 + 2x_8 + x_9) - (x_1 + 2x_2 + x_3)$$

$$G_y = (x_3 + 2x_6 + x_9) - (x_1 + 2x_4 + x_7)$$

Horizontal edge Vertical edge

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

Figure Sobel masks for edge detection

- The following shows the a 3 x3 region of image and various masks used to compute the gradient at point z_5

ii. SECOND DERIVATE can be approximated by the *Laplacian* mask given in Figure . The drawbacks of *Laplacian* are its sensitivity to noise and incapability to detect the direction of the edge. On the other hand, because it is the second derivate it produces double peak (positive and negative impulse). By Laplacian, we can detect whether the pixel lies in the dark or bright side of the edge. This property can be used in image segmentation.

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

Figure : Mask for Laplacian (second derivate).

a) Laplacian Operator:

- It is a linear operator and is the simplest isotropic derivative operator, which is defined as,

$$\text{Laplacian} \quad \nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial^2 x} + \frac{\partial^2 f(x, y)}{\partial^2 y} \quad \text{-----}(1)$$

Where,

$$\frac{\partial^2 f(x, y)}{\partial^2 y} \approx f(x, y + 1) + f(x, y - 1) - 2f(x, y) \quad \text{-----(2)-}$$

$$\frac{\partial^2 f(x, y)}{\partial^2 x} \approx f(x + 1, y) + f(x - 1, y) - 2f(x, y) \quad \text{-----}(3)$$

Sub (2) & (3) in (1), we get,

$$\nabla f^2 \cong [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)] \text{ -----(4)}$$

The above digital laplacian equation is implemented using filter mask,

0	1	0
1	-4	1
0	1	0

The digital laplacian equation including the diagonal terms is given as,

$$\nabla f^2 \cong \left[\begin{array}{l} f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) + f(x+1, y+1) + f(x+1, y-1) \\ + f(x-1, y-1) + f(x-1, y+1) - 8f(x, y) \end{array} \right]$$

the corresponding filter mask is given as,

1	1	1
1	-8	1
1	1	1

Similarly , two other implementations of the laplacian are,

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

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

3. Explain Edge linking and boundary detection in detail

Edge linking

Edge detection algorithm is followed by linking procedures to assemble edge pixels into meaningful edges.

■ Basic approaches

1. Local Processing
2. Global Processing via the Hough Transform
3. Global Processing via Graph-Theoretic Techniques

1. Local processing

Analyze the characteristics of pixels in a small neighborhood (say, 3x3, 5x5) about every edge pixels (x,y) in an image. All points that are similar according to a set of predefined criteria are linked, forming an edge of pixels that share those criteria

Criteria

1. The strength of the response of the gradient operator used to produce the edge pixel
 - an edge pixel with coordinates (x₀,y₀) in a predefined neighborhood of (x,y) is similar in magnitude to the pixel at (x,y) if
$$|\nabla f(x,y) - \nabla f(x_0,y_0)| \leq E$$
2. The direction of the gradient vector
 - an edge pixel with coordinates (x₀,y₀) in a predefined neighborhood of (x,y) is similar in angle to the pixel at (x,y) if
$$|\alpha(x,y) - \alpha(x_0,y_0)| < A$$
3. 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.**
4. The process is repeated at every location in the image
5. A record must be kept
6. Simply by assigning a different gray level to each set of linked edge pixels.
7. find rectangles whose sizes makes them suitable candidates for license plates
8. Use horizontal and vertical Sobel operators ,eliminate isolated short segments
- link conditions:

gradient value > 25

gradient direction differs < 15°

2. Global Processing via the Hough Transform

Suppose that we have an image consisting of several samples of a straight line, Hough [1962] proposed a method (commonly referred as *Hough transform*) for finding the line (or lines) among these samples (or edge pixels). Consider a point (x_i, y_i). There is an infinite number of lines passing through this point, however, they all can be described by

$$y_i = a \cdot x_i + b$$

This means that all the lines passing (x_i, y_i) are defined by two parameters, *a* and *b*. The equation can be rewritten as

$$b = -x_i \cdot a + y_i$$

Now, if we consider *b* as a function of *a*, where x_i and y_i are two constants,

the parameters can be represented by a single line in the ab -plane. The Hough transform is a process where each pixel sample in the original xy -plane is transformed to a line in the ab -plane. Consider two pixels (x_1, y_1) and (x_2, y_2) . Suppose that we draw a line L across these points. The transformed lines of (x_1, y_1) and (x_2, y_2) in the ab -plane intersect each other in the point (a', b') , which is the description of the line L , see Figure . In general, the more lines cross point (a, b) the stronger indication that there is a line $y = a \cdot x + b$ in the image.

To implement the Hough transform a two-dimensional matrix is needed, see Figure. In each cell of the matrix there is a counter of how many lines cross that point. Each line in the ab -plane increases the counter of the cells that are along its way. A problem in this implementation is that both the slope (a) and intercept (b) approach infinity as the line approaches the vertical. One way around this difficulty is to use the normal representation of a line:

$$x \cdot \cos \alpha + y \cdot \sin \alpha = d$$

Here d represents the shortest distance between origin and the line. α represents the angle of the shortest path in respect to the x -axis. Their corresponding ranges are $[0, \sqrt{2}D]$, and $[-90, 90]$, where D is the distance between corners in the image. Although the focus has been on straight lines, the Hough transform is applicable to any other shape. For example, the points lying on the circle

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2$$

can be detected by using the approach just discussed. The basic difference is the presence of three parameters (c_1, c_2, c_3) , which results in a 3-dimensional parameter space.

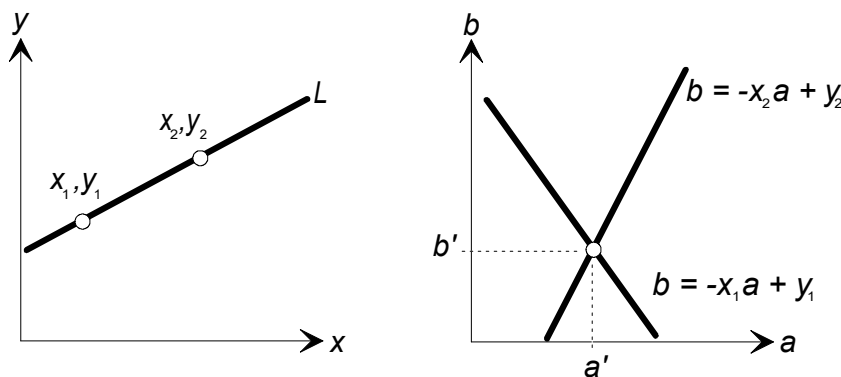


Figure : Hough transform: xy -plane (left); ab -plane (right).

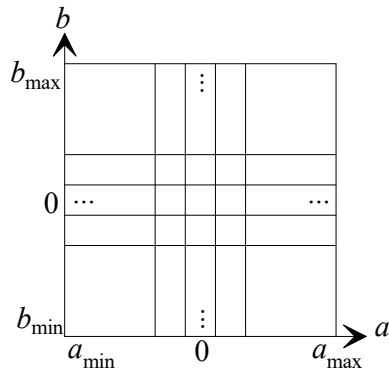


Figure : Quantization of the parameter plane for use in Hough transform.

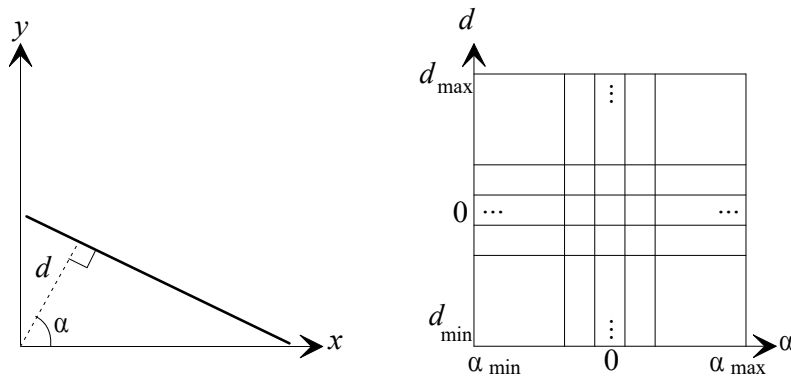


Figure : Normal representation of a line (left); quantization of the d -plane into cells.

HOUGH TRANSFORM STEPS:

1. Compute the gradient of an image and threshold it to obtain a binary image.
2. Specify subdivisions in the $\rho\theta$ -plane.
3. Examine the counts of the accumulator cells for high pixel concentrations.
4. Examine the relationship (principally for continuity) between pixels in a chosen cell.
5. based on computing the distance between disconnected pixels identified during traversal of the set of pixels corresponding to a given accumulator cell.
6. a gap at any point is significant if the distance between that point and its closet neighbor exceeds a certain threshold.
7. link criteria:
 - 1). the pixels belonged to one of the set of pixels linked according to the highest count
 - 2). no gaps were longer than 5 pixels

4. Explain Thresholding in detail.

Thresholding

- Suppose that the image consists of one or more objects and background, each having distinct gray-level values. The purpose of *thresholding* is to separate these areas from each other by using the information given by the histogram of the image.
- If the object and the background have clearly uniform gray-level values, the object can easily be detected by splitting the histogram using a suitable threshold value.

- **The threshold is considered as a function of:**

$$T = g(i, j, x_{i,j}, p(i, j))$$

where i and j are the coordinates of the pixel, $x_{i,j}$ is the intensity value of the pixel, and $p(i, j)$ is a function describing the neighborhood of the pixel (e.g. the average value of the neighboring pixels).

- **A threshold image is defined as:**

$$f(x, y) = \begin{cases} 1 & \text{if } x_{i,j} > T \\ 0 & \text{if } x_{i,j} \leq T \end{cases}$$

Pixels labeled 1 corresponds to objects and Pixels labeled 0 corresponds to background

- Multilevel thresholding classifies a point $f(x, y)$ belonging to

- to an object class if $T_1 < f(x, y) \leq T_2$
- to another object class if $f(x, y) > T_2$
- to background if $f(x, y) \leq T_1$

- When T depends on

- only $f(x, y)$: only on gray-level values \Rightarrow Global threshold
- both $f(x, y)$ and $p(x, y)$: on gray-level values and its neighbors \Rightarrow Local threshold
- the spatial coordinates x and y \Rightarrow dynamic or Adaptive threshold

An example of thresholding and the corresponding histogram is shown below

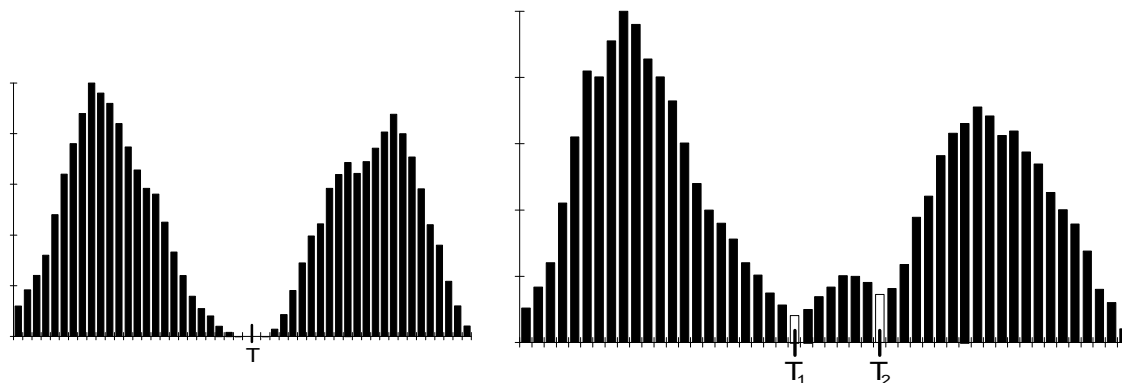


Figure : Histogram partitioned with one threshold (single threshold), and with two thresholds (multiple thresholds).

The thresholding can be classified according to the parameters it depends on:

Thresholding:	Effective parameters:
Global	$x_{i,j}$
Local	$x_{i,j}, p(i,j)$
Dynamic	$x_{i,j}, p(i,j), i, j$

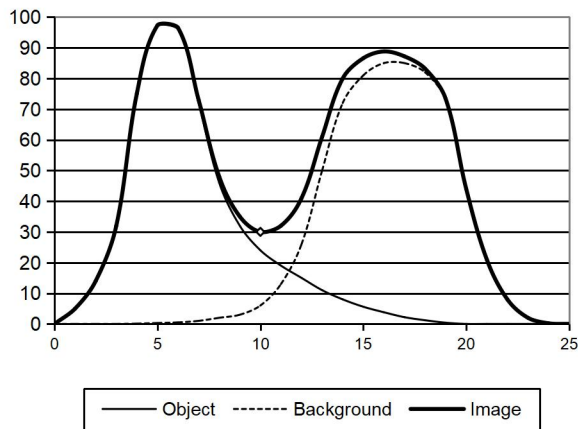


Figure : Example of histogram with two overlapping peaks.

1. Basic Global Thresholding:

This is based on visual inspection of histogram

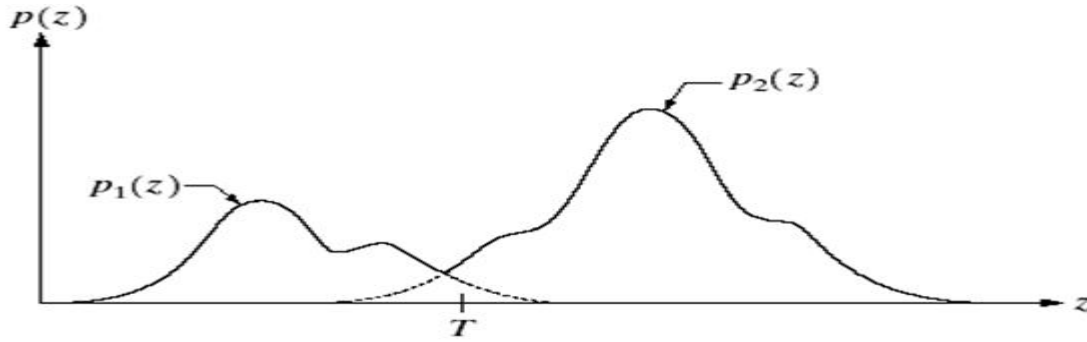
- Select an initial estimate for T .
- Segment the image using T . This will produce two groups of pixels: G_1 consisting of all pixels with gray level values $> T$ and G_2 consisting of pixels with gray level values $\leq T$
- Compute the average gray level values μ_1 and μ_2 for the pixels in regions G_1 and G_2
- Compute a new threshold value
- $T = 0.5 (\mu_1 + \mu_2)$
- Repeat steps 2 through 4 until the difference in T in successive iterations is smaller than a predefined parameter T_0 .

2. Basic Adaptive Thresholding:

- Subdivide original image into small areas.
- Utilize a different threshold to segment each sub images.
- Since the threshold used for each pixel depends on the location of the pixel in terms of the sub images, this type of thresholding is adaptive.

3. Optimal Global and Adaptive Thresholding

- Consider the following two probability density functions



- Assume the larger PDF to correspond to the back ground level and the smaller PDF to gray levels of objects in the image. The overall gray-level variation in the image,

$$p(z) = P_1 p_1(z) + P_2 p_2(z)$$

$$P_1 + P_2 = 1$$

Where P1 is the pbb that the pixel is an object pixel, P2 is the pbb that the pixel is a background pixel.

- The probability of error classifying a background point as an object point is,as

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

This is the area under the curve of P2(z) to the left of the threshold

- The probability of error classifying a object point as an background point is,as

$$E_2(T) = \int_T^{\infty} p_1(z) dz$$

This is the area under the curve of P1(z) to the right of the threshold

- Then the overall probability of error is $E(T) = P_2 E_1(T) + P_1 E_2(T)$

- To find the minimum error,

$$\frac{dE(T)}{dT} = \frac{d(P_2 E_1(T) + P_1 E_2(T))}{dT} = 0$$

- The result is

$$P_1 p_1(T) = P_2 p_2(T)$$

- If the two PDFs are Gaussian then,

$$\begin{aligned} p(z) &= P_1 p_1(z) + P_2 p_2(z) \\ &= \frac{P_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(z-\mu_1)^2}{2\sigma_1^2}} + \frac{P_2}{\sqrt{2\pi}\sigma_2} e^{-\frac{(z-\mu_2)^2}{2\sigma_2^2}} \end{aligned}$$

where

μ_1 and σ_1^2 are the mean and variance of the Gaussian density of one object

μ_2 and σ_2^2 are the mean and variance of the Gaussian density of the other object

Substituting equation (2) in (1), the solution for the threshold T is given as ,

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

$$\text{where } A = \sigma_1^2 - \sigma_2^2$$

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

$$C = \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 2\sigma_1^2\sigma_2^2 \ln(\sigma_2P_1 / \sigma_1P_2)$$

- Since the quadratic equation has two possible solutions, two threshold values are required to obtain the optimal solution.
- If the variances are equal, a single threshold is sufficient given as,

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

Threshold selection Based on Boundary Characteristics:

- The threshold value selection depends on the peaks in the histogram
- The threshold selection is improved by considering the pixels lying on or near the boundary between the objects and background
- A three level image can be formed using the gradient, which is given as,

Light object of dark background

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

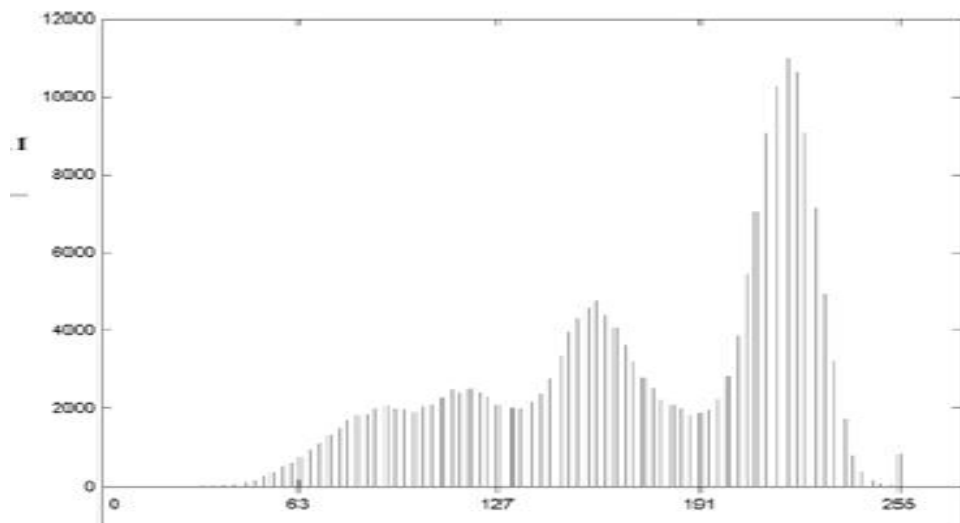
- **all pixels that are not on an edge are labeled 0**
- **all pixels that are on the dark side of an edge are labeled +**
- **all pixels that are on the light side an edge are labeled -**
- The transition from light back ground to dark object is represented by the occurrence of ‘-’ followed by ‘+’
- The transition from object to the back ground is represented by the occurrence of ‘+’ followed by ‘-’
- Thus a horizontal or vertical scan line containing a section of an object has the following structure: (...)(-,+)(0 or +)(+,-)(...)

5. Write short notes on Region growing segmentation. Or Explain Region splitting and Merging (May / June 2009)

Region growing is a procedure that groups pixels or sub regions 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 neighboring pixels that have properties similar to the seed

Region growing (or *pixel aggregation*) starts from an *initial region*, which can be a group of pixels (or a single pixel) called **seed points**. The region is then extended by including new pixels into it. The included pixels must be neighboring pixels to the region and they must meet a **uniform criterion**.

criteria: (i) The absolute gray-level difference between any pixel and the seed has to be less than 65 and (ii) the pixel has to be 8-connected to at least one pixel in that region (if more, the regions are merged) .This is selected from the following histogram



The growing continues until a **stopping rule** takes effect. The region growing has several practical problems:

- how to select the seed points
- growing rule (or uniform criterion)
- stopping rule.

All of these parameters depend on the application. In the infra-red images in military applications the interesting parts are the areas which are hotter compared to their surroundings, thus a natural choice for the seed points would be the brightest pixel(s) in the image. In interactive applications the seed point(s) can be given by the user who gives (by pointing with the mouse, for example) any pixels inside the object, and the computer then finds the desired area by region growing algorithm.

