SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

**SCHOOL OF BIO AND CHEMICAL ENGINEERING**

**DEPARTMENT OF BIOMEDICAL ENGINEERING**

# UNIT – I – Medical Image Processing – SBMA1603

# UNIT-1

# IMAGE FUNDAMENTALS

## 1.1 Introduction

The digital image processing deals with developing a digital system that performs operations on a digital image. An image is nothing more than a two dimensional signal. It is defined by the mathematical function f(x,y) where x and y are the two co-ordinates horizontallyand vertically and the amplitude of f at any pair of coordinate (x, y) is called the intensity or gray level of the image at that point. When x, y and the amplitude values of f are all finite discretequantities, we call the image a digital image. The field of image digital image processing refersto the processing of digital image by means of a digital computer.

A digital image is composed of a finite number of elements, each of which has a particular location and values of these elements are referred to as picture elements, image elements, pels and pixels.

Digital image processing focuses on two major tasks

- Improvement of pictorial information for human interpretation
- Processing of image data for storage, transmission and representation for autonomous machine perception.

Digital image processing deals with manipulation of digital images through a digital computer. It is a subfield of signals and systems but focus particularly on images. DIP focuses on developing a computer system that is able to perform processing on an image. The input of thatsystem is a digital image and the system process that image using efficient algorithms, and gives an image as an output. The most common example is Adobe Photoshop. It is one of the widelyused applications for processing digital images.

## 1.2 Structure of Human Eye:

The eye is nearly a sphere with average approximately 20 mm diameter. The eye is enclosed with three membranes

a) The cornea and sclera - it is a tough, transparent tissue that covers the anterior surface of the eye. Rest of the optic globe is covered by the sclera

b) The choroid –It contains a network of blood vessels that serve as the major source of nutrition to the eyes. It helps to reduce extraneous light entering in the eye.

It has two parts

Iris Diaphragms- it contracts or expands to control the amount of light that enters the eyes

Ciliary body

c) Retina – it is innermost membrane of the eye. When the eye is properly focused, light from an object outside the eye is imaged on the retina. There are various light receptors over the surface of the retina. The two major classes of the receptors are-

1) cones- it is in the number about 6 to 7 million. These are located in the central portion of the

retina called the fovea. These are highly sensitive connected to its own nerve end. Cone vision is called photopicor bright light vision

2) Rods – these are very much in number from 75 to 150 million and are distributed over the entire retinal surface. The large area of distribution and the fact that several roads are connected

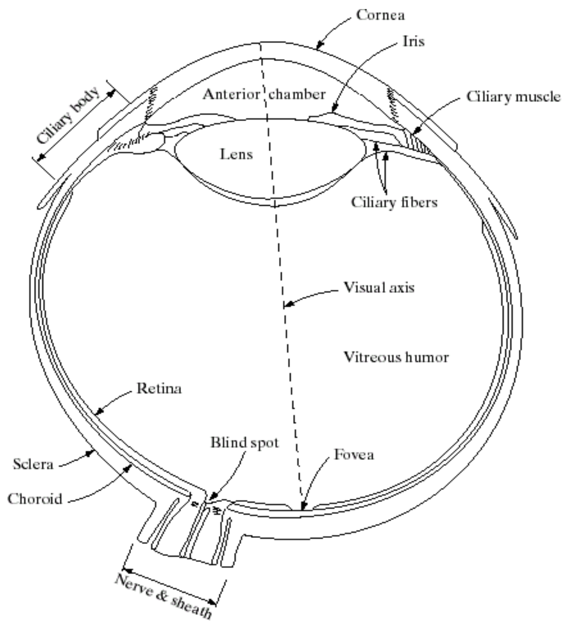



**Fig 1.1 Structure of Eye**

to a single nerve give a general overall picture of the field of view. They are not involved in the color vision and are sensitive to low level of illumination. Rod vision is called is scotopic or dim light vision. The absent of reciprocators is called blind spot.

**1.3 Image Formation in the Eye:**

- The major difference between the lens of the eye and an ordinary optical lens in that the former is flexible.
- The shape of the lens of the eye is controlled by tension in the fiber of the ciliary body. To focus on the distant object the controlling muscles allow the lens to become thicker in order to focus on object near the eye it becomes relatively flattened.
- The distance between the center of the lens and the retina is called the focal length and it varies from 17mm to 14mm as the refractive power of the lens increases from its minimum to its maximum**.**
- When the eye focuses on an object farther away than about 3m.the lens exhibits its lowest refractive power. When the eye focuses on a nearly object. The lens is most strongly refractive.
- The retinal image is reflected primarily in the area of the fovea. Perception then takes

place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

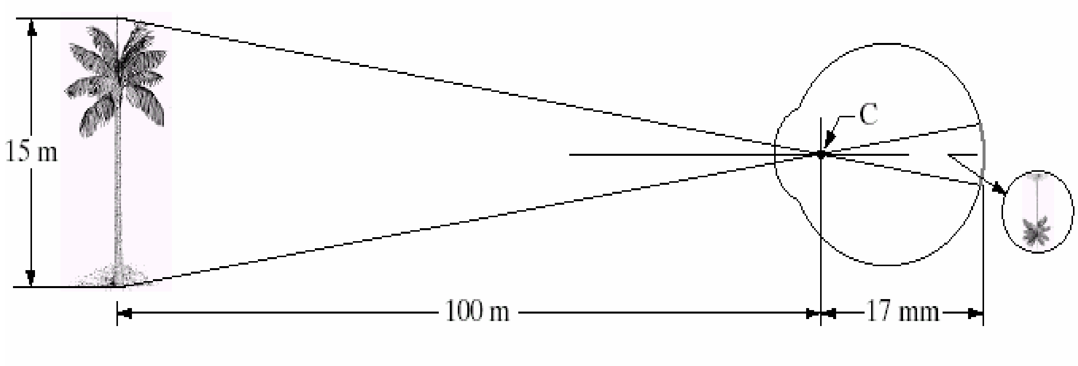
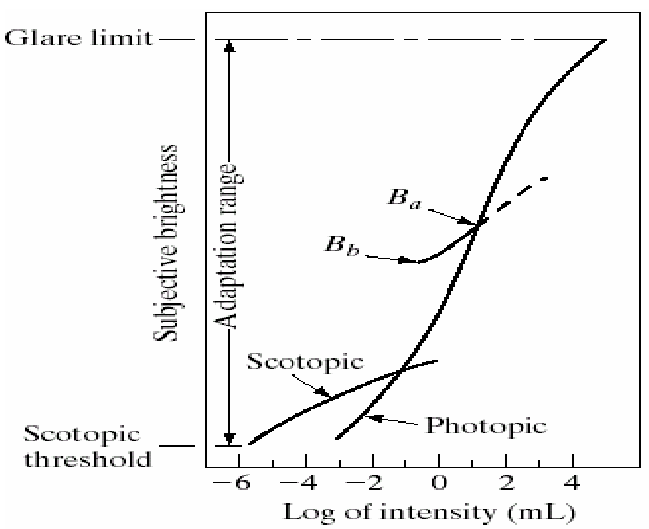



**Fig 1.2 Image Formation in eye**

- Focal length of the eye: 17 to 14 mm
- Let *h* be the height in mm of that object in the retinal image,then 15/100 = *h* / 17 , *h* = 2.55mm
- The retinal image is reflected primarily in the area of the fovea.

**1.4 Brightness Adaption and Discrimination:**

Digital image are displayed as a discrete set of intensities. The range of light intensity levels to which the human visual system can adopt is enormous- on the order of $10^{10}$-from scotopic threshold to the glare limit. Experimental evidences indicate that subjective brightness is a logarithmic function of the light intensity incident on the eye.



- The curve represents the range of intensities to which the visual system can adopt. But the visual system cannot operate over such a dynamic range simultaneously. Rather, it is accomplished by change in its overcall sensitivity called brightness adaptation.
- For any given set of conditions, the current sensitivity level to which of the visual system is called brightness adoption level , Ba in the curve. The small intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level. It is restricted at level Bb, at and below which all stimuli are perceived as indistinguishable blacks. The upper portion of the curve is not actually restricted. Wholesimply raise the adaptation level higher than Ba.
- The ability of the eye to discriminate between changes in light intensity at any specific adaptation level is also of considerable interest.
- Take a flat, uniformly illuminated area large enough to occupy the entire field of view of the subject. It may be a diffuser such as an opaque glass, that is illuminated from behind by a light source whose intensity, I can be varied. To this field is added an increment of illumination ΔI in the form of a short duration flash that appears as circle in the center of the uniformly illuminated field.
- If ΔI is not bright enough, the subject cannot see any perceivable changes.

### 1.5 Fundamental Steps in Digital Image Processing:

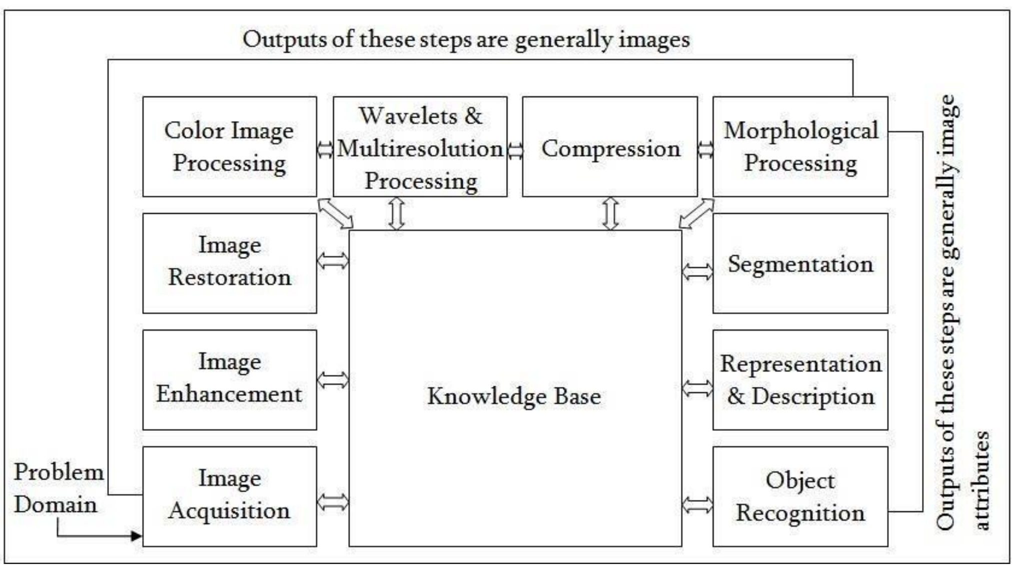



**Fig 1.3 Fundamental Steps of Digital Image Processing**

There are two categories of the steps involved in the image processing

(1) Methods whose outputs are input are images.

(2) Methods whose outputs are attributes extracted from those images.

There are some fundamental steps but as they are fundamental, all these steps may have sub-steps. The fundamental steps are described below with a neat diagram.

**Image Acquisition:** This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling etc.

**Image Enhancement:** Image enhancement is 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. Such as, changing brightness & contrast etc.

**Image Restoration:** Image restoration is an area that also deals with improving the appearanceof 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 modelsof image degradation.

**Color Image Processing:** Color image processing is an area that has been gaining its importancebecause of the significant increase in the use of digital images over the Internet. This may includecolor modeling and processing in a digital domain etc.

**Wavelets and Multiresolution Processing:** Wavelets are the foundation for representing images in various degrees of resolution. Images subdivision successively into smaller regions for data compression and for pyramidal representation.

**Compression:** Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data. It has to major approaches a) Lossless Compression b) Lossy Compression.

**Morphological Processing:** Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

**Segmentation:** 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.

**Representation and Description:** Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundaryof a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

**Object recognition:** Recognition is the process that assigns a label, such as, “vehicle” to an object based on its descriptors.

**Knowledge Base:** Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conductedin seeking that information. The knowledge base also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

### 1.6 Components of Image processing system:

**Image Sensors:** With reference to sensing, two elements are required to acquire digital image. The first is a physical device that is sensitive to the energy radiated by the object we wish to image and second is specialized image processing hardware.

**Specialize image processing hardware:** It consists of the digitizer just mentioned, plus hardware that performs other primitive operations such as an arithmetic logic unit, which performs arithmetic such addition and subtraction and logical operations in parallel on images

**Computer:** It is a general purpose computer and can range from a PC to a supercomputer depending on the application. In dedicated applications, sometimes specially designed computer are used to achieve a required level of performance.

**Software:** It consist of specialized modules that perform specific tasks a well-designed package also includes capability for the user to write code, as a minimum, utilizes the specialized module. More sophisticated software packages allow the integration of these modules.

**Mass storage:** This capability is a must in image processing applications. An image of size 1024 x1024pixels ,in which the intensity of each pixel is an 8- bit quantity requires one megabytes of storage space if the image is not compressed .Image processing applications falls into three principal categories of storage

i) Short term storage for use during processing

ii) On line storage for relatively fast retrieval

iii) Archival storage such as magnetic tapes and disks

**Image displays:** Image displays in use today are mainly color TV monitors. These monitors are driven by the outputs of image and graphics displays cards that are an integral part of computer system

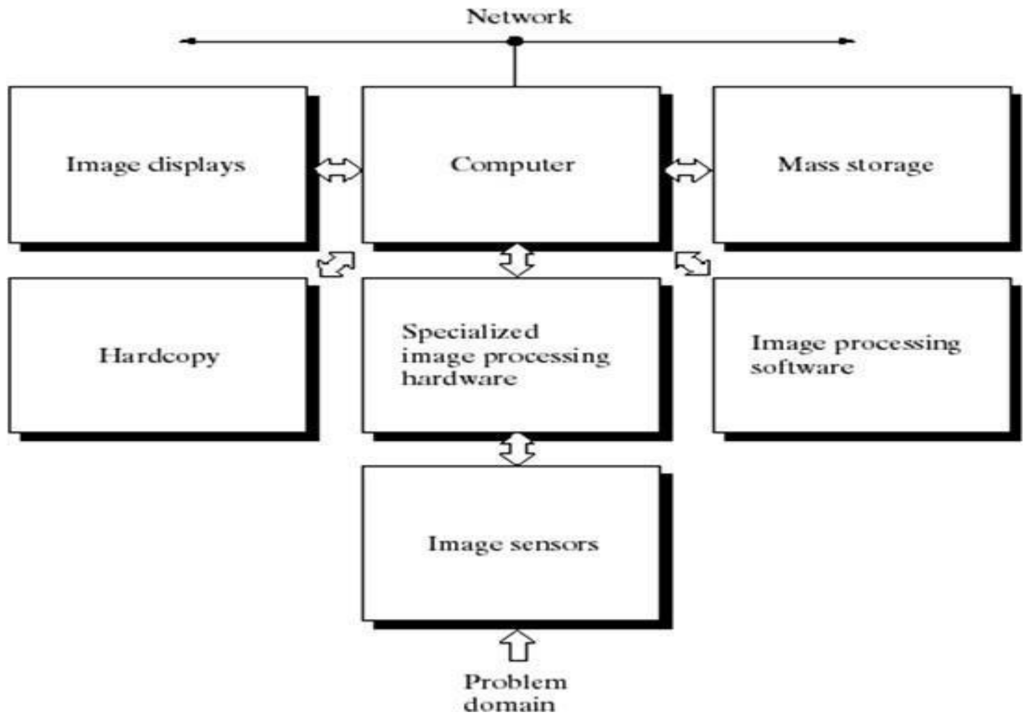



**Fig1.4 Components of Image processing system**

**Hardcopy devices :** The devices for recording image includes laser printers, film cameras, heat sensitive devices inkjet units and digital units such as optical and CD ROM disk. Films provide the highest possible resolution, but paper is the obvious medium of choice for written applications.

**Networking:** It 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 bandwidth.

**1.7 Image Formation Model:**

An image is denoted by a two dimensional function of the form f{x, y}. The value or amplitude of f at spatial coordinates {x,y} is a positive scalar quantity whose physical meaningis determined by the source of the image.

When an image is generated by a physical process, its values are proportional to energyradiated by a physical source. As a consequence, f(x,y) must be nonzero and finite; that is o<f(x,y) <co.

The function f(x,y) may be characterized by two components-

The amount of the source illumination incident on the scene being viewed.

The amount of the source illumination reflected back by the objects in the sceneThese are called illumination and reflectance components and are denoted by i (x,y) and r (x,y) respectively. The functions combine as a product to form f(x,y).

We call the intensity of a monochrome image at any coordinates (x,y) the gray level (l) of the image at that point l= f (x, y.)

L min ≤ l ≤ Lmax

Lmin is to be positive and Lmax must befinite Lmin = imin rmin

Lamx = imax rmax

The interval [Lmin, Lmax] is called gray scale. Common practice is to shift this interval numerically to the interval [0, L-l] where l=0 is considered black and l= L-1 is considered whiteon the gray scale. All intermediate values are shades of gray of gray varying from black to white.

## 1.8 Image Sampling and Quantization:

- To create a digital image, we need to convert the continuous sensed data into digital from. This involves two processes – sampling and quantization. An image may be continuous with respect to the x and y coordinates and also in amplitude.To convert it into digital form we have to sample the function in both coordinatesand in amplitudes.

- **Digitalizing the coordinate values is called sampling**
- **Digitalizing the amplitude values is called quantization**

- There is a continuous the image along the line segment AB.To simple this function, we take equally spaced samples along line AB. The location of each samples is given by a vertical tick back (mark) in the bottom part. The samples are shown as block squares superimposed on function the set of these discrete locations gives the sampled function.

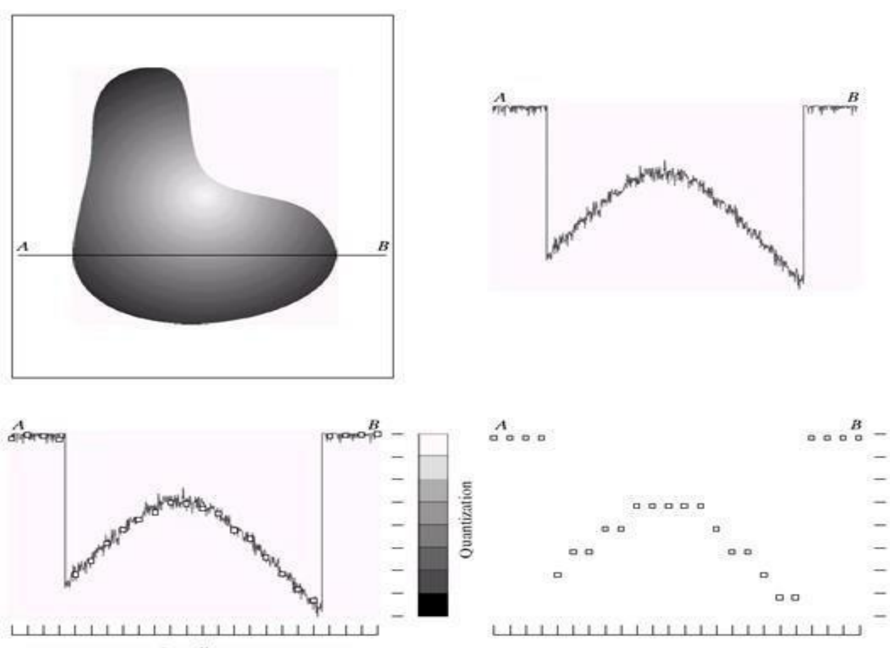



**Fig 1.5 Sampling and Quantization**

- In order to form a digital, the gray level values must also be converted (quantized) into discrete quantities. So we divide the gray level scale into eight discrete levelsranging from block to white. The vertical tick mark assign the specific value assigned to each of the eight level values.
  - The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment it made depending on the vertical proximity of a simple to a vertical tick mark.
  - Starting at the top of the image and covering out this procedure line by line produces a two dimensional digital image.
  - if a signal is sampled at more than twice its highest frequency component, then it can be reconstructed exactly from its samples.
  - But, if it is sampled at less than that frequency (called under sampling), then **aliasing** will result.
  - This causes frequencies to appear in the sampled signal that were not in the original signal.
  - Note that subsampling of a digital image will cause under sampling if the subsampling rate is less than twice the maximum frequency in the digital image.
  - **Aliasing** can be prevented if a signal is filtered to eliminate high frequencies so that its highest frequency component will be less than twice the sampling rate.
  - Gating function: exists for all space (or time) and has value zero everywhere except for a finite range of space/time. Often used for theoretical analysis of signals. But, a gating signal is mathematically defined and contains unbounded frequencies.
  - A signal which is periodic, $x(t) = x(t+T)$ for all $t$ and where $T$ is the period, has a finite maximum frequency component. So it is a bandlimited signal.
  - Sampling at a higher sampling rate (usually twice or more) than necessary to prevent aliasing is called oversampling.

