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 ELECTRICAL & ELECTRONICS ENGINEERING

DEPARTMENT OF ECE

UNIT – I – DIGITAL IMAGE FUNDAMENTALS – SECA 3009

UNIT 1

DIGITAL IMAGE FUNDAMENTALS - INTRODUCTION

Image Representation, Components of Digital Image Processing Systems, Image Sensing and Acquisition, Elements of Visual Perception, Image formation model, Image Sampling and Quantization, Relationship between pixels.

1.1 INTRODUCTION

An image is defined as a two-dimensional function, F(X, Y), where x and y are spatial coordinates, and the amplitude of F at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the intensity values of f are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images 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 value. These elements are called picture elements, image elements, pels, and pixels.

1.2 IMAGE REPRESENTATION

There are two principal ways to represent digital images. Assume that an image f(x, y) is sampled so that the resulting digital image has M rows and N columns. The values of the coordinates (x, y) now become discrete quantities. If we consider integer values for these discrete coordinates, then it becomes a digital representation. Thus, the values of the coordinates at the origin are (x, y) = (0, 0). The next coordinate values along the first row of the image are represented as (x, y) = (0, 1).
Figure 1 shows the coordinate convention used.

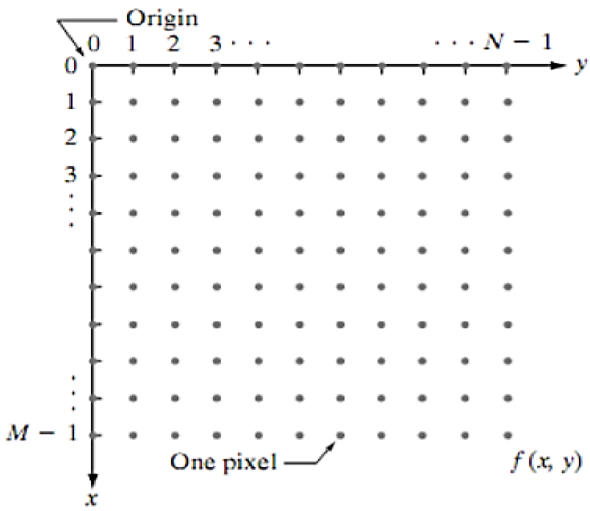


Figure 1.1: General representation of an image pixels

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

Figure 1.2: Image pixel represented as spatial coordinates

1.3 COMPONENTS OF AN IMAGE PROCESSING SYSTEM

Figure 3 shows the fundamental parts DIP framework. To obtain advanced pictures we need an actual gadget which is delicate to the energy transmitted by the article (picture catching interaction). The second, called a digitizer, is a gadget for changing over the yield of the actual detecting gadget into a digital picture. Particular picture handling equipment comprises of the digitizer just referenced, in addition to equipment that performs activities, like a number arithmetic& logic unit (ALU). The computer in a DIP framework is a computer or a supercomputer. Computerized storage for picture preparing applications falls into three head classifications: (1) short-term storage for use during processing, (2) on-line storage for relatively fast recall, and (3) archival storage, characterized by infrequent access. Capacity is estimated in bytes (eight pieces), Kbytes (1,000 bytes), Mbytes (1,000,000 bytes), Gbytes (which means giga, or one billion, bytes), and Tbytes (which means tera, or one trillion, bytes). Online storage is done using CD or optical-discs.

Image displays in use today are mainly color (preferably flat screen) TV monitors. Monitors are driven by the outputs of image and graphics display cards that are an integral part of the computer system. Hardcopy devices for recording images include laser printers, film cameras, heat-sensitive devices, inkjet units, and digital units, such as optical and CDROM disks. Software for image processing consists of specialized modules that perform specific tasks. Sophisticated software packages allow the integration of those modules and general-purpose software commands from at least one computer language.

Because of the large amount of data inherent in image processing applications, the key consideration in image transmission is bandwidth. In dedicated networks, this typically is not a problem, but communications with remote sites via the Internet are not always as efficient. This limitation can be avoided by using fiber optic based transmission and utilization of other broadband technologies.