The growing rule can be the differences of pixels. The value of the pixel in consideration can be compared to the value of its nearest neighbor inside the region.

If the difference does not exceed a predefined threshold the pixel is attended to the region. Another criterion is to examine the variance of the region, or gradient of the pixel. Several growing criteria worth consideration are summarized in the following:

Global criteria:

Difference between $x_{i,j}$ and the average value of the region.

The *variance* σ^2 of the region if $x_{i,j}$ is attended to it.

Local criteria:

Difference between $x_{i,j}$ and the nearest pixel inside the region.

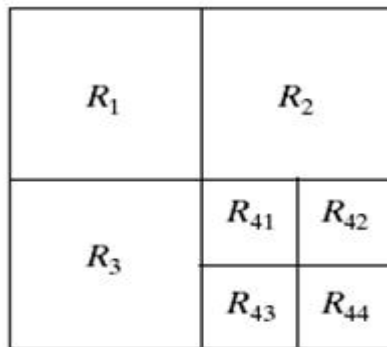
The *variance* of the local neighborhood of pixel $x_{i,j}$. If the variance is low enough (meaning that the pixel belongs to a homogenous area) it will be attended to the region.

Gradient of the pixel. If it is small enough (meaning that the pixel does not belong to an edge) it will be attended to the region.

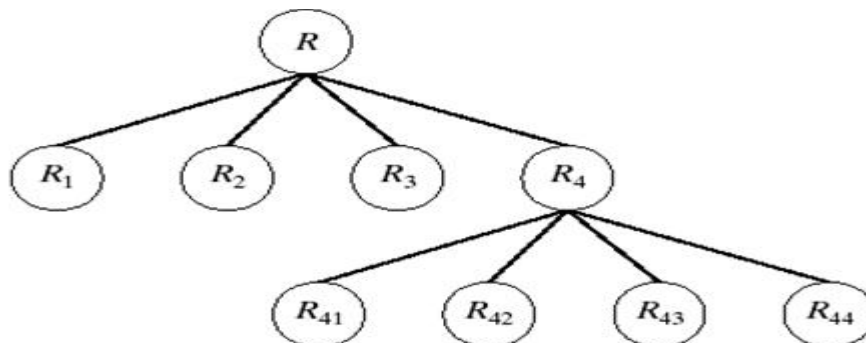
The growing will stop automatically when there are no more pixels that meet the uniform criterion. The stopping can also take effect whenever the size of the region gets too large (assuming a priori knowledge of the object's preferable size). The shape of the object can also be used as a stopping rule, however, this requires some form of pattern recognition and prior knowledge of the object's shape.

Splitting and merging

- The quad tree segmentation is a simple but still powerful tool for image decomposition.
- In this technique , an image is divided into various sub images of disjoint regions and then merge the connected regions together.
- **Steps involved in splitting and merging**
 - Split into 4 disjoint quadrants any region R_i for which $P(R_i)=FALSE$.
 - Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)=TRUE$.
 - Stop when no further merging or splitting is possible.
- The following shows the partitioned image,



- Corresponding quad tree



➤ An **example of quad tree-based** split and merge is shown below

- In the first phase the image is decomposed into four sub blocks (a), which are then further divided, except the top leftmost block (b).
- At the third phase the splitting phase is completed (c).
- At the merging phase the neighboring blocks that are uniform and have the same color are merged together.
- The segmentation results to only two regions (d), which is ideal in this favorable example.

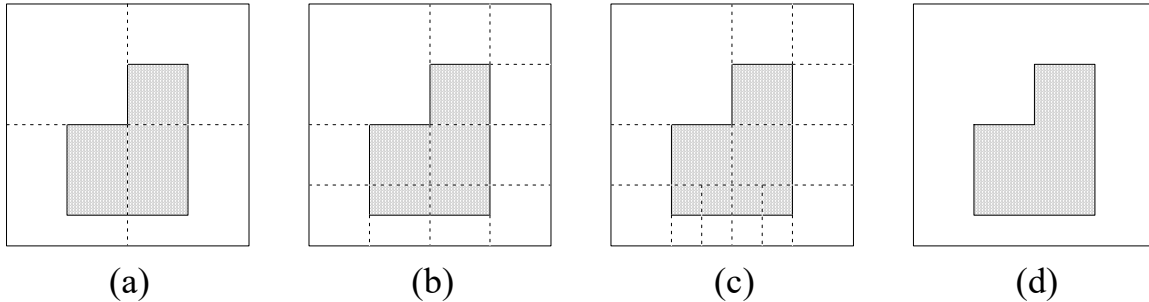


Figure : Example of quadtree based splitting and merging.

6. Explain in detail about Segmentation by Morphological Watersheds

The aim of Segmentation is to separate regions wrt brightness, color, reflectivity, texture, etc. Segmentation based on three principal concepts: (a) detection of discontinuities, (b) thresholding, and (c) region processing.

1. Basic Concepts

The concept of watersheds is based on visualizing an image in three dimensions: two spatial coordinates versus gray levels. In such a "topographic" interpretation, we consider three types of points:

- a. Points belonging to a regional minimum
- b. Catchment basin / watershed of a regional minimum

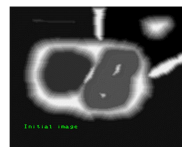
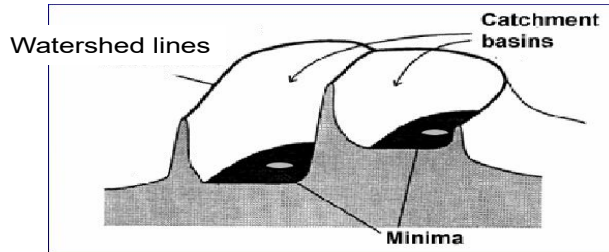
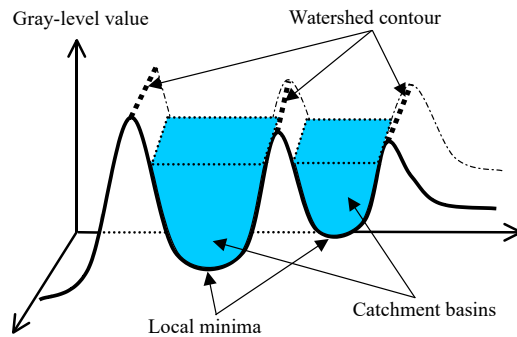
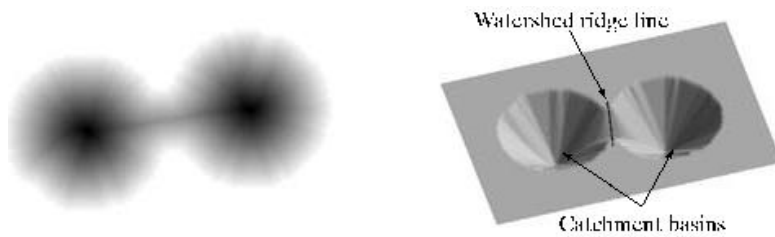
Points at which a drop of water will certainly fall to a single minimum

- c. Divide lines / Watershed lines

Points at which a drop of water will be equally likely to fall to more than one minimum. Crest lines on the topographic surface

This technique is to identify all the third type of points for segmentation

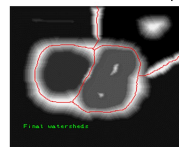
1. Piercing holes in each regional minimum of I
2. The 3D topography is flooded from below gradually
3. When the rising water in distinct catchment basins is about to merge, a dam is built to prevent the merging
4. The dam boundaries correspond to the watershed lines to be extracted by a Watershed segmentation algorithm



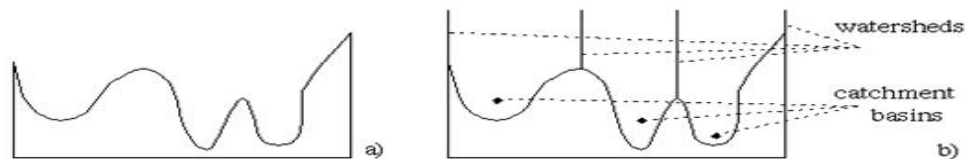
Initial image



Topographic surface



Final watersheds



One-dimensional example of watershed segmentation. (a) Gray level profile of image data. (b) Watershed segmentation – local minima of gray level (altitude) yield catchment basins, local maxima define the watershed lines.

The principal objective of segmentation algorithms based on these concepts is to find the watershed lines. The basic idea is simple: Suppose that a hole is punched in each regional minimum and that the entire topography is flooded from below by letting water

rise through the holes at a uniform rate. When the rising water in distinct catchment basins is about to merge, a dam is built to prevent the merging. The flooding will eventually reach a stage when only the tops of the dams are visible above the water line. These dam boundaries correspond to the divide lines of the watersheds. Therefore, they are the (continuous) boundaries extracted by a watershed segmentation algorithm.



These ideas can be explained further with the aid of Fig.. Figure (a) shows a simple gray-scale image and Fig. (b) is a topographic view, in which the height of the "mountains" is proportional to gray-level values in the input image. For ease of interpretation, the backsides of structures are shaded. This is not to be confused with gray-level values; only the general topography of the three-dimensional representation is of interest. In order to prevent the rising water from spilling out through the edges of the structure, we imagine the perimeter of the entire topography (image) being enclosed by dams of height greater than the highest possible mountain, whose value is determined by the highest possible gray-level value in the input image.

Suppose that a *hole* is punched in each regional minimum [shown as dark areas in Fig. (b)] and that the entire topography is flooded from below by letting water rise through the holes *at* a uniform rate. Figure (c) shows the first stage of flooding, where the "water" shown in light gray, has covered only areas that correspond to the very dark background in the image. In Figs. (d) and (e) we see that the water now has risen into the first and second catchment basins, respectively. As the water continues to rise, it will eventually overflow

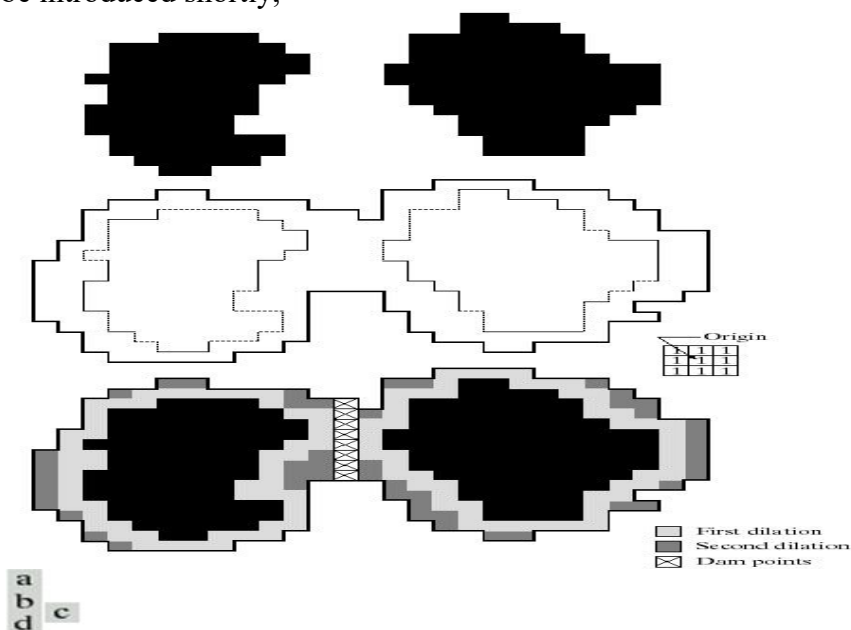
from one catchment basin into another. The first indication of this is shown in (f). Here, water from the left basin actually overflowed into the basin on the right and a short "dam" (consisting of single pixels) was built to prevent water from merging at that level of flooding (the details of dam building are discussed in the following section). The effect is more pronounced as water continues to rise, as shown in Fig. (g). This figure shows a longer dam between the two catchment basins and another dam in the top part of the right basin. The latter dam was built to prevent merging of water from that basin with water from areas corresponding to the background. This process is continued until the maximum level of flooding (corresponding to the highest gray-level value in the image) is reached. The final dams correspond to the watershed lines, which are the desired segmentation result. The result for this example is shown in Fig. (h) as a dark, one-pixel-thick path superimposed on the original image. Note the important property that the watershed lines form a connected path, thus giving continuous boundaries between regions.

One of the principal applications of watershed segmentation is in the extraction of nearly uniform (blob like) objects from the background. Regions characterized by small variations in gray levels have small gradient values. Thus, in practice, we often see watershed segmentation applied to the gradient of an image, rather than to the image itself. In this formulation, the regional minima of catchment basins correlate nicely with the small value of the gradient corresponding to the objects of interest.

2. Dam Construction

Dam construction is based on binary images, which are members of 2-D integer space Z^2 . The simplest way to construct dams separating sets of binary points is to use morphological dilation.

The basics of how to construct dams using dilation are illustrated in Fig. Figure (a) shows portions of two catchment basins at flooding step $n - 1$ and Fig. (b) shows the result at the next flooding step, n . The water has spilled from one basin to the other and, therefore, a dam must be built to keep this from happening. In order to be consistent with notation to be introduced shortly,



(a) Two partially flooded catchment basins at stage $n - 1$ of flooding. (b) Flooding at stage n , showing that water has spilled between basins (for clarity, water is shown in white rather than black). (c) Structuring element used for dilation. (d) Result of dilation and dam construction.

let M_1 and M_2 denote the sets of coordinates of points in two regional minima. Then let the set of coordinates of points in the *catchment basin* associated with these two minima at stage $n - 1$ of flooding be denoted by $C_{n-1}(M_1)$ and $C_{n-1}(M_2)$ respectively.

These are the two black regions shown in Fig. (a). Suppose that each of the connected components in Fig.(a) is dilated by the structuring element shown in Fig. (c), subject to two conditions:

- (1) The dilation has to be constrained to q (this means that the center of the structuring element can be located only at points in q during dilation), and
- (2) The dilation cannot be performed on points that would cause the sets being dilated to merge (become a single connected component).

Figure (d) shows that a first dilation pass (in light gray) expanded the boundary of each original connected component. Note that condition (1) was satisfied by every point during dilation, and condition (2) did not apply to any point during the dilation process; thus the boundary of each region was expanded uniformly.

Let the union of these two sets be denoted by $C[n - 1]$, There are two connected components in Fig. (a) and only one connected component in Fig. (b), This connected component encompasses the earlier two components, shown dashed. The fact that two connected components have become a *single* component indicates that water between the two catchment basins has merged at flooding step n . Let this connected component be denoted q . Note that the two components from step $n - 1$ can be extracted from q by performing the simple AND operation $C_{n-1}(M_1) \cap q$ and $C_{n-1}(M_2) \cap q$. We note also that all points belonging to an individual catchment basin form a single connected component.

In the second dilation (shown in medium gray), several points failed condition (1) while meeting condition (2), resulting in the broken perimeter shown in the figure. It also is evident that the only points in q that satisfy the two conditions under consideration describe the one-pixel-thick connected path shown crossed-hatched in Fig. (d). This path constitutes the desired separating dam at stage n of flooding. Construction of the dam at this level of flooding is completed by setting all the points in the path just determined to a value greater than the maximum gray-level value of the image. The height of all dams is generally set at 1 plus the maximum allowed value in the image. This will prevent water from crossing over the part of the completed dam as the level of flooding is increased. It is important to note that dams built by this procedure, which are the desired segmentation boundaries, are connected components. In other words, this method eliminates the problems of broken segmentation lines,

Although the procedure just described is based on a simple example, the method used for more complex situations is exactly the same, including the use of the 3 X 3 symmetric structuring elements shown in Fig. (c).

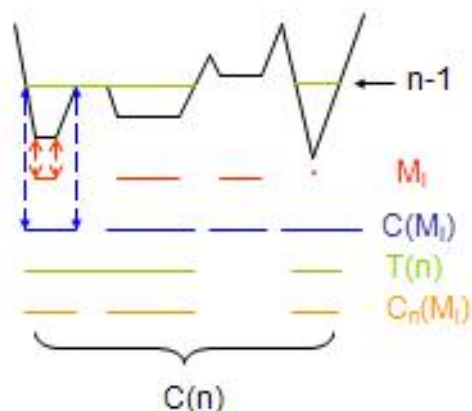
-
1. At each step of the algorithm, the binary image is obtained in the following manner
 1. Initially, the set of pixels with minimum gray level are 1, others 0.
 2. In each subsequent step, we flood the 3D topography from below and the pixels covered by the rising water are 1s and others 0s. M_1, M_2 :
 - Sets of coordinates of points in the two regional minima
 2. $C_{n-1}(M_1), C_{n-1}(M_2)$
 - Sets of coordinates of points in the catchment basins associated with M_1, M_2 at stage $n-1$ of flooding (catchment basins up to the flooding level)
 3. $C[n-1]$

- Union of $C_{n-1}(M_1), C_{n-1}(M_2)$
4. At flooding step $n-1$, there are two connected components. At flooding step n , there is only one connected component
- This indicates that the water between the two catchment basins has merged at flooding step n
 - Use “ q ” to denote the single connected component
5. Steps
- Repeatedly dilate $C_{n-1}(M_1), C_{n-1}(M_2)$ by the 3×3 structuring element shown, subject to the following condition
 - Constrained to q (center of the structuring element can not go beyond q during dilation)
6. The dam is constructed by the points on which the dilation would cause the sets being dilated to merge.
- Resulting one-pixel thick connected path
7. Setting the gray level at each point in the resultant path to a value greater than the maximum gray value of the image. Usually $\max+1$

7. Explain Watershed Segmentation Algorithm in detail

Watershed Transform

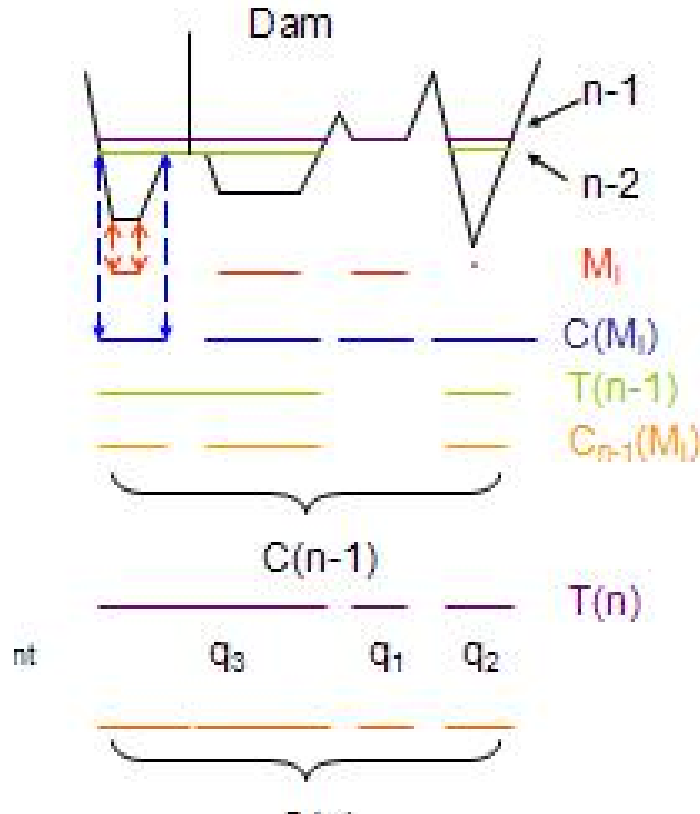
1. Denote $C_n(M_i)$ as the set of coordinates of points in the catchment basin associated with minimum M_i at flooding stage n .
 - $C_n(M_i) = C(M_i) \cap T[n]$
 - $C_n(M_i) \subseteq T[n]$
2. Denote $C[n]$ as the union of the flooded catchment basin portions at stage n :
 - Initialization
 - Let $C[\min+1] = T[\min+1]$
3. At each step n , assume $C[n-1]$ has been constructed. The goal is to obtain $C[n]$ from $C[n-1]$



Denote $Q[n]$ as the set of connected components in $T[n]$.