**1.9 Zooming and Shrinking of Digital Images:**

Zooming may be said oversampling and shirking may be called as under sampling these techniques are applied to a digital image. These are two steps of zooming

i )Creation of new pixel locations

ii) Assignment of gray level to those new locations.

- In order to perform gray –level assignment for any point in the overly, we look for the closest pixel in the original image and assign its gray level to the new pixel in the grid. This method rowan as nearest neighbor interpolation
- **Pixel replication** - Is a special case of nearest neighbor interpolation, it is applicable if we want to increase the size of an image an integer number of times.
- For eg. - To increase the size of image as double. We can duplicate each column. This doubles the size of the image horizontal direction. To increase assignment of each of each

vertical direction we can duplicate each row. The gray level assignment of each pixel is determined by the fact that new locations are exact duplicates of old locations.

- **Drawbacks**
  - (i) Although nearest neighbor interpolation is fast, it has the undesirable featurethat it produces a check board that is not desirable
  - (ii)
- **Bilinear interpolation-**Using the four nearest neighbor of a point .let (x,y) denote the coordinate of a point in the zoomed image and let v(x1,y1) denote the gray levels assigned to it .for bilinear interpolation. The assigned gray levels is given by

$$V(x1,y1)\text{-}ax1+by1+cx1y1+d$$

Where the four coefficient are determined from the four equation in four unknowns that canbe writing using the four nearest neighbor of point (x1,y1).

- **Shrinking** is done in the similar manner .the equivalent process of the pixel replication is row–column deletion .shrinking leads to the problem of aliasing.

## 1.10 Basic relationship between pixels:

### 1) Neighbor of a pixel

There are three kinds of neighbors of a pixel:

i. N4(p) 4-neighbors: the set of horizontal and vertical neighbors

ii. ND(p) diagonal neighbors: the set of 4 diagonal neighbors

iii. N8(p) 8-neighbors: union of 4-neighbors and diagonal neighbors

A pixel p at coordinate (x,y) has four horizontal and vertical neighbor whose coordinate can be given by

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

This set of pixel called the 4-neighbours of p is denoted by n4 (p), Each pixel is a unit distance from (x,y) and some of the neighbors of P lie outside the digital image of (x,y) is on the border if the image .The four diagonal neighbor of P have coordinated

neighbors of P denoted by ns (p)

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

### 2) Adjacency

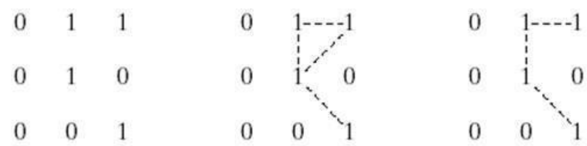
Two pixels that are neighbors and have the same grey-level (or some other specified similarity criterion) are adjacent .Pixels can be 4-adjacent, diagonally adjacent, 8- adjacent, or m-adjacent.

***m*-adjacency (mixed adjacency):**

Two pixels *p* and *q* of the same value (or specified similarity) are *m*-adjacent if either

i. *q* and *p* are 4-adjacent, or

ii. *p* and *q* are diagonally adjacent and do not have any common 4-adjacent neighbors. They cannot be both (i) and (ii).



(a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

- Path:
  - The length of the path
  - Closed path
- Connectivity in a subset *S* of an image
  - Two pixels are connected if there is a path between them that lies completely within *S*.
- Connected component of *S*:
  - The set of all pixels in *S* that are connected to a given pixel in *S*.
- Region of an image
- Boundary, border or contour of a region
- Edge: a path of one or more pixels that separate two regions of significantly different gray levels.

**3) Distance measures**

a. Distance function: a function of two points, *p* and *q*, in space that satisfies three criteria

(a) $D(p,q) = 0$

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

(c) $D(p,z) = D(p,q) + D(q,z)$

b. The Euclidean distance $D_e(p, q)$

$$D_e(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$$

c. The city-block (Manhattan) distance $D4(p, q)$

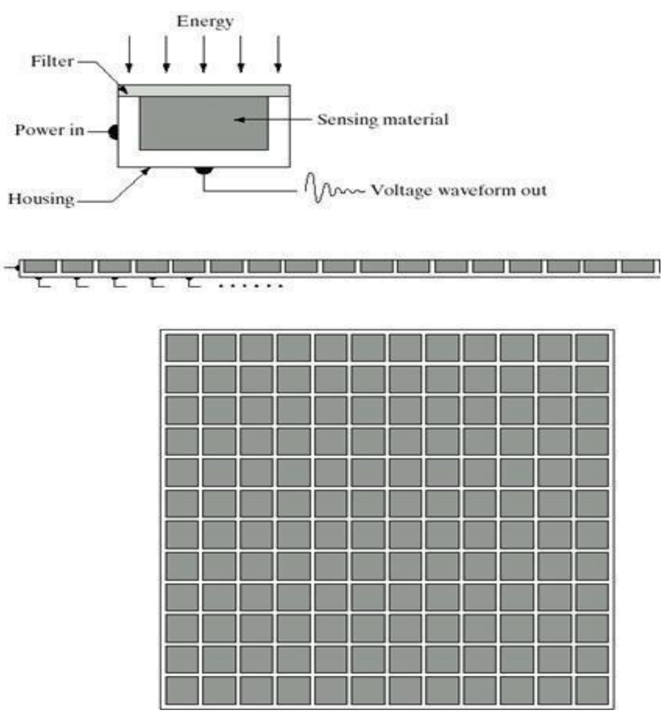
$$D_4(p, q) = | x - s | + | y - t |$$

d. The chessboard distance $D8(p, q)$

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

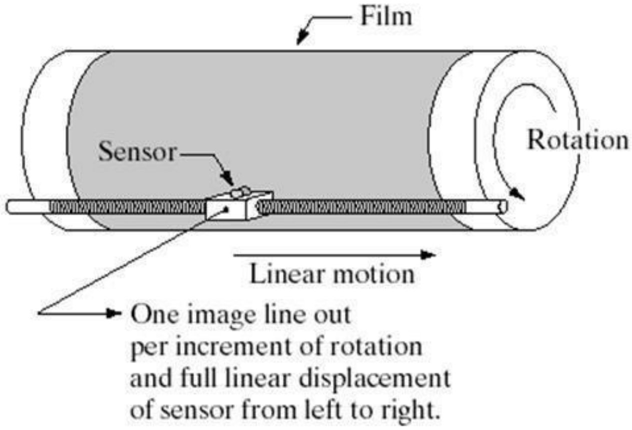
**1.11 Image sensing and acquisition:**

- Transform of illumination energy into digital images**.** The incoming energy is transformed intoa voltage by the combination of input electrical power and sensor material.
- Output voltage waveform = response of the sensor(s)
- A digital quantity is obtained from each sensor by digitizing itsresponse. Ex: Photodiode made of silicon
- Output voltage waveform proportional tolight Filter in front: increase selectivity



(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.

**Image acquisition using a single sensor**



Combining a single sensor with motion to generate a 2-D image.

**Image acquisition using sensor arrays:**

- Illumination source reflected from a scene element
- Imaging system collects the incoming energy and focus it onto an image plane (sensor array)
- Response of each sensor proportional to the integral of the light energy projected
- Sensor output: analog signal →digitized
- CCD cameras: widely used in modern applications: private consumers, industry, astronomy...CCD: Charge Couple Device

**References**

1. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing Using Matlab", 3rd Edition Tata McGraw Hill Pvt. Ltd., 2011.
2. Anil Jain K. "Fundamentals of Digital Image Processing", PHI Learning Pvt. Ltd., 2011.

**Question Bank**

| S.No | **PART - A** |
|---|---|
| 1. | Outline the steps involved in DIP. |
| 2. | Specify the elements of DIP system. |
| 3. | Categorize digital storage. |
| 4. | Mention the types of light receptors. |
| 5. | Differentiate photopic and scotopic vision. |
| 6. | Give an equation for Euclidean distance measure. |
| 7. | Brief about Gray level? |

8. State sampling and quantization
9. What do you meant by Zooming of digital images?
10. Define Aliasing.

S.No **PART-B**

1. Analyse how image is formed in the eye with a neat structure of human eye
2. With a neat block diagram, detail about the fundamental steps in image processing.
3. Explain the components of image processing system with block diagram.
4. Construct suitable examples for the various relationships between pixels and describe them in detail.
5. Discuss in detail about image sensing and acquisition.
6. Illustrate how the image is digitized by sampling and quantization process
7. Explain image zooming and shrinking.
8. Describe the functions of elements of digital image processing system.
9. Integrate the need for various steps in digital image processing. Explain their prominence of each block.
10. Demonstrate the image formation model.

# SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

**SCHOOL OF BIO AND CHEMICAL ENGINEERING**

**DEPARTMENT OF BIOMEDICAL ENGINEERING**

# UNIT – II – Medical Image Processing – SBMA1603

# UNIT-2

# IMAGE ENHANCEMENT

## 2.1 Enhancement by Point Processing

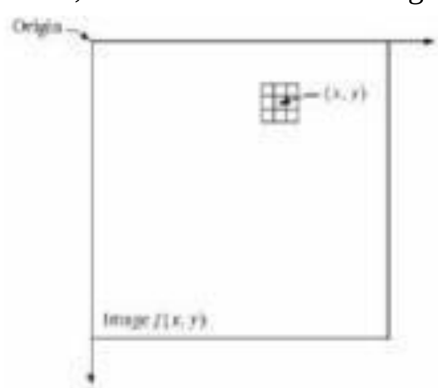
The principal objective of enhancement is to process an image so that the result is more suitable than the original image for a specific application. Image enhancement approaches fall into two board categories

- Spatial domain methods
- Frequency domain methods

The term spatial domain refers to the image plane itself and approaches in this categories are based on direct manipulation of pixel in an image. Spatial domain process are denoted by the expression

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

f(x,y)- input image T- operator on f, defined over some neighborhood of f(x,y) g(x,y)-




processed image. The neighborhood of a point (x,y) can be explain by using as square or rectangular sub image area centered at (x,y).

The center of sub image is moved from pixel to pixel starting at the top left corner. The operator T is applied to each location (x,y) to find the output g at that location . The process utilizes only the pixel in the area of the image spanned by the neighborhood.

It is the simplest form of the transformations when the neighborhood is of size IxI. In this case g depends only on the value of f at (x,y) and T becomes a gray level transformation function of the forms

$$S=T(r)$$

r- Denotes the gray level of f(x,y)

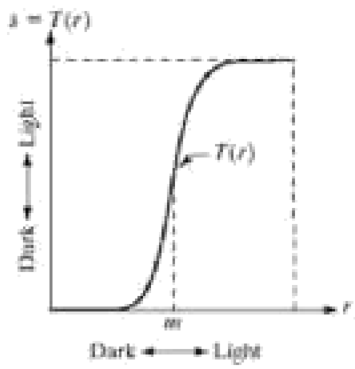
s- Denotes the gray level of g(x,y) at any point (x,y)

Because enhancement at any point in an image deepens only on the gray level at that point,technique in this category are referred to as point processing.

There are basically three kinds of functions in gray level transformation –

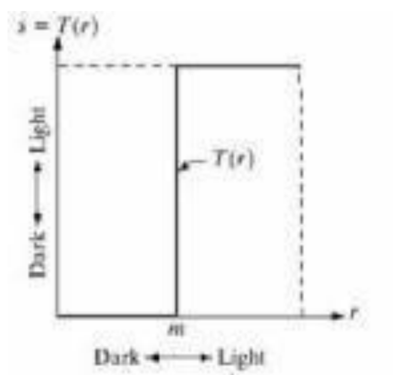
**Point Processing:**

**Contract stretching -**It produces an image of higher contrast than the original one. The operation is performed by darkening the levels below m and brightening the levels above m in the original image.




In this technique the value of r below m are compressed by the transformation function into a narrow range of s towards black .The opposite effect takes place for the values of r above m.

**Thresholding function:** It is a limiting case where T(r) produces a two levels binaryimage. The values below m are transformed as black and above m are transformed as white.




**2.2 Basic Gray Level Transformation:**

These are the simplest image enhancement techniques.

**2.2.1 Image Negative:** The negative of in image with gray level in the range [0, l-1] is obtained byusing the negative transformation.

The expression of the transformation is

$$s = L - 1 - r$$

Reverting the intensity levels of an image in this manner produces the equivalent of a photographic negative. This type of processing is practically suited for enhancing white or gray details embedded in dark regions of an image especially when the black areas are dominant in size.

**2.2.2 Log transformations:**

The general form of the log transformation is

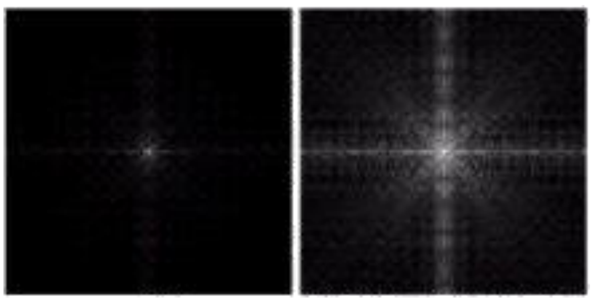
$$s = c \log(1 + r)$$

Where c- constant

$R \geq o$

This transformation maps a narrow range of gray level values in the input image into a wider range of output gray levels. The opposite is true for higher values of input levels. We would use this transformations to expand the values of dark pixels in an image while compressing the higherlevel values. The opposite is true for inverse log transformation. The log transformation function has an important characteristic that it compresses the dynamic range of images with largevariations in pixel values.

Eg- Fourier spectrum



**2.2.3 Power Law Transformation:**

Power law transformations has the basic form

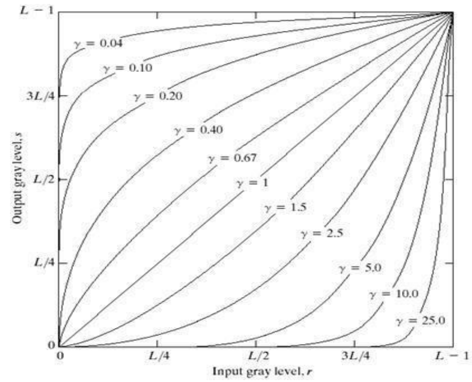
$$s = cr^{\gamma}$$

Where $c$ and $\gamma$ are positive constants.

It also sometime written as:

$$s = c(r + \varepsilon)^{\gamma}$$

Power law curves with fractional values of y map a narrow range of dark input values into a wider range of output values, with the opposite being true for higher values of input gray levels.We may get various curves by varying values of $\gamma$.




A variety of devices used for image capture, printing and display respond according to a powerlaw. The process used to correct this power law response phenomenon is called gamma correction.

For eg-CRT devices have intensity to voltage response that is a power function.

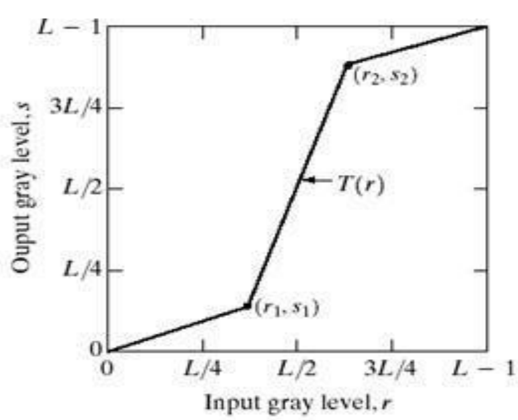
Gamma correction is important if displaying an image accurately on a computer screen is of concern. Images that are not corrected properly can look either bleached out or too dark. Colorphenomenon also uses this concept of gamma correction. It is becoming more popular due to useof images over the internet. It is important in general purpose contract manipulation. To make an image black we use $\gamma > 1$ and $\gamma < 1$ for white image.

### 2.2.4 Piece wise linear transformation functions

The principal advantage of piecewise linear functions is that these functions can bearbitrarily complex. But their specification requires considerably more user input.

- **Contrast Stretching**

It is the simplest piecewise linear transformation function. We may have various low contrast images and that might result due to various reasons such as lack of illumination, problemin imaging sensor or wrong setting of lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of gray levels in the image being processed.



The location of points (r1, s1) and (r2, s2) control the shape of the curve.

a) If r1=r2 and s1=s2, the transformation is a linear function that deduces no change in gray levels.

b) If r1=s1, s1=0 , and s2=L-1, then the transformation become a thresholding function that creates a binary image

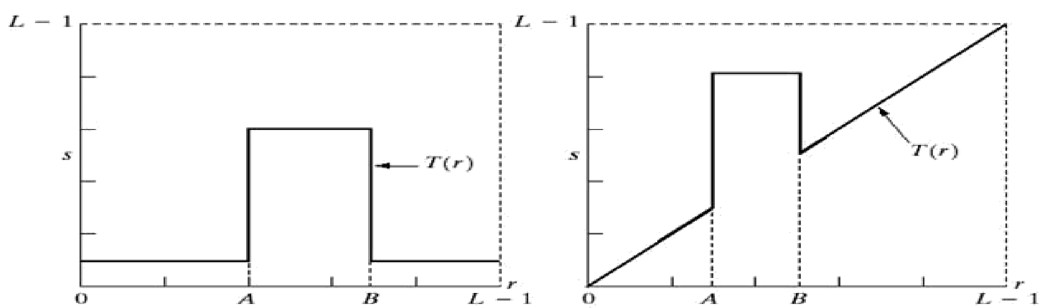
c) Intermediate values of (r1, s1) and (r2, s2) produce various degrees of spread in the grayvalue of the output image thus effecting its contract. Generally r1≤ r2 and s1 ≤ s2 so that the function is single valued and monotonically increasing.

- **Gray Level Slicing**

Highlighting a specific range of gray levels in an image is often desirable
For example when enhancing features such as masses of water in satellite image and enhancingflaws in x- ray images.
There are two ways of doing this-

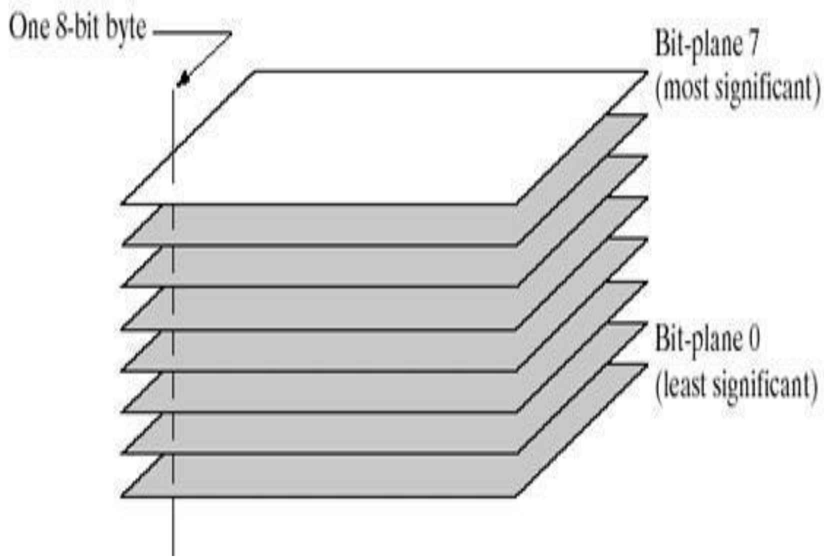
(1) One method is to display a high value for all gray level in the range. Of interest and a lowvalue for all other gray level.