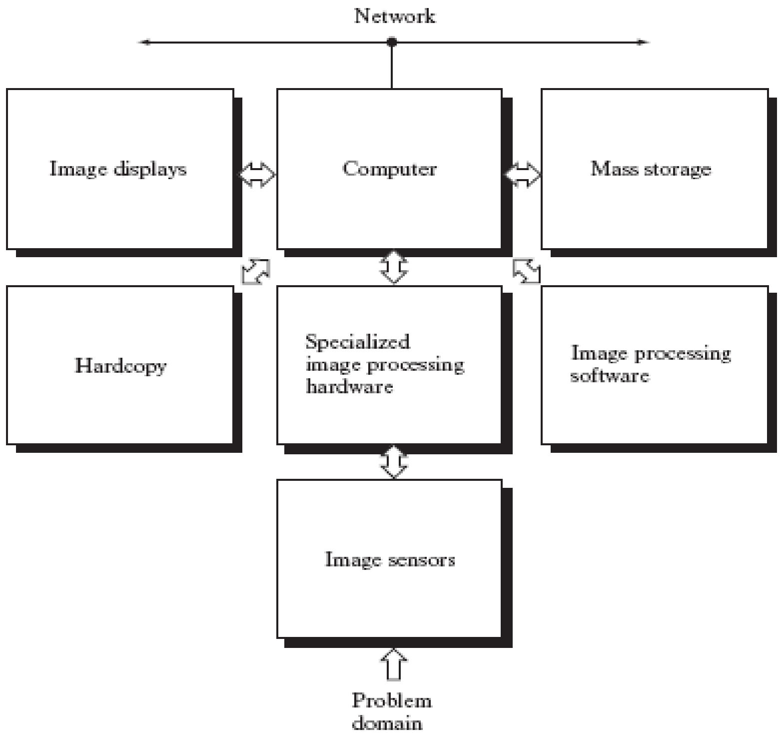


Figure 1.3: Components of a digital Image Processing System

1.3.1 Elements of Visual Perception

Although the field of digital image processing is built on a foundation of mathematical and probabilistic formulations, human intuition and analysis play a central role in the choice of one technique versus another, and this choice often is made based on subjective, visual judgments. Hence, developing a basic understanding of human visual perception is necessary. In particular, our interest is in the mechanics and parameters related to how images are formed and perceived by humans.

1.3.2 Structure of Human eye

Figure 4 shows a simplified horizontal cross section of the human eye. The eye is nearly a sphere, with an average diameter of approximately 20 mm. Three membranes enclose the eye: the cornea and sclera outer cover; the choroid; and the retina. The cornea is a tough, transparent tissue that covers the anterior surface of the eye. Continuous with the cornea, the sclera is an opaque membrane that encloses the remainder of the optic globe. The choroid lies directly below the sclera. This membrane contains a network of blood vessels that serve as the major source of nutrition to the eye. Even superficial injury to the choroid, often not deemed serious, can lead to severe eye damage as a result of inflammation that restricts blood flow. The choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optic globe. At its anterior extreme, the choroid is divided into the ciliary body and the iris. The latter contracts or expands to control the amount of light that enters the eye. The central opening of the iris (the pupil) varies in diameter from approximately 2 to 8 mm. The front of the iris contains the visible pigment of the eye, whereas the back contains a black pigment. The lens is made up of concentric layers of fibrous cells and is suspended by fibers that attach to the ciliary body. It contains 60 to 70% water, about 6% fat, and more protein than any other tissue in the eye. The lens is colored by a slightly yellow pigmentation that increases with age. In extreme cases, excessive clouding of the lens, caused by the affliction commonly referred to as cataracts, can lead to poor color discrimination and loss of clear vision. The lens absorbs approximately 8% of the visible light spectrum, with relatively higher absorption at shorter wavelengths. Both infrared and ultraviolet light are absorbed appreciably by proteins with in the lens structure and, in excessive amounts, can damage the eye.

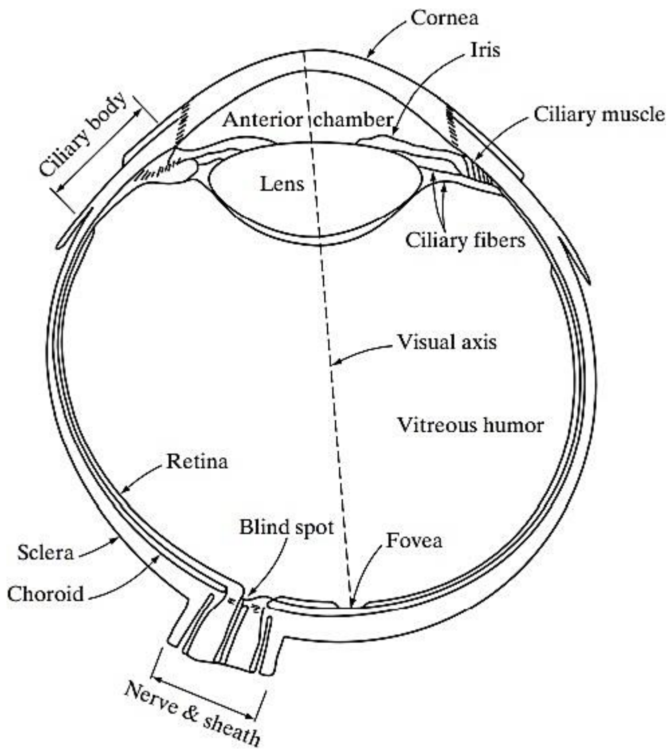


Figure 1.4: Simplified diagram of a cross section of the human eye.