4. For each $q \in Q[n]$, there are three possibilities
 1. $q \cap C[n-1]$ is empty (q_1)
 - A new minimum is encountered
 - q is incorporated into $C[n-1]$ to form $C[n]$

2. $q \cap C[n-1]$ contains one connected component of $C[n-1]$ (q_2)
 - q is incorporated into $C[n-1]$ to form $C[n]$
3. $q \cap C[n-1]$ contains more than one connected components of $C[n-1]$ (q_3)
 - A ridge separating two or more catchment basins has been encountered
 - A dam has to be built within q to prevent overflow between the catchment basins
4. Repeat the procedure until $n=\max+1$



WATERSHED SEGMENTATION ALGORITHM

1. Let $M_1, M_2, M_3, \dots, M_n$ be the sets of coordinates of points in the regional minima of the image $g(x,y)$
2. $C(M_i)$ be the coordinates of points of the catchment basin associated with regional minima M_i

$$T[n] = \{ (s,t) \mid g(s,t) < n \}$$

$T[n]$ = Set of points in $g(x,y)$ which are lying below the plane $g(x,y) = n$

n = Stage of flooding, varies from $\min+1$ to $\max+1$

\min = minimum gray level value

\max = maximum gray level value

3. Let $C_n(M_i)$ be the set of points in the catchment basin associated with M_i that are flooded at stage n .

$$C_n(M_i) = C(M_i) \cap T[n]$$

$C_n(M_i) = 1$ at location (x,y) if $(x,y) \in C(M_i)$ AND $(x,y) \in T[n]$, otherwise it is 0.

4. $C[n]$ be the union of flooded catchment basin portions at the stage n

$$C[n] = \bigcup_{i=1}^R C_n(M_i)$$

$$C[\max + 1] = \bigcup_{i=1}^R C(M_i)$$

5. Algorithm keeps on increasing the level of flooding, and during the process $C_n(M_i)$ and $T[n]$ either increase or remain constant.

Algorithm initializes $C[\min + 1] = T[\min + 1]$, and then proceeds recursively assuming that at step n $C[n-1]$ has been constructed.

6. Let Q be set of connected components in $T[n]$. For each connected component $q \in Q[n]$, there are three possibilities:

- (a) $q \cap C[n-1]$ is empty.
- (b) $q \cap C[n-1]$ contains one connected component of $C[n-1]$.
- (c) $q \cap C[n-1]$ contains more than one connected component of $C[n-1]$.

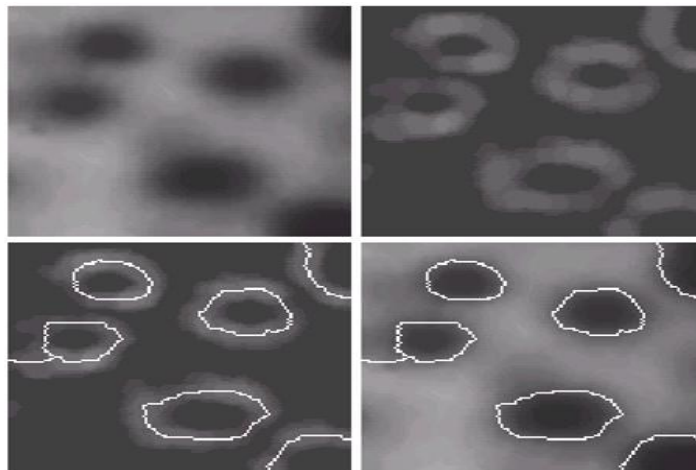
7. Condition (a) occurs when a new minima is encountered, in this case q is added to set $C[n-1]$ to form $C[n]$.

8. Condition (b) occurs when q lies within a catchment basin of some regional minima, in that case

9. Condition (c) occurs when ridge between two catchment basins is hit and further flooding will cause the waters from two basins will merge, so a dam must be built within q .

a b
c d

(a) Image of blobs. (b) Image gradient.
(c) Watershed lines.
(d) Watershed lines superimposed on original image.



Consider the image and its gradient, shown in Figs. (a) and (b), respectively. Application of the watershed algorithm just described yielded the watershed lines (white paths) of the gradient image shown in Fig. (c). These segmentation boundaries are shown superimposed on the original image in Fig. (d). As noted at the beginning of this section, (the segmentation boundaries have the important property of being connected paths).

A morphological region growing approach.

Seed points:

- local minima points

Growing method:

- Dilation

Predicates

- Similar gradient values

Sub-region boundary

- Dam building

To avoid over-segmentation

- Use markers

Disadv: Concept is difficult to understand

8. Explain Morphological Image Processing in details

Morphological Image Processing is an important tool in the Digital Image processing, since that science can rigorously quantify many aspects of the geometrical structure of the way that agrees with the human intuition and perception.

Morphologic image processing technology is based on geometry. It emphasizes on studying geometry structure of image. We can find relationship between each part of image. When processing image with morphological theory. Accordingly we can comprehend the structural character of image in the morphological approach an image is analyzed in terms of some predetermined geometric shape known as structuring element.

Morphological processing is capable of removing noise and clutter as well as the ability to edit an image based on the size and shape of the objects of interest. Morphological Image Processing is used in the place of a Linear Image Processing, because it sometimes distort the underlying geometric form of an image, but in Morphological image Processing, the information of the image is not lost.

In the Morphological Image Processing the original image can be reconstructed by using Dilation, Erosion, Opening and Closing operations for a finite no of times. The major objective of this paper is to reconstruct the class of such finite length Morphological Image Processing tool in a suitable mathematical structure using Java language.

Binary Image:-

The image data of Binary Image is Black and White. Each pixel is either '0' or '1'. A digital Image is called Binary Image if the grey levels range from 0 and 1.

Ex: A Binary Image shown below is,

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

(A 4* 5 Binary Image)

DILATION:-



Dilation - grow image regions

Dilation causes objects to dilate or grow in size. The amount and the way that they grow depends upon the choice of the structuring element [3]. Dilation makes an object larger by adding pixels around its edges.

The Dilation of an Image 'A' by a structuring element 'B' is written as $A \oplus B$. To compute the Dilation, we position 'B' such that its origin is at pixel co-ordinates (x , y) and apply the rule.

$$g(x, y) = \begin{cases} 1 & \text{if 'B' hits 'A'} \\ 0 & \text{Otherwise} \end{cases}$$

Repeat for all pixel co-ordinates. Dilation creates new image showing all the location of a structuring element origin at which that structuring element HITS the Input Image. In this it adds a layer of pixel to an object, there by enlarging it. Pixels are added to both the inner and outer boundaries of regions, so Dilation will shrink the holes enclosed by a single region and make the gaps between different regions smaller. Dilation will also tend to fill in any small intrusions into a region's boundaries.

The results of Dilation are influenced not just by the size of the structuring element but by its shape also.

Dilation is a Morphological operation; it can be performed on both Binary and Grey Tone Images. It helps in extracting the outer boundaries of the given images.

For Binary Image:-

Dilation operation is defined as follows,

$$D(A, B) = A \oplus B$$

Where,

A is the image

B is the structuring element of the order 3 * 3.

Many structuring elements are requested for Dilating the entire image.

EROSION:-



Erosion - shrink image regions

Erosion causes objects to shrink. The amount of the way that they shrink depend upon the choice of the structuring element. Erosion makes an object smaller by removing or Eroding away the pixels on its edges [3].

The Erosion of an image 'A' by a structuring element 'B' is denoted as $A \ominus B$. To compute the Erosion, we position 'B' such that its origin is at image pixel co-ordinate (x, y) and apply the rule.

$$g(x, y) = \begin{cases} 1 & \text{if 'B' Fits 'A',} \\ 0 & \text{otherwise} \end{cases}$$

Repeat for all x and y or pixel co-ordinates. Erosion creates new image that marks all the locations of a Structuring elements origin at which that Structuring Element Fits the input image. The Erosion operation seems to strip away a layer of pixels from an object, shrinking it in the process. Pixels are eroded from both the inner and outer boundaries of regions. So, Erosion will

enlarge the holes enclosed by a single region as well as making the gap between different regions larger. Erosion will also tend to eliminate small extrusions on a regions boundaries.

The result of erosion depends on Structuring element size with larger Structuring elements having a more pronounced effect & the result of Erosion with a large Structuring element is similar to the result obtained by iterated Erosion using a smaller structuring element of the same shape.

Erosion is the Morphological operation, it can be performed on Binary and Grey images. It helps in extracting the inner boundaries of a given image.

For Binary Images:-

Erosion operation is defined as follows,

$$E(A, B) = A \ominus B$$

Where, A is the image

B is the structuring element of the order 3×3 .

Many structuring elements are required for eroding the entire image.

OPENING:-



Opening - structured removal of image region boundary pixels

It is a powerful operator, obtained by combining Erosion and Dilation. "Opening separates the Objects". As we know, Dilation expands an image and Erosion shrinks it [3]. Opening generally smoothes the contour of an image, breaks narrow Isthmuses and eliminates thin Protrusions [1].

The Opening of an image 'A' by a structuring element 'B' is denoted as $A \circ B$ and is defined as an Erosion followed by a Dilation, and is written as [3],

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

Opening operation is obtained by doing Dilation on Eroded Image. It is to smoothen the curves of the image. Opening spaces objects that are too close together, detaches objects that are touching and should not be, and enlarges holes inside objects.

Opening involves one or more Erosions followed by one Dilation.

CLOSING:-



Closing - structured filling in of image region boundary pixels

It is a powerful operator, obtained by combining Erosion and Dilation. "Closing, join the Objects" [3]. Closing also tends to smooth sections of contours but, as opposed to Opening, it generally fuses narrow breaks and long thin Gulf's, eliminates small holes and fills gaps in the contour [1].

The Closing of an image 'A' by a structuring element 'B' is denoted as $A \bullet B$ and defined as a Dilation followed by an Erosion; and is written as [3],

$A \bullet B = (A \oplus B) \ominus B$. Closing is obtained by doing Erosion on Dilated image. Closing joins broken objects and fills in unwanted holes in objects.

Closing involves one or more Dilations followed by one Erosion.

FUTURE SCOPE:-

The Morphological Image Processing can be further applied to a wide spectrum of problems including:

- Medical image analysis: Tumor detection, measurement of size and shape of internal organs, Regurgitation, etc.
- Robotics: Recognition and interpretation of objects in a scene, motion control and execution through visual feedback
- Radar imaging: Target detection and identification.

UNIT V IMAGE COMPRESSION AND RECOGNITION

Need for data compression, Huffman, Run Length Encoding, Shift codes, Arithmetic coding, JPEG standard, MPEG. Boundary representation, Boundary description, Fourier Descriptor, Regional Descriptors – Topological feature, Texture - Patterns and Pattern classes - Recognition based on matching.

PART – A

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. Define Data Compression.

The term data compression refers to the process of reducing the amount of data required to represent a given quantity of information.

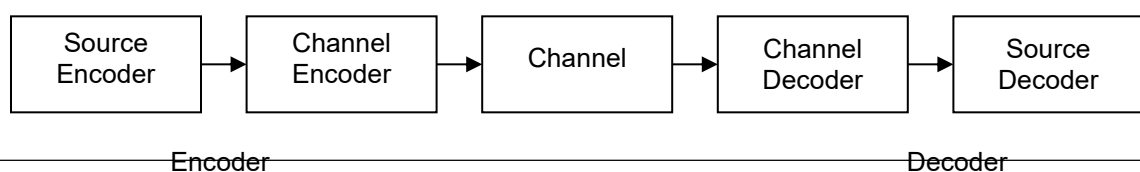
3. What are two main types of Data compression?

1. **Lossless compression** can recover the exact original data after compression. It is used mainly for compressing database records, spreadsheets or word processing files.
2. **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 digitized voice where losses outside visual perception can be tolerated.

4. What is the need for Compression?

1. 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 decompress it when it is retrieved.
2. 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.
3. 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. Show the block diagram of a general compression system model.



6. Define Relative data redundancy.

The relative data redundancy R_D of the first data set can be defined as

$$R_D = 1 - \frac{1}{C_R}$$

Where C_R , commonly called the compression ratio, is $C_R = \frac{n_1}{n_2}$

7. What are the three basic data redundancies can be identified and exploited?

The three basic data redundancies that can be identified and exploited are,

- i) Coding redundancy
- ii) Interpixel redundancy
- iii) Psychovisual redundancy

8. 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.

9. 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

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.

Encoder has 2 components they are 1. source encoder 2. 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 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?

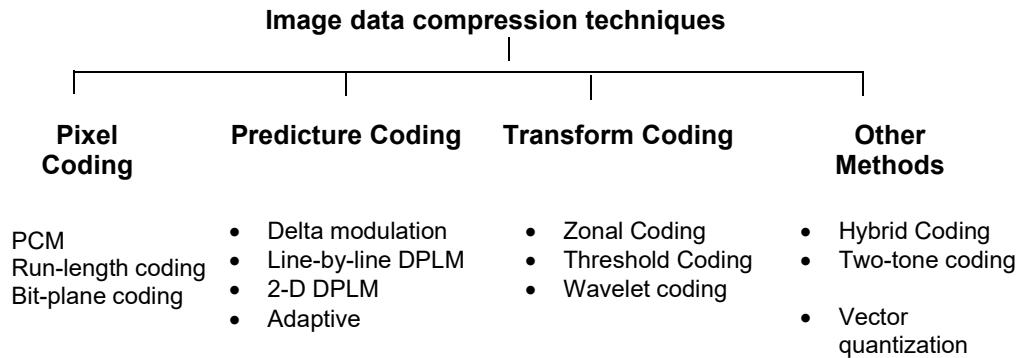
1. Source decoder- has two components

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

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

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

15. What are the basic image data compression techniques?



16. What is meant by Error-free compression?

Error-free compression (or) lossless compression is the only acceptable means of data reduction. Error free compression techniques generally are composed of two relatively independent operations:

- (1) Devising an alternative representation of the image in which its interpixel redundancies
- (2) The coding the representation to eliminate coding redundancies.

17. What are the techniques that are followed in lossless compression?

The coding techniques that are followed lossless compression.

- (1) Variable –length coding
- (2) LZW coding
- (3) Bit-plane coding
- (4) Lossless predictive coding.

18. What is use of variable-length coding?

The simplest approach to error-free image compression is to reduce only coding redundancy to reduce this redundancy it is required to construct a variable-length code that assigns the shortest possible code words to the most probable gray level. This is called as variable length coding.

19. What are coding techniques that are followed in variable length coding?

Some of the coding techniques that are used in variable length coding are

- (1) Huffman coding
- (2) Arithmetic coding

20. 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.

An effective alternative to constant area coding is to represent each row of an image (or) bit plane by a sequence of lengths that describe successive runs of black and white pixels. This technique referred to as run-length coding.

21. Define compression ratio& bit rate

$$\text{Bit rate} = \frac{\text{size of the compressed file}}{\text{pixels in the image}} = \frac{C}{N} \quad (\text{bits per pixel})$$

Where C is the number of bits in the compressed file, and N (=XY) is the number of pixels in the image. If the bit rate is very low, compression ratio might be a more practical measure:

$$\text{Compression ratio} = \frac{\text{size of the original file}}{\text{size of the compressed file}} = \frac{N \cdot k}{C}$$

Where k is the number of bits per pixel in the original image.

22. 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

23. 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.

24. Define Huffman coding

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

25. Define Block code and instantaneous code

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

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

26. Define uniquely decodable code and B2 code

A code word that is not a combination of any other codeword is said to be uniquely decodable code.

Each code word is made up of continuation bit c and information bit which are binary numbers. This is called B2 code or B code. This is called B2 code because two information bits are used for continuation bits

27. 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.

28. 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.

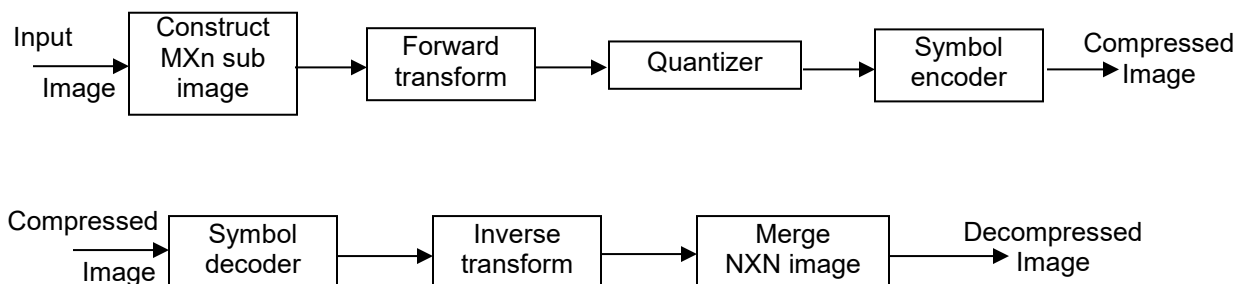
29. Define Bit-plane Coding.

One of the effective techniques for reducing an image's interpixel redundancies is to process the image's bit plane individually. This techniques is called bit-plane coding is based on the concept of decomposing a multilevel image into a series of binary image and compressing each binary image.

30. Define Transform coding.

Compression techniques that are based on modifying the transform an image. In transform coding, a reversible, linear transform is used to map the image into a set of transform coefficients which are then quantized and coded.

31. Show the block diagram of encoder& decoder of transform coding.



32. Define wavelet coding.

Wavelet coding is based on the idea that the coefficients of a transform that decorrelates the pixels of an image can be coded more efficiently than the original pixels themselves.

33. What is meant by 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, One of the most popular and comprehensive continuous tone, still frame compression standard is the JPEG standard.

34. 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.

35. What are the basic steps in JPEG?

The Major Steps in JPEG Coding involve:

- 1.DCT (Discrete Cosine Transformation)
- 2.Quantization
- 3.Zigzag Scan
- 4.DPCM on DC component
- 5.RLE on AC Components
- 6.Entropy Coding

36. 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

37. Define I-frame& P-frame& B-frame

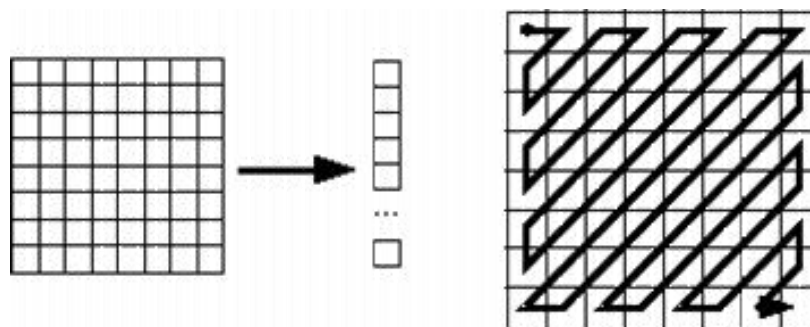
I-frame is Intraframe or Independent frame. An I-frame is compressed independently of all frames. It resembles a JPEG encoded image. It is the reference point for the motion estimation needed to generate subsequent P and P-frame.

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

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.

38. 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



PART B

1.Explain compression types

Now consider an encoder and a decoder as shown in Fig.. When the encoder receives the original image file, the image file will be converted into a series of binary data, which is called the bit-stream. The decoder then receives the encoded bit-stream and decodes it to form the decoded image. If the total data quantity of the bit-stream is less than the total data quantity of the original image, then this is called image compression.

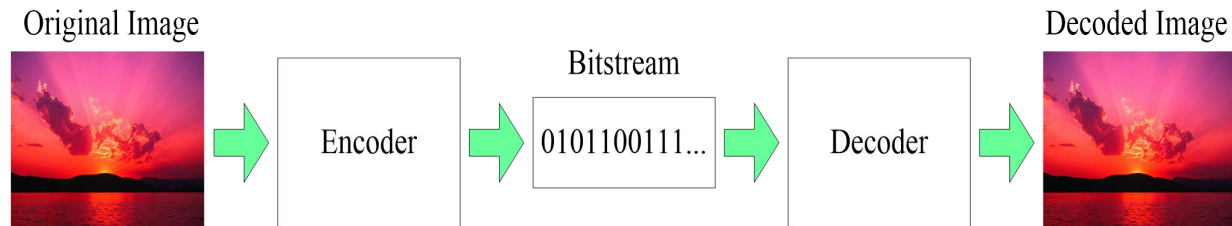


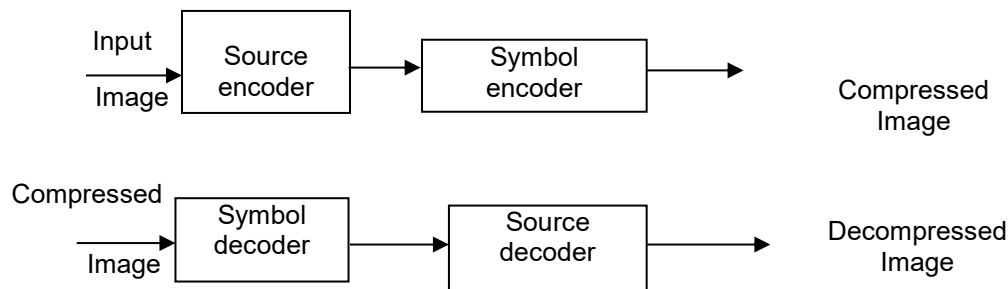
Fig. The basic flow of image compression coding

1.**Lossless compression** can recover the exact original data after compression. It is used mainly for compressing database records, spreadsheets or word processing files.