(2) Second method is to brighten the desired ranges of gray levels but preserve the backgroundand gray level tonalities in the image.




- **Bit Plane Slicing**

Sometimes it is important to highlight the contribution made to the total image appearanceby specific bits. Suppose that each pixel is represented by 8 bits. Imagine that an image is composed of eight 1-bit planes ranging from bit plane 0 for the least significant bit to bit plane 7 for the most significant bit. In terms of 8-bit bytes, plane 0 contains all the lowest order bits in theimage and plane 7 contains all the high order bits. High order bits contain the majority of visually significant data and contribute to more subtle details in the image. Separating a digital image into its bits planes is useful for analysing the relative importance played by each bit of the image. It helps in determining the adequacy of the number of bits used to quantize each pixel. It is also useful for imagecompression.




**2.3 Histogram Processing:**

The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function of theForm H(rk)=nk

where rk is the kth gray level and nk is the number of pixels in the image having the level rk..

A normalized histogram is given by the equation

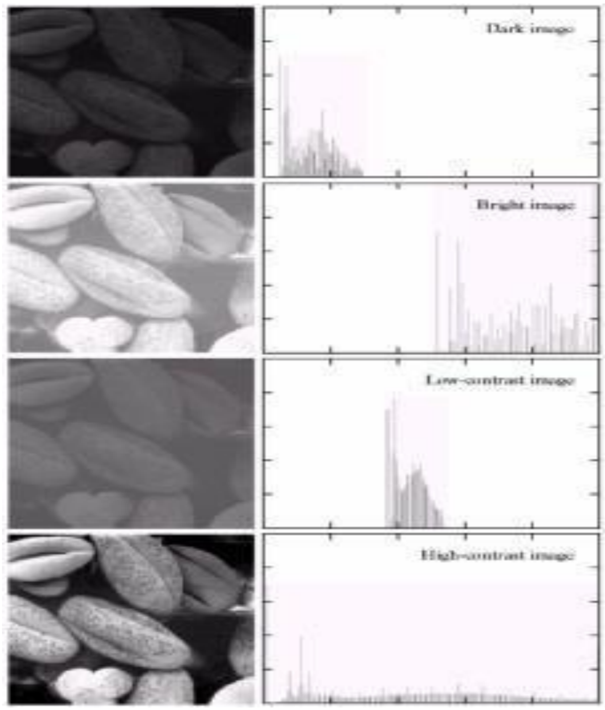
p (rk) = nk/n for k=0,1,2,…..,L-1

P(rk) gives the estimate of the probability of occurrence of gray level rk. The sum of all components of a normalized histogram is equal to 1.The histogram plots are simple plots of

H (rk) = nk versus rk.

In the dark image the components of the histogram are concentrated on the low (dark) side of thegray scale. In case of bright image the histogram components are biased towards the high side of the gray scale. The histogram of a low contrast image will be narrow and will be centred towards the middle of the gray scale.

The components of the histogram in the high contrast image cover a broad range of the grayscale. The net effect of this will be an image that shows a great deal of gray levels details and has high dynamic range.




## 2.4 Histogram Equalization

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

Let there be a continuous function with r being gray levels of the image to be enhanced. The range of r is [0, 1] with r=0 repressing black and r=1 representing white.
The transformation function is of the form

$$S=T(r) \text{ where } 0<r<1$$

It produces a level s for every pixel value r in the original image.The transformation function is assumed to fulfil two conditions

- T(r) is single valued and monotonically increasing in the interval $0<T(r)<1$ for $0<r$, 1
- The transformation function should be single valued so that the inverse transformations should exist. Monotonically increasing condition preserves the increasing order from black to white in theoutput image.

- The second conditions guarantee that the output gray levels will be in the same range as the input levels. The gray levels of the image may be viewed as random variables in the interval [0.1] .The most fundamental descriptor of a random variable is its probability density function (PDF) Pr(r) and Ps(s) denote the probability density functions of random variables r and s respectively. Basic results from an elementary probability theory states that if Pr(r) and Tr are known and T-1(s) satisfies conditions (a),

Let $p_r(r)$ and $p_s(s)$ be a probability density functions. If we assume $p_r(r)$ and T(r) are know, the

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right|$$

Image processing interests on the following formulation, where the right side is the cumulative distribution function

$$s = T(r) = (L-1)\int_0^r p_r(w)dw$$

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = (L-1)\frac{d}{dr}\left[\int_0^r p_r(w)dw\right] = (L-1)p_r(r)$$

Use the previous formulation yields, a uniform probability density function

$$p_s(s) = p_r(r)\frac{dr}{ds} = p_r(r)\left[\frac{1}{(L-1)p_r(r)}\right] = \frac{1}{L-1}$$

Histogram equalization determine the transformation that seek to produce an output image that has a uniform histogram.

For discrete values we deal with probability and summations instead of probability densityfunctions and integrals.

The transformation function is

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j; k = 0,1,2,..,L-1$$

Function that seeks to produce an output image that has a uniform histogram. It is a good approach when automatic enhancement is needed. Thus, an output image is obtained by mapping each pixel with level rk in the input image into a corresponding pixel with level sk. Equalization automatically determines a transformation.

## 2.5 Enhancement Using Arithmetic/Logic Operations

- These operations are performed on a pixel by basis between two or more images excluding not operation which is performed on a single image. It depends on the hardware and/or software that the actual mechanism of implementation should be sequential, parallel or simultaneous.

- Logic operations are also generally operated on a pixel by pixel basis. Only AND, OR and NOT logical operators are functionally complete. Because all other operators can be implemented by using these operators. While applying the operations on gray scale images, pixel values are processed as strings of binary numbers. The NOT logic operation performsthe same function as the negative transformation.

- Image Masking is also referred to as region of Interest (Ro1) processing. This is done to highlight a particular area and to differentiate it from the rest of the image. Out of the four arithmetic operations, subtraction and addition are the most useful for image enhancement.

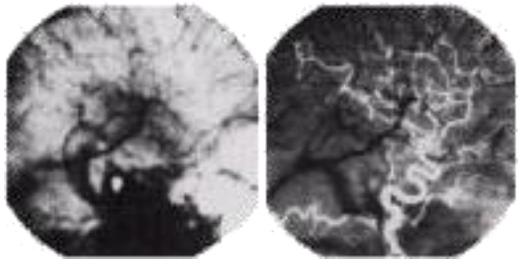
### 2.5.1 Image Subtraction

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

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

It is obtained by computing the difference between all pairs of corresponding pixels fromf and h.

- The key usefulness of subtraction is the enhancement of difference between images. This concept is used in another gray scale transformation for enhancement known as bit plane slicing. The higher order bit planes of an image carry a significant amount of visually relevant detail while the lower planes contribute to fine details.
- It we subtract the four least significant bit planes from the image. The result will be nearly identical but there will be a slight drop in the overall contrast due to less variability in the gray level values of image.
- The use of image subtraction is seen in medical imaging area named as mask mode radiography. The mask h (x,y) is an X-ray image of a region of a patient's body this image is captured by using as intensified TV camera located opposite to the x-ray machine then aconsistent medium is injected into the patient's blood storm and then a series of image aretaken of the region same as h(x,y).The mask is then subtracted from the series of incoming image. This subtraction will give the area which will be the difference between f(x,y) andh(x,y) this difference will be given as enhanced detail in the output image.



- This produces a move shoving now the contrast medium propagates through various arteries of the area being viewed. Most of the image in use today is 8- bit image sothe values of the image lie in the range 0 to 255.The value in the difference image can lie from -255 to 255. For these reasons we have to do some sort of scaling to display the results
- There are two methods to scale an image

  (i) Add 255 to every pixel and then divide at by 2.This gives the surety that pixel values will be in the range 0 to 255 but it is not guaranteed whether it will cover the entire8 – bit range or not. It is a simple method and fast to implement but will not utilize the entire gray scale range to display the results.

  (ii) Another approach is

  (a) Obtain the value of minimum difference

  (b) Add the negative of minimum value to the pixels in the differenceimage (this will give a modified image whose minimum value will be 0)

  (c) Perform scaling on the difference image by multiplying each pixel bythe quantity 255/max.
- This approach is complicated and difficult to implement. Image subtraction is used insegmentation application also.

### 2.5.2 Image Averaging

Consider a noisy image g(x,y) formed by the addition of noise η(x,y) to the original image f(x,y)

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

where $f(x, y)$: an original image

$\eta(x, y)$ : the addition of noise

One simple way to reduce this granular noise is to take several identical pictures and average them,thus smoothing out the randomness.

Assuming that at every point of coordinate (x,y) the noise is uncorrelated and has zero average value. The objective of image averaging is to reduce the noise content by adding a set of noise images, {gi(x,y)}.If in image formed by image averaging K different noisy images

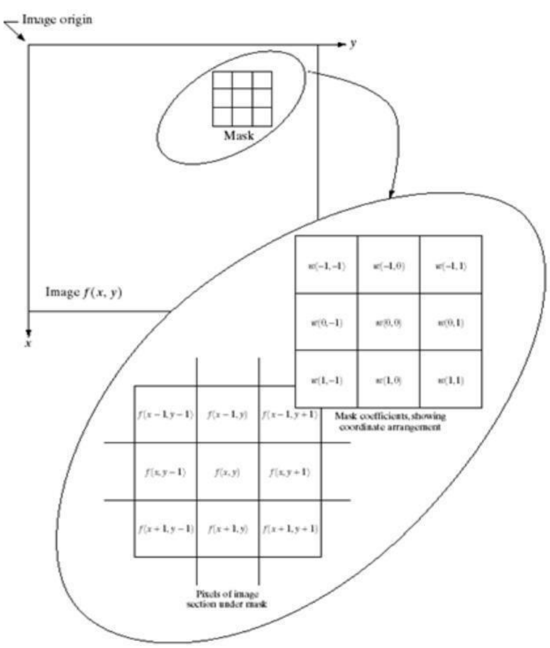
$$\overline{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

$$E\{\overline{g}(x,y)\} = f(x,y)$$

E{g(x,y)} = f(x,y) means that g(x,y) approaches f(x,y) as the number of noisy image used in the averaging processes increases. Image averaging is important in various applications such as in thefield of astronomy where the images are low light levels.

### 2.6 Basic of Spatial Filtering

Spatial filtering is an example of neighborhood operations; in this the operations are doneon the values of the image pixels in the neighborhood and the corresponding value of a sub imagethat has the same dimensions as of the neighborhood. This sub image is called a filter, mask, kernel, template or window; the values in the filter sub image are referred to as coefficients rather than pixel. Spatial filtering operations are performed directly on the pixel values (amplitude/gray scale)of the image. The process consists of moving the filter mask from point to point in the image. At each point (x,y) the response is calculated using a predefined relationship.




| | | |
|---|---|---|
| $w_1$ | $w_2$ | $w_3$ |
| $w_4$ | $w_5$ | $w_6$ |
| $w_7$ | $w_8$ | $w_9$ |

For linear spatial filtering the response is given by a sum of products of the filter coefficient and the corresponding image pixels in the area spanned by the filter mask.

The results R of liner filtering with the filter mask at point (x,y) in the image is

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

The sum of products of the mask coefficient with the corresponding pixel directly under the mask. The coefficient w (0,0) coincides with image value f(x,y) indicating that mask it centered at (x,y).when the computation of sum of products takes place.

For a mask of size MxN we assume m=2a+1 and n=2b+1, where a and b are non negative integers. It shows that all the masks are of add size.

In the general liner filtering of an image of size f of size M*N with a filter mask of size m*m is given by the expression

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

Where $a = (m-1)/2$ and $b = (n-1)/2$

To generate a complete filtered image this equation must be applied for x=0, 1, 2, M-1 and y=0, 1, 2 , N-1. Thus the mask processes all the pixels in the image. The process of linear filtering is similar to frequency domain concept called convolution. For this reason, linear spatial filtering often is referred to as convolving a mask with an image. Filter mask are sometimes called convolution mask.

**R= W1, Z1, +W2, Z2 +…. + Wmn Zmn**

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

An important point in implementing neighborhood operations for spatial filtering is the issue of what happens when the center of the filter approaches the border of the image.

i) To limit the excursion of the center of the mask to be at distance of less than (n-1) /2 pixels form the border. The resulting filtered image will be smaller than the original but all the pixels will be processed with the full mask.

ii) Filter all pixels only with the section of the mask that is fully contained in the image. It will create bands of pixels near the border that will be processed with a partial mask.

iii) Padding the image by adding rows and columns of o's & or padding by replicating rows and columns. The padding is removed at the end of the process.

### 2.6.1 Smoothing Spatial Filters

These filters are used for blurring and noise reduction blurring is used in pre-processing steps such as removal of small details from an image prior to object extraction and bridging of small gaps in lines or curves.

**Smoothing Linear Filters**

The output of a smoothing liner spatial filter is simply the average of the pixel contained in the neighborhood of the filter mask. These filters are also called **averaging filters** or low pass filters. The operation is performed by replacing the value of every pixel in the image by the average of the gray levels in the neighborhood defined by the filter mask. This process reduces sharp transitions in gray levels in the image.

A major application of smoothing is noise reduction but because edge are also provided using sharp transitions so smoothing filters have the undesirable side effect that they blur edges . It also removes an effect named as false contouring which is caused by using insufficient number of gray levels in the image. Irrelevant details can also be removed by these kinds of filters, irrelevant means which are not of our interest. A spatial averaging filter in which all coefficients are equal is sometimes referred to as a **"box filter".**

**A weighted average filter** is the one in which pixel are multiplied by different coefficients.

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

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

Box filter

$\frac{1}{16} \times$

| | | |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 4 | 2 |
| 1 | 2 | 1 |

Weighted average

The general implementation for filtering an *M*X*N* image with a weighted averaging filter of size *m*x*n* is given by

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

where $m = 2a+1, \quad n = 2b+1.$

**2.6.2 Order Statistics Filter**

- These are nonlinear spatial filter whose response is based on ordering of the pixels contained in the image area compressed by the filter and the replacing the value of the center pixel with value determined by the ranking result.
- The best example of this category is median filter. In this filter the values of the center pixel is replaced by median of gray levels in the neighborhood of that pixel. Median filters are quite popular because, for certain types of random noise, they provide excellent noise- reduction capabilities, with considerably less blurring than linear

smoothing filters.

- These filters are particularly effective in the case of impulse or salt and pepper noise. It iscalled so because of its appearance as white and black dots superimposed on an image.
- The median £ of a set of values is such that half the values in the set less than or equal to £and half are greater than or equal to this. In order to perform median filtering at a point in an image, we first sort the values of the pixel in the question and its neighbors, determine theirmedian and assign this value to that pixel.
- We introduce some additional order-statistics filters. Order-statistics filters are spatial filters whose response is based on ordering (ranking) the pixels contained in the image area encompassed by the filter. The response of the filter at any point is determined by the rankingresult

**Median filter**

The best-known order-statistics filter is the median filter, which, as its name implies, replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel:

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

The original value of the pixel is included in the computation of the median. Median filters are quite popular because, for certain types of random noise, they provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Median filters are particularly effective in the presence of both bipolar and unipolar impulse noise. In fact,the median filter yields excellent results for images corrupted by this type of noise.

**Max and min filters**

Although the median filter is by far the order-statistics filter most used in image processing. It is by no means the only one. The median represents the 50th percentile of a ranked set of numbers, but the reader will recall from basic statistics that ranking lends itself to many other possibilities. For example, 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. Also, because pepper noise has very low values, it is reduced by this filter as a result of the max selection process in the sub image area S. The 0th percentile filter is the Min filter.

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

## 2.7 Sharpening Spatial Filters

- The principal objective of sharpening is to highlight fine details in an image or to enhance details that have been blurred either in error or as a natural effect of particular method for image acquisition.
- The applications of image sharpening range from electronic printing and medical imagingto industrial inspection and autonomous guidance in military systems.
- As smoothing can be achieved by integration, sharpening can be achieved by spatial differentiation. The strength of response of derivative operator is proportional to the degree of discontinuity of the image at that point at which the operator is applied. Thus image differentiation enhances edges and other discontinuities and deemphasizes the areas with slow varying grey levels.
- It is a common practice to approximate the magnitude of the gradient by using absolute values instead of square and square roots.

A basic definition of a first order derivative of a one dimensional function *f(x)* is the difference.

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

The second-order derivative of a one-dimensional function $f(x)$ is

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$$

Development of the Laplacian method

The two dimensional Laplacian operator for continuous functions:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

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

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

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

Laplacian highlights gray-level discontinuities in an image and deemphasize the regions of slow varying gray levels. This makes the background a black image. The background texture can be recovered by adding the original and Laplacian images.

- To sharpen an image, the Laplacian of the image is subtracted from the original image.

$$g(x,y) = \begin{cases} f(x,y) - \nabla_2 f \\ f(x,y) + \nabla_2 f \end{cases}$$

If the center coefficient of the Laplacian mask is negative.
If the center coefficient of the Laplacian mask is positive.

| | | |
|---|---|---|
| 0 | 1 | 0 |
| 1 | −4 | 1 |
| 0 | 1 | 0 |

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | −8 | 1 |
| 1 | 1 | 1 |

| | | |
|---|---|---|
| 0 | −1 | 0 |
| −1 | 4 | −1 |
| 0 | −1 | 0 |

| | | |
|---|---|---|
| −1 | −1 | −1 |
| −1 | 8 | −1 |
| −1 | −1 | −1 |

### 2.7.1 Unsharp Masking and High Boost Filtering

Unsharp masking means subtracting a blurred version of an image form the image itself. Where f(x,y) denotes the sharpened image obtained by unsharp masking and f(x,y) is ablurred version of (x,y)

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

A slight further generalization of unsharp masking is called high boost filtering. A highboost filtered image is defined at any point (x,y) as

$$f_{hb}(x,y) = Af(x,y) - \overline{f}(x,y)$$

## 2.8 Image Enhancement by Frequency Domain

### Basis of Filtering in Frequency Domain

Basic steps of filtering in frequency Domain

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

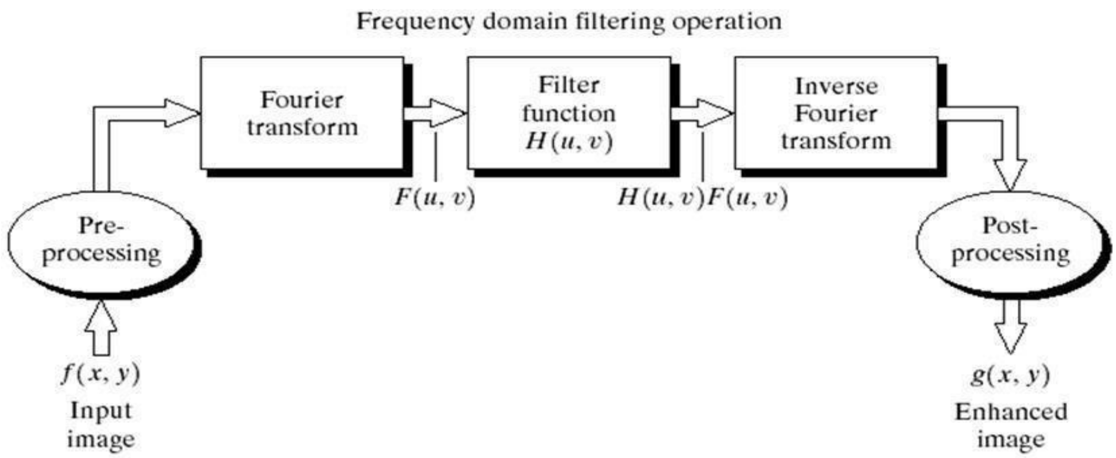
ii) Compute F(u,v), Fourier Transform of the image

iii) Multiply f(u,v) by a filter function H(u,v)

iv) Compute the inverse DFT of Result of (iii)

v) Obtain the real part of result of (iv)

vi) Multiply the result in (v) by $(-1)^{x+y}$




H (u,v) called a filter because it suppresses certain frequencies from the image while leaving others unchanged.