The innermost membrane of the eye is the retina, which lines the inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina. There are two classes of receptors: cones and rods. The cones in each eye number between 6 and 7 million. They are located primarily in the central portion of the retina, called the fovea, and are highly sensitive to color. Humans can resolve fine details with these cones largely because each one is connected to its own nerve end. Muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea. Cone vision is called photopic or bright-light vision. The number of rods is much larger: Some 75 to 150 million are distributed over the retinal surface. The larger area of distribution and the fact that several rods are connected to a single nerve end reduce the amount of detail discernible by these receptors. Rods serve to give a general, overall picture of the field of view. They are not involved in color vision and are sensitive to low levels of illumination. For example, objects that appear brightly colored in daylight when seen by moonlight appear as colorless forms because only the rods are stimulated. This phenomenon is known as scotopic or dim-light vision. Figure 5 shows the density of rods and cones for a cross section of the right eye passing through the region of emergence of the optic nerve from the eye. The absence of receptors in this area results in the so-called blind spot. Except f or this region, the distribution of receptors is radially symmetric about the fovea. Receptor density is measured in degrees from the fovea (that is, in degrees off axis, as measured by the angle formed by the visual axis and a line passing through the center of the lens and intersecting the retina). Note in Figure 5 that cones are most dense in the center of the retina (in the center area of the fovea). Note also that rods increase in density from the center out to approximately 20° off axis and then decrease in density out to the extreme periphery of the retina. The fovea itself is a circular indentation in the retina of about 1.5 mm in diameter.

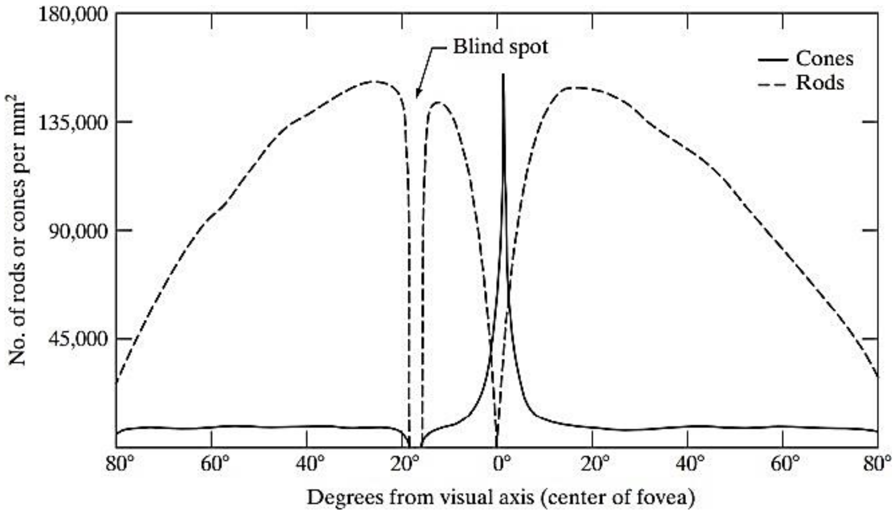


Figure 1.5: Distribution of rods and cones in the retina.

1.4 IMAGE FORMATION IN THE EYE

In an ordinary photographic camera, the lens has a fixed focal length, and focusing at various distances is achieved by varying the distance between the lens and the imaging plane, where the film (or imaging chip in the case of a digital camera) is located. In the human eye, the converse is true; the distance between the lens and the imaging region (the retina) is fixed, and the focal length needed to achieve proper focus is obtained by varying the shape of the lens. The fibers in the ciliary body accomplish this, flattening or thickening the lens for distant or near objects, respectively. The distance between the center of the lens and the retina along the visual axis is approximately 17 mm. The range of focal lengths is approximately 14 mm to 17 mm, the latter taking place when the eye is relaxed and focused at distances greater than about 3 m. The geometry in Figure. 6 illustrates how to obtain the dimensions of an image formed on the retina. For example, suppose that a person is looking at a tree 15 m high at a distance of 100 m. Letting h denote the height of that object in the retinal image, the geometry of Figure 6 yields or As indicated, the retinal image is focused primarily on the region of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that ultimately are decoded by the brain.