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

Applications: Medical imagery

A.The block diagram of encoder& decoder of Lossless compression



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

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

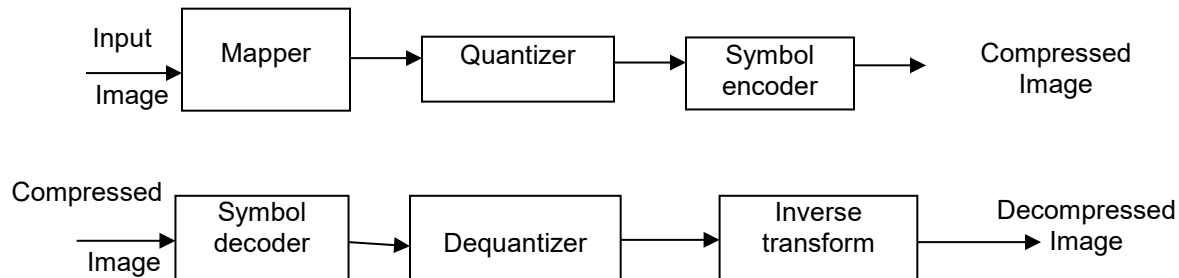
Source decoder- has two components

- a) Symbol decoder- This performs inverse operation of symbol encoder.
- b) Inverse mapping- This performs inverse operation of mapper.

Lossless Compression technique

Variable length coding (Huffmann,arithmetic), LZW coding,Bit Plane coding,Lossless Predictive coding

B.The block diagram of encoder& decoder of Lossy compression



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

1. Information loss is tolerable
- 2.Many-to-1 mapping in compression eg. quantization
- 3.Applications: commercial distribution (DVD) and rate constrained environment where lossless methods can not provide enough compression ratio

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.

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.

Lossy Compression technique

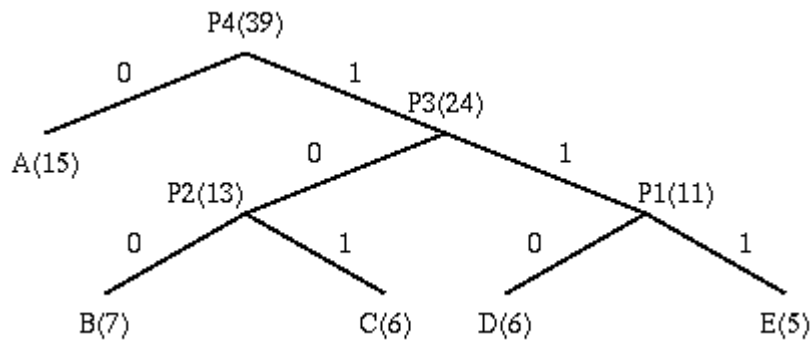
Lossy predictive coding (DPCM,Delta modulation), Transform coding, Wavelet coding.

2.Explain in detail the Huffman coding procedure with an example. (May / June 2009), (May / June 2006) (Variable length coding)

Huffman coding is based on the frequency of occurrence of a data item (pixel in images). The principle is to use a lower number of bits to encode the data that occurs more frequently. Codes are stored in a **Code Book** which may be constructed for each image or a set of images. In all cases the code book plus encoded data must be transmitted to enable decoding.

A bottom-up approach

1. Initialization: Put all nodes in an OPEN list, keep it sorted at all times (e.g., ABCDE).
2. Repeat until the OPEN list has only one node left:
 - (a) From OPEN pick two nodes having the lowest frequencies/probabilities, create a parent node of them.
 - (b) Assign the sum of the children's frequencies/probabilities to the parent node and insert it into OPEN.
 - (c) Assign code 0, 1 to the two branches of the tree, and delete the children from OPEN.



Symbol	Count	$\log(1/p)$	Code	Subtotal (# of bits)
A	15	1.38	0	15
B	7	2.48	100	21
C	6	2.70	101	18
D	6	2.70	110	18
E	5	2.96	111	15

TOTAL (# of bits): 87

a)Huffmann coding:

- i) Order the given symbols in the decreasing probability
- ii) Combine the bottom two or the least probability symbols into a single symbol that replaces in the next source reduction
- iii) Repeat the above two steps until the source reduction is left with two symbols per probability
- iv) Assign '0' to the first symbol and '1' to the second symbol

- v) Code each reduced source starting with the smallest source and working back to the original source.

1. Average Information per Symbol or Entropy:

$$H(S) = \sum_{i=1}^n p_i I(s_i) = -\sum_{i=1}^n p_i \log_2 p_i \text{ Bits/symbol}$$

2. Average length of the code is $L_{avg} = \sum_{i=1}^N p_i L_i$

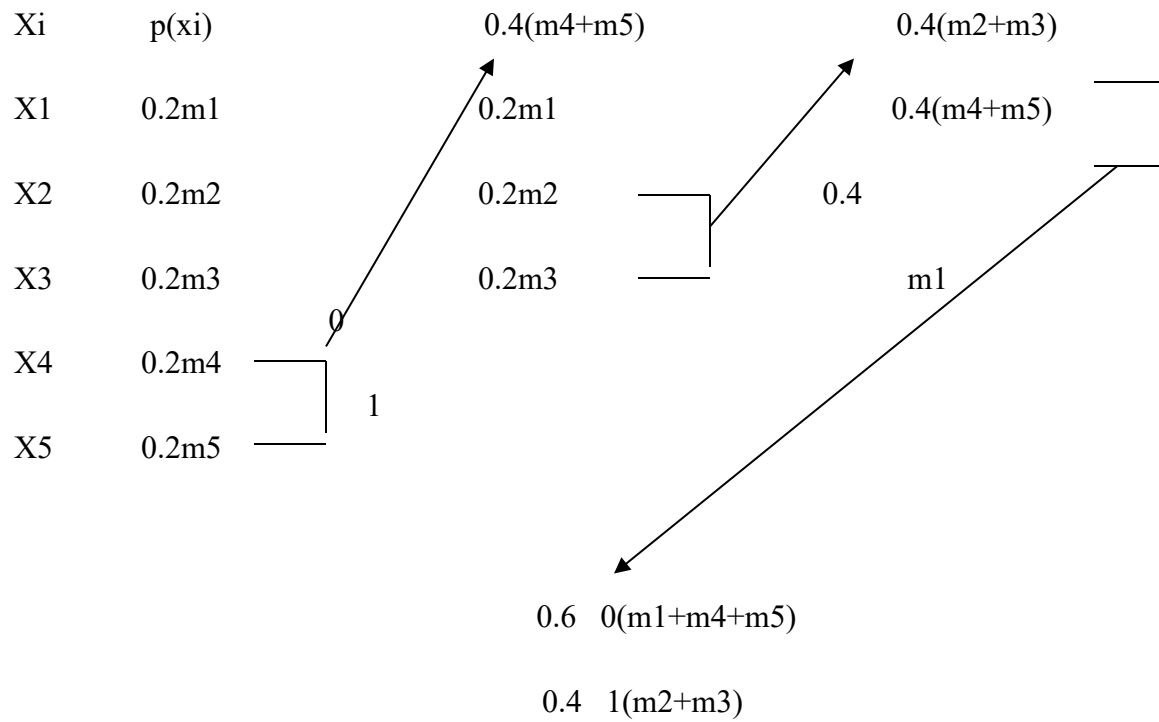
3. Efficiency is given as $\eta = \frac{H(s)}{L_{avg}}$

Pi = probability of occurrence of the ith symbol,

Li = length of the ith code word,

N = number of symbols to be encoded

Problem: A DMS x has five equally likely symbols P (x1) = p(x2) = P(X3) =p(x4)= P(x5) =0.2 construct the Huffman code cu calculate the efficiency.



X_i	code	Length
x_1	00	2
x_2	10	2
x_3	11	2
x_4	0 0 0	3
x_5	0 0 1	3

$$H(x) = - \sum_{i=1}^5 p(x_i) \log_2 p(x_i)$$

$$= - [(0.2 \log_2(0.2) \times 5)]$$

$$H(x) = 2.32$$

$$L = \sum_{i=1}^5 p(x_i) n_i$$

$$= (0.2) (2+2+2+3+3)$$

$$L = 2.4$$

$$\text{Therefore } \%q = H(x) / L \log_2(M) = 2.32 / 2.4 \log_2(2) = 96.7\%$$

Properties of Huffman coding:

1. Optimum code for a given data set requires two passes.
2. Code construction complexity $O(N \log N)$.
3. Fast lookup table based implementation.
4. Requires at least one bit per symbol.
5. Average codeword length is within one bit of zero-order entropy (Tighter bounds are known): $H \leq R < H+1$ bit
6. Susceptible to bit errors.

b)The Shannon-Fano Algorithm

Algorithm:

1. List the source symbols in order of decreasing probability.
2. partition the set in to two sets that are close to equiprobables as possible and assign 0 to the upper set, 1 to the loser set.
3. continue this process, each time partitioning the sets with as nearly equal probabilities as possible until further portioning is not possible.

Example A DMS x has four symbols x_1, x_2, x_3 cu x_4 with $P(x_1) = 1/2$, $P(x_2) = 1/4$ and $P(x_3) = P(x_4) = 1/8$ construct the Shannon-fano code for x ./ show that his code has the optimum property that $n_i = I(x_i)$ cu the code efficiency is 100 percent.

X_i	$p(x_i)$		code
X_1	0.5	0	1
X_2	0.25	1 0	2
X_3	0.125	1 1 0	3
X_4	0.125	1 1 1	3

$$H(x) = \sum_{i=1}^4 p(x_i) \log_2(1/p(x_i))$$

$$= p(x_1) \log_2(1/p(x_1)) + p(x_2) \log_2(1/p(x_2)) + p(x_3) \log_2(1/p(x_3)) + p(x_4) \log_2(1/p(x_4))$$

$$= -[0.5 \log_2(0.5) + 0.25 \log_2(0.25) + 0.125 \log_2(0.125) + 0.125 \log_2(0.125)]$$

$$= -[-0.5000 - 0.5000 - 0.3750 - 0.3750]$$

$$= -[-1.75] = 1.75 \text{ bits/symbol.}$$

$$L = \sum_{i=1}^4 p(x_i) n_i$$

$$= 0.5(1) + 0.25(2) + 0.125(3) + 0.125(3)$$

$$L = 1.75$$

$$Q = H(x)/L \log_2(2) = 1.75/1.75 \log_2(2) = 1$$

Therefore $q = 100\%$

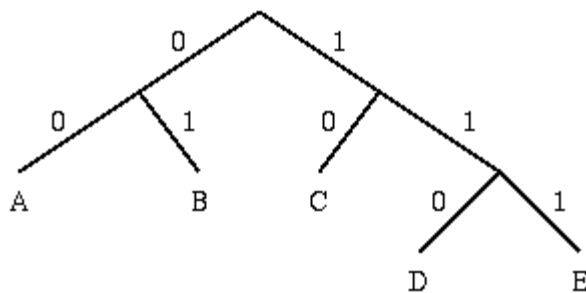
This is a basic information theoretic algorithm. A simple example will be used to illustrate the algorithm:

Symbol	A	B	C	D	E
Count	15	7	6	6	5

Encoding for the Shannon-Fano Algorithm:

A top-down approach

1. Sort symbols according to their frequencies/probabilities, e.g., ABCDE.
2. Recursively divide into two parts, each with approx. same number of counts.



3. Describe arithmetic coding of images with an example. (May / June 2006)

Arithmetic Coding:

Basic idea: Map input sequence of symbols into one single codeword

Codeword is determined and updated incrementally with each new symbol (symbol-by-symbol coding)

At any time, the determined codeword uniquely represents all the past occurring symbols.

Codeword is represented by a half-open subinterval $[L_c, H_c] = [0, 1]$. The half-open subinterval gives the set of all codewords that can be used to encode the input symbol sequence, which consists of all past input symbols. Any real number within the subinterval $[L_c, H_c]$ can be assigned as the codeword representing the past occurring symbols.

Algorithm:

Let

$S = \{s_0, \dots, s_{(N-1)}\}$ – source alphabet

$p_k = P(s_k)$ – probability of symbol s_k , $0 \leq k \leq (N - 1)$

$[L_c, H_c]$ – Interval assigned to symbol s_k , where $p_k = H_{s_k} - L_{s_k}$

- The subinterval limits can be computed as:

$$L_{s_k} = \sum_{i=0}^{k-1} p_i = P_{k-1}; \quad 0 \leq k \leq (N-1)$$

$$H_{s_k} = \sum_{i=0}^k p_i = P_k; \quad 0 \leq k \leq (N-1)$$

Encoding Procedure:

1. Coding begins by dividing interval $[0, 1]$ into N non-overlapping intervals with lengths equal to symbol probabilities p_k
2. $L_c = 0$; $H_c = 1$
3. Calculate code subinterval length: $\text{length} = H_c - L_c$
4. Get next input symbol s_k
5. Update the code subinterval
 $L_c = L_c + \text{length} \cdot L_{s_k}$
 $H_c = L_c + \text{length} \cdot H_{s_k}$
6. Repeat from Step 3 until all the input sequence has been encoded

Decoding procedure:

1. $L_c = 0$; $H_c = 1$
2. Calculate code subinterval length: $\text{length} = H_c - L_c$
3. Find symbol subinterval , $0 \leq k \leq (N - 1)$ such that
4. Output symbol s_k
5. Update the code subinterval
 $L_c = L_c + \text{length} \cdot L_{s_k}$
 $H_c = L_c + \text{length} \cdot H_{s_k}$
6. Repeat from Step 2 until last symbol is decoded.

In order to determine when to stop the decoding:

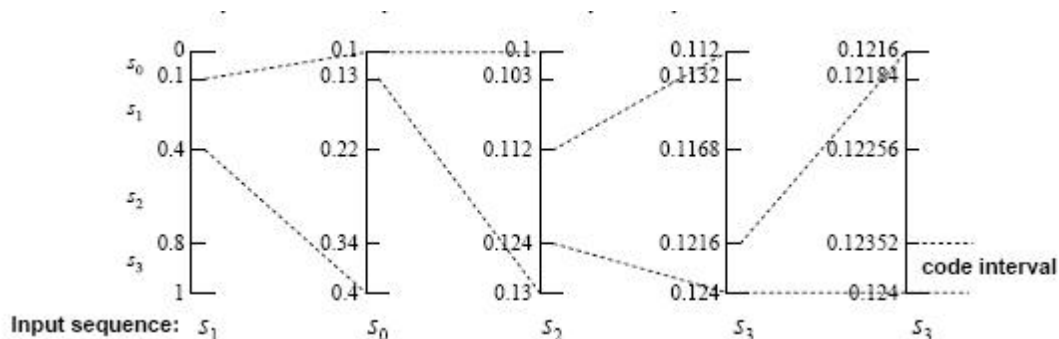
" A special end-of-sequence symbol can be added to the source alphabet.

" If fixed-length block of symbols are encoded, the decoder can simply keep a count of the number of decoded symbols

Example:

1. Construct code words using arithmetic coding for the sequence s_0, s_1, s_2, s_3 with probabilities $\{0.1, 0.3, 0.4, 0.2\}$

Source symbol s_k	Probability P_k	Symbol Subinterval $[L_k, H_k)$
s_0	0.1	$[0, 0.1)$
s_1	0.3	$[0.1, 0.4)$
s_2	0.4	$[0.4, 0.8)$
s_3	0.2	$[0.8, 1)$



Answer:

Iteration # I	Encoded symbol s_k	Code Subinterval $[L_c, H_c)$
1	s_1	$[0.1, 0.4)$
2	s_0	$[0.1, 0.13)$
3	s_2	$[0.112, 0.124)$
4	s_3	$[0.1216, 0.124)$
5	s_3	$[0.12352, 0.124)$

4. Explain the lossless Bit plane coding or shift coding

Bit-Plane Coding:

The concept of decomposing a multilevel image into a series of binary images and compressing each binary image via one of the several well known binary compression methods is called as Bit plane coding.

Bit Plane Decomposition: An m -bit gray scale image can be converted into m binary images by bit-plane slicing. These individual images are then encoded using run-length coding. Code the bit planes separately, using RLE (flatten each plane row-wise into a 1D array), Golomb coding, or any other lossless compression technique.

- Let I be an image where every pixel value is n -bit long
- Express every pixel in binary using n bits

- Form n binary matrices (called bitplanes), where the i -th matrix consists of the i -th bits of the pixels of I .

Example: Let I be the following 2×2 image where the pixels are 3 bits long

```

101 110
111 011

```

The corresponding 3 bit planes are:

```

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

```

However, a small difference in the gray level of adjacent pixels can cause a disruption of the run of zeroes or ones.

Eg: Let us say one pixel has a gray level of 127 and the next pixel has a gray level of 128. In binary: $127 = 01111111$ & $128 = 10000000$. Here the MSB's of the two binary codes for 127 and 128 are different, bit plane 7 will contain a zero-valued pixel next to a pixel of value 1, creating a 0 to 1 or 1 to 0 transitions at that point.

Therefore a small change in gray level has decreased the run-lengths in all the bit-planes.

GRAY CODE APPROACH

- Gray coded images are free of this problem which affects images which are in binary format.
- In gray code the representation of adjacent gray levels will differ only in one bit (unlike binary format where all the bits can change).
- Thus when gray levels 127 and 128 are adjacent, only the 7th bit plane will contain a 0 to 1 transition, because the gray codes that correspond to 127 and 128 are as follows,
- In gray code:

$$127 = 01000000$$

$$128 = 11000000$$

- Thus, small changes in gray levels are less likely to affect all m bit planes.
- Let $g_{m-1} \dots g_1 g_0$ represent the gray code representation of a binary number.
- Then:

$$g_i = a_i \oplus a_{i+1} \quad 0 \leq i \leq m-2$$

$$g_{m-1} = a_{m-1}$$

To convert a binary number $b_1 b_2 b_3 \dots b_{n-1} b_n$ to its corresponding binary reflected Gray code. Start at the right with the digit b_n . If the b_{n-1} is 1, replace b_n by $1-b_n$; otherwise, leave it unchanged. Then proceed to b_{n-1} .

Continue up to the first digit b_1 , which is kept the same since it is assumed to be a $b_0 = 0$. The resulting number is the reflected binary Gray code.

Example:

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

Decoding a gray coded image:

The MSB is retained as such,i.e.,

$$a_i = g_i \oplus a_{i+1} \quad 0 \leq i \leq m-2$$

$$a_{m-1} = g_{m-1}$$

5. Explain the Run-length Coding

Run-length Coding:

1. Run-length Encoding, or RLE is a technique used to reduce the size of a repeating string of characters.
2. This repeating string is called a run, typically RLE encodes a run of symbols into two bytes , a count and a symbol.
3. RLE can compress any type of data
4. RLE cannot achieve high compression ratios compared to other compression methods
5. It is easy to implement and is quick to execute.
6. Run-length encoding is supported by most bitmap file formats such as TIFF, BMP and PCX
7. The most common approaches for determining the value of a run are (a) to specify the value of the first run of each row, or (b) to assume that each row begins with a white run, whose run length may in fact be zero.
8. Similarly, in terms of entropy, the approximation run – length entropy of the image is ,

$$H_{RL} = \frac{H_0 + H_1}{L_0 + L_1}$$

Where, L0 and L1 denotes the average values of black and white run lengths, H0 is the entropy of black runs, and H1 is the entropy of white runs.

9. The above expression provides an estimate of the average number of bits per pixel required to code the run lengths in a binary image using a variable-length code.