### 2.8.1 Smoothing Frequency Domain Filters

**Low pass filtering:**

Edges and other sharp transition of the gray levels of an image contribute significantly to the high frequency contents of its Fourier transformation. Hence smoothing is achieved in the frequency domain by attenuation a specified range of high frequency components in the transformof a given image.

Basic model of filtering in the frequency domain is

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

where *F*(*u*,*v*): the Fourier transform of the image to be smoothed

*H*(*u*,*v*): a filter transfer function

Objective is to find out a filter function H (u,v) that yields G (u,v) by attenuating the highfrequency component of F (u,v)

There are three types of low pass filters

1. Ideal
2. Butterworth
3. Gaussian

**Ideal Low Pass Filter**

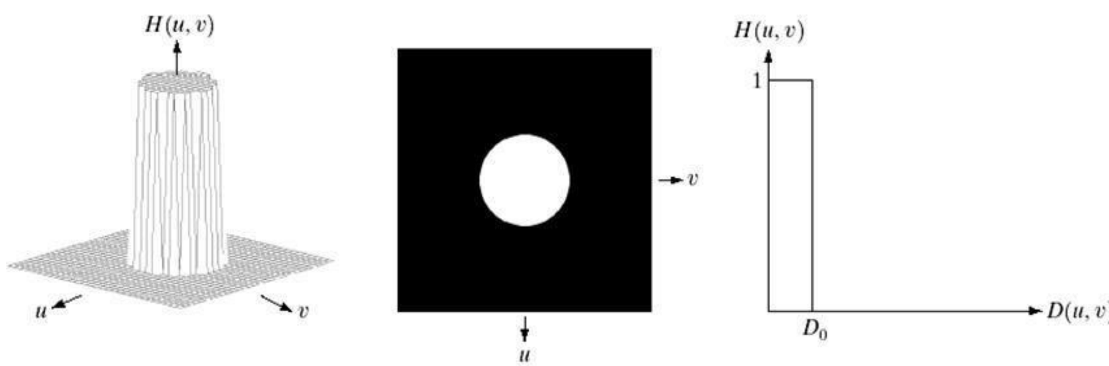
It is the simplest of all the three filters. It cuts of all high frequency component of the Fourier transform that are at a distance greater that a specified distance D0 form the origin of the transform. It is called a two – dimensional ideal low pass filter (ILPF) and has the transfer

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

function

Where *D* (*u*,*v*) : the distance from point (*u*,*v*) to the center of the frequency rectangle

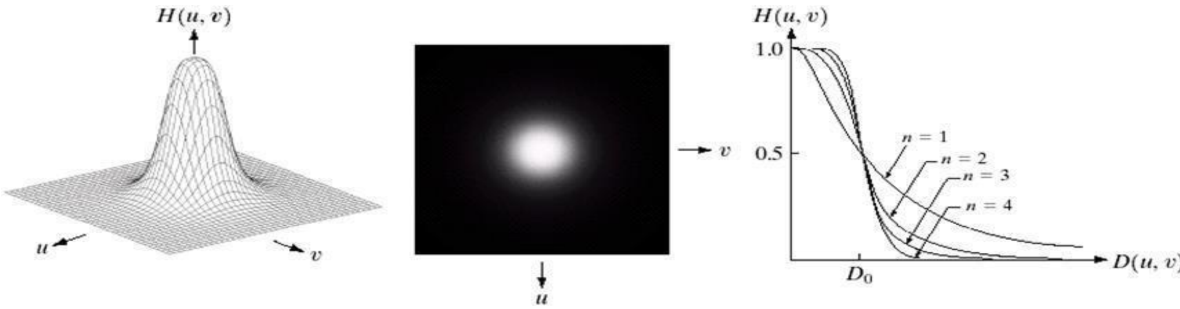
$$D(u,v) = \left[(u - M/2)^2 + (v - N/2)^2\right]^{1/2}$$




**Butterworth Low Pass Filter**

It has a parameter called the filter order. For high values of filter order it approaches the form of the ideal filter whereas for low filter order values it reach Gaussian filter. It may be viewed as a transition between two extremes. The transfer function of a Butterworth low pass filter (BLPF)of order n with cut off frequency at distance Do from the origin is defined as

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




(a) Perspective plot of a Butterworth lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross sections of orders 1 through 4.

Most appropriate value of n is 2.It does not have sharp discontinuity unlike ILPF that establishes a clear cut-off between passed and filtered frequencies. Defining a cut-off frequency is a main concern in these filters. This filter gives a smooth transition in blurring as a function of increasing cut-off frequency. A Butterworth filter of order 1 has no ringing. Ringing increases as a function of filter order. (Higher order leads to negative values)

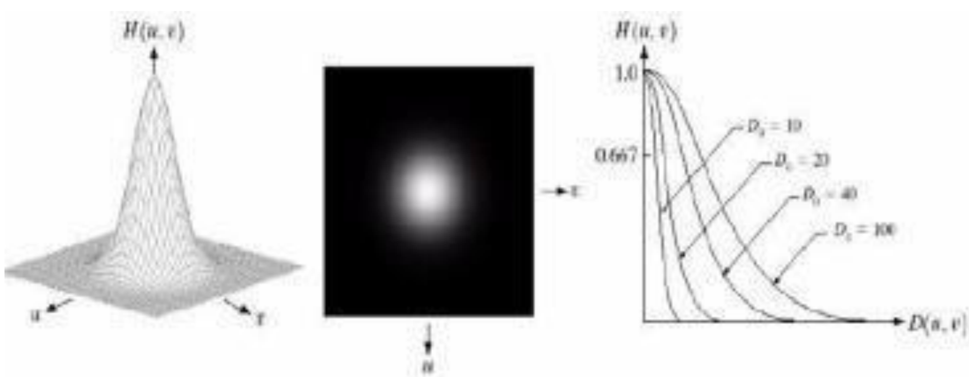
**Gaussian Low Pass Filter**

The transfer function of a Gaussian low pass filter is

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

D(u,v)- the distance of point (u,v) from the center of the transform $\sigma$ = D0- specified cut off frequency

The filter has an important characteristic that the inverse of it is also Gaussain.



### 2.8.2 Sharpening Frequency Domain

**High pass filtering:**

Image sharpening can be achieved by a high pass filtering process, which attenuatesthe low frequency components without disturbing high-frequency information. These are radially symmetric and completely specified by a cross section.

If we have the transfer function of a low pass filter the corresponding high pass filtercan be obtained using the equation

$$Hhp\ (u,v)=1-Hlp\ (u,v)$$

**Ideal High Pass Filter**

This filter is opposite of the Ideal Low Pass filter and has the transfer function of the form

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

**Butterworth High Pass Filter**

The transfer function of Butterworth High Pass filter of order n is given by the equation

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

**Gaussian High Pass Filter**

The transfer function of a Gaussain High Pass Filter is given by the equation

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

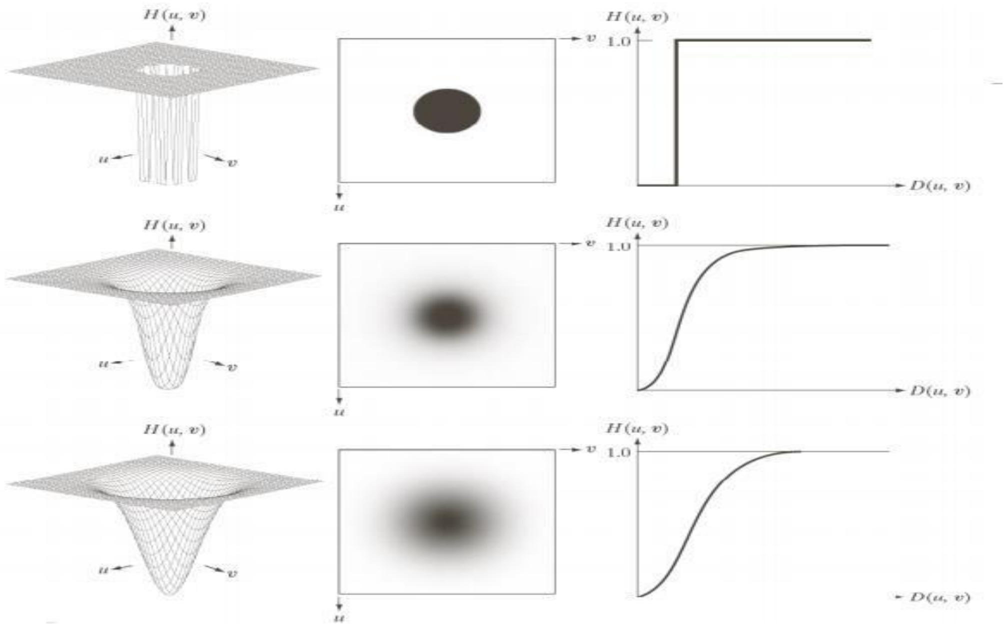



**FIGURE** Top row: Perspective plot, image representation, and cross section of a typical ideal highpass filter. Middle and bottom rows: The same sequence for typical Butterworth and Gaussian highpass filters.

**TABLE**
Highpass filters. $D_0$ is the cutoff frequency and $n$ is the order of the Butterworth filter.

| Ideal | Butterworth | Gaussian |
|---|---|---|
| $H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \le D_0 \\ 0 & \text{if } D(u,v) > D_0 \end{cases}$ | $H(u,v) = \dfrac{1}{1 + [D_0/D(u,v)]^{2n}}$ | $H(u,v) = 1 - e^{-D^2(u,v)/2D_0^2}$ |

## 2.9 Discrete Fourier Transform and the Frequency Domain

Any function that periodically reports itself can be expressed as a sum of sines and cosines of different frequencies each multiplied by a different coefficient, this sum is called Fourier series. Even the functions which are non-periodic but whose area under the curve if finite can also be represented in such form; this is now called Fourier transform.

A function represented in either of these forms and can be completely reconstructed via an inverse process with no loss of information.

### 1-D Fourier Transformation and its Inverse

If there is a single variable, continuous function f(x), then Fourier transformation F (u) may be given as

$$\mathcal{F}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\exp(-j2\pi ux)dx \qquad j = \sqrt{-1}$$

And the reverse process to recover f(x) from F(u) is

$$\mathcal{F}^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u)\exp[j2\pi ux]du$$

Equation (a) and (b) comprise of Fourier transformation pair.

Fourier transformation of a discrete function of one variable f(x), x=0, 1, 2, m-1 is given by

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

to obtain f(x) from F(u)

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

Now each of the m terms of F (u) is called a frequency component of transformation

The above two equation (e) and (f) comprise of a discrete Fourier transformation pair. According to Euler's formula

$e^{jx} = \cos x + j \sin x$Substituting these value

to equation (e)

$$F(u) = \Sigma f(x) [\cos 2\pi ux/N + j\sin 2\pi ux/N] \text{ for } u=0, 1, 2, \ldots\ldots, N-1$$

The Fourier transformation separates a function into various components, based on frequency components. These components are complex quantities.

$$F(u) = R(u) + jI(u) \qquad F(u) = |F(U)| e^{j\phi(u)}$$

or

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2} \qquad \phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right]$$

F(u) in polar coordinates

**Fourier Transformation and its Inverse**

The Fourier Transform of a two dimensional continuous function f(x,y) (an image)of size M * N is given by

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

**F {f(x, y)}**

Inverse Fourier transformation is given by equation

$$\mathcal{F}^{-1}\{F(u,v)\} = f(x,y) = \int_{-\infty}^{\infty} \int F(u,v) \exp[j2\pi(ux+vy)] dudv$$

Where (u,v) are frequency variables.

Preprocessing is done to shift the origin of F (u,v) to frequency coordinate (m/2,n/2) which is thecenter of the M*N area occupied by the 2D-FT. It is known as frequency rectangle.

It extends form u =0 to M-1 and v=0 to N-1. For this, we multiply the input image by $(-1)^{x+y}$ prior to compute the transformation

$$\mathcal{F}\{f(x,y)(-1)^{x+y}\} = F(u-M/2, v-N/2)$$

Ƒ (.) denotes the Fourier transformation of the argumentValue of transformation at (u,v)=(0,0) is

$$F(0,0) = 1/MN \, \Sigma\Sigma f(x,y)$$

## 2.10 Discrete Fourier Transform and its Properties

In the two-variable case the discrete Fourier transform pair is

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

for $u = 0, 1, 2, \ldots, M-1$, $v = 0, 1, 2, \ldots, N-1$, and

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

for $x = 0, 1, 2, \ldots, M-1$ and $y = 0, 1, 2, \ldots, N-1$.

When images are sampled in a squared array, i.e. M=N, we can write

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

for $u, v = 0, 1, 2, \ldots, N-1$, and

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

2 Dimencional DFT

$$F(u,v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)\, e^{-j2\pi\left[\frac{um}{M} + \frac{vn}{N}\right]}$$

$$f(m,n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F[u,v]\, e^{j2\pi\left[\frac{um}{M} + \frac{vn}{N}\right]}$$

where u, v are frequency variables [Transform].
x, y are spatial variable [Image].

* Fourier spectrum $= |F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)}$
* Phase angle $\phi[u,v] = \tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right]$
* Power spectrum
  $P(u,v) = [F(u,v)]^2 = R^2(u,v) + I^2(u,v)$.

## Properties of 2D- DFT.

1) Periodicity

To prove : $DFT\left[f(m+N, n+N)\right] = F(u,v)$

Solution :-

$$= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m+N, n+N)\, e^{-j2\pi\left[\frac{u}{m}(m+N) + \frac{v}{n}(m+N)\right]}.$$

Put $m+N = m'$, $n+N = n'$

$$= \frac{1}{MN} \sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} f(m', n')\, e^{-j2\pi\left[\frac{um'}{m} + \frac{vn'}{n}\right]} = F[u,v]$$

$$DFT\left[f(m+N, n+N)\right] = F(u,v) .$$

Hence proved.

2) Time Shift.

To prove:-

$$DFT\{f(m-m_0, n-n_0)\} = F[u,v]$$

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m-m_0, n-n_0)\, e^{-j2\pi\left[\frac{m-m_0}{M}u + \frac{n-n_0}{N}v\right]}$$

Put $m-m_0 = m'$, $n-n_0 = n'$

$$\frac{1}{MN}\sum_{m'=0}^{M-1}\sum_{n'=0}^{N-1} f(m', n')\, e^{-j2\pi\left[\frac{um'}{M} + \frac{vn'}{N}\right]}$$

$$= f(u,v)$$

$$DFT\{f(m-m_0, n-n_0\} = F[u,v]$$

Hence proved.

3) Rotational property.

To Prove,

$$f(r,\theta+\theta_0) \Longleftrightarrow F[\omega, \phi+\theta_0]$$

Solution;

Consider Polar Coordinates

$x = r\cos\theta$; $y = r\sin\theta$, $u = \omega\cos\phi$, $v = \omega\sin\phi$.

then $f(x,y) = f(r,\theta)$

$F[u,v] = F[\omega,\phi]$

If $f(x,y)$ and $F[u,v]$ rotates by the same angle.

$$F[r,\theta+\theta_0] \Longleftrightarrow F[\omega,\phi+\theta_0]$$

Hence proved.

4)| Scaling Property

To Prove

$$DFT\left[f(am, bn)\right] = F\left[u/a, v/b\right]$$

Solution:-

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\, e^{-j2\pi\left[\frac{um}{M}+\frac{vn}{N}\right]}$$

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(am,bn)\, e^{-j2\pi\left[\frac{um}{M}+\frac{vn}{N}\right]}.$$

$$am = m' \quad bn = n'$$

$$m = m'/a \quad n = n'/a.$$

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m'n')\, e^{-j2\pi\left[\frac{um'}{aM}+\frac{vn'}{bN}\right]}$$

$$F\left[u/a, v/b\right]$$

$$DFT\left[f(am, bn)\right] = F\left[u/a, v/b\right]$$

Hence Proved.

5) Frequency shift.

To prove.

$$F(k-k_0, l-l_0) = f(m,n).$$

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} F(u,v)\, e^{-j2\pi\left[\frac{um}{M}+\frac{vn}{N}\right]}.$$

↳ Inverse DFT.

$$= \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} F(k-k_0, l-l_0)\, e^{j2\pi\left[\frac{(k-k_0)m}{M}+\frac{(l-l_0)n}{N}\right]}.$$


**References**

1. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing Using Matlab", 3rd Edition Tata McGraw Hill Pvt. Ltd., 2011.
2. Anil Jain K. "Fundamentals of Digital Image Processing", PHI Learning Pvt. Ltd., 2011.

**Question Bank**

| S.No | PART-A |
|---|---|
| 1. | Give the Properties of two-dimensional DFT |
| 2. | Specify the objective of image enhancement technique. |
| 3. | What is contrast stretching? |
| 4. | Estimate the link between spatial and frequency domain filtering. |
| 5. | Define image subtraction. |
| 6. | Specify the need for image averaging. |
| 7. | Write the expression for negative and log transformation. |
| 8. | What is meant by bit plane slicing? |
| 9. | Recall the term histogram. |
| 10. | Write the steps involved in frequency domain filtering. |
| 11. | Give the mask used for high boost filtering. |
| | **PART-B** |
| 12. | Illustrate the 2D Fourier transform and its pair. State and prove their property. |
| 13. | State and prove the properties of DFT. |
| 14. | Compose about the various grey level transformations with examples and plot the graph of the transformation functions. |
| 15. | Describe piecewise-linear transformation. |
| 16. | Estimate the constraints of histogram equalization and technique of histogram processing in detail. |
| 17. | Explain spatial filtering in image enhancement. |
| 18. | Analyse the performance of image smoothing filter with its model in the spatial domain. |
| 19. | i) Distinguish between spatial & frequency domain image enhancement<br>ii) Classify the performance of the sharpening filters |
| 20. | Illustrate image enhancement in frequency domain using low pass and high pass filtering. |
| 21. | Demonstrate the image enhancement techniques by point processing, |

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

**SCHOOL OF BIO AND CHEMICAL ENGINEERING**

**DEPARTMENT OF BIOMEDICAL ENGINEERING**

# UNIT – III – Medical Image Processing – SBMA1603

# UNIT-3

# IMAGE RESTORATION

## Introduction

- Restoration improves image in some predefined sense. It is an objective process. Restoration attempts to reconstruct an image that has been degraded by using a priori knowledge of the degradation phenomenon. These techniques are oriented toward modeling the degradationand then applying the inverse process in order to recover the original image.
- Image Restoration refers to a class of methods that aim to remove or reduce the degradations that have occurred while the digital image was being obtained.
- All natural images when displayed have gone through some sort of degradation:
  a) During display mode
  b) Acquisition mode, or
  c) Processing mode
- The degradations may be due to
  a) Sensor noise
  b) Blur due to camera misfocus
  c) Relative object-camera motion
  d) Random atmospheric turbulence
  e) Others

### 3.1 A Model of Image Restoration Process

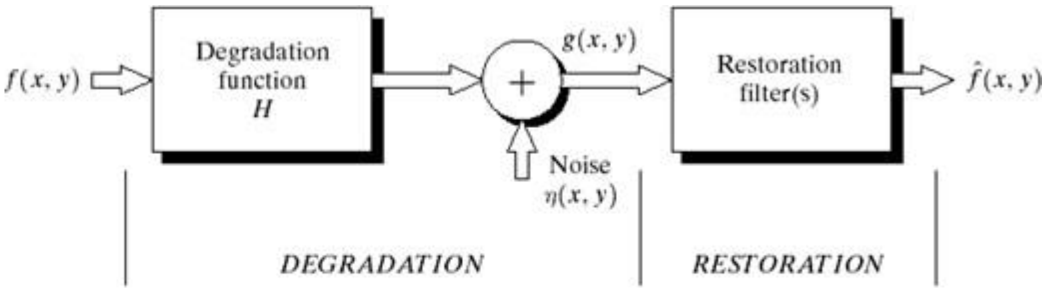
- Degradation process operates on a degradation function that operates on an input image withan additive noise term.
- Input image is represented by using the notation f(x,y), noise term can be represented as η(x,y).These two terms when combined gives the result as g(x,y).
- If we are given g(x,y), some knowledge about the degradation function H or J and some knowledge about the additive noise teem η(x,y), the objective of restoration is to obtain an estimate *f'(x,y)* of the original image. We want the estimate to be as close as possible to the original image. The more we know about h and η, the closer f(x,y) will be to f'(x,y).
- If it is a linear position invariant process, then degraded image is given in the spatial domainby

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

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

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

- The terms in the capital letters are the Fourier Transform of the corresponding terms in thespatial domain.