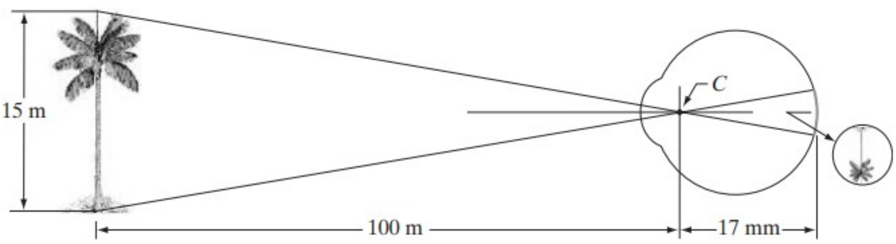


Figure 1.6: Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens

1.4.1 Brightness Adaptation and Discrimination

Because digital images are displayed as a discrete set of intensities, the eye's ability to discriminate between different intensity levels is an important consideration in presenting image processing results. The range of light intensity levels to which the human visual system can adapt is enormous—on the order of — from the scotopic threshold to the glare limit. Experimental evidence indicates that subjective brightness (intensity as perceived by the human visual system) is a logarithmic function of the light intensity incident on the eye.

1.5 IMAGE SENSING AND ACQUISITION

Most of the images in which we are interested are generated by the combination of an "illumination" source and the reflection or absorption of energy from that source by the elements of the "scene" being imaged. We enclose illumination and scene in quotes to emphasize the fact that they are considerably more general than the familiar situation in which a visible light source illuminates a common everyday 3-D (three-dimensional) scene. For example, the illumination may originate from a source of electromagnetic energy such as radar, infrared, or X-ray system. But, as noted earlier, it could originate from less traditional sources, such as ultrasound or even a computer-generated illumination pattern. Similarly, the scene elements could be familiar objects, but they can just as easily be molecules, buried rock formations, or a human brain. Depending on the nature of the source, illumination energy is reflected from, or transmitted through, objects. An example in the first category is light reflected from a planar surface. An example in the second category is when X-rays pass through a patient's body for the purpose of generating a diagnostic X-ray film. In some applications, the reflected or transmitted energy is focused onto a photoconverter (e.g., a phosphor screen), which converts the energy into visible light. Electron microscopy and some applications of gamma imaging use this approach. Figure 7 shows the three principal sensor arrangements used to transform illumination energy into digital images. The idea is simple: Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected. The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response. In this section, we look at the principal modalities for image sensing and generation.

1.5.1 Image Acquisition Using a Single Sensor

Figure 7 (a) shows the components of a single sensor. Perhaps the most familiar sensor of this type is the photodiode, which is constructed of silicon materials and whose output voltage waveform is proportional to light. The use of a filter in front of a sensor improves selectivity. For example, a green (pass) filter in front of a light sensor favors light in the green band of the color spectrum. As a consequence, the sensor output will be stronger f or green light than for other components in the visible spectrum. In order to generate a 2 -D image using a single sensor, there has to be relative displacements in both the x- and y-directions between the sensor and the area to be imaged.

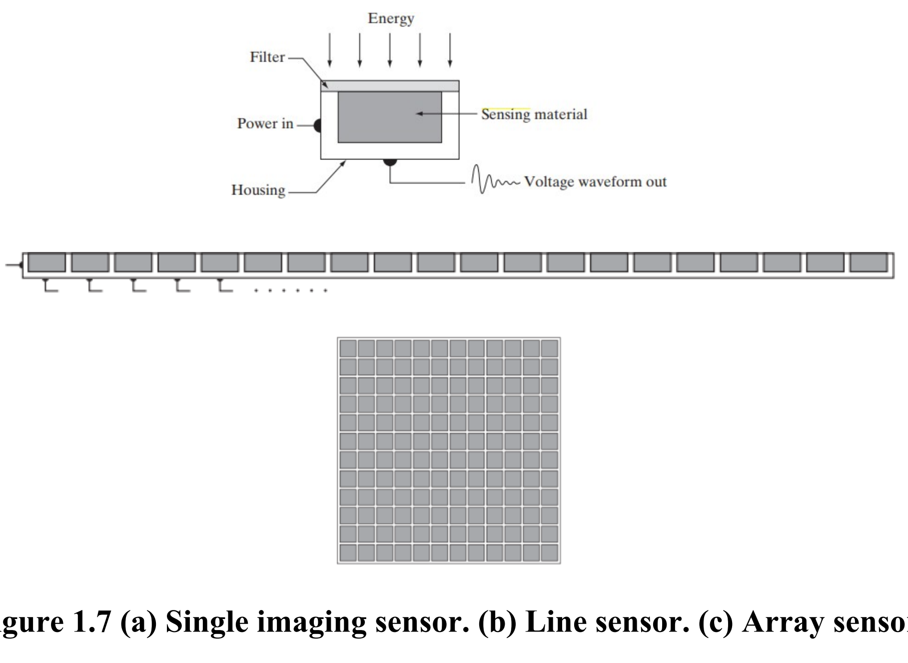


Figure 1.7 (a) Single imaging sensor. (b) Line sensor. (c) Array sensor.