Example:

i.WWWWWWWWWWWWWBWWWWWWWWWWWWWWBBBWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWBWWWWWWWWWWWWWWWWWWWW

RLE coding: 12W1B12W3B24W1B14W

ii) 1111100011111100000 –RLE coding - 5W3B6W5B – 5365

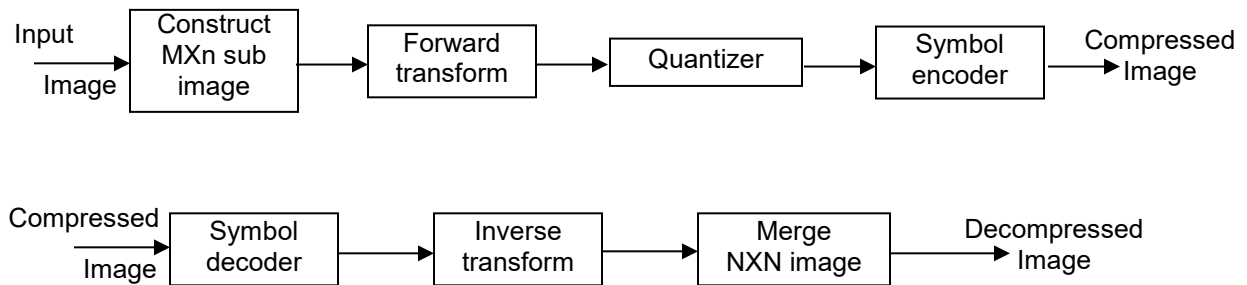
iii) 0000011100 – RLE coding – 5B3BW2B – 532 – 0532 (first '0' indicates starting with black)

6. Explain Transform coding. (May / June 2009), (May / June 2009)

Transform Coding

Here the compression technique is based on modifying the transform of an image.

The block diagram of encoder& decoder of transform coding.



Explanation:

1. The NxN output image is subdivided into sub images of size nxn, which are then transformed to generate $(N/n)^2$ sub image transform arrays each of size $n \times m$.

2. Transformation is done to decorrelate the pixels of each sub image or to pack as much information as possible into the smallest number of transform coefficients.

- The following transforms are used in the encoder block diagram
 - a. KLT
 - b. DCT
 - c. WHT
 - d. DFT

Mostly used is DCT or DFT

3. Quantization is done to eliminate the coefficients which carry least information. These omitted coefficients will have small impacts on the quality of the reconstructed sub images.

4. Finally, the quantized coefficients are encoded.

5. The decoder is used to extract the code words, which are applied to the dequantizer, which is used to reconstruct the set of quantized transform coefficients.

6. Then the set of transform coefficients that represents the image is produced from the quantized coefficients.

7. These coefficients are inverse transformed and give the lossy version of the image. Here the compression ratio is low.

STEPS:

i. Subimage size selection:

Normally 4x4, 8x8 and 16x16 image sizes are selected

ii. Transform Selection:

Consider an image $f(x,y)$ of size $N \times N$, then the transformed image is given as,

Transformed image = input image x Transformation function

$$\text{i.e., } T(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) g(x, y, u, v)$$

where, $g(x, y, u, v)$ = transform function

The original signal at the decoder is given as,

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

Where $h(x, y, u, v) = g^{-1}(x, y, u, v)$

a. In case of DFT:

$$g(x, y, u, v) = \frac{1}{N^2} e^{-j2\pi \left(\frac{ux+vy}{N} \right)}$$

$$h(x, y, u, v) = e^{j2\pi \left(\frac{ux+vy}{N} \right)}$$

b. In case of WHT:

$$g(x, y, u, v) = h(x, y, u, v) = \frac{1}{N} (-1)^{\sum_{i=0}^{m-1} [b_i(x) p_i(u) + b_i(y) p_i(v)]}$$

Where $N = 2^m$, $b_k(z)$ is the k^{th} bit from right to left in the binary representation of z .

c. In case of K L transform:

$$v = \Phi^{*T} u$$

Where, u = input image

Φ_k = k^{th} column of Eigen vectors

V = transformed image

d. In case of **Discrete Cosine Transform (DCT)**

$$g(x, y, u, v) = h(x, y, u, v) = \alpha(u) \alpha(v) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cdot \cos \left[\frac{(2y+1)v\pi}{2N} \right]$$

wh

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

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}} & \text{for } v = 0 \\ \sqrt{2/N} & \text{for } v = 1, 2, \dots, N-1 \end{cases}$$

iii. Bit Allocation.

The overall process of truncating, quantizing, and coding the coefficients of a transformed sub image is commonly called bit allocation

a. Zonal coding

The process in which the retained coefficients are selected on the basis of maximum variance is called zonal coding.

Let N_t be the number of transmitted samples and the zonal mask is defined as an array which takes the unity value in the zone of largest N_t variances of the transformed samples.

$$\text{i.e., } m(k, l) = \begin{cases} 1, (k, l) \in I_t \\ 0, \text{else,} \end{cases}$$

Here the pixel with maximum variance is replaced by 1 and a 0 in all other locations. Thus the shape of the mask depends on where the maximum variances pixels are located. It can be of any shape.

Threshold coding

The process in which the retained coefficients are selected on the basis of maximum magnitude is called threshold coding.

1. Here we encode the N_t coefficients of largest amplitudes.
2. The amplitude of the pixels are taken and compared with the threshold value
3. If it is greater than the threshold η , replace it by 1 else by 0.

Threshold mask m_η is given by,

$$m_\eta(k, l) = \begin{cases} 1, (k, l) \in I'_t \\ 0, \text{elsewhere} \end{cases} \\ \text{where, } I'_t = \{(k, l); |V(k, l)| > \eta\}$$

4. Then the samples retained are quantized by a suitable uniform quantizer followed by an entropy coder.

There are three basic ways to threshold a transformed sub image.

- a. A single global threshold can be applied to all sub images – here the level of compression differs from image to image depending on the number of coefficients.
- b. A different threshold can be used for each sub image – The same number of coefficients is discarded for each sub image.

The threshold can be varied as a function of the location of each coefficients within the sub image.

Here,

$$\hat{T}(u, v) = \text{round} \left[\frac{T(u, v)}{Z(u, v)} \right]$$

$$Z = \begin{bmatrix} Z(0,0) & Z(0,1) & \dots & Z(0,n-1) \\ Z(1,0) & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ Z(n-1,0) & \dots & \dots & Z(n-1,n-1) \end{bmatrix}$$

Where, $\hat{T}(u, v)$ is a threshold and quantized approximation of $T(u, v)$ and $Z(u, v)$ is an element of the transform normalization array.

The denormalized array is given as,

$$T'(u, v) = \hat{T}(u, v) z(u, v)$$

The decompressed sub image is given as,

$$= \frac{1}{T'(u, v)}$$

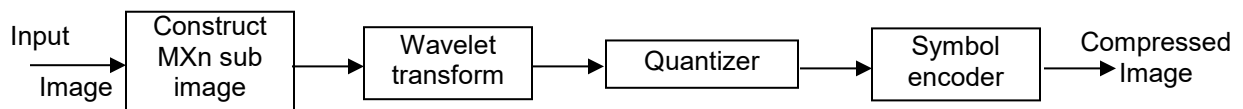
And $\hat{T}(u, v)$ assumes integer value k if and only if,

$$kc - \frac{c}{2} \leq T(u, v) \leq kc + \frac{c}{2}$$

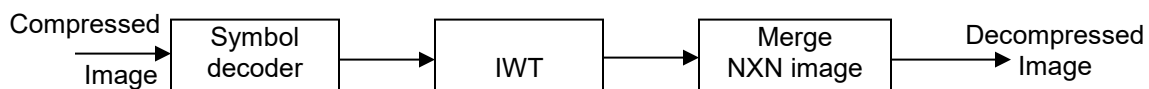
If $Z(u, v) > 2T(u, v)$, then $\hat{T}(u, v) = 0$ and the transform coefficient is completely truncated or discarded.

7. Describe on the Wavelet coding of images. (May / June 2009) or Explain the lossy compression wavelet coding.

ENCODER



DECODER



Explanation:

1. Input image is applied to wavelet transform, which is used to convert the original input image into horizontal, vertical and diagonal decomposition coefficients with zero mean and laplacian like distribution.
2. These wavelets pack most of the information into a small number of coefficients and the remaining coefficients can be quantized or truncated to zero so that the coding redundancy is minimized.
3. Runlength, huffmann, arithmetic, bit plane coding are used for symbol encoding
4. Decoder does the inverse operation of encoding

The difference between wavelet and transform coding is that the sub image processing is not needed here. Wavelet transforms are computationally efficient and inherently local.

Wavelet selection:

The four Discrete wavelet transforms are,

1. Haar wavelets – simplest
2. Daubechies wavelets – most popular
3. Symlet wavelet
4. Bio orthogonal wavelet

Wavelets

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale.

These basis functions or baby wavelets are obtained from a single prototype wavelet called the mother wavelet, by dilations or contractions (scaling) and translations (shifts).

wavelet Transform

Wavelet Transform is a type of signal representation that can give the frequency content of the signal at a particular instant of time.

1. Wavelet transform decomposes a signal into a set of basis functions.
2. These basis functions are called **wavelets**
3. Wavelets are obtained from a single prototype wavelet $\psi(t)$ called **mother wavelet by dilations and shifting**:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$$

where a is the scaling parameter and b is the shifting parameter

The 1-D wavelet transform is given by :

$$W_f(a,b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}(t) dt$$

The inverse 1-D wavelet transform is given by:

$$x(t) = \frac{1}{C} \int_0^{\infty} \int_{-\infty}^{\infty} W_f(a,b) \psi_{a,b}(t) db \frac{da}{a^2}$$

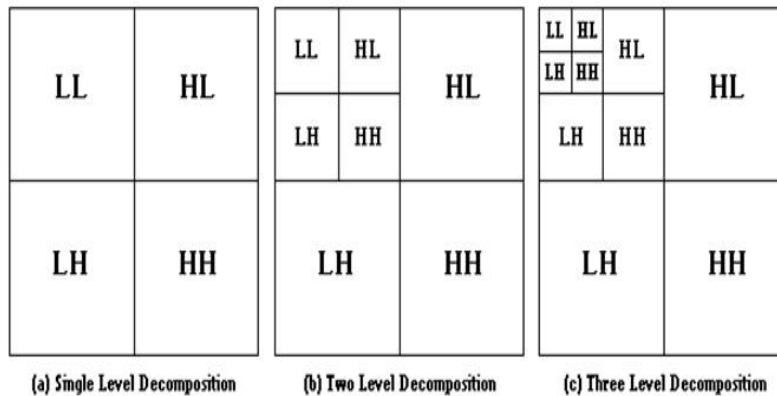
$$\text{where } C = \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty$$

Wavelet analysis has advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes

Types of wavelet transform

1. The discrete wavelet transform
2. The continuous wavelet transform

Input signal is filtered and separated into low and high frequency components



LL: Horizontal Lowpass & Vertical Lowpass , LH: Horizontal Lowpass & Vertical Highpass
 HL: Horizontal Highpass & Vertical Lowpass , HH: Horizontal Highpass & Vertical Highpass

Quantizer Design:

The largest factor effecting wavelet coding compression and reconstruction error is coefficient quantization.

The effectiveness of the quantization can be improved by ,(i) introducing an enlarged quantization interval around zero, called a **dead zone** and (ii) adapting the size of the quantization interval from scale to scale.

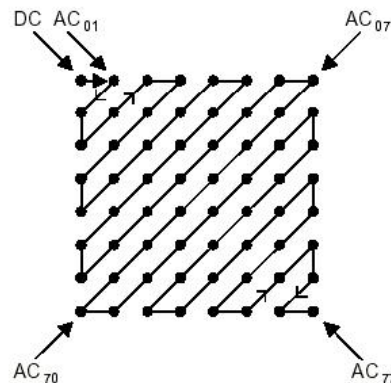
8. Write notes on JPEG standard with neat diagram. (May / June 2009), (May / June 2006)

1. JPEG: It defines three different coding systems:

1. A lossy baseline coding system, adequate for most compression applications
2. An extended coding system for greater compression, higher precision or progressive reconstruction applications
3. A lossless independent coding system for reversible compression

Details of JPEG compression Algorithm

- 1) Level shift the original image
- 2) Divide the input image in to 8x8 blocks
- 3) Compute DCT for each block (matrix) (8x8)
- 4) Sort elements of the 8x8 matrix



- 5) Process one block at a time to get the output vector with trailing zeros
Truncated

The quantized coefficient is defined as

$$F(u, v)_{Quantization} = \text{round} \left(\frac{F(u, v)}{Q(u, v)} \right) \text{ and the reverse the process}$$

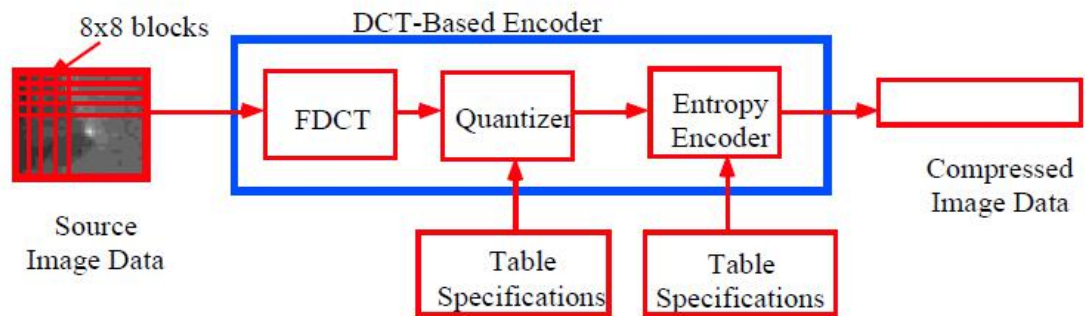
can be achieved by

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

6) Delete unused portion of the output vector and encode it using Huffman Coding.

7) Store the size, number of blocks for use in the decoder

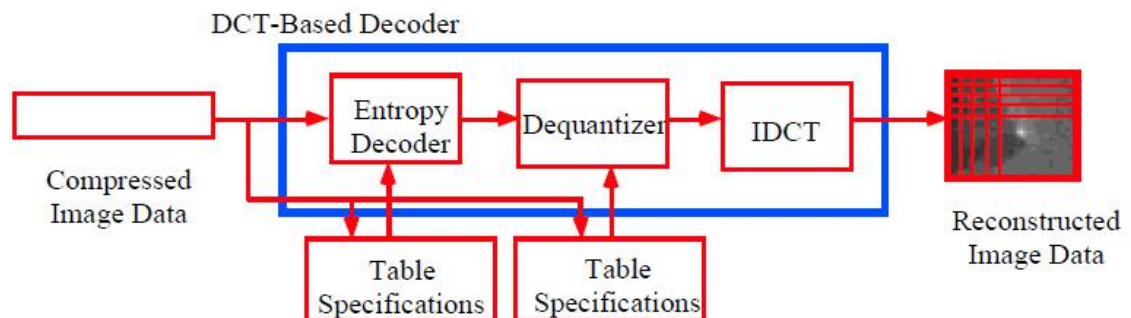
JPEG Encoder Block Diagram



Details of JPEG Decompression Algorithm

- 1) Compute the reverse order for the output vector
- 2) Perform Huffman decoding next.
- 3) Restore the order of the matrix
- 4) Denormalize the DCT and perform block processing to reconstruct the Original image.
- 5) Level shift back the reconstructed image

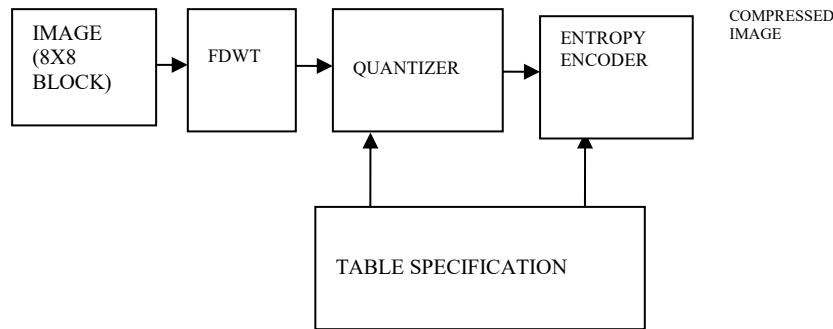
JPEG Decoder Block Diagram



2. JPEG 2000

JPEG 2000 Encoder Algorithm:

- The first step of the encoding process is to DC level shift the pixels of the image by subtracting 2^{m-1} , where 2^m is the number of gray levels of the image
- If the input image is a color image, then RGB value is converted into YUV and these components are level shifted individually.
- After the image has been level shifted, its components are divided into tiles.
- Tiles are rectangular arrays of pixels that contain the same relative portion of all the components. Tile component is the basic unit of the original or reconstructed image.
- A wavelet transform is applied on each tile. The tile is decomposed into different resolution levels.
- The decomposition levels are made up of sub bands of coefficients that describe the frequency characteristics of local areas of the tile components, rather than across the entire image component.
- Thus 1D-DWT of the rows and columns of each tile component is then computed.



Main structure of JPEG2000 encoder.

- This involves six sequential lifting and scaling operations.

$$Y(2n+1) = X(2n+1) + \alpha[X(2n) + X(2n+2)], i_o - 3 \leq 2n+1 \leq i_1 + 3$$

$$Y(2n) = X(2n) + \beta[Y(2n-1) + Y(2n+1)], i_o - 2 \leq 2n \leq i_1 + 2$$

$$Y(2n+1) = Y(2n+1) + \gamma[Y(2n) + Y(2n+2)], i_o - 1 \leq 2n+1 \leq i_1 + 1$$

$$Y(2n) = Y(2n) + \delta[Y(2n-1) + Y(2n+1)], i_o \leq 2n \leq i_1$$

$$Y(2n+1) = (-k) \cdot Y(2n-1), i_o - 1 \leq 2n+1 \leq i_1$$

$$Y(2n) = Y(2n) / K, i_o \leq 2n \leq i_1$$

X = input tile component

Y = resulting transform coefficients

i_o and i_1 = represent the position of the tiles component within a component.

Lifting Parameter:

$$\alpha = -1.586134342, \beta = -0.052980118, \gamma = 0.882911075 \text{ and } \delta = 0.433506852$$

Scaling parameter:

$$K = 1.230174105$$

- i. After the completion of scaling and lifting operations, the even indexed value $Y(2n)$ represents the FWT LP filtered output and odd indexed values of $Y(2n+1)$ correspond to the FWT HP filtered output.
- j. The sub bands of coefficients are quantized and collected into rectangular arrays of “code blocks..Thus after transformation, all coefficients are quantized.
- k. Quantization is the process by which the coefficients are reduced in precision. Each of the transform coefficients $a_b(u,v)$ of the sub band b is quantized to the value $q_b(u,v)$ according to the formula,

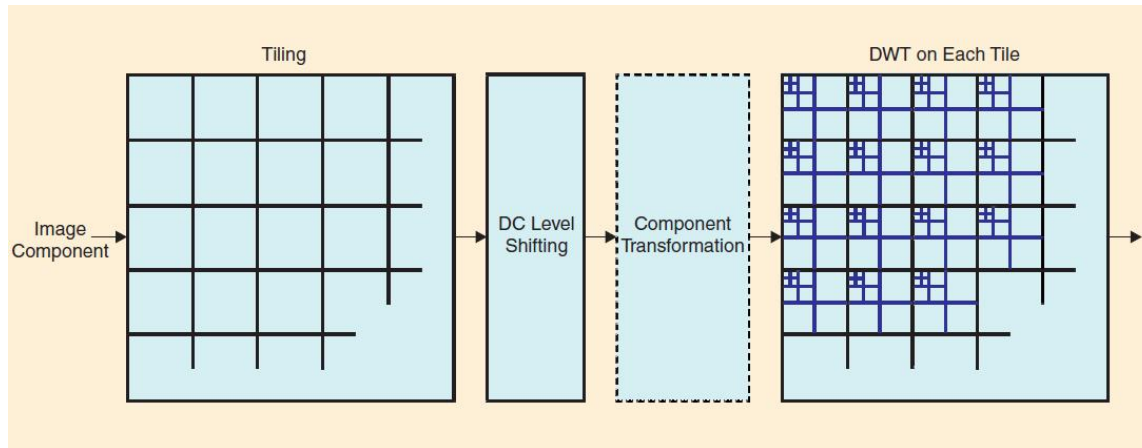
$$q_b(u,v) = \text{sign}(a_b(u,v)) \left\lfloor \frac{|a_b(u,v)|}{\Delta_b} \right\rfloor$$

Where, $\Delta_b = \text{quantization step}$ $2^{R_b - \varepsilon_b} \left[1 + \frac{\mu_b}{2^{11}} \right]$

R_b = nominal dynamic range of subband b

ε_b, μ_b = number of bits allocated to the exponent and mantissa of the subband's coefficients.