### 3.2 Restoration Using Spatial Filtering

When the only degradation present in an image is noise, i.e.

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

Or

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

The noise terms are unknown so subtracting them from g(x,y) or G(u,v) is not a realistic approach. In case of periodic noise it is possible to estimate N (u,v) from the spectrum G(u,v). So N (u,v) can be subtracted from G(u,v) to obtain an estimate of original image. Spatial filteringcan be done when only additive noise is present.

The following techniques can be used to reduce the noise effect:

#### 3.2.1 Mean Filter

- **Arithmetic Mean Filter**

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

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

This operation can be using a convolution mask in which all coefficients have value 1/mn. A mean filter smoothes local variations in image Noise is reduced as a result of blurring. For every pixel in the image, the pixel value is replaced by the mean value of its neighboring pixels NxM witha weight $W_k$=1/mn. This will resulted in a smoothing effect in the image.

- **Geometric mean filter**

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

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

Here, each restored pixel is given by the product of the pixel in the subimage window, raised to the power 1/mn. A geometric mean filters but it to loose image details in the process.

- **Harmonic mean filter**

  The harmonic mean filtering operation is given by the expression

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

The harmonic mean filter works well for salt noise but fails for pepper noise. It does well with Gaussian noise also.

- **Contraharmonic Mean:**

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

*Q* is the *order* of the filter and adjusting its value changes the filter's behaviour.

Positive values of *Q* eliminate pepper noise
Negative values of *Q* eliminate salt noise

### 3.2.2 Order statistics filter

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

- **Median filter**

It is the best order statistic filter; it replaces the value of a pixel by the median of gray levels in the Neighborhood of the pixel.

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

The original of the pixel is included in the computation of the median of the filter are quite possible because for certain types of random noise, the provide excellent noise reduction capabilities with considerably less blurring then smoothing filters of similar size. These are effective for bipolar and unipolor impulse noise.

- **Max and Min Filters**

  Using the l00th percentile of ranked set of numbers is called the ***max filter*** and is given bythe equation

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

It is used for finding the brightest point in an image. Pepper noise in the image has very low values;it is reduced by max filter using the max selection process in the sublimated area sky.

The 0th percentile filter is ***min filter***

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

This filter is useful for finding the darkest point in image. Also, it reduces salt noise of the min operation.

- **Midpoint Filter**

  The midpoint filter simply computes the midpoint between the maximum and minimumvalues in the area encompassed by the filter

$$\hat{f}(x,y) = \frac{1}{2}\left[\max_{(s,t)\in S_{xy}} \{g(s,t)\} + \min_{(s,t)\in S_{xy}} \{g(s,t)\}\right]$$

It comeliness the order statistics and averaging .This filter works best for randomly distributed noise like Gaussian or uniform noise.

- **Alpha-Trimmed Mean Filter:**

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

We can delete the *d/2* lowest and *d/2* highest grey levels.So *gr(s, t)* represents the remaining *mn –d* pixels.efficiently handling mixture noise, e.g., a combination of salt-and-pepper and Gaussian noise.

### 3.3. Noise reduction using Frequency domain filtering

#### 3.3.1 Band Reject Filters

It removes a band of frequencies about the origin of the Fourier transformer.

- **Ideal Band reject Filter**

  An ideal band reject filter is given by the expression

$$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \prec D_0 - \frac{W}{2} \\ 0 & \text{if } D_0 - \frac{W}{2} \le D(u,v) \le D_0 + \frac{W}{2} \\ 1 & \text{if } D(u,v) \succ D_0 + \frac{W}{2} \end{cases}$$

D (u, v) - the distance from the origin of the centered frequency rectangle.

W- Width of the band

Do- radial center

- **Butterworth Band reject Filter**

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

- **Gaussian Band reject Filter**

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

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

### 3.3.2 Band Pass Filters

The function of a band pass filter is opposite to that of a band reject filter. It allows a specific frequency band of the image to be passed and blocks the rest of frequencies. The transfer function of a band pass filter can be obtained from a corresponding band reject filter with transfer function Hbr (u, v) by using the equation-

**Hbp (u, v) =1-Hbr (u, v)**

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

### 3.3.3 Notch Filters

This type of filters rejects frequencies I predefined in neighborhood above a centre frequency These filters are symmetric about origin in the Fourier transform the transfer function of ideal notch reject filter of radius do with centre at (…) and by symmetry at (……) is

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

$$D_1(u,v)=\left[\left(u-\frac{M}{2}-u_0\right)^2+\left(v-\frac{N}{2}-v_0\right)^2\right]^{\frac{1}{2}}$$

$$D_1(u,v)=\left[\left(u-\frac{M}{2}+u_0\right)^2+\left(v-\frac{N}{2}+v_0\right)^2\right]^{2}$$

The center of the frequency rectangle has been shifted to the point (M/2, N/2)

➢ **Butterworth notch reject filter** of order n is given by

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

➢ **Gaussian notch reject filter** is given by

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

- **Notch pass filter**: These filter become high pass rather than suppress. The frequencies contained in the notch areas. These filters will perform exactly the opposite function as the notch reject filter.

The transfer function of this filter may be given as

**Hnp (u, v) =1-Hnr (u, v)**

Hnp(u,v)- transfer function of the pass filter Hnr(u,v)- transfer function of a notch reject filter.

### 3.4 Inverse Filtering

It is a process of restoring an image degraded by a degradation function H. This function can be obtained by any method.

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

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

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

It shows an interesting result that even if we know the depredation function we cannot recover the under rated image exactly because N (u,v) is not known .

If the degradation value has zero or very small values then the ratio N (u,v)/H(u,v) could easily dominate the estimate F(u,v).

## 3.5 Least Mean Square Filter

**Minimum Mean Square Error (Wiener) Filtering**

This filter incorporates both degradation function and statistical behavior of noise into the restoration process.
The main concept behind this approach is that the images and noise are considered as random variables and the objective is to find an estimate f of the uncorrupted image f such that the mean sequence error between then is minimized.

$$\hat{f}(x) = \sum_{s=-\infty}^{\infty} h_W(x-s)g(s),$$

This error measure is given by

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

Where e( ) is the expected value of the argument
Assuming that the noise and the image are uncorrelated (means zero average value) one or other has zero mean values
The minimum error function of the above expression is given in the frequency ……. is given by the expression.

$$H_W(u,v) = \frac{H^*(u,v)S_g(u,v)}{|H(u,v)|^2 S_g(u,v) + S_m(u,v)} = \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_m(u,v)/S_g(u,v)}$$

Product of a complex quantity with its conjugate is equal to the magnitude of …… complex quantity squared. This result is known as wiener Filter The filter was named so because of the name of its inventor N Wiener. The term in the bracket is known as minimum mean square error filter or least square error filter.
H*(u,v)-degradation function .
H*(u,v)-complex conjugate of H(u,v)
H(u,v) H(u,v)
Sn(u,v)=IN(u,v)I$^2$- power spectrum of the noise
Sf(u,v)=IF(u,v)$^2$- power spectrum of the underrated image
H(u,v)=Fourier transformer of the degraded function
G(u,v)=Fourier transformer of the degraded image
The restored image in the spatial domain is given by the inverse Fourier transformed of the frequency domain estimate F(u,v).
Mean square error in statistical form can be approveiment by the function

$$H_W(u,v) = \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}$$

**References**


1. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, “Digital Image Processing Using Matlab”, 3rd Edition Tata McGraw Hill Pvt. Ltd., 2011.
2. Anil Jain K. “Fundamentals of Digital Image Processing”, PHI Learning Pvt. Ltd., 2011.

**Question Bank**

| S.No | **PART-A** |
|---|---|
| 1. | What is meant by Image Restoration? |
| 2. | Differentiate enhancement from restoration. |
| 3. | Construct the image restoration model. |
| 4. | Classify the types of noise models. |
| 5. | Compare maximum and minimum filter. |
| 6. | Classify order statistic filter. |
| 7. | What is inverse filtering? |
| 8. | What is meant by least mean square filter? |
| 9. | Define Notch filter. |
| 10. | Define Averaging Filters |

| S.No | **PART-B** |
|---|---|
| 1. | Illustrate image degradation model /restoration process in detail. |
| 2. | Enumerate noise reduction in frequency domain. |
| 3. | Elaborate image restoration using spatial filtering. |
| 4. | Derive a wiener filter for image restoration and specify its advantages over inverse filter. |
| 5. | Compare and Contrast restoration using mean and median filter. |
| 6. | Explain the use of wiener filter or least mean square filter in image restoration. |
| 7. | Explain image degradation and restoration process. |
| 8. | Demonstrate inverse filtering for removal of blur caused by any motion and describe how it restore the image. |

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

**SCHOOL OF BIO AND CHEMICAL ENGINEERING**

**DEPARTMENT OF BIOMEDICAL ENGINEERING**

# UNIT – IV – Medical Image Processing – SBMA1603

# UNIT-4

# IMAGE SEGMENTATION AND COLOR IMAGE

## 4.1 Detection of Discontinuities

- **Edge and Line Detection**

  There are three kinds of discontinuities of intensity: points, lines and edges. The most commonway to look for discontinuities is to scan a small mask over the image. The mask determines which kind of discontinuity to look for.

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

| | | |
|---|---|---|
| $w_1$ | $w_2$ | $w_3$ |
| $w_4$ | $w_5$ | $w_6$ |
| $w_7$ | $w_8$ | $w_9$ |

$$|R| \geq T$$

where $T$ : a nonnegative threshold

- Only slightly more common than point detection is to find a one pixel wide line in an image.
- For digital images the only three point straight lines are only horizontal, vertical, or diagonal (+ or $-45^0$).

Horizontal

| | | |
|---|---|---|
| −1 | −1 | −1 |
| 2 | 2 | 2 |
| −1 | −1 | −1 |

+45°

| | | |
|---|---|---|
| −1 | −1 | 2 |
| −1 | 2 | −1 |
| 2 | −1 | −1 |

Vertical

| | | |
|---|---|---|
| −1 | 2 | −1 |
| −1 | 2 | −1 |
| −1 | 2 | −1 |

−45°

| | | |
|---|---|---|
| 2 | −1 | −1 |
| −1 | 2 | −1 |
| −1 | −1 | 2 |

First-order derivatives:

– The gradient of an image $f(x,y)$ at location $(x,y)$ is defined as the vector:

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

The magnitude of this vector:

$$\nabla f = \text{mag}(\nabla \mathbf{f}) = \left[G_x^2 + G_y^2\right]^{1/2}$$

The direction of this vector:

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_x}{G_y}\right)$$

| | | | | |
|---|---|---|---|---|
| −1 | 0 | | 0 | −1 |
| 0 | 1 | | 1 | 0 |

Roberts

| | | | | | | |
|---|---|---|---|---|---|---|
| −1 | −1 | −1 | | −1 | 0 | 1 |
| 0 | 0 | 0 | | −1 | 0 | 1 |
| 1 | 1 | 1 | | −1 | 0 | 1 |

Prewitt

| | | | | | | |
|---|---|---|---|---|---|---|
| −1 | −2 | −1 | | −1 | 0 | 1 |
| 0 | 0 | 0 | | −2 | 0 | 2 |
| 1 | 2 | 1 | | −1 | 0 | 1 |

Sobel

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 1 | 1 | | −1 | −1 | 0 |
| −1 | 0 | 1 | | −1 | 0 | 1 |
| −1 | −1 | 0 | | 0 | 1 | 1 |

Prewitt

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 1 | 2 | | −2 | −1 | 0 |
| −1 | 0 | 1 | | −1 | 0 | 1 |
| −2 | −1 | 0 | | 0 | 1 | 2 |

Sobel

Prewitt and Sobel masks for detecting diagonal edges.

### 4.2 Region-Based Segmentation

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

**Basic Formulation:** Let R represent the entire image region. We may view segmentation as a process that partitions R into n subregions, $R_1, R_2 \ldots, R_n$, such that

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

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

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

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

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

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

Condition (a) indicates that the segmentation must be complete; that is, every pixel must be in a region.

Condition (b) requires that points in a region must be connected in some predefined sense.

Condition (c) indicates that the regions must be disjoint.

Condition (d) deals with the properties that must be satisfied by the pixels in a segmented

region—for example P ($R_i$) = TRUE if all pixels in $R_i$, have the same gray level.
Finally, condition (c) indicates that regions $R_i$ and $R_j$ are different in the sense of predicate P.

**4.2.1 Region Growing:**

As its name implies, region growing is a procedure that group's pixels or sub regions into larger regions based on predefined criteria. The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighboring pixels that have properties similar to the seed (such as specific ranges of gray level or color). When a priori information is not available, the procedure is to compute at every pixel the same set of properties that ultimately will be used to assign pixels to regions during the growing process. If the result of these computations shows clusters of values, the pixels whose properties place them near the centroid of these clusters can be used as seeds. The selection of similarity criteria depends not only on the problem under consideration, but also on the type of image data available. For example, the analysis of land-use satellite imagery depends heavily on the use of color. This problem would be significantly more difficult, or even impossible, to handle without the inherent information available in color images. When the images are monochrome, region analysis must be carried out with a set of descriptors based on gray levels and spatial properties (such as moments or texture). Basically, growing a region should stop when no more pixels satisfy the criteria for inclusion in that region. Criteria such as gray level, texture, and color, are local in nature and do not take into account the "history" of region growth. Additional criteria that increase the power of a region growing algorithm utilize the concept of size, likeness between a candidate pixel and the pixels grown so far (such as a comparison of the gray level of a candidate and the average gray level of the grown region), and the shape of the region being grown. The use of these types of descriptors is based on the assumption that a model of expected results is at least partially available. Figure (a) shows an X-ray image of a weld (the horizontal dark region) containing several cracks and porosities (the bright, white streaks running horizontally through the middle of the image). We wish to use region growing to segment the regions of the weld failures. These segmented features could be used for inspection, for inclusion in a database of historical studies, for controlling an automated welding system, and for other numerous applications.

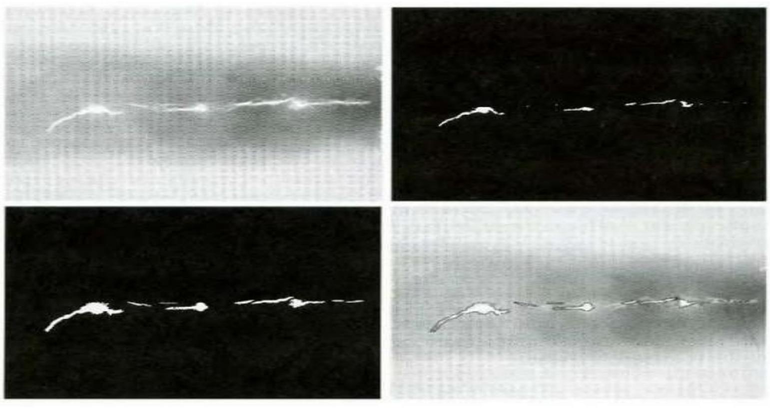


**Fig.** (a) Image showing defective welds, (b) Seed points, (c) Result of region growing, (d) Boundaries of segmented ; defective welds (in black).

**Fig.1**

The first order of business is to determine the initial seed points. In this application, it is known that pixels of defective welds tend to have the maximum allowable digital value B55 in this case). Based on this information, we selected as starting points all pixels having values of 255. The points thus extracted from the original image are shown in Fig. Note that many of the points are clustered into seed regions. The next step is to choose criteria for region growing.

In this particular example we chose two criteria for a pixel to be annexed to a region:

(1) The absolute gray-level difference between any pixel and the seed had to be less than 65. This number is based on the histogram shown in Fig.2 and represents the difference between 255 and the location ofthe first major valley to the left, which is representative of the highest gray level value in the dark weld region.

(2) To be included in one of the regions, the pixel had to be 8-connected to at least one pixel in thatregion. If a pixel was found to be connected to more than one region, the regions were merged. Figure.1 (c) shows the regions that resulted by starting with the seeds in Fig.2 (b) and utilizing the criteria defined in the previous paragraph. Superimposing the boundaries of these regions on the original image [Fig.1 (d)] reveals that the region-growing procedure did indeed segment the defective welds with an acceptable degree of accuracy. It is of interest to note that it was not necessary to specify any stopping rules in this case because the criteria for region growing were sufficient to isolate the features of interest.

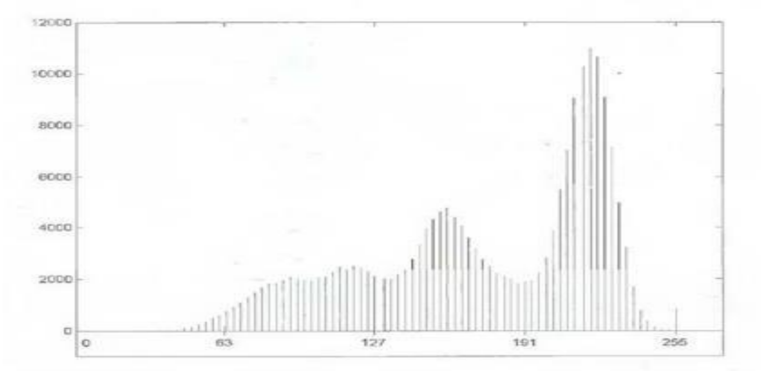



**Fig. .2 Histogram of Fig. .1 (a)**

### 4.2.2 Region Splitting and Merging

The procedure just discussed grows regions from a set of seed points. An alternative is to subdivide an image initially into a set of arbitrary, disjointed regions and then merge and/or split the regions in an attempt to satisfy the conditions. A split and merge algorithm that iteratively works toward satisfying these constraints is developed. Let R represent the entire image region and select a predicate P. One approach for segmenting R is to subdivide it successively into smaller and smaller quadrant regions so that, for any region $R_i$, P $(R_i)$ = TRUE. We start with the entire region. If P(R) = FALSE, we divide the image into quadrants. If P is FALSE for any quadrant, we subdivide that quadrant into sub quadrants, and so on. This particular splitting technique has a convenient representation in the form of a so-called quadtree (that is, a tree in which nodes have exactly four descendants), as illustrated in Fig.3. Note that the root of the tree corresponds to the entire image and that each node corresponds to a subdivision. In thiscase, only $R_4$ was subdivided further.