Figure 8 shows an arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Because mechanical motion can be controlled with high precision, this method is an inexpensive (but slow) way to obtain high-resolution images. Other similar mechanical arrangements use a flat bed, with the sensor moving in two linear directions. These types of mechanical digitizers sometimes are referred to as microdensitometers. Another example of imaging with a single sensor places a laser source coincident with the sensor. Moving mirrors are used to control the outgoing beam in a scanning pattern and to direct the reflected laser signal onto the sensor. This arrangement can be used also to acquire images using strip and array sensors, which are discussed in the following two sections.

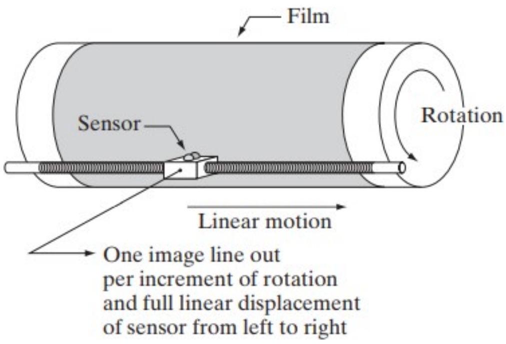


Figure 1.8: Combining a single sensor with motion to generate a 2-D image.

1.5.2 Image Acquisition Using Sensor Strips

A geometry that is used much more frequently than single sensors consist of an in-line arrangement of sensors in the form of a sensor strip, as Figure 7(b) shows. The strip provides imaging elements in one direction Motion perpendicular to the strip provides imaging in the other direction, as shown in Figure 9(a). This is the type of arrangement used in most flatbed scanners. Sensing devices with 4000 or more in-line sensors are possible. In-line sensors are used routinely in airborne imaging applications, in which the imaging system is mounted on an aircraft that flies at a constant altitude and speed over the geographical area to be imaged. One-dimensional imaging sensor strips that respond to various bands of the electromagnetic spectrum are mounted perpendicular to the direction of flight. The imaging strip gives one line of an image at a time, and the motion of the strip completes the other dimension of a two- dimensional image. Lenses or other focusing schemes are used to project the area to be scanned onto the sensors. Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional ("slice") images of 3-D objects, as Figure 9(b) shows. A rotating X-ray source provides illumination and the sensors opposite the source collect the X-ray energy that passes through the object (the sensors obviously have to be sensitive to X-ray energy). This is the basis for medical and industrial computerized axial tomography (CAT) imaging.