- l. Encode the quantized coefficients using Huffman or arithmetic coding.



JPEG 2000 Decoder Algorithm:

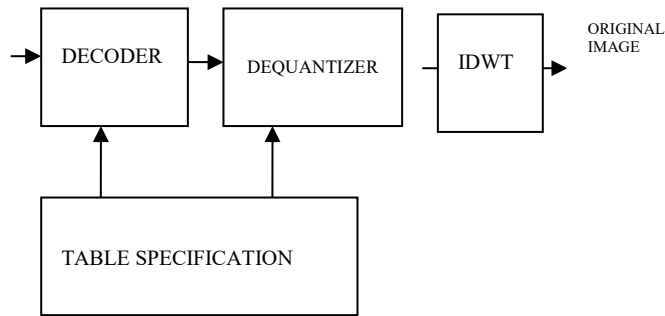
- a. Decode the bit modeled or arithmetically bit streams
- b. Dequantize the coefficients using,

$$R_{qb}(u,v) = \begin{cases} (\overline{q_b}(u,v) + 2^{M_b - N_b(u,v)}) \cdot \Delta_b, & \overline{q_b}(u,v) > 0 \\ (\overline{q_b}(u,v) - 2^{M_b - N_b(u,v)}) \cdot \Delta_b, & \overline{q_b}(u,v) < 0 \\ 0, & \overline{q_b}(u,v) = 0 \end{cases}$$

Where , $R_{qb}(u, v)$ = dequantized transform coefficients

Nb(u,v) = number of decoded bit planes for $\overline{q_b}(u, v)$

Nb = encoder would have encoded Nb bit planes for a particular subband.



Main structure of JPEG 2000 decoder.

- c. The dequantized coefficients are then inverse transformed by column and by row using inverse forward wavelet transform filter bank or using the following lifting based operations

$$X(2n+1) = (-1/k) \cdot Y(2n+1), i_o - 2 \leq 2n+1 \leq i_1 + 2$$

$$X(2n) = X(2n) - \delta[X(2n-1) + X(2n+1)], i_o - 3 \leq 2n \leq i_1 + 3$$

$$X(2n+1) = X(2n+1) - \gamma[X(2n) + X(2n+2)], i_o - 2 \leq 2n+1 \leq i_1 + 2$$

$$X(2n) = X(2n) - \beta[X(2n-1) + X(2n+1)], i_o - 1 \leq 2n \leq i_1 + 1$$

$$X(2n+1) = X(2n+1) - \alpha[X(2n) + X(2n+2)], i_o \leq 2n+1 \leq i_1$$

- d. Perform DC level shifting by adding 2^{m-1}

9.. Briefly discuss the MPEG compression standard. (May / June 2009)or Explain the Video compression standards.

Video Compression standards:

Video compression is used for compression of moving picture frames.

Based on the applications, video compression standards are grouped into two categories,

- (1) Video tele conferencing standards
- (2) Multimedia standards

a. Video teleconferencing standards:

The International Telecommunication Union (ITU) has a number of video conferencing compression standards such as, H.261, H.262, H.263 and H.320

1.H.261 is used to support full motion video transmission over T1 lines

2.H.263 is designed for very low bit rate video in the range of 10 to 30 Kbits/sec

3.H.320 is used for ISND

b.Multimedia Standards

These are used for video on demand, digital HDTV broadcasting and image/video database services.The principal multimedia compression standards are MPEG-1, MPEG-2 and MPEG-4

1. MPEG-1 is an entertainment quality coding standard for the storage and retrieval of video on digital media such as CDROMs.It supports a bit rate of 1.5 Mbit/s.Thus it is used for compression of low resolution.

2. MPEG-2 is used for cable TV distribution and narrow channel satellite broadcasting with a bit rate of 2 to 10 Mbit/s. Thus is used for higher resolution standards, mainly for studio quality audio and video compression.

3. MPEG-4 provides 5 and 64 Kbit/s for mobile and PSTN and upto 4 Mbit/s for TV and film applications.

4. MPEG -4 provides (a) improved video compression efficiency; (b) Content based interactivity and (c) universal access including increased robustness.

MPEG Algorithm:

MPEG standard consists of both video and audio compression. The MPEG algorithm relies on two basic techniques

Block based motion compensation

DCT based compression

1.The video signal is sequence of still picture frames. From the picture, slices are formed in the raster scan order as shown below,

2.For each slice the macro blocks are obtained. The macro block consists of four blocks of 8 x 8 pixels. Thus the micro block has a size of 16 x 16 pixels. If the macro block is in RGB format, then it is converted into YCrCb format, where Cr and Cb represent chrominance signals. Once the sequence of macro blocks is formed, coding can take place.

3.The motion is estimated by predicting the current frame on the basis of certain previous and/or forward frame. The information sent to the decoder consists of the compressed DCT coefficients of the residual block together with the motion vector. There are three types of pictures in MPEG:

Intra-frames (I)

Predicted frames (P)

Bidirectionally predicted frames (B)

Figure demonstrates the position of the different types of pictures. Every Nth frame in the video sequence is an I-picture, and every Mth frame a P-picture. Here N=12 and M=4. The rest of the frames are B-pictures.

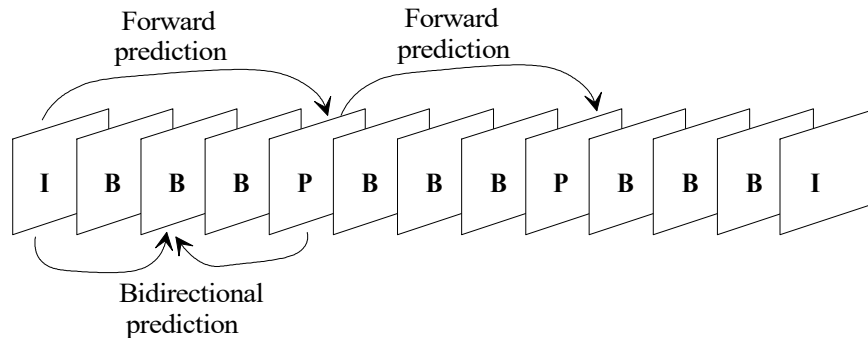


Figure: Interframe coding in MPEG.

Compression of the picture types:

1. Intra frame (I –frame) are coded as still images by DCT algorithm. They provide access points for random access, but only with moderate compression. It is used as the reference point for the motion estimation.

2. Predicted frame (P-frame) are coded with reference to a past picture. The current frame is predicted on the basis of the previous I- or P-picture. The residual (difference between the prediction and the original picture) is then compressed by DCT.

3. Bidirectional frame (B-frame) is the compressed difference between the current frame and a prediction of it based on the previous I- or P- frame and next P-frame. Thus the decoder must have access to both the past and future reference frames. The encoded frames are therefore recorded before transmission and the decoder reconstructs and displays them in the proper sequence. Bidirectional pictures are never used as reference.

1. After the pictures are divided into 16x16 macro blocks, each consisting of four 8x8 elementary blocks, the DCT of the macro block is calculated whose coefficients are quantized.

The choice of the prediction method is chosen for each macroblock separately.

2. The intra-coded blocks are quantized differently from the predicted blocks:

Intra-coded blocks contain information in all frequencies and are quantized differently from the predicted blocks. The predicted blocks, contain mostly high frequencies and can be quantized with more coarse quantization tables.

Motion estimation and compensation:

1. The prediction block in the reference frame is not necessarily in the same coordinates than the block in the current frame.

2. Because of motion in the image sequence, the most suitable predictor for the current block may exist anywhere in the reference frame.

3. The motion estimation specifies where the best prediction (best match) is found.

4. Motion compensation consists of calculating the difference between the reference and the current block.

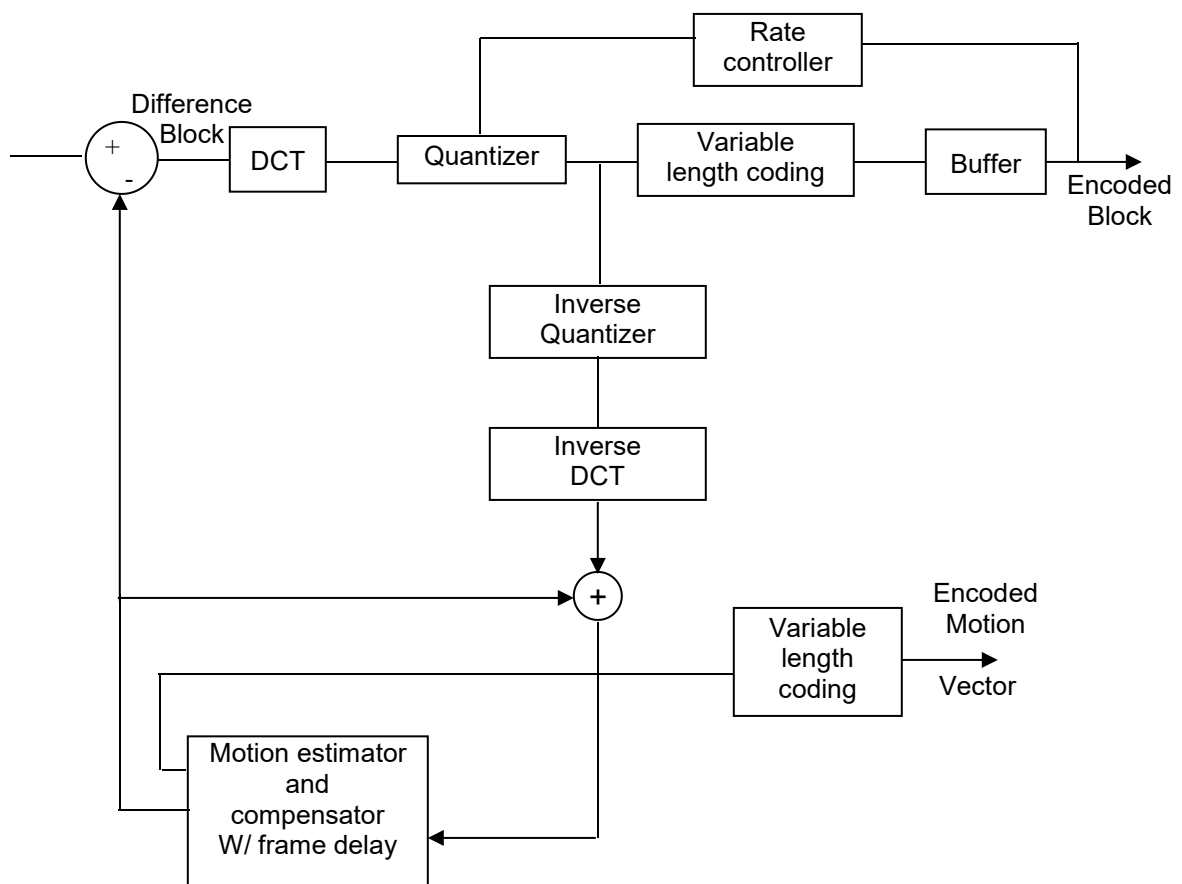
5. Finally the quantized DCT coefficients are then run – length encoded in a zig-zag order and then Huffman coded to compress the data.

6. Also the encoder is designed to generate a bit stream that matches the capacity of the intended video channel. This is done by using a rate controller which adjusts the quantization parameters as a function of the occupancy of the output buffer.

7. As the buffer becomes fuller the quantization is made coarser, so that fewer bits stream into buffer.

Three basic types of encoded output frames:

1. **Intraframe or independent frame (I-frame).** An I frame is compressed independently of all previous and future video frames. Of the three possible encoded output frames, it most highly resembles a JPEG encoded image. Moreover, it is the reference point for the motion estimation needed to generate subsequent P- and B-frames.



I- frames provide the highest degree of random access, ease of editing, and greatest resistance to the propagation of transmission error. As a result, all standards require their periodic insertion into the compressed code stream.

2. **Predictive frame (P-frame).** A P-frame is the compressed difference between the current frame and a prediction of its based on the previous I-or P-frame. The difference is formed in the leftmost summer of fig. The prediction is motion compensated and typically involves sliding the decoded block in the lower part of fig. around its immediate neighborhood in the current frame and computing a measure of correlation (such as the sum of the square of the pixel-by-pixel differences). In fact, the process is often carried out in subpixel increments (such as sliding the subimage $\frac{1}{4}$ pixels at a time), which necessitates interpolating pixel values prior to computing the correlation measure. The computed motion vector is variable length coded and transmitted as an integral part of the encoded data stream. Motion estimation is carried out on the macro block level.

3. **Bidirectional frame (B- 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 and next- P-frame. Accordingly, the decoder must have access to both past and future reference frames. The encoded frames are therefore reordered before transmission; the decoder reconstructs and displays them in the proper sequence.

10. Explain the Basics of Vector Quantization in detail.

Definition:

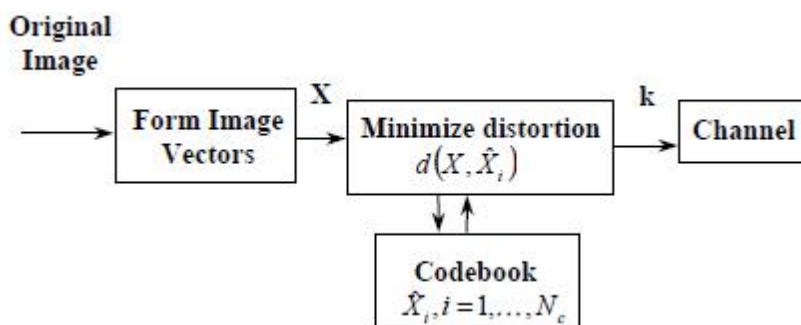
Mapping of accurate data's into inaccurate data's is known as Quantization. If this procedure is extended for vector data , then the procedure is called as **Vector Quantization**.

Vector Quantization is a process which can map n-dimensional vectors in the vector space R into a finite set of vectors Y in the vector space R.

$$Y = \{y_i, i = 1, 2, 3, \dots, n\}$$

Set of all code words is called as code book

Vector Quantization Transmitter:



1. Here, original image is decomposed into n dimensional image vectors, vectors can be block of pixel values or a 3-dimensional vector formed from the RGB color components.

2. Each vector is compared with a collection of representative code vectors,

$\hat{X}_i, i = 1, 2, 3, \dots, N_c$ taken from a previously generated codebook.

3. Best match code vector is chosen using a minimum distortion rule

-Choose \hat{X}_k such that $d(X, \hat{X}_k) \leq d(X, \hat{X}_j), \text{ for } j = 1, 2, 3, \dots, N_c$

denotes the distortion incurred in replacing the original vector X with the code vector \hat{X}

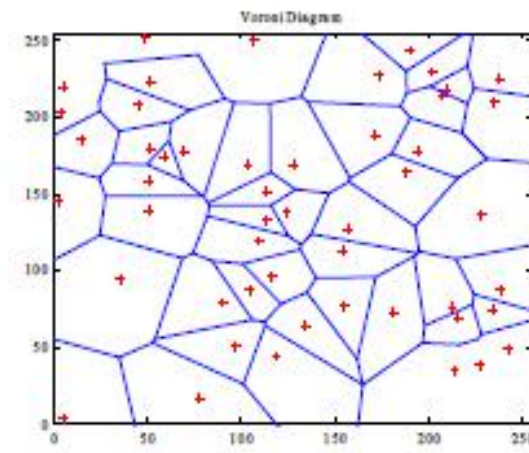
4. Distance measure is done using square of the Euclidean distance between two vectors

$$d(X, \hat{X}) = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2$$

\hat{x}_i = i th component of the code word.

x = i th component of the input vector.

VQ transforms the vectors of data into indexes that represents clusters of vectors



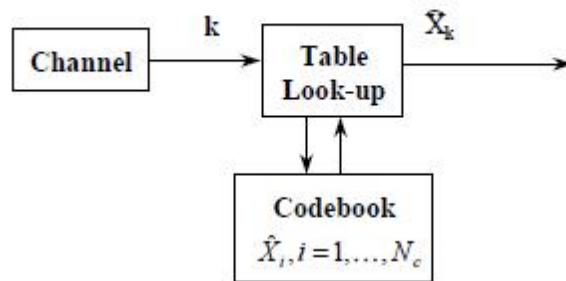
Voronoi region is the nearest neighbor region associated with each codeword Y_i .

After the minimum distortion code vector has been found, the index k is transmitted using $\log_2 N_c$ bits.

Vector Quantization Receiver:

1. VQ decoder uses a duplicate codebook and a table lookup to reproduce the image

2. Compression is obtained by using a codebook with relatively few code vectors compared to the original image vectors



Linde-Buzo-Gray (LBG) Algorithm is one of the suboptimal code book design procedure

Start with:

1. a set of training vectors
2. initial codebook
3. distortion measure d
4. fractional distortion change threshold

Initialize:

1. iteration counter l to 1
2. average distortion over all training vectors, $D(0)$ to a very large number

Based on the minimum euclidean distance, the input vector is attached to the code word and average distortion is computed. Then revise the code words by computing the average of each cluster.

Drawbacks:

The size of the codebook increases for higher compression ratio's and the search time required to find the minimal distortion code also increases.

Code book Design:

Consider an 8 bit image - Compressed to 1 bit per pixel, If $n = 4$ (2x2 blocks), codebook should contain only 16 vectors, 16 vectors will represent 256 possible image vectors.

If block size is increased to 4x4, the codebook is 256 codevectors, The total number of vectors is 256

Resulting mean-square quantization error becomes smaller as the block size increases.

Thus the Number of codebook vectors increase

11. Explain Sub band coding in detail

In sub band coding, the spectrum of the input is decomposed into a set of bandlimited components, which is called sub bands. Ideally, the sub bands can be assembled back to reconstruct the original spectrum without any error. Fig. shows the block diagram of two-band filter bank and the decomposed spectrum. At first, the input signal will be filtered into low pass and high pass components through analysis filters. After filtering, the data amount of the low pass and high pass components will become twice that of the original signal; therefore, the low pass and high pass components must be down sampled to reduce the data quantity. At the receiver, the received data must be up sampled to approximate the original signal. Finally, the up

sampled signal passes the synthesis filters and is added to form the reconstructed approximation signal.

After sub band coding, the amount of data does not reduce in reality. However, the human perception system has different sensitivity to different frequency band. For example, the human eyes are less sensitive to high frequency-band color components, while the human ears is less sensitive to the low-frequency band less than 0.01 Hz and high-frequency band larger than 20 KHz. We can take advantage of such characteristics to reduce the amount of data. Once the less sensitive components are reduced, we can achieve the objective of data compression.

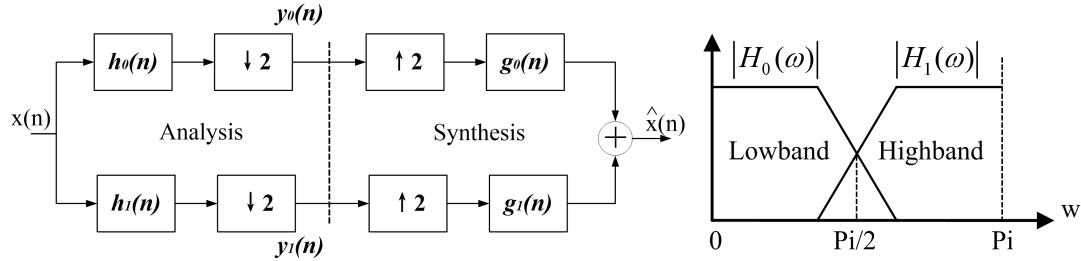


Fig Two-band filter bank for one-dimension sub band coding and decoding

Now back to the discussion on the DWT. In two dimensional wavelet transform, a two-dimensional scaling function, $\phi(x, y)$, and three two-dimensional wavelet function $\psi^H(x, y)$, $\psi^V(x, y)$ and $\psi^D(x, y)$, are required. Each is the product of a one-dimensional scaling function $\phi(x)$ and corresponding wavelet function $\psi(x)$.

$$\phi(x, y) = \phi(x)\phi(y) \quad \psi^H(x, y) = \psi(x)\phi(y)$$

$$\psi^V(x, y) = \phi(y)\psi(x) \quad \psi^D(x, y) = \psi(x)\psi(y)$$

where ψ^H measures variations along columns (like horizontal edges), ψ^V responds to variations along rows (like vertical edges), and ψ^D corresponds to variations along diagonals.