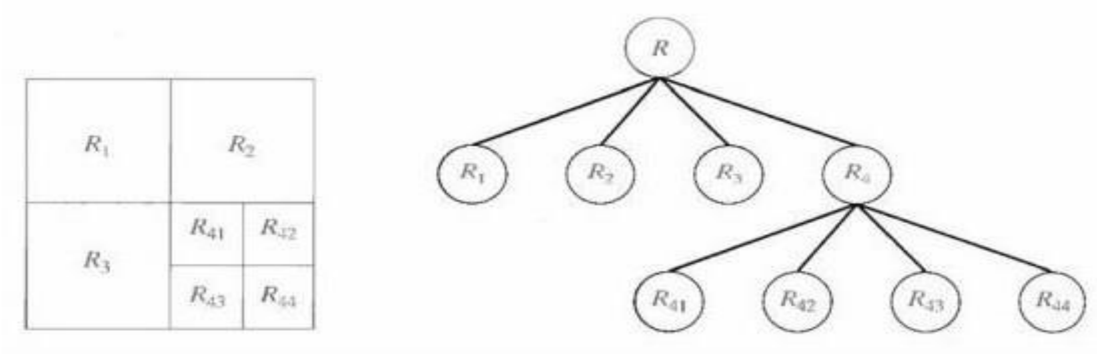



**Fig. .3 (a) Partitioned image (b) Corresponding quadtree.**

If only splitting were used, the final partition likely would contain adjacent regions with identical properties. This drawback may be remedied by allowing merging, as well as splitting. Satisfying the constraints, requires merging only adjacent regions whose combined pixels satisfy the predicate P. That is, two adjacent regions Rj and Rk are merged only if P (Rj U Rk) = TRUE. The preceding discussion may be summarized by the following procedure, in which, at any step we

1. Split into four disjoint quadrants any region Ri, for which P (Ri) = FALSE.
2. Merge any adjacent regions Rj and Rk for which P (Rj U Rk) = TRUE.
3. Stop when no further merging or splitting is possible.

Several variations of the preceding basic theme are possible. For example, one possibility is to split the image initially into a set of blocks. Further splitting is carried out as described previously, but merging is initially limited to groups of four blocks that are descendants in the quadtree representation and that satisfy the predicate P. When no further merging of this type are possible, the procedure is terminated by one final merging of regions satisfying step 2. At this point, the merged regions may be of different sizes. The principal advantage of this approach is that it uses the same quadtree for splitting and merging, until the final merging step.

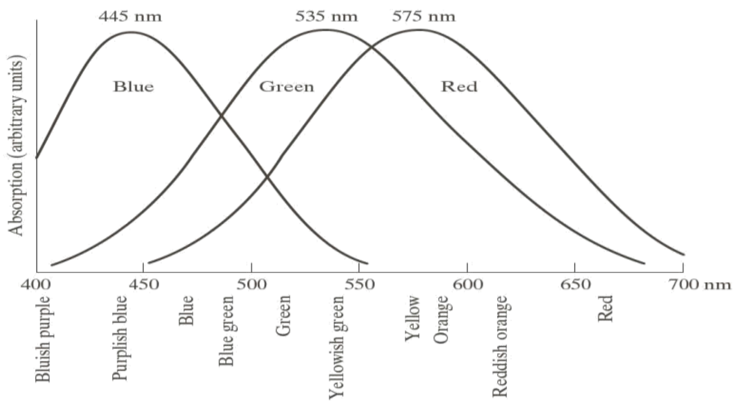
**4.3 Color image processing**

In 1666, Isaac Newton discovered that when a beam of sunlight passes through a glass prism, the emerging beam is split into a spectrum of colors. The human visual system can distinguish hundreds of thousands of different color shades and intensities, but only around 100 shades of grey. Therefore, in an image, a great deal of extra information may be contained in the color, and this extra information can then be used to simplify image analysis, e.g. object identification and extraction based on color. Three independent quantities are used to describe any particular color. The *hue* is determined by the dominant wavelength. Visible colors occur between about 400nm (violet) and 700nm (red) on the electromagneticspectrum,

**Color Fundamentals**

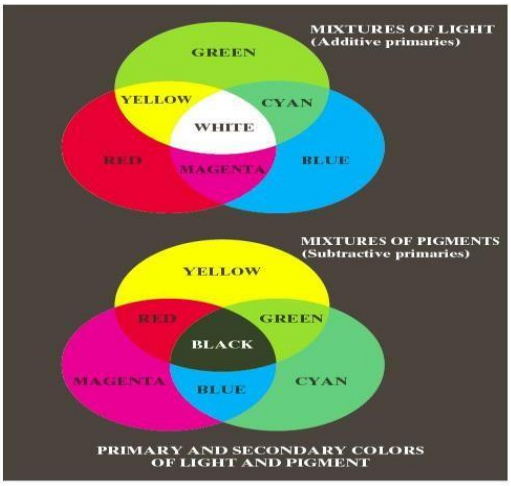
- ► 6 to 7 million cones in the human eye can be divided into three principal sensing categories, corresponding roughly to red, green, and blue.
  65%: red 33%: green 2%: blue (blue cones are the most sensitive)
- ► The characteristics generally used to distinguish one color from another are brightness, hue, and saturation.

- ► **Brightness**: the achromatic notion of intensity.
- ► **Hue**: dominant wavelength in a mixture of light waves, represents dominant color asperceived by an observers**.**
- ► **Saturation**: relative purity or the amount of white light mixed with its hue.



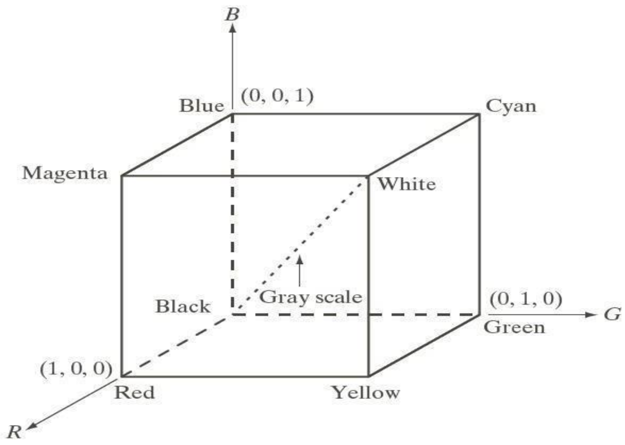

### 4.4 Color models:

The purpose of a color model is to facilitate the specification of colors in some standard. In essence, a color model is a specification of a coordinate system and a subspace within that system where each color is represented by a single point. Most color models are oriented either toward specific hardware or toward applications. Red, green, and blue, three primary colors. Cone cells in human eye areresponsible for color vision.




**The RGB Model**

In the RGB model each color appears in its primary spectral components of red, green and blue. The model is based on a Cartesian coordinate system RGB values are at 3 corners. Cyan magenta and yellow are at three other corners. Black is at the origin. White is the corner furthest from the origin. Different colors are points on or inside the cube represented by RGB vectors. Images represented in the RGB color model consist of three component images – one for each primary color. When fed into a monitor these images are combined to create a composite color image. The number of bits used to represent each pixel is referred to as the color depth. A 24-bit image is often referred to as a full-color image as it allows16, 777,216 colors .The RGB model is used for color monitors and most video cameras.




**The CMY Model:**

The CMY (cyan- magenta- yellow) model is a subtractive model appropriate to absorption of colors, for example due to pigments in paints. Whereas the RGB model asks what is added to black to get a particular color, theCMY model asks what is subtracted from white. In this case, the primaries are cyan, magenta and yellow, with red, green and blue as secondary colors. When a surface coated with cyan pigment is illuminated by white light, no red light is reflected, and similarly for magenta and green, and yellow and blue. The CMY model is used by printing devices and filters. Equal amounts of the pigment primaries, cyan, magenta, and yellow should produce black. In practice, combining these colors for printing produces a muddy- looking black.

To produce true black, the predominant color in printing, the fourth color, black, is added, giving rise to the CMYK color model. The primary colors (R, G, B) can be added to produce the secondary colors.

- Red plus blue can generate magenta
- Green plus blue can generate cyan
- Red plus green can generate yellow

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

**The HIS Model:**

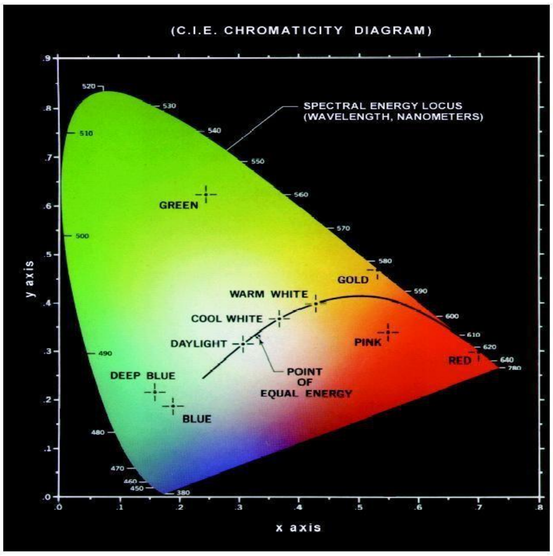
Human eye distinguish one color from the other based on hue, saturation, and brightness.

-  Hue, is a color that is evoked by a single wavelength of light in the visible spectrum, or by a relatively narrow band of wavelengths; hue represents dominant color as perceived by the observer.

- Saturation (purity) refers to the relative amount of white light mixed with a hue; it is inversely proportional to the amount of white light added; to desaturate a color of given intensity in asubtractive system (such as watercolor), one can add white, black, gray.
- Brightness embodies the achromatic notion of intensity.
- Hue and saturation together are called chromaticity. For any particular color, the amounts of red, green, and blue needed to form it are called tristimulus values, and they are denoted as X, Y, and Z.A color is specified by its trichromatic coefficients, defined as

$$x = \frac{X}{X+Y+Z}$$
$$y = \frac{Y}{X+Y+Z}$$
$$z = \frac{Z}{X+Y+Z}$$
$$x + y + z = 1$$

**CIE chromaticity diagram:** It is a function of *x* (red) and *y* (green),
*z* can be derived by *z* =1-*x*-*y*




Some notes on CIE chromaticity diagram

- Any color located on the boundary of the chromaticity chart is fully saturated; any point not on theboundary but within the diagram represents some mixture of spectrum colors
- The point of equal energy represents the standard white light; its saturation is zero
- As a point leaves the boundary and approaches the point of equal energy, more white light is addedto the color and it becomes less saturated
- A straight line segment joining any two points in the diagram defines all the different color variations that can be obtained by combining these two colors additively

The HSImodel uses three measures to describe colors, hue, saturation, and intensity

- Their relationship can be represented in a cylindrical coordinate system
- angle around the central vertical axis corresponds to "hue",
- the distance from the axis corresponds to "saturation",
- The distance along the axis corresponds to "value"

**Conversion from RGB to HIS**

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

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

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

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

**Conversion from HSI to RGB**

- RG sector ($0 <= H < 120^\circ$)

$$B = I(1-S) \quad R = I\left[1 + \frac{S\cos H}{\cos(60^\circ - H)}\right] \quad G = 3I - (R+B)$$

- GB sector ($120^\circ <= H < 240^\circ$)

$$H = H - 120^\circ$$

$$R = I(1-S) \quad G = I\left[1 + \frac{S\cos H}{\cos(60^\circ - H)}\right] \quad B = 3I - (R+G)$$

- BR sector ($240^\circ <= H <= 360^\circ$)

$$H = H - 240^\circ$$

$$G = I(1-S) \quad B = I\left[1 + \frac{S\cos H}{\cos(60^\circ - H)}\right] \quad R = 3I - (G+B)$$

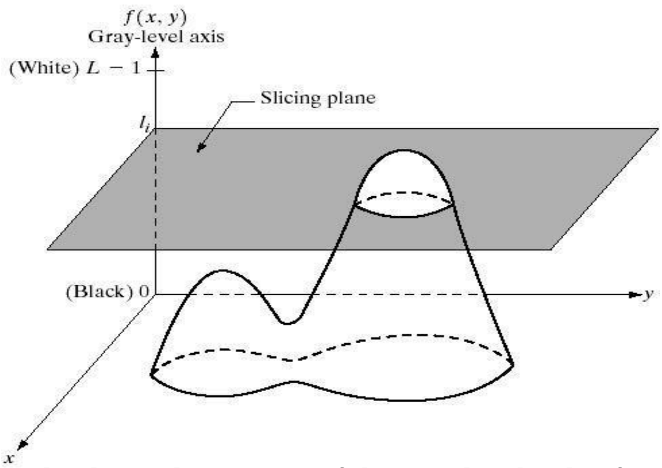
### 4.5 Pseudocolour Image Processing:

- Pseudocolour image processing consists of assigning colors to grey values based on a specificcriterion
- The principle use of Pseudocolour image processing is for human visualization
- Humans can discern between thousands of color shades and intensities, compared to only about two dozen or so shades of grey

**Intensity Slicing**

Intensity slicing and color coding is one of the simplest kinds of pseudocolour image processing

- First we consider an image as a 3D function mapping spatial coordinates to intensities (that we canconsider heights)
- Now consider placing planes at certain levels parallel to the coordinate plane
- If a value is one side of such a plane it is rendered in one color, and a different color if on the other side. Intensity slicing and color coding is one of the simplest kinds of pseudocolour image processing




Intensity slicing and color coding is one of the simplest kinds of Pseudocolour image processing In general intensity slicing can be summarized as:

Let [*0*, *L*-1] represent the grey scale

- Let $l_0$ represent black [$f(x, y) = 0$] and let $l_{L-1}$ represent white [$f(x, y) = L-1$]
- Suppose *P* planes perpendicular to the intensity axis are defined at levels $l_1, l_2, \ldots, l_p$
- Assuming that $0 < P < L-1$, then the *P* planes partition the grey scale into $P + 1$ intervals $V_1, V_2, \ldots, V_{P+1}$

Grey level color assignments can then be made according to the relation:

$$f(x,y) = c_k \qquad \text{if } f(x,y) \in V_k$$

Where $c_k$ is the color associated with the $k^{th}$ intensity level $V_k$ defined by the partitioning planes at $l = k - 1$and $l = k$

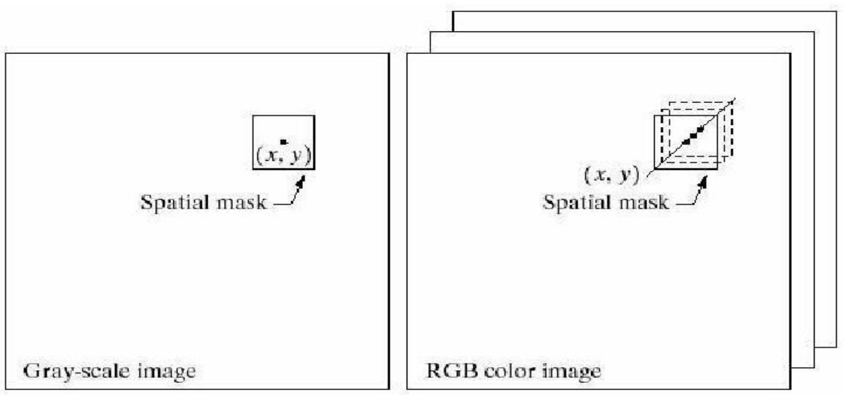
## 4.6 Full-Color Image Processing

- Full-color image processing approaches fall into two major categories.
  (1) Process each component image individually and form composite processed color imagefrom the individually processed components
  (2) Work with color pixels directly; color pixels really are vectors:

$$\mathbf{c}(x,y) = \begin{pmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{pmatrix} = \begin{pmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{pmatrix}$$

- Note: the results of individual color component processing are not always equivalent to directprocessing in color vector space.
- Processing is equivalent if:
  - (1) the process is applicable to both scalars and vectors;
  - (2) the operation on each component of a vector is independent of the other components.
- Result for per-color-component and vector-based processing is equivalent.

**Neighborhood averaging**




## 4.7 Morphological Image Processing

**Preliminaries**

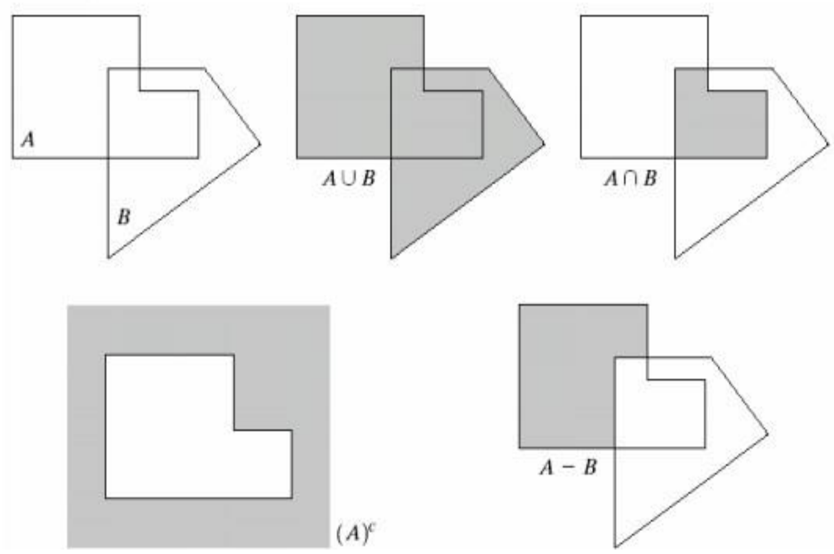
**Some Basic Concepts form Set Theory**

• If every element of a set A is also an element of another set B, then A is said to be a subsetof B, denoted as **A ⊆ B**

• The union of two sets A and B, denoted by **C = A ∪ B**

• The intersection of two sets A and B, denote by **D= A ∩ B**

• Disjoint or mutually exclusive **A ∩ B = ∅**




a b c
d e

(a) Two sets $A$ and $B$. (b) The union of $A$ and $B$. (c) The intersection of $A$ and $B$. (d) The complement of $A$. (e) The difference between $A$ and $B$.