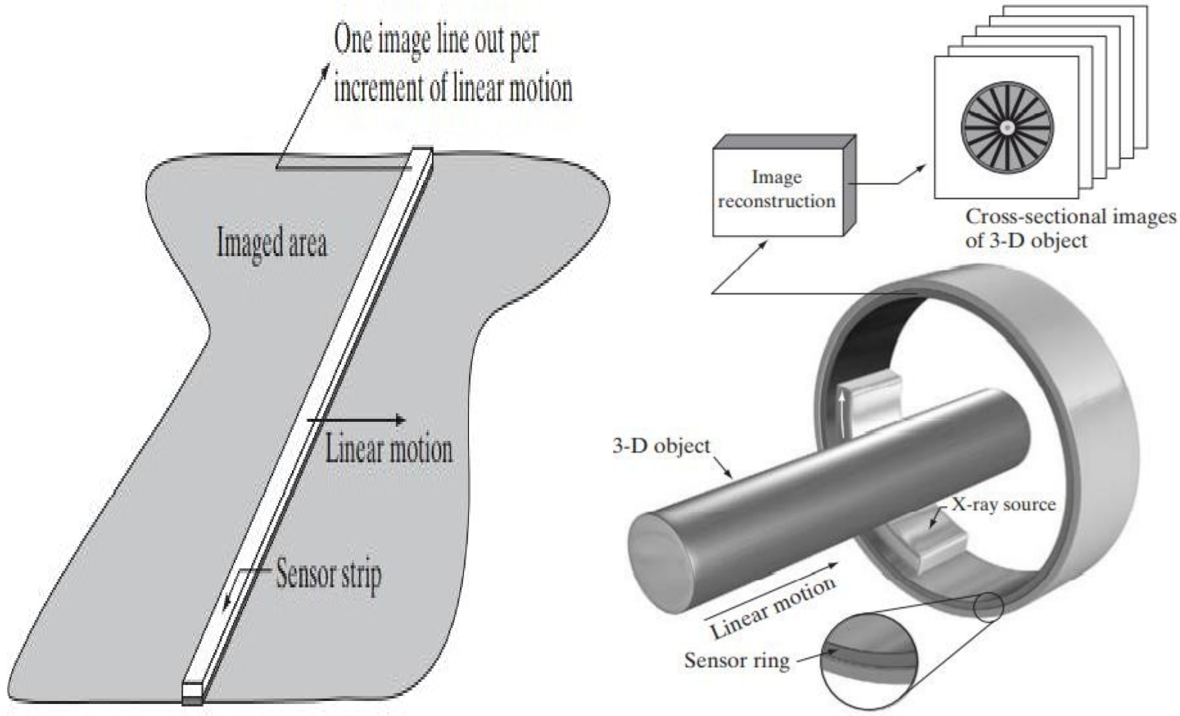


Figure 1.9 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

1.5.3 Image Acquisition Using Sensor Arrays

Figure 10 (c) shows individual sensors arranged in the form of a 2-D array. Numerous electromagnetic and some ultrasonic sensing devices frequently are arranged in an array format. This is also the predominant arrangement found in digital cameras. A typical sensor for these cameras is a CCD array, which can be manufactured with a broad range of sensing properties and can be packaged in rugged arrays of elements or more. CCD sensors are used widely in digital cameras and other light sensing instruments. The response of each sensor is proportional to the integral of the light energy projected onto the surface of the sensor, a property that is used in astronomical and other applications requiring low noise images. Noise reduction is achieved by letting the sensor integrate the input light signal over minutes or even hours. Because the sensor array in Figure 10 (c) is two-dimensional, its key advantage is that a complete image can be obtained by focusing the energy pattern onto the surface of the array. Motion obviously is not necessary, as is the case with the sensor arrangements discussed in the preceding two sections. The principal manner in which array sensors are used is shown in Figure 10. This figure shows the energy from an illumination source being reflected from a scene element (as mentioned at the beginning of this section, the energy also could be transmitted through the scene elements). The first function performed by the imaging system in Figure 10 (c) is to collect the incoming energy and focus it onto an image plane. If the illumination is light, the front end of the imaging system is an optical lens that projects the viewed scene onto the lens focal plane, as Figure 10(d) shows. The sensor array, which is coincident with the focal plane, produces outputs proportional to the integral of the light received at each sensor. Digital and analog circuitry sweep these outputs and convert them to an analog signal, which is then digitized by another section of the imaging system. The output is a digital image, as shown diagrammatically in Figure 10 (e).