Similar to the one-dimensional discrete wavelet transform, the two-dimensional DWT can be implemented using digital filters and samplers. With separable two-dimensional scaling and wavelet functions, we simply take the one-dimensional DWT of the rows of $f(x, y)$, followed by the one-dimensional DWT of the resulting columns. Fig. shows the block diagram of two-dimensional DWT

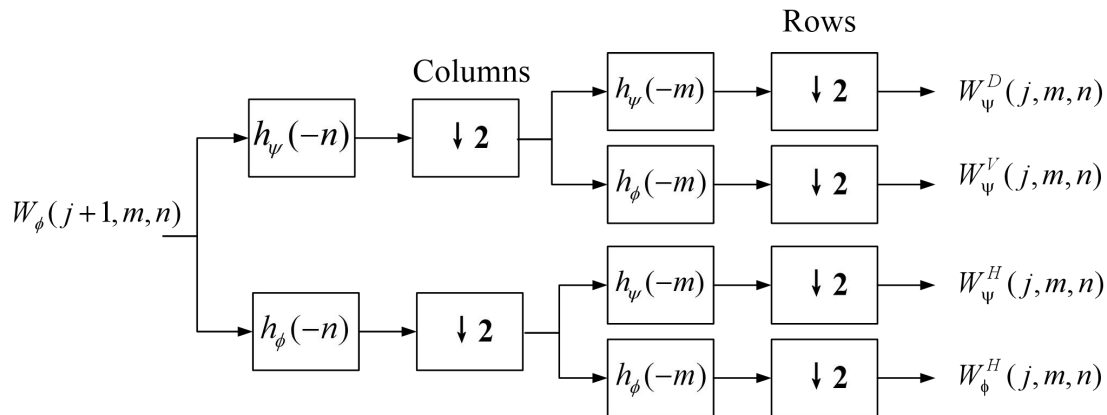


Fig. The analysis filter bank of the two-dimensional DWT

As in the one-dimensional case, image $f(x, y)$ is used as the first scale input, and output four quarter-size sub-images — W_ϕ , W_ψ^H , W_ψ^V , and W_ψ^D as shown in the middle of Fig. The approximation output $W_\phi^H(j, m, n)$ of the filter banks in Fig. can be tied to other input analysis filter bank to obtain more sub images, producing the two-scale decomposition as shown in the left of Fig. Fig. shows the synthesis filter bank that reverses the process described above.

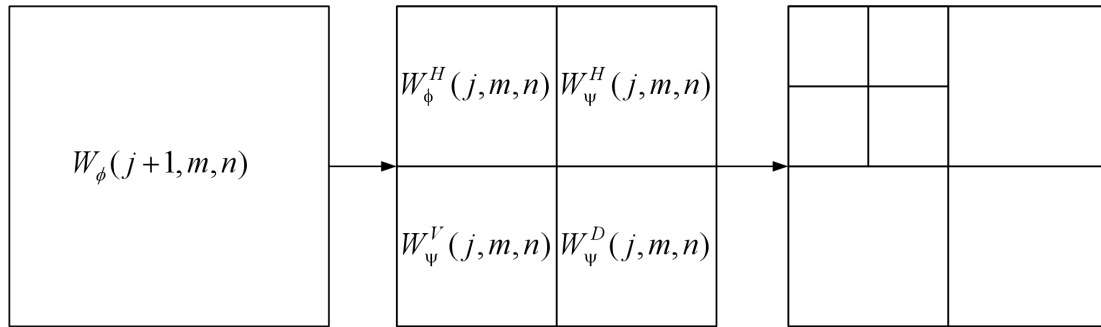
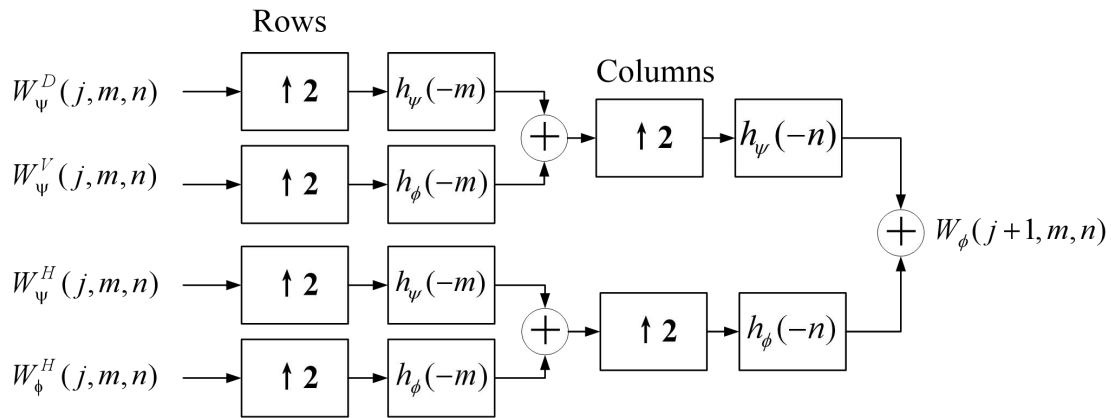


Fig. Two-scale of two-dimensional decomposition



The synthesis filter bank of the two-dimensional DWT

12. Explain the following terms 1. Boundary descriptors 2. Regional descriptors 3. Texture descriptors

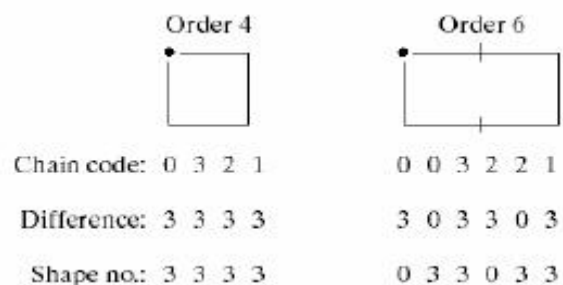
I. Simple Descriptors

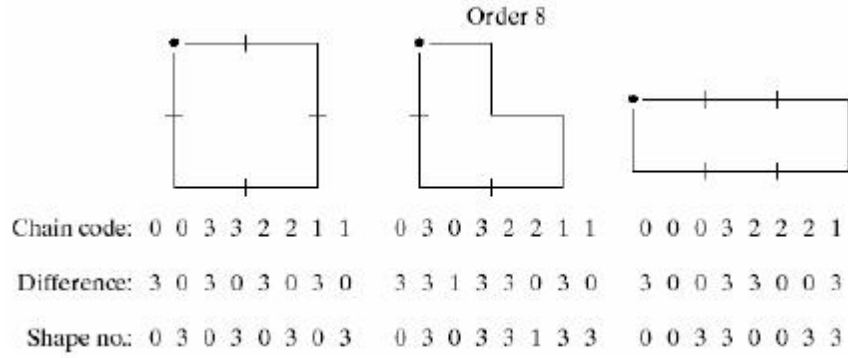
The following are few measures used as simple descriptors in boundary descriptors,

- Length of a boundary
 - Diameter of a boundary
 - Major and minor axis
 - Eccentricity and
 - Curvature
- **Length of a boundary** is defined as the number of pixels along a boundary. Eg. for a chain coded curve with unit spacing in both directions the number of vertical and horizontal components plus $\sqrt{2}$ times the number of diagonal components gives its exact length
- **The diameter of a boundary B** is defined as
 $\text{Diam}(B) = \max[D(p_i, p_j)]$
 D-distance measure
 p_i, p_j - points on the boundary
- The line segment connecting the two extreme points that comprise the diameter is called **the major axis of the boundary**
- **The minor axis of a boundary** is defined as the line perpendicular to the major axis
- **Eccentricity of the boundary** is defined as the ratio of the major to the minor axis
- **Curvature** is the rate of change of slope.

II. Shape number

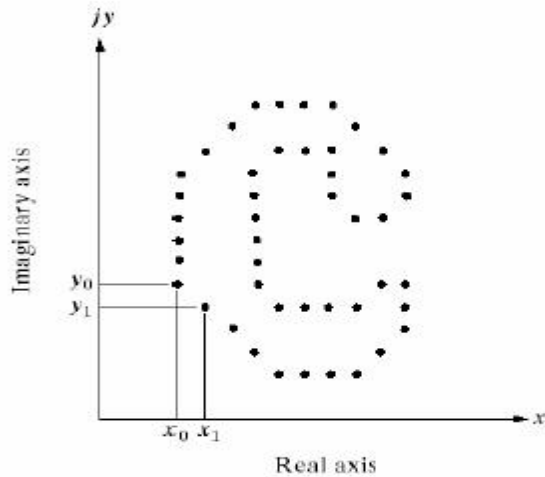
- Shape number is defined as the **first difference of smallest magnitude**.
- **The order n** of a shape number is the number of digits in its representation.





III. Fourier Descriptors

- Consider the following digital boundary and its representation as a complex sequence in the xy plane



- It is seen that the boundary is depicted by the coordinate pairs such as $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots, (x_{K-1}, y_{K-1})$.
- Thus the boundary is represented as a sequence of coordinate pairs,

$$S(k) = [x(k), y(k)]$$
- Also each coordinate pair is expressed as a complex number,

$$S(k) = x(k) + jy(k), k = 0, 1, 2, \dots, K-1$$
- **The Discrete Fourier transform of $S(k)$ is given as,**

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K}, u = 0, 1, 2, \dots, K-1$$

These complex coefficients $a(u)$ are called the Fourier descriptors of the boundary.

- **The Inverse Fourier transform of these coefficients restores $s(k)$ given as,**

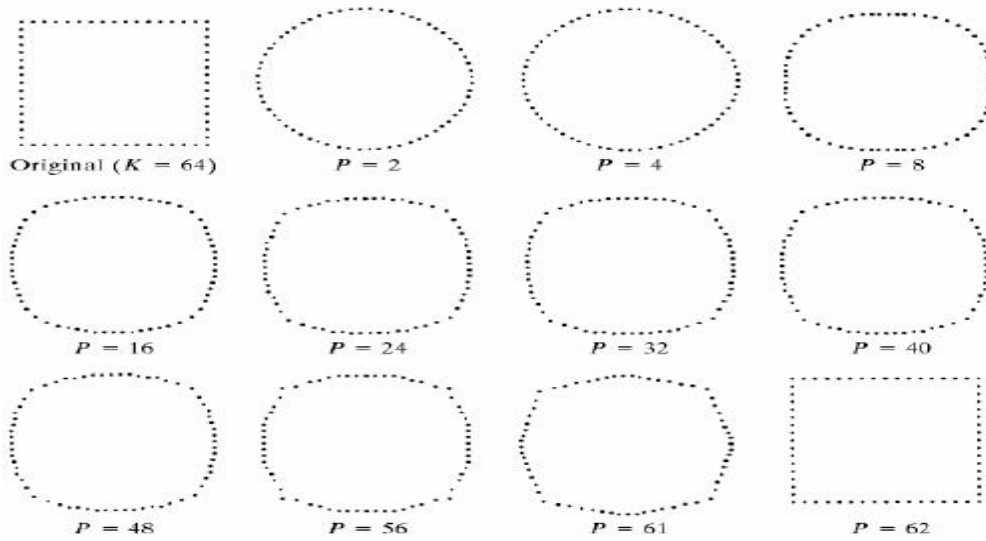
$$s(k) = \sum_{u=0}^{K-1} a(u) e^{j2\pi k u / K}, k = 0, 1, 2, \dots, K-1$$

- Instead of using all the Fourier coefficients, only the first P coefficients are used.
(i.e.,) $a(u) = 0$, for $u > P-1$

Hence the resulting approximation is given as,

$$\hat{s}(k) = \sum_{u=0}^{P-1} a(u) e^{j2\pi k u / K}, k = 0, 1, 2, \dots, K-1$$

- Thus, when P is too low, then the reconstructed boundary shape is deviated from the original..
- **Advantage: It reduces a 2D to a 1D problem.**
- The following shows the reconstruction of the object using Fourier descriptors. As the value of P increases, the boundary reconstructed is same as the original.

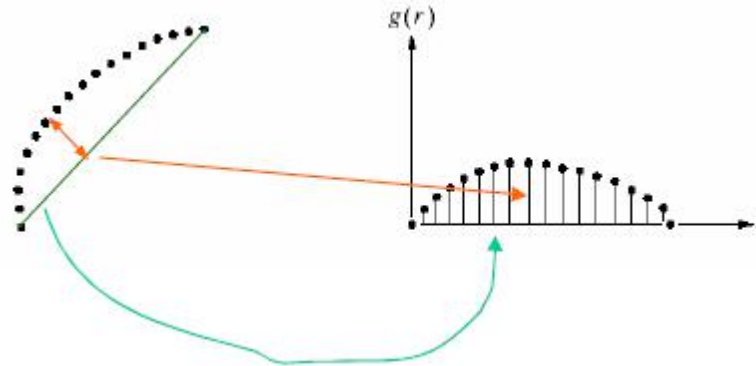


Basic properties of Fourier descriptors

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k) e^{j\theta}$	$a_r(u) = a(u) e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy} \delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u) e^{-j2\pi k_0 u / K}$

IV. Statistical moments

- It is used to describe the shape of the boundary segments, such as the mean, and higher order moments
- Consider the following boundary segment and its representation as a 1D function



- Let v be a discrete random variable denoting the amplitude of g and let $p(v_i)$ be the corresponding histogram, Where $i = 0, 1, 2, \dots, A-1$ and A = the number of discrete amplitude increments in which we divide the amplitude scale, then the n th moment of v about the mean is ,

$$\mu_n(v) = \sum_{i=0}^{A-1} (v_i - m)^n p(v_i), i = 0, 1, 2, \dots, A-1$$

Where m is the mean value of v and , μ_2 is its variance ,

$$m = \sum_{i=0}^{A-1} v_i p(v_i)$$

- Another approach is to normalize $g(r)$ to unit area and consider it as a histogram, i.e., $g(r)$ is the probability of occurrence of value r_i and the moments are given as,

$$\mu_n(v) = \sum_{i=0}^{K-1} (r_i - m)^n g(r_i), i = 0, 1, 2, \dots, K-1$$

Where,

$$m = \sum_{i=0}^{K-1} r_i g(r_i)$$

K is the number of points on the boundary.

- The advantage of moments over other techniques is that the implementation of moments is straight forward and they also carry a 'physical interpretation of boundary shape'

Regional and simple descriptors

The various **Regional Descriptors** are as follows,

- Simple Descriptors**
- Topological Descriptors**
- Texture**

Which are explained as follows,

a) Simple Descriptors

The following are **few measures used as simple descriptors in region descriptors**

- **Area**

- **Perimeter**
 - **Compactness**
 - **Mean and median of gray levels**
 - **Minimum and maximum of gray levels**
 - **Number of pixels with values above and below mean**
- The Area of a region is defined as the number of pixels in the region.
 - The Perimeter of a region is defined as the length of its boundary
 - Compactness of a region is defined as $(\text{perimeter})^2/\text{area}$. It is a dimensionless quantity and is insensitive to uniform scale changes.

b) Topological Descriptors

- Topological properties are used for global descriptions of regions in the image plane.
- **Topology is defined as the study of properties of a figure that are unaffected by any deformation, as long as there is no tearing or joining of the figure.**
- The following are the **topological properties**
 - **Number of holes**
 - **Number of connected components**
 - **Euler number**
- **Euler number is defined as $E = C - H$,**
Where, C is connected components
H is the number of holes
- The Euler number can be applied to straight line segments such as polygonal networks
- Polygonal networks can be described by the number of vertices V, the number of edges A and the number of faces F as,
$$E = V - A + F$$
- Consider the following region with two holes

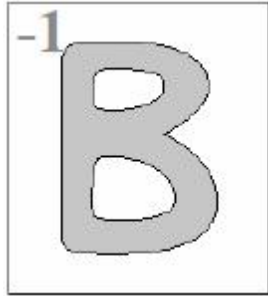


Here connected component $C = 1$

Number of holes $H = 2$

Euler number $E = C - H = 1 - 2 = -1$

- Consider the following letter B with two holes

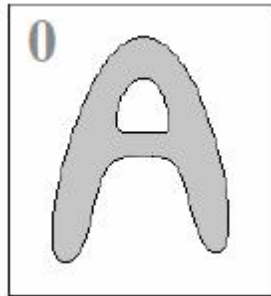


Here connected component $C = 1$

Number of holes $H = 2$

Euler number $E = C - H = 1 - 2 = -1$

- Consider the following letter A with one hole



Here connected component $C = 1$

Number of holes $H = 1$

Euler number $E = C - H = 1 - 1 = 0$

c) Texture

- **Texture refers to repetition of basic texture elements known as texels or texture primitives or texture elements**
- Texture is one of the regional descriptors.
- It provides measures of properties such as smoothness, coarseness and regularity.
- There are 3 approaches used to describe the texture of a region.

They are:

- Statistical
- Structural
- Spectral

i. **statistical approach**

- Statistical approaches describe smooth, coarse, grainy characteristics of texture. This is the simplest one compared to others. It describes texture using statistical moments of the gray-level histogram of an image or region.
- The following are **few measures used as statistical approach in Regional descriptors**
- Let z be a random variable denoting gray levels and let $p(z_i)$ be the corresponding histogram,
Where $I = 0, 1, 2, \dots, L-1$ and L = the number of distinct gray levels, then the n th moment of z about the mean is ,

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i), i = 0, 1, 2, \dots, L-1$$

Where m is the mean value of

z and it defines the average gray level of each region, is given as,

$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$

- The second moment is a measure of smoothness given as,

$$R = 1 - \frac{1}{1 + \sigma^2(z)}, \text{ gray level contrast}$$

R = 0 for areas of constant intensity, and R = 1 for large values of $\sigma^2(z)$

- The normalized relative smoothness is given as,

$$R = 1 - \frac{1}{1 + \frac{\sigma^2(z)}{(L-1)^2}}$$

- The third moment is a measure of skewness of the histogram given as,

$$\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$$

- The measure of uniformity is given as,

$$U = \sum_{i=0}^{L-1} p^2(z_i)$$

- An average entropy measure is given as,

$$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

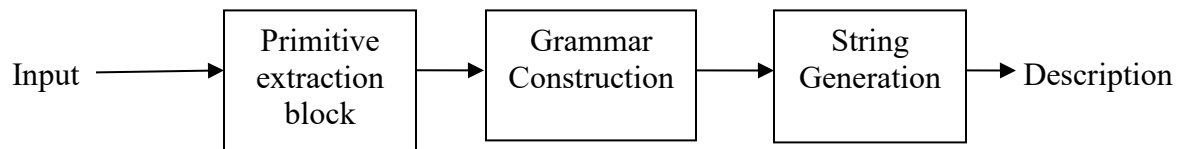
Entropy is a measure of variability and is zero for constant image.

- The above measure of texture are computed only using histogram and have limitation that they carry no information regarding the relative position of pixels with respect to each other.
- In another approach, not only the distribution of intensities is considered but also the position of pixels with equal or nearly equal intensity values are considered.
- Let **P** be a position operator and let **A** be a k x k matrix whose element a_{ij} is the number of times with gray level z_i occur relative to points with gray level z_j , with $1 \leq i, j \leq k$.
- Let **n** be the total number of point pairs in the image that satisfy **P**. If a matrix **C** is formed by dividing every element of **A** by **n**, then C_{ij} is an estimate of the joint probability that a pair of points satisfying **P** will have values (z_i, z_j) . The matrix **C** is called the **gray-level co-occurrence matrix**.

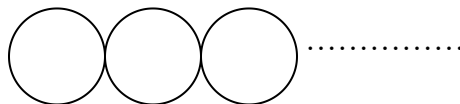
ii. Structural approach

- Structural approach deals with the arrangement of image primitives such as description of texture based on regularly spaced parallel lines.