The complement of a set A is the set of elements not contained in A

$$A^c = \{\omega | \omega \notin A\}$$

The difference of two sets A and B, denoted A - B, is defined as

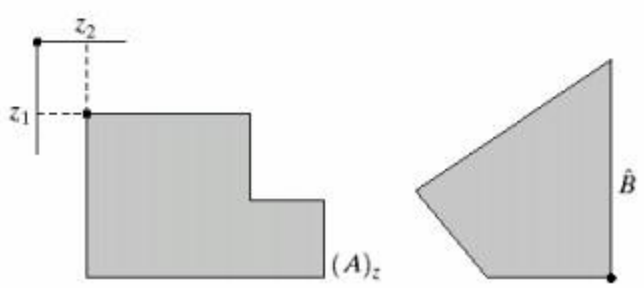
$$A - B = \{\omega | \omega \in A, \omega \notin B\} = A \cap B^c$$

The reflection of set B, denoted ,is defined as

$$\hat{B} = \{\omega | \omega = -b, for \quad b \in B\}$$

- The translation of set A by point z = (z 1,z 2), denoted (A)z, ,is defined as

$$(A)_z = \{c | c = a + z, for \quad a \in A\}$$




a b

(a) Translation of $A$ by $z$.
(b) Reflection of $B$. The sets $A$ and $B$ are from

**TABLE** The three basic logical operations.

| $p$ | $q$ | $p$ AND $q$ (also $p \cdot q$) | $p$ OR $q$ (also $p + q$) | NOT ($p$) (also $\bar{p}$) |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 |
| 1 | 1 | 1 | 1 | 0 |

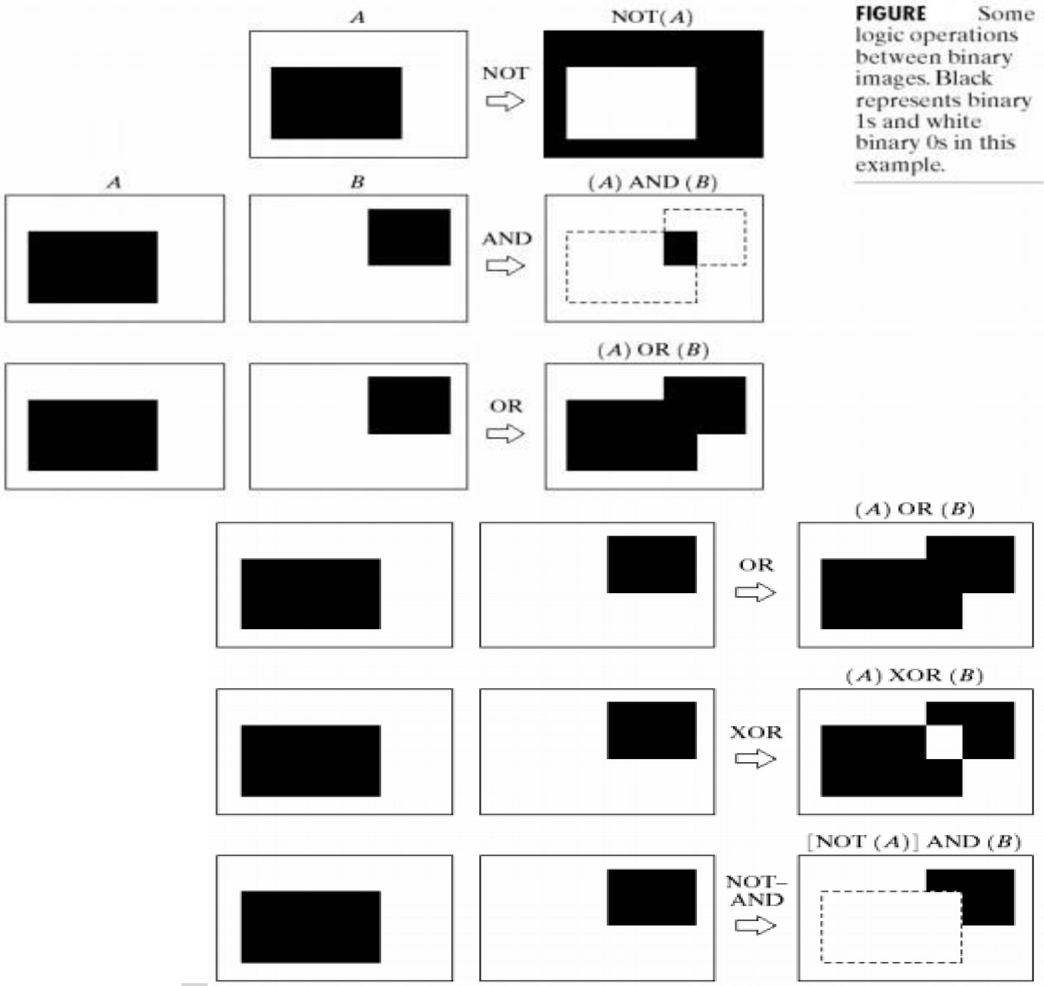



**FIGURE** Some logic operations between binary images. Black represents binary 1s and white binary 0s in this example.

### 4.7.1 Dilation and Erosion

Dilation and erosion are basic morphological processing operations. They are defined in terms of more elementary set operations, but are employed as the basic elements of many algorithms. Both dilation and erosion are produced by the interaction of a set called a structuring element with a set of pixels of interest in theimage.

The structuring element has both a shape and an origin.

Let $A$ be a set of pixels and let $B$ be a structuring element. Let $(\hat{B})_s$ be the reflection of $B$ about its origin and followed by a shift by $s$. Dilation, written $A \oplus B$, is the set of all shifts that satisfy the following:

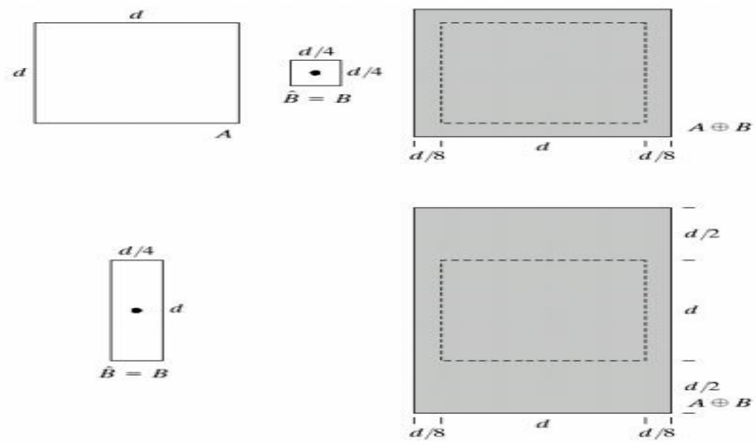
$$A \oplus B = \{s | (\hat{B})_s \cap A \neq \emptyset\}$$

Equivalently,

$$A \oplus B = \{s | \left((\hat{B})_s \cap A\right) \subseteq A\}$$

## Morphological Dilation

Any pixel in the output image touched by the · in the structuring element is set to ON when any point of the structuring element touches a ON pixel in the original image. This tends to close up holes in an image by expanding the ON regions. It also makes objects larger. Note that the result depends upon both the shape of the structuring element and the location of its origin.




a b c
d e

**FIGURE** (a) Set $A$. (b) Square structuring element (dot is the center). (c) Dilation of $A$ by $B$, shown shaded. (d) Elongated structuring element. (e) Dilation of $A$ using this element.

## Morphological Erosion

The erosion of A by B, denoted $A \Theta B$ ,is defined as

$$A \Theta B = \left\{ z \middle| (B)_z \subseteq A \right\} \qquad (9.2-3)$$

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

Any pixel in the output image touched by the · in the structuring element is set to ON when every point of the structuring element touches a ON pixel in the original image. This tends to makes objects smaller by removing pixels.

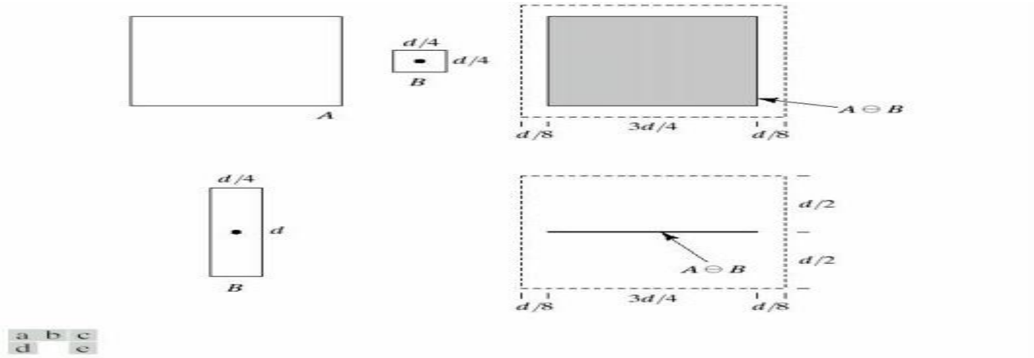



**FIGURE** (a) Set A. (b) Square structuring element. (c) Erosion of A by B, shown shaded. (d) Elongated structuring element. (e) Erosion of A using this element.

## References

1. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing Using Matlab", 3rd Edition Tata McGraw Hill Pvt. Ltd., 2011.
2. Anil Jain K. "Fundamentals of Digital Image Processing", PHI Learning Pvt. Ltd., 2011.
3. William K. Pratt, "Introduction to Digital Image Processing", CRC Press, 2013.

## Question Bank

### PART-A

| S.No | |
|---|---|
| 1. | What are the two properties used for establishing similarity of edge pixels? |
| 2. | Formulate how the derivatives are obtained in edge detection |
| 3. | What is line detection? |
| 4. | What is Hue of saturation? |
| 5. | What do you meant by Color model? |
| 6. | List the hardware oriented color models? |
| 7. | Compare full color and pseudo color? |
| 8. | Define intensity slicing. |
| 9. | What is Chromaticity |
| 10. | What is segmentation? |
| 11. | Mention the applications of segmentation. |
| 12. | List out the steps involved in splitting and merging. |
| 13. | What is the condition to be met by the partitions in region based segmentation? |

### PART-B

| S.No | |
|---|---|
| 1. | Discuss about region based image segmentation techniques. |
| 2. | Evaluate about pseudocolor image processing with appropriate examples. |
| 3. | Explain about full color image processing. |
| 4. | Interpret in detail about various color models. |
| 5. | What is edge detection? Summarize the types of edge detection operations. |
| 6. | Describe about Point and Line detection. |
| 7. | Illustrate how region growing, region splitting and merging approaches are used for image segmentation. |
| 8. | Demonstrate the Dilation operation in image morphology with examples? |
| 9. | Demonstrate the Erosion operation in image morphology with examples? |
| 10. | Describe dilation and erosion morphological transformations on a binary image |

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

**SCHOOL OF BIO AND CHEMICAL ENGINEERING**

**DEPARTMENT OF BIOMEDICAL ENGINEERING**

# UNIT – V – Medical Image Processing – SBMA1603

# UNIT-5

# IMAGE COMPRESSION

Image compression address the problem of reducing the amount of data required to represent a digital image with no significant loss of information**.** Interest in image compression dates back more than 25 years. The field is now poised significant growth through the practical application of the theoretic work that began in 1940s, when C.E. Shannon and others first formulated the probabilistic view of information and its representation, transmission and compression.

Images take a lot of storage space:

- 1024 x 1024 x 32 x bits images requires 4 MB
- suppose you have some video that is 640 x 480 x 24 bits x 30 frames per second , 1 minute of video would require 1.54 GB

  Many bytes take a long time to transfer slow connections – suppose we have 56,000 bps
- 4MB will take almost 10 minutes, - 1.54 GB will take almost 66 hours

  Storage problems, plus the desire to exchange images over the Internet, have led to a large interest in imagecompression algorithms.

**Definition:** 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.

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

**Need for Compression:**

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

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

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

**5.2 DATA REDUNDANCY:** Data are the means by which information is conveyed. Various amounts of data can be used to convey the same amount of information. Example: Four different representation of the same information (number five)

1) A picture (1001, 632 bits);
2) A word “five” spelled in English using the ASCII character set (32 bits);
3) A single ASCII digit (8bits);
4) Binary integer (3bits)

Compression algorithms remove redundancy

If more data are used than is strictly necessary, then we say that there is redundancy in the dataset.

- Let $n_1$ and $n_2$ denote the number of information carrying units in two data sets that represent the same information
- The relative redundancy $R_D$ is define as : $$R_D = 1 - \frac{1}{C_R}$$
- where $C_R$ is the compression ratio $$C_R = \frac{n_1}{n_2}$$

When $n_2 << n_1$, $C_R \rightarrow$ large value and $R_D \rightarrow 1$. Larger values of C indicate better compression

- If $n_1 = n_2$, $C_R=1$ and $R_D=0$ *no redundancy*
- If $n_1 >> n_2$, $C_R=0$ and $R_D= -\infty$ *high redundancy*
- If $n_1 << n_2$, $C_R= \infty$ and $R_D=1$ *undesirable*
- *A* compression ratio of 10 (10:1) means that the first data set has 10 information carrying units (say, bits) for every 1 unit in the second (compressed) data set.

Data compression is defined as the process of encoding data using a representation that reduces the overall size of data. This reduction is possible when the original dataset contains some type of redundancy. Digital image

compression is a field that studies methods for reducing the total number of bits required to represent an image. This can be achieved by eliminating various types of redundancy that exist in the pixel values. In general, three basic redundancies exist in digital images that follow

**REDUNDANCY IN DIGITAL IMAGES**

–**Coding redundancy**- usually appear as results of the uniform representation of each pixel

–**Spatial/Temporal redundancy**-because the adjacent pixels tend to have similarity in practical.

– **Irrelevant Information**-Image contain information which are ignored by the human visual system

### 5.2.1 Coding Redundancy:

Our quantized data is represented using code words. The code words are ordered in the same way as the intensities that they represent; thus the bit pattern 00000000, corresponding to the value 0, represents the darkest points in an image and the bit pattern 11111111, corresponding to the value 255, represents the brightest points.

- if the size of the code word is larger than is necessary to represent all quantization levels, then we havecoding redundancy

An 8-bit coding scheme has the capacity to represent 256 distinct levels of intensity in an image. But if there are only 16 different grey levels in an image, the image exhibits coding redundancy because it could be represented using a 4-bit coding scheme. Coding redundancy can also arise due to the use of fixed-length code words.

Grey level histogram of an image also can provide a great deal of insight into the construction of codes to reduce the amount of data used to represent it.

Let us assume, that a discrete random variable $rk$ in the interval (0,1) represents the grey levels of an image and that each $rk$ occurs with probability $Pr(rk)$. Probability can be estimated from the histogram of an image using

$$Pr\ (rk) = \mathrm{h}k\ /n \text{ for } k = 0, 1 \ldots\ldots L-1 \qquad (3)$$

Where $L$ is the number of grey levels and $hk$ is the frequency of occurrence of grey level $k$ (the number of times that the $k$th grey level appears in the image) and $n$ is the total number of the pixels in the image. If the number of the bits used to represent each value of $rk$ is $l\ (rk)$, the average number of bits required to represent each pixel is:

$$L_{avg} = \sum_{k=0}^{L-1} l(r_k) P_r(r_k)$$

Example:

| $r_k$ | $p_r(r_k)$ | **Code 1** | $l_1(r_k)$ | **Code 2** | $l_2(r_k)$ |
|---|---|---|---|---|---|
| $r_0 = 0$ | 0.19 | 000 | 3 | 11 | 2 |
| $r_1 = 1/7$ | 0.25 | 001 | 3 | 01 | 2 |
| $r_2 = 2/7$ | 0.21 | 010 | 3 | 10 | 2 |
| $r_3 = 3/7$ | 0.16 | 011 | 3 | 001 | 3 |
| $r_4 = 4/7$ | 0.08 | 100 | 3 | 0001 | 4 |
| $r_5 = 5/7$ | 0.06 | 101 | 3 | 00001 | 5 |
| $r_6 = 6/7$ | 0.03 | 110 | 3 | 000001 | 6 |
| $r_7 = 1$ | 0.02 | 111 | 3 | 000000 | 6 |

$$\begin{aligned} L_{av} &= \sum_{k=0}^{7} l(r_k)(p(r_k)) \\ &= 2(019) + 2(0.25) + 3(0.16) + \ldots + 6(0.02) \\ &= 2.7\ bits \end{aligned}$$

$$C_R = \frac{3}{2} = 1.11$$

$$R_D = 1 - \frac{1}{1.11} = 0.099$$

### 5.2.2 Interpixel Redundancy:

Consider the images shown in Figs. 1.1(a) and (b). As Figs. 1.1(c) and (d) show, these images have virtually identical histograms. Note also that both histograms are trimodal, indicating the presence of three dominant ranges of gray-level values. Because the gray levels in these images are not equally probable, variable-length coding can be used to reduce the coding redundancy that would result from a straight or natural binary encoding of their pixels. The coding process, however, would not alter the level of correlation between the pixels within the images. In other words, the codes used to represent the gray levels of each image have nothing to do with the correlation between pixels. These correlations result from the structural or geometric relationships between the objects in the image.

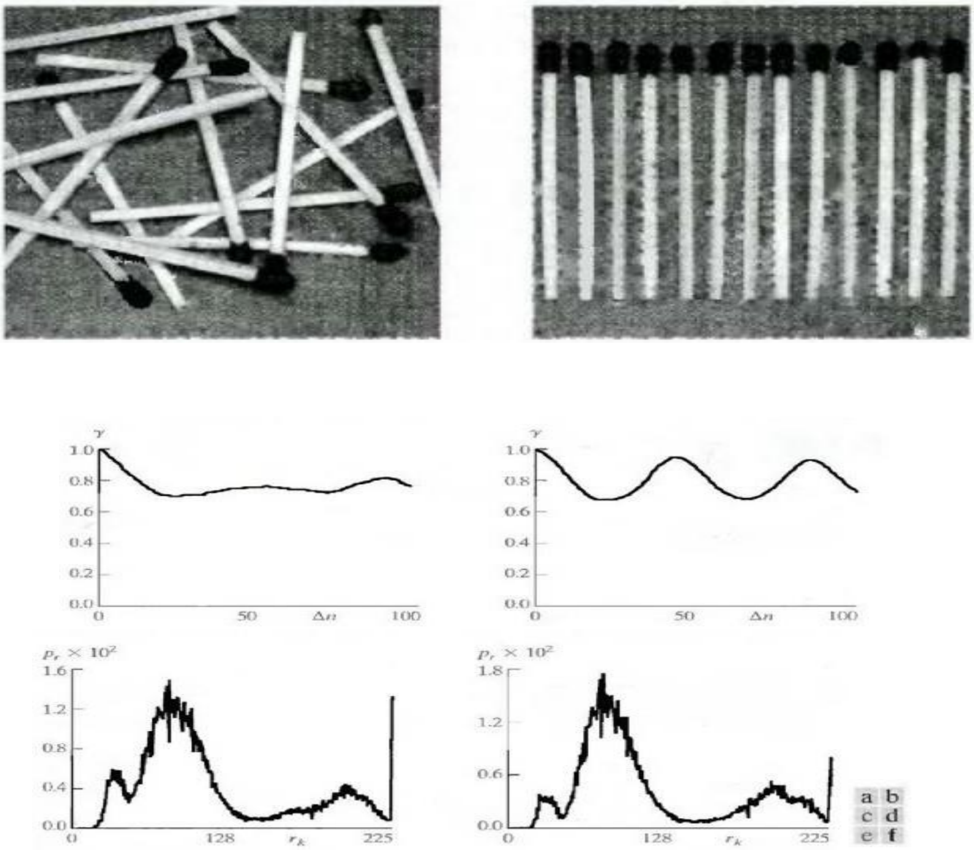



**Fig.1.1 Two images and their gray-level histograms and normalized autocorrelation coefficients along one line.**

Figures 1.1(e) and (f) show the respective autocorrelation coefficients computed along one line of each image.

$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)}$$

where

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x, y) f(x, y + \Delta n).$$

The scaling factor in Eq. above accounts for the varying number of sum terms that arise for each integer value of n. Of course, n must be strictly less than N, the number of pixels on a line. The variable x is the coordinate of the line used in the computation. Note the dramatic difference between the shape of the functions shown in Figs. 1.1(e) and(f). Their shapes can be qualitatively related to the structure in the images in Figs. 1.1(a) and (b).This relationship is particularly noticeable in Fig. 1.1 (f), where the high correlation between pixels separated by 45 and 90 samples can be directly related to the spacing between the vertically oriented matches of Fig. 1.1(b). In addition, the adjacent pixels of both images are highly correlated. When n is 1, $\gamma$ is 0.9922 and 0.9928 for the images of Figs.
1.1 (a) and (b), respectively. These values are typical of most properly sampled television images.

These illustrations reflect another important form of data redundancy—one directly related to the interpixel correlations within an image. Because the value of any given pixel can be reasonably predicted from the value ofits neighbors, the information carried by individual pixels is relatively small. Much of the visual contribution of a single pixel to an image is redundant; it could have been guessed on the basis of the values of its neighbors. A variety of names, including spatial redundancy, geometric redundancy, and interframe redundancy, have been coined to refer to these interpixel dependencies. We use the term interpixel redundancy to encompass them all.