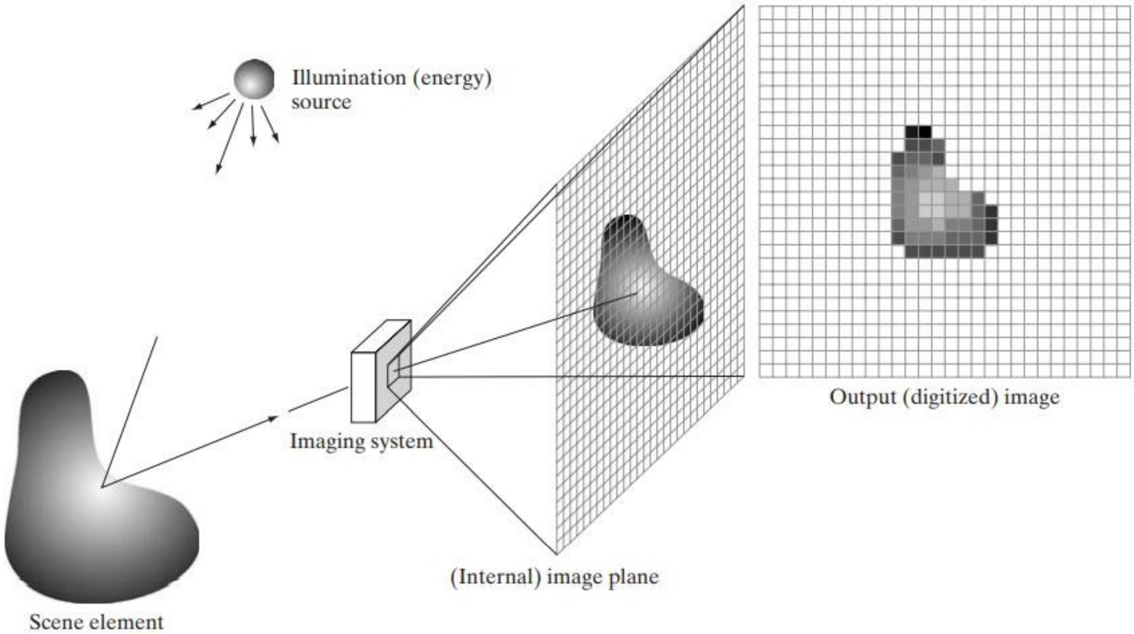


Figure 1.10: An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image

1.6 IMAGE SAMPLING AND QUANTIZATION

The basic idea behind sampling and quantization is illustrated in Figure 11. Figure 11(a) shows a continuous image f that we want to convert to digital form. An image may be continuous with respect to the x- and y-coordinates, and also in amplitude. To convert it to digital form, we have to sample the function in both coordinates and in amplitude. Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization. The one-dimensional function in Figure 1.11 (b) is a plot of amplitude (intensity level) values of the continuous image along the line segment AB in Figure 11 (a). The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in Figure 11(c). The spatial location of each sample is indicated by a vertical tick mark in the bottom part of the figure. The samples are shown as small white squares superimposed on the function. The set of these discrete locations gives the sampled function. However, the values of the samples still span (vertically) a continuous range of intensity values. In order to form a digital function, the intensity values also must be converted (quantized) into discrete quantities. The right side of Figure 11 (c) shows the intensity scale divided into eight discrete intervals, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight intensity intervals. The continuous intensity levels are quantized by assigning one of the eight values to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown in Figure 1.11 (d). Starting at the top of the image and carrying out this procedure line by line produces a two-dimensional digital image. It is implied in Figure 1.11 that, in addition to the number of discrete levels used, the accuracy achieved in quantization is highly dependent on the noise content of the sampled signal. Sampling in the manner just described assumes that we have a continuous image in both coordinate directions as well as in amplitude. In practice, the method of sampling is determined by the sensor arrangement used to generate the image.

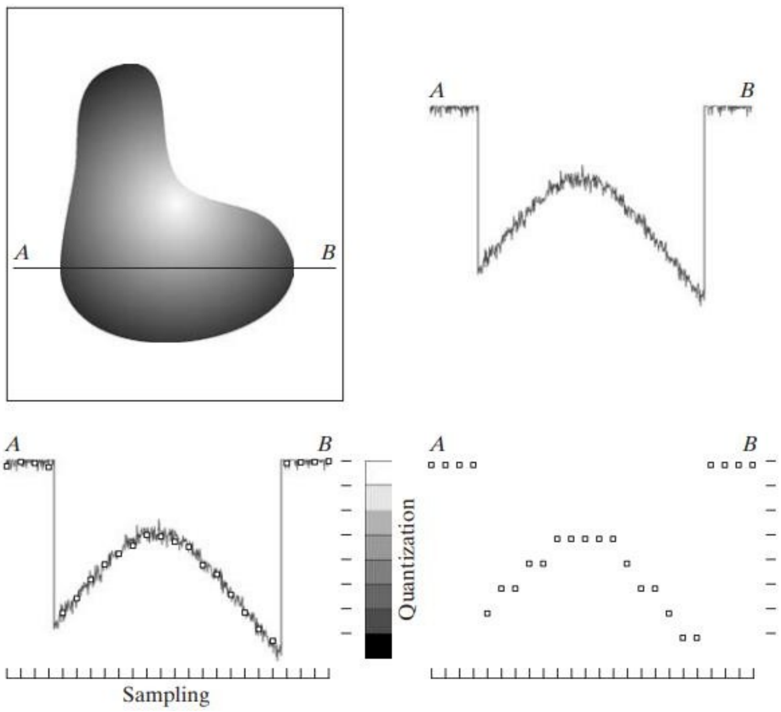


Figure 1.11: Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

When an image is generated by a single sensing element combined with mechanical motion, as in Figure 7, the output of the sensor is quantized in the manner described above. However , spatial sampling is accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data. Mechanical motion can be made very exact so, in principle, there is almost no limit as to how fine we can sample an image using this approach. In practice, limits on sampling accuracy are determined by other factors, such as the quality of the optical components of the system. When a sensing strip is used f or image acquisition, the number of sensors in the strip establishes the sampling limitations in one image direction. Mechanical motion in the other direction can be controlled more accurately, but it makes little sense to try to achieve sampling density in one direction that exceeds the sampling limits established by the number of sensors in the other. Quantization of the sensor outputs completes the process of generating a digital image. When a sensing array is used for image acquisition, there is no motion and the number of sensors in the array establishes the limits of sampling in both directions. Quantization of the sensor outputs is as before. Figure 1.12 illustrates this concept. Figure 1.12(a) shows a continuous image projected onto the plane of an array sensor. Figure 1.12(b) shows the image after sampling and quantization. Clearly, the quality of a digital image is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization.