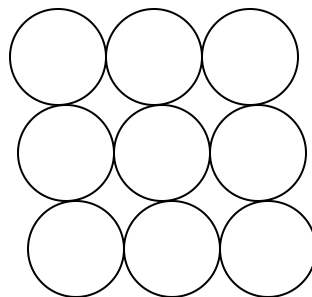
- This technique uses the structure and inter relationship among components in a pattern.
- The block diagram of structural approach is shown below,



- The input is divided into simple subpatterns and given to the primitive extraction block. Good primitives are used as basic elements to provide compact description of patterns
- Grammar construction block is used to generate a useful pattern description language
- If **a** represents a circle, then **aaa..** means circles to the right, so the rule **aS** allows the generation of picture as given below,



- Some new rules can be added,
 $S \rightarrow dA$ (**d** means 'circle down')
 $A \rightarrow lA$ (**l** means 'circle to the left')
 $A \rightarrow l$
 $A \rightarrow dA$
 $S \rightarrow a$
 The string **aaadlldaa** can generate 3 x 3 matrix of circles.



iii. Spectral approaches

- Spectral approach is based on the properties of the Fourier spectrum and is primarily to detect global periodicity in an image by identifying high energy, narrow peaks in spectrum.
- There are 3 features of Fourier spectrum that are useful for texture description. They are:
 - Prominent peaks in spectrum gives the principal direction of texture patterns.
 - The location of peaks in frequency plane gives fundamental spatial period of patterns.
 - Eliminating any periodic components by our filtering leaves

non- periodic image elements

- The spectrum in polar coordinates are denoted as $S(r, \theta)$, where S is the spectrum function and r and θ are the variables in this coordinate system
- For each direction θ , $S(r, \theta)$ may be considered as a 1D function $S_\theta(r, \theta)$ defined as

$$S(r) = \sum_{\theta=0}^{\pi} S_\theta(r)$$

- For each direction r , $S(r, \theta)$ may be considered as a 1D function $S_r(r, \theta)$ defined as

$$S(\theta) = \sum_{r=1}^R S_r(\theta)$$

- By varying above co-ordinates we can generate two 1D functions, $S(r)$ and $S(\theta)$, that gives the spectral – energy description of texture for an entire image or region.

13.Explain Object Recognition in detail

Object Recognition

The automatic recognition of *objects* or *patterns* is one of the important image analysis tasks. The approaches to pattern recognition are divided into two principal areas:

- Decision-theoretic methods: deal with patterns described using quantitative descriptors, such as length, area, and texture.
- Structural methods: deal with patterns best described by qualitative descriptors (symbolic information), such as the relational descriptors.

Patterns and Pattern Classes

- A pattern is an arrangement of descriptors (or features).
- A pattern class is a family of patterns that share some common properties. Pattern classes are denoted w_1, w_2, \dots, w_N where N is the number of classes.
- Pattern recognition by machine involves techniques for assigning patterns to their respective classes—automatically and with as little human intervention as possible.

The object or pattern recognition task consists of two steps:

- feature selection (extraction)
- matching (classification)

There are three common pattern arrangements used in practice:

- Numeric vectors (for quantitative descriptions)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Strings and trees (for structural descriptions)

$x = \text{abababa} \dots$

Recognition Based on Decision-Theoretic Methods

These methods are based on the use of decision functions. Let $x = (x_1, x_2, \dots, x_n)^T$ represent an n -dimensional pattern vector. For N known pattern classes w_1, w_2, \dots, w_N the idea here is to find N decision functions $d_1(x), d_2(x), \dots, d_N(x)$ with the property that, if an unknown pattern x is said to belong to the i^{th} pattern class if, upon substitution of x into all decision functions, $d_i(x)$ yields the largest numerical value.

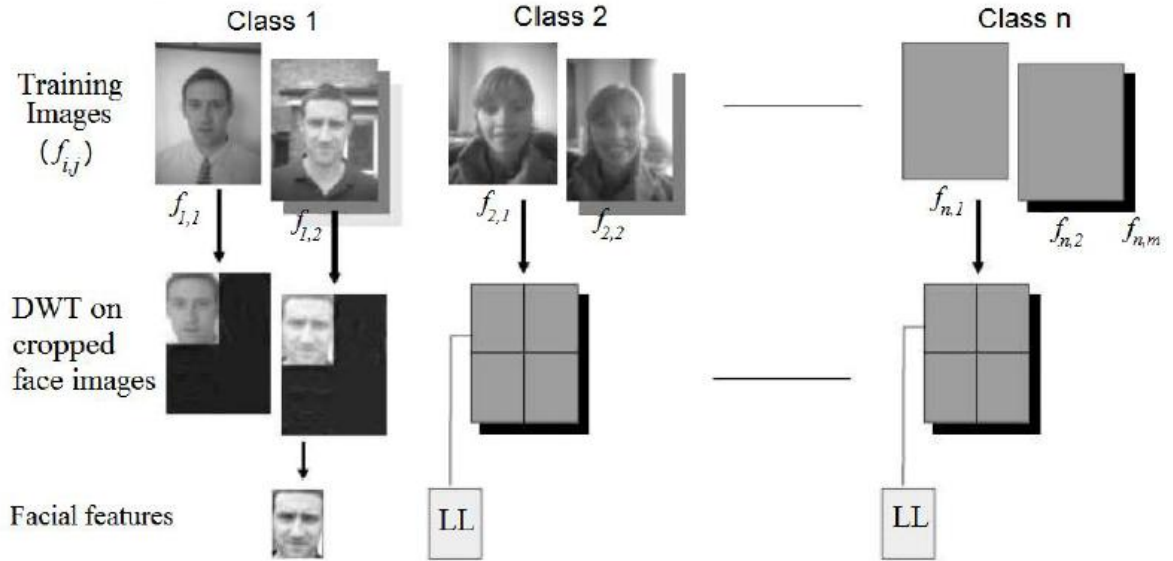
Matching

Recognition techniques based on matching represent each class by a prototype pattern vector. Set of patterns of known classes is called the training set. Set of patterns of unknown classes is called the testing set. An unknown pattern is assigned to the class to which it is closest in terms of a predefined metric. The simplest approach is the minimum-distance classifier, which, as its name implies, computes the (Euclidean) distance between the unknown and each of the prototype vectors. Then, it chooses the smallest distance to make a decision.

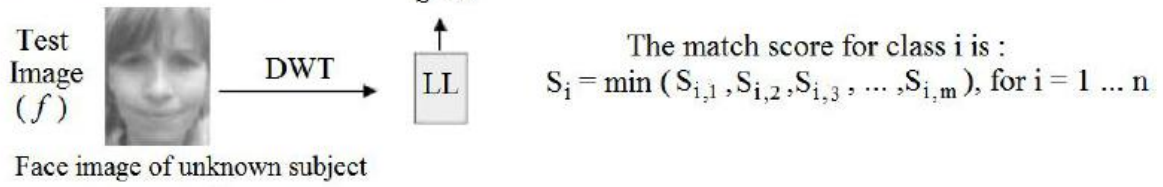
Wavelet-based Face Recognition Application

In the enrolment stage, each face image in the training set is transformed to the wavelet domain to extract its pattern vector (i.e. subband). The choice of an appropriate subband varies depending on the operational circumstances of the face recognition application. The decomposition level is predetermined based on the efficiency and accuracy requirements and the size of the face image. In the recognition stage, a minimum-distance classification method was used to classify the unknown face

Enrolment Process



Identification Process



Let the set $F = \{f_{i,1}, f_{i,2}, f_{i,3}, \dots, f_{i,m}\}$ be a training set of face images of n subjects, where each subject i has m images. In the enrolment stage, wavelet transform is applied on each training image so that a set $W_k(F)$ of multi-resolution decomposed images result. A new set $LL_k(F)$ of all k -level LL -subbands will be obtained from the transformed face images in the set $W_k(F)$. The new set $LL_k(F)$ forms the set of features for the training images. Thus, the training face image I of subject i ($f_{i,1}$) is expressed by its feature vector $LL_{k,i,1}$. The collection of feature vectors $LL_{k,i,1}, LL_{k,i,2}, \dots, LL_{k,i,m}$ represents the stored template of subject i . In a similar manner, 3 new sets $HL_k(F)$, $LH_k(F)$, $HH_k(F)$ can be created at the same decomposition level k from the k th-level HL -, LH -, HH -subbands respectively.

In the recognition phase, a minimum-distance classifier is used to classify the input face image. When a probe face image is introduced to the system, it is decomposed by wavelet transform, and a certain subband (e.g. LL_k) is chosen to represent the feature vector of the probe image. A match score $S_{i,j}$ can now be computed between the probe feature vector and each of the feature vectors j of the subject i in the feature set $LL_k(F)$. Then, the identity (i.e. class) of the training image which gives the minimum score is assigned to the probe image:

$$S_i = \min (S_{ij}) \quad (j = 1, \dots, m)$$

Many similarity measures can be used for the minimum-distance classifier, for example CityBlock or Euclidean distance functions.

QUESTION BANK

UNIT I -DIGITAL IMAGE FUNDAMENTALS

Part A

1. **What is mach band effect? (Nov /Dec 2005) , (May / June 2007)**
2. **What is a band limited function? (Nov /Dec 2005)**
3. **Which type of light receptor is responsible for (a) photopic (b) scotopic vision? (May / June 2006)**
4. **What are the stages of digital image processing system? (May / June 2009)**
5. **What are the applications of digital image processing?**
6. **What is Weber ratio? (May/June 2009), (May / June 2007)**
7. **Define intensity ratios and log ratios. (Nov / Dec 2005)**
8. **What is digital image?**
9. **Explain the terms illumination and reflectance for an image model?**
10. **What is the role of point spread function in image processing? (May / June 2006)**
11. **Write the significance of image sampling and quantization. (May / June 2009)**
12. **What are the practical limitations in sampling and reconstruction? (May / June 2006)**
13. **Stat the circular convolution theorem of DFT. (May / June 2009)**

14. **What are the properties of unitary transform? (May / June 2009)**
15. State 2D sampling theory.
16. What is the need for quantization?
17. State and explain the convolution property of Fourier transform?
18. Explain the terms aliasing and fold over frequencies?
19. What is circulant matrix? Give an example.
20. **Explain the decorrelation property of KL transform. (May / June 2006)**
21. **Compare SVD and KL transform. (May / June 2009)**
22. **Write N x N discrete cosine transform. (May / June 2009)**
23. Define the SVD transform of images.
24. What is cosine transform?
25. What is the special feature of KL transform?
26. Define unitary transform
27. Name the working principle of digital camera
28. Define dither
29. Name the color models
30. Differentiate brightness and contrast
31. Differentiate hue and saturation
32. Give the HSI model

Part-B

1. **Explain about the element of digital image processing (May / June 2009) (May / June 2007)**
2. **Explain the different elements of visual perception in detail. (May / June 2009)**
3. **Illustrate the concept of Brightness adaptation. (May / June 2007)**
4. Explain luminance and brightness and contrast with respect to human vision?
5. **Write short notes on Non uniform sampling and Quantization. (May / June 2007)**
6. **Discuss in detail the process of uniform sampling and Quantization. (May / June 2007)**
7. **State the properties of optimum mean square quantizer. (Nov / Dec 2005)**
8. **Describe the properties of 2D Fourier transform and its application in image processing. (May/June 2009)**
9. **Discuss the effects of non uniform sampling and quantization. (May/June 2009)**
10. **State and prove 2D sampling theorem. (May/June 2009)**
11. **Derive the expression for the signal to noise ratio of Lloyd max quantizer. (May/June 2009)**
12. **State the condition for a matrix to be (1) unitary (2) orthogonal. Give an example of a matrix which is both orthogonal and unitary. (May / June 2006)**
13. **Explain how a matrix which is represented using singular value decomposition. (May / June 2006)**
14. **Explain the principle of vector quantization of images. (May / June 2006)**
15. **Explain Weber's Law and Lateral inhibition with reference to human visual system. (May / June 2006)**
16. **Describe Lloyd max quantizer for image given the probability density function? (May / June 2006)**
17. **Explain any four properties of DFT. (Nov / Dec 2005)**

18. Explain the properties of DCT. (Nov / Dec 2005)
19. Explain the salient features of DCT. (May / June 2007)
20. Find the DFT for the set of values $\{1,0,1,2,0,0,1\}$ (May / June 2009)
21. List the properties of 2D-DFT. (May / June 2009) (May / June 2007)
22. Explain how the basis vector of KL transform of a $N \times 1$ random vector is obtained. (May / June 2006)
23. Explain about the steps in digital image processing
24. Explain the working principle of digital camera(VIDICON)
25. Write short notes on Color image fundamentals
26. Explain about dither

UNIT II- IMAGE ENHANCEMENT

Part A

1. What is meant by histogram of a digital image? (May / June 2006), (May / June 2009), (May / June 2007)
2. List the use of LPF, HPF and BPF spatial filters. (May / June 2009)
3. Explain any one technique for zooming of an image. (May / June 2006)
4. What is image averaging?
5. Write the transfer function for Butterworth filter. (May/June 2009)
6. What is mean by contrast stretching? (May/June 2009)
7. What is Median Filtering? (May / June 2009) (May / June 2007)
8. Why the image enhancement is needed in image processing technique. (May / June 2009)
9. If $y(m) = (2,3,8,4,2)$ and $N = (-1,0,1)$. Find the medial filter output. (Nov / Dec 2005)
10. What is meant by histogram specification?
11. What are the sources of degradation in image?
12. What is pseudo inverse filter?
13. Explain the importance of wiener filter?
14. Explain the need for smoothing and sharpening filtering?
15. Define the ceptrum of a signal
16. What is pseudo inverse operator?
17. What is Geometric mean filter?
18. What is Harmonic mean filter?
19. What is a Contra harmonic mean filter?
20. What is Color image enhancement?

Part-B

2. Explain how image subtraction and image averaging is used to enhance the image. (May / June 2009)
3. Explain the various sharpening filters in spatial domain. (May / June 2009)
4. Discuss in detail the significance of Homomorphic filtering in image enhancement. (May / June 2009) (May / June 2007)
5. Write short notes on Point operations, Transform operations and Pseudo coloring with respect to image enhancement. (May / June 2009) (May / June 2006)
6. Explain any two spatial operations. (May / June 2006)

7. Write short notes on Contrast stretching and Grey level slicing. (May / June 2009)
8. Explain in detail about Sharpening filter, Edge Detection, Smoothing filter and Noise removal. (May / June 2009)
9. What is histogram of an image? How does it modify an image in equalization and specification techniques? (May / June 2009)
10. Explain Enhancement using point operations. (Nov / Dec 2005) (May / June 2006)
11. Write short notes on Histogram equalization and Histogram modification. (Nov / Dec 2005) (May / June 2009) (May / June 2007)
12. Write short notes on Directional Smoothing and color image enhancement. (Nov / Dec 2005) (May / June 2006)
13. Explain how a high pass filtered image can be generated using a low pass filter. (May / June 2006)
14. How is the digital negative of an image generated? Where is this used? (May / June 2006)
15. Discuss the role of non linear filters in image enhancement. (May / June 2007)
16. Explain any one method of pseudo color image processing. (May / June 2007)
17. Write short notes on Geometric mean, Harmonic mean, Contra harmonic mean filters

UNIT III- IMAGE RESTORATION

Part A

1. What is pseudo inverse?
2. What is Blind Image Restoration? (May/June 2009)
3. What is Constrained Restoration? (May/June 2009)
4. What is speckle noise?
5. Write the transfer function of pseudo inverse filter. (May / June 2006)
6. Explain any two reasons for image degradation. (May / June 2006)
7. What are the factors that influence image degradation ? (May / June 2009)
8. Distinguish between image enhancement and image restoration. (May / June 2007)
9. What is inverse filtering?
10. What is the basic equation used for constrained restoration?
11. Define Geometric transformations
12. What are the spatial transformations?
13. Differentiate unconstrained restoration and constrained restoration
14. What is meant by Lagrange multiplier?
15. Classify Image restoration techniques.
16. What is the expression for the error measure for Weiner filter?
17. Give the expression for Weiner filtering and give its other names for Weiner filter.
18. What is the PDF of Gaussian Noise?

Part-B

1. Explain in detail the constrained least squares restoration. (May / June 2009)

2. **Write notes on inverse filtering as applied to image restoration. (May / June 2009)**
3. **Give the degradation model for continuous function. (May / June 2009)**
4. **With the aid of block diagram, describe the digital image restoration system and explain the image observation models**
5. **Explain in detail the function of Wiener filter. (May / June 2009)**
6. **Write short notes on Inverse filter and Pseudo inverse filter. (Nov / Dec 2005)**
7. **Write short notes on Wiener filter characteristics and Geometric mean filter. (Nov / Dec 2005)**
8. **Derive the Wiener filter equation. Under what conditions does it become the pseudo inverse filter? (May / June 2006)**
9. **Explain the image degradation model and various approaches of image restoration technique. (May / June 2009)**
10. **Explain Constrained and Unconstrained Restoration in detail. (May / June 2009)**
11. **Describe the principle of wiener filtering in image Restoration.**
12. Explain the method of removal of blur caused by uniform linear motion?
13. Discuss the differences between Geometric transformations and spatial transformations techniques?

UNIT IV- IMAGE SEGMENTATION

Part A

1. **How do you detect isolated points in an image? (May/June 2009)**
2. **What is meant by Edge detection? (May / June 2009)**
3. **Define Image Segmentation (May / June 2009)**
4. What are the different image segmentation techniques?
5. What are the two properties that are followed in image segmentation?
6. What are the three types of discontinuity in digital image?
7. What is edge?
8. Define zero crossing property of edge detection.
9. What is meant by gradient operators?
10. Define Laplacian.
11. Write about linking edge points.
12. **How do you detect isolated points in an image? (May/June 2009)**
13. What is the role of gradient operator and laplacian operator in segmentation?
14. What are the types of thresholding?
15. What is global, Local and dynamic or adaptive threshold?
16. Define region growing?
17. Specify the steps involved in splitting and merging?
18. What is meant by region – Based segmentation?
19. Give an applications of segmentation
20. What is the morphological operation?
21. What is watershed segmentation?

Part-B

1. **Explain on the Region based segmentation techniques. (May / June 2009)**

2. **Explain edge detection.**
3. **Explain Edge linking and boundary detection in detail**
4. **Explain Thresholding in detail.**
5. Write short notes on Region growing segmentation. Or Explain Region splitting and Merging (May / June 2009)
6. **Explain segmentation by morphological operation in detail**
7. **Explain water segmentation algorithm in detail**

UNIT V- IMAGE COMPRESSION

Part A

1. **What is need for image data compression? (May / June 2009)**
2. **List the types of coding in image processing. (May / June 2009)**
3. **List the basic types of data redundancy. (May / June 2009)**
4. **Give the lossless predictive model. (May / June 2009)**
5. **Give an example of a variable length encoder. What is its advantage? (May / June 2006)**
6. **What are the standards used for compression of still images and images with motion? (May / June 2006)**
7. **What are the advantages of image data compression? (May / June 2009)**
8. **Mention some image data compression techniques. (May / June 2009)**
9. **What is the average length of code if the probability of symbols is given as {0.4, 0.3, 0.1, 0.1, 0.06, 0.04}? (Nov / Dec 2005)**
10. **What is known as bit plane coding? (May / June 2007)**
11. **Mention the significant features of Wavelet transform. (May / June 2007)**
12. Define entropy coding
13. Draw the general block diagram for an image compression model
14. What is zonal coding?
15. State Huffman coding algorithm
16. What is the basic principle of image compression in spatial domain?
17. Mention any two image compression standards used?
18. Mention the limitations of bit plane coding.
19. Compare jpeg and jpeg 2000
20. Compare the Huffman and arithmetic coding
21. Define Shift codes

Part-B

1. **Explain in detail the Huffman coding procedure with an example. (May / June 2009), (May / June 2006)**
2. **Describe arithmetic coding of images with an example. (May / June 2006)**
3. **Explain the Pixel coding with example. (May / June 2009)**
4. **Describe on the Wavelet coding of images. (May / June 2009)**
5. **Explain in detail the method of Zonal and Threshold coding. (May / June 2009) , (May / June 2006)**
6. **Briefly discuss the MPEG compression standard. (May / June 2009)**
7. **Explain image compression and how it is achieved through wavelet transform. (May / June 2009)**

8. **Write notes on JPEG standard (jpeg and jpeg 2000)with neat diagram. (May / June 2009), (May / June 2006)**
9. **Apply Huffman coding procedure to the following message ensemble and determine average length of encoded message and coding efficiency. The symbols (x1,x2,x3,x4,x5,x6,x7,x8) are emitted with probabilities of (0.22,0.2,0.18,0.15,0.08,0.05,0.02). (May / June 2009)**
10. **Explain the run length coding scheme. (May / June 2006)**
11. **Explain the image compression standards**
12. What are the building blocks of an image encoder? State their function?
13. **Explain Transform coding. (May / June 2009), (May / June 2009)**
14. Explain Vector Quantization with neat diagram
15. Compute the Huffman's code and IB code following gray level distribution.

levels(r)	W	W2	W3	W4	W5	W6
probability p(r)	0.4	0.3	0.2	0.5	0.3	0.2