In order to reduce the interpixel redundancies in an image, the 2-D pixel array normally used for human viewing and interpretation must be transformed into a more efficient (but usually "nonvisual") format. For example, the differences between adjacent pixels can be used to represent an image. Transformations of this type (that is, those that remove interpixel redundancy) are referred to as mappings. They are called reversible mappings if the original image elements can be reconstructed from the transformed data set.

**5.2.3 Psychovisual Redundancy:**

The brightness of a region, as perceived by the eye, depends on factors other than simply the light reflected by the region. For example, intensity variations (Mach bands) can be perceived in an area of constant intensity. Such phenomena result from the fact that the eye does not respond with equal sensitivity to all visual information. Certain information simply has less relative importance than other information in normal visual processing. This information is said to be psychovisually redundant. It can be eliminated without significantly impairing the quality of image perception.

That psychovisual redundancies exist should not come as a surprise, because human perception of the information in an image normally does not involve quantitative analysis of every pixel value in the image. In general, an observer searches for distinguishing features such as edges or textural regions and mentally combines them into recognizable groupings. The brain then correlates these groupings with prior knowledge in order to complete the image interpretation process. Psychovisual redundancy is fundamentally different from the redundancies discussed earlier. Unlike coding and interpixel redundancy, psychovisual redundancy is associated with real or quantifiable visual information. Its elimination is possible only because the information itself is not essential for normal visual processing. Since the elimination of psychovisually redundant data results in a loss of quantitative information, it iscommonly referred to as quantization.

This terminology is consistent with normal usage of the word, which generally means the mapping of a broad range of input values to a limited number of output values. As it is an irreversible operation (visual information is lost), quantization results in lossy data compression.

## 5.3 IMAGE COMPRESSION MODELS

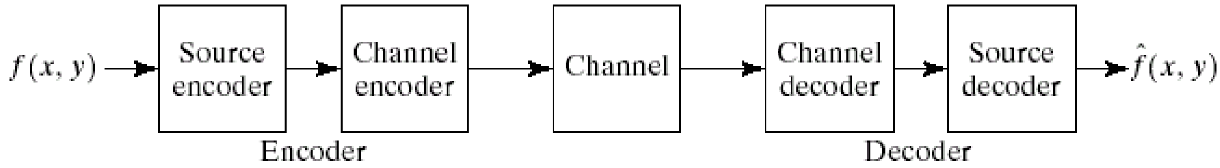
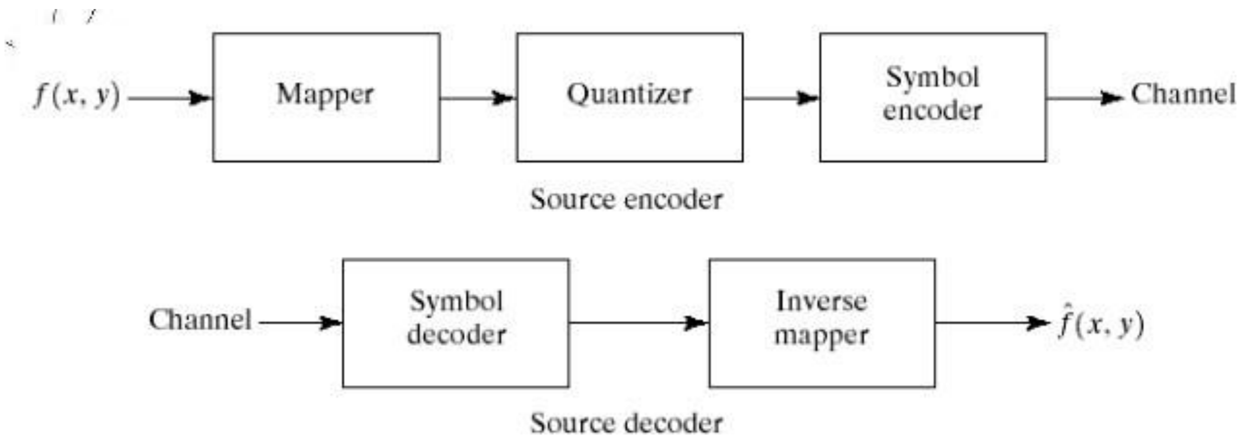



Figure shows, a compression system consists of two distinct structural blocks: an encoder and a decoder. An input image f(x, y) is fed into the encoder, which creates a set of symbols from the input data. After transmission over the channel, the encoded representation is fed to the decoder, where a reconstructed output image f^(x, y) is generated. In general, f^(x, y) may or may not be an exact replica of f(x, y). If it is, the system is error free or information preserving; if not, some level of distortion is present in the reconstructed image. Both the encoder and decoder shown in Fig. 3.1 consist of two relatively independent functions or sub blocks. The encoder is made up ofa source encoder, which removes input redundancies, and a channel encoder, which increases the noise immunity of the source encoder's output. As would be expected, the decoder includes a channel decoder followed by a sourcedecoder. If the channel between the encoder and decoder is noise free (not prone to error), the channel encoder and decoder are omitted, and the general encoder and decoder become the source encoder and decoder, respectively.




**The Source Encoder and Decoder:**

- **Source Encoder**

  Reduces/eliminates any coding, interpixel or psychovisual redundancies. The Source Encoder contains 3processes:

  •**Mapper**

  Transforms the image into array of coefficients reducing interpixel redundancies. This is a reversible process which is not lossy. May or may not reduce directly the amount of data required to represent the image.

  •**Quantizer**

This process reduces the accuracy and hence psychovisual redundancies of a given image. This process isirreversible and therefore lossy. It must be omitted when error-free compression is desired.

**•Symbol Encoder**

This is the source encoding process where fixed or variable-length code is used to represent mapped and quantized data sets. This is a reversible process (not lossy). Removes coding redundancy by assigning shortest codes for the most frequently occurring output values.

- **Source Decoder** contains two components.

  **•Symbol Decoder:** This is the inverse of the symbol encoder and reverse of the variable-length coding isapplied.

  **•Inverse Mapper:** Inverse of the removal of the interpixel redundancy.

  •The only lossy element is the Quantizer which removes the psycho visual redundancies causing irreversibleloss. Every Lossy Compression methods contain the Quantizer module.

  •If error-free compression is desired the quantizer module is removed.

**The Channel Encoder and Decoder:**

The channel encoder and decoder play an important role in the overall encoding-decoding process when thechannel is noisy or prone to error. They are designed to reduce the impact of channel noise by inserting a controlled form of redundancy into the source encoded data. As the output of the source encoder contains little redundancy, it would be highly sensitive to transmission noise without the addition of this "controlled redundancy." One of the most useful channel encoding techniques was devised by R. W. Hamming (Hamming [1950]). It is based on appending enough bits to the data being encoded to ensure that some minimum number of bits must change between valid code words. Hamming showed, for example, that if 3 bits of redundancy are added to a 4-bit word, so that the distance between any two valid code words is 3, all single-bit errors can be detected and corrected. (By appending additional bits of redundancy, multiple-bit errors can be detected and corrected.) The 7- bit Hamming (7, 4) code word $h_1$, $h_2$, $h_3$…., $h_6$, $h_7$ associated with a 4-bit binary number $b_3b_2b_1b_0$ is

$$h_1 = b_3 \oplus b_2 \oplus b_0 \qquad h_3 = b_3$$
$$h_2 = b_3 \oplus b_1 \oplus b_0 \qquad h_5 = b_2$$
$$h_4 = b_2 \oplus b_1 \oplus b_0 \qquad h_6 = b_1$$
$$h_7 = b_0$$

$$c_1 = h_1 \oplus h_3 \oplus h_5 \oplus h_7$$
$$c_2 = h_2 \oplus h_3 \oplus h_6 \oplus h_7$$
$$c_4 = h_4 \oplus h_5 \oplus h_6 \oplus h_7.$$

Where denotes the exclusive OR operation. Note that bits h1, h2, and h4 are even- parity bits for the bit fields b3 b2 b0, b3b1b0, and b2b1b0, respectively. (Recall that a string of binary bits has even parity if the number of bits with a value of 1 is even.) To decode a Hamming encoded result, the channel decoder must check the encoded value for odd parity over the bit fields in which even parity was previously established. A single-bit error is indicated by a nonzero parity word c4c2c1, where

If a nonzero value is found, the decoder simply complements the code word bit position indicated by the parityword. The decoded binary value is then extracted from the corrected code word as h3h5h6h7.

## 5.4 ELEMENTS OF INFORMATION THEORY

### Measuring Information

The generation of information is modeled as a probabilistic process. Random event E occurs with probability P(E)

$$I(E) = \log \frac{1}{P(E)} = -\log P(E)$$

The base of the logarithm determines the units used to measure the information. If the base 2 is selected theresulting information unit is called bit. If P(E)=0.5 (two possible equally likely events) the information is one bit.

### The Information Channel

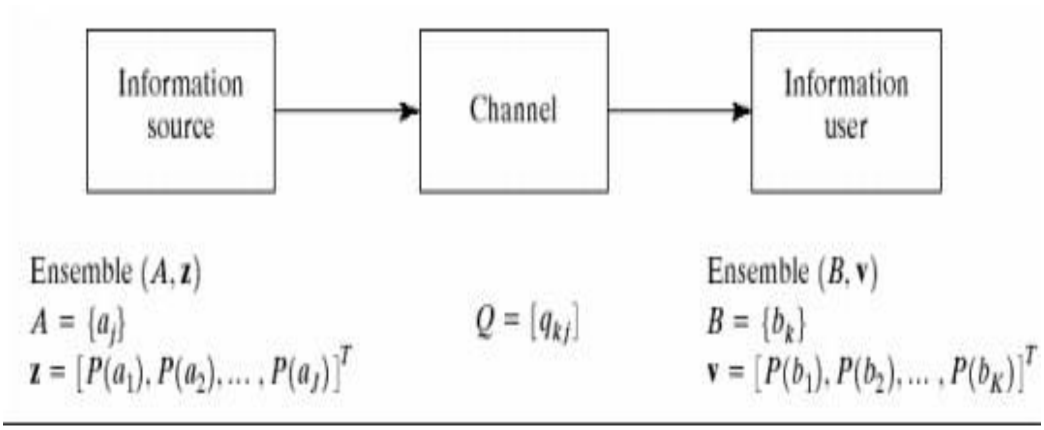
- The information channel: the physical medium that links the source to the user (tel. line, wire, electromagnetic waves, etc.)
- The source
  - Modeled as a discrete random variable
  - Source alphabet $A=\{a_j\}$
  - Symbols (letters) $a_j$ with probabilities $P(a_j)$
  - The average self-information obtained from $k$ outputs is

    $$-kP(a_1)\log P(a_1) - kP(a_2)\log P(a_2) - \ldots - kP(a_J)\log P(a_J) =$$
    $$= -k\sum_{j=1}^{J} P(a_j)\log P(a_j)$$

  - The average information per source output is

    $$H(\mathbf{z}) = -\sum_{j=1}^{J} P(a_j)\log P(a_j)$$

  - (*uncertainty* or *entropy*)




- The channel
  - The information in the output of the channel is a discrete random variable
  - Channel alphabet $B=\{b_k\}$ with letters $b_k$ with probabilities $P(b_k)$
  - The probabilities are related to the source **z** by conditional probabilities

$$P(b_k) = \sum_{j=1}^{J} P(b_k \mid a_j) \log P(a_j)$$

  - Channel (transition) matrix

$$\mathbf{Q} = \begin{bmatrix} P(b_1 \mid a_1) & P(b_1 \mid a_2) & \ldots & P(b_1 \mid a_J) \\ P(b_2 \mid a_1) & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ P(b_K \mid a_1) & P(b_K \mid a_2) & \ldots & P(b_K \mid a_J) \end{bmatrix}$$

$\mathbf{Q} = [q_{kj}]$, where $q_{kj} = P(b_k \mid a_j)$

- The capacity of the channel
  - The conditional entropy function for each $b_k$

$$H(\mathbf{z} \mid b_k) = -\sum_{j=1}^{J} P(a_j \mid b_k) \log P(a_j \mid b_k)$$

  - The averaged value over all $b_k$ (*equivocation*)

$$H(\mathbf{z} \mid \mathbf{v}) = \sum_{k=1}^{K} H(\mathbf{z} \mid b_k) P(b_k) = -\sum_{j=1}^{J} \sum_{k=1}^{K} P(a_j, b_k) \log P(a_j \mid b_k)$$

  - The mutual information of **z** and **v**

$$I(\mathbf{z}, \mathbf{v}) = H(\mathbf{z}) - H(\mathbf{z} \mid \mathbf{v}) = \sum_{j=1}^{J} \sum_{k=1}^{K} P(a_j, b_k) \log \frac{P(a_j, b_k)}{P(a_j) P(b_k)} = \sum_{j=1}^{J} \sum_{k=1}^{K} P(a_j) q_{kj} \log \frac{q_{kj}}{\sum_{j=1}^{J} P(a_j) q_{kj}}$$

  - The capacity

$$C = \max_{\mathbf{z}} [I(\mathbf{z}, \mathbf{v})]$$

## 5.5. Fundamental Coding Theorems

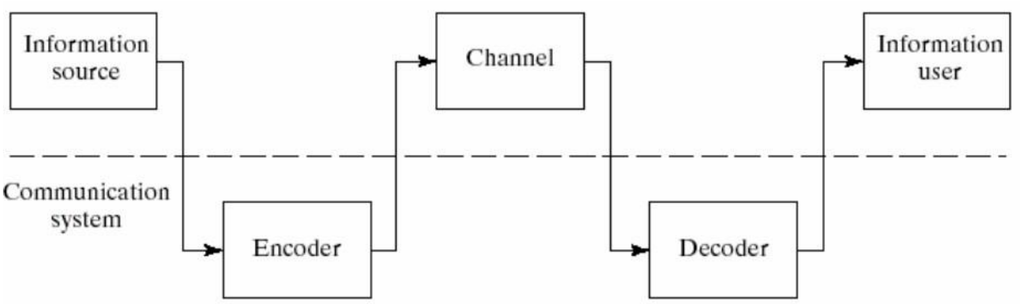



**Fig: A Communication System Model**

### 5.5.1 Noiseless Coding Theorem

***The Noiseless Coding Theorem or Shannon's First Theorem or Shannon's Source Coding Theorem for Lossless Data Compression***

When both the information channel and communication system are error-free

Defines the minimum average codeword length per source symbol that can be achieved.

Aim: to represent source as compact as possible.

Let the information source (A,**z**), with statistically independent source symbols, output an n-tuple of symbols from source alphabet A. Then, the source output takes on one of the $J^n$ possible values, denoted by, $\alpha_i$, from

$$A = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{J^n}\}$$

*Probability of a given* $\alpha_i$, $P(\alpha_i)$ *is related to single* − *symbol probabilities as*

$$P(\alpha_i) = P(a_{j1})P(a_{j2})\ldots P(a_{jn})$$
$$\mathbf{z}' = \{P(\alpha_1), P(\alpha_2), \ldots, P(\alpha_{J^n})\}$$

*Entropy of the source is given by*

$$H(\mathbf{z}') = -\sum_{i=1}^{J^n} P(\alpha_i)\log P(\alpha_i)$$
$$H(\mathbf{z}') = nH(\mathbf{z})$$

Hence, the entropy of the zero-memory source is n times the entropy of the corresponding single symbol source. Such a source is called the ***$n^{th}$ extension*** of single-symbol source.

$$Self\ information\ of\ \alpha_i\ is\ \log\frac{1}{P(\alpha_i)}.$$

$$\log\frac{1}{P(\alpha_i)} \le l(\alpha_i) < \log\frac{1}{P(\alpha_i)} + 1$$

*$\alpha_i$ is therefore represented by a codeword whose length is the smallest integer exceeding the self - information of $\alpha_i$.*

$$P(\alpha_i)\log\frac{1}{P(\alpha_i)} \le P(\alpha_i)l(\alpha_i) < P(\alpha_i)\log\frac{1}{P(\alpha_i)} + P(\alpha_i)$$

$$\sum_{i=1}^{n^J} P(\alpha_i)\log\frac{1}{P(\alpha_i)} \le \sum_{i=1}^{n^J} P(\alpha_i)l(\alpha_i) < \sum_{i=1}^{n^J} P(\alpha_i)\log\frac{1}{P(\alpha_i)} + 1$$

$$H(\mathbf{z}') \le L_{avg} < H(\mathbf{z}') + 1$$

$$where\ L'_{avg} = \sum_{i=1}^{J^n} P(\alpha_i)l(\alpha_i)$$

$$\frac{H(\mathbf{z}')}{n} \le \frac{L'_{avg}}{n} < \frac{H(\mathbf{z}')+1}{n}$$

$$H(\mathbf{z}) \le \frac{L'_{avg}}{n} < H(\mathbf{z}) + \frac{1}{n}$$

$$\lim_{n\to\infty}\left[\frac{L'_{avg}}{n}\right] = H(\mathbf{z})$$

Shannon's source coding theorem for lossless data compression states that ***for any code used to represent the symbols from a source, the minimum number of bits required to represent the source symbols on an average must be atleast equal to the entropy of the source.***

$$H(\mathbf{z}) \le \frac{L'_{avg}}{n} < H(\mathbf{z}) + \frac{1}{n}$$

*The efficiency $\eta$ of any encoding strategy can be defined as*

$$\eta = \frac{nH(\mathbf{z})}{L'_{avg}}$$

$$= \frac{H(\mathbf{z}')}{L'_{avg}}$$

### 5.5.2 Noisy Coding Theorem

***The Noisy Coding Theorem or Shannon's Second Theorem***

When the channel is noisy or prone to error

Aim: to encode information so that the communication is made reliable and the error is minimized.

Use of repetitive coding scheme

Encode nth extension of source using K-ary code sequences of length r, $K^r \leq J^n$.

Select only φ of the $K^r$ possible code sequences as valid codewords.

A zero-memory information source generates information at a rate equal to its entropy.

The nth extension of the source provides information at a rate of $\frac{H(\mathbf{z}')}{n}$ information units per symbol.

If the information is coded, the maximum rate of coded information is log(φ /r) and occurs when the φ valid codewords used to code the source are equally probable.

Hence, a code of size φ and block length r is said to have a rate of

$$R = \log \frac{\varphi}{r}$$

information units per symbol.

The noisy coding theorem thus states that ***for any R<C, where C is the capacity of the zero-memory channel with matrix Q, there exists an integer r, and code of block length r and rate R such that the probability of a block decoding error is less than or equal to ε for any ε>0***.

That is, the probability of error can be made arbitrarily small so long as the coded message rate is less than the capacity of the channel.

**References**


1. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing Using Matlab", 3rd Edition Tata McGraw Hill Pvt. Ltd., 2011.
2. Anil Jain K. "Fundamentals of Digital Image Processing", PHI Learning Pvt. Ltd., 2011.
3. William K. Pratt, "Introduction to Digital Image Processing", CRC Press, 2013.


**Question Bank**

**PART-A**

S.No

1. What is image compression?
2. Investigate the performance metrics for evaluating image compression.
3. List the need for Compression?
4. Define compression ratio.
5. What is Redundancy?

6. Validate the types of data redundancy.
7. What is the operation of source encoder?
8. What is the function of channel encoder?
9. Categorize video compression standards.
10. Specify the fundamental coding theorem.
11. What is meant by inverse mapping?

| S.No | **PART-B** |
| --- | --- |
| 1. | What is data redundancy? Illustrate various types of data redundancy in detail. |
| 2. | Demonstrate in detail about Image compression model? |
| 3. | Discuss in detail source encoder and decoder, |
| 4. | Analyze Shannon's first theorem for noiseless coding theorem. |
| 5. | Apply and analyze Shannon's second theorem for noisy coding theorem. |
| 6. | Evaluate fundamental coding theorem. |
| 7. | Summarize the different types of redundancy. |
| 8. | Compare and contrast noiseless and noisy coding theorem. |