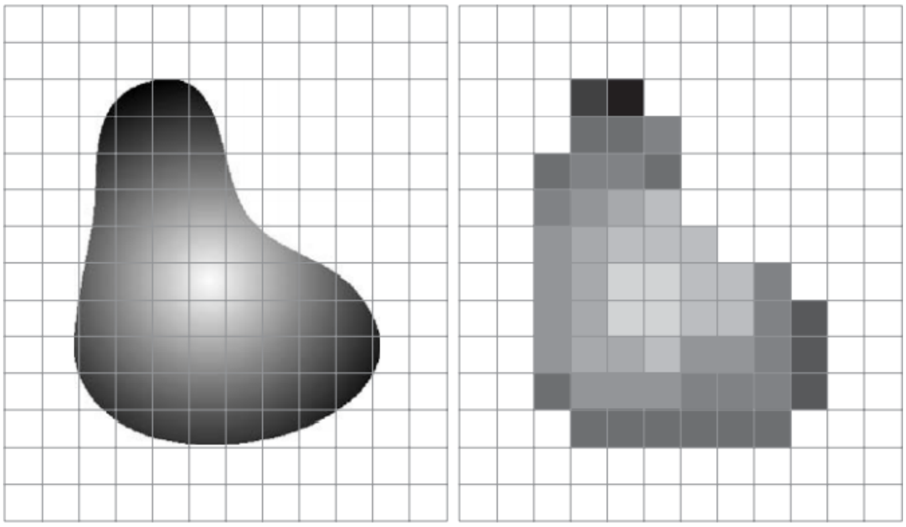


Figure 1.12 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

1.7 BASIC RELATIONSHIPS BETWEEN PIXELS NEIGHBORS OF A PIXEL

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

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

This set of pixels, called the 4-neighbors of p, is denoted by N4(p).

Each pixel is a unit distance from (x, y), and some of the neighbor locations of p lie outside the digital image if (x, y) is on the border of the image.

The four diagonal neighbors of p have coordinates

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

and are denoted by ND(p). These points, together with the 4-neighbors, are called the 8-neighbors of p, denoted by N8(p).

Adjacency, Connectivity, Regions and Boundaries

Let V be the set of intensity values used to define adjacency. In a binary image, V = {1} if we are referring to adjacency of pixels with value 1. In a gray-scale image, the idea is the same, but set V typically contains more elements. For example, in the adjacency of pixels with a range of possible intensity values 0 to 255, set V could be any subset of these 256 values. We consider three types of adjacency:

(a) 4-adjacency. Two pixels p and q with values from V are 4-adjacent if q is in the set

(b) 8-adjacency. Two pixels p and q with values from V are 8-adjacent if q is in the set

(c) m-adjacency (mixed adjacency). Two pixels p and q with values from V are m-adjacent if

i) q is in or

ii) q is in ND(p) and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

Mixed adjacency is a modification of 8-adjacency. It is introduced to eliminate the ambiguities that often arise when 8-adjacency is used. A (digital) path (or curve) from pixel p with coordinates (x, y) to pixel q with coordinates is a sequence of distinct pixels with coordinates

(x0, y0), (x1, y1), … , (xn, yn)

where (x0, y0) = (x, y), (xn, yn) = (s, t) and pixels (xi, yi) and (xi-1, yi-1) are adjacent f or 1<= i <= n. In this case, n is the length of the path. . If (x0, y0) = (xn, yn) the path is a closed path

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

Let R be a subset of pixels in an image. We call R a region of the image if R is a connected set. Two regions, and are said to be adjacent if their union forms a connected set. Regions that are not adjacent are said to be disjoint. We consider 4- and 8-adjacency when referring to regions.

Distance Measures

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

i. D(p, q) >= 0 (D(p, q) = 0 iff p =q),

ii. D(p, q) = D(q, p) and

iii. D(p, z) < = D(p, q) + D(q, z).

The Euclidean distance between p and q is defined as

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

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

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

The distance (called the city-block distance) between p and q is defined as

In this case, the pixels having a D4 distance from (x, y) less than or equal to some value r form a diamond centered at (x, y). For example, the pixels with D4 distance <=2 from (x, y) (the center point) form the following contours of constant distance:

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

The pixels with D4=1 are the 4-neighbors of (x, y).

The distance (called the chessboard distance) between p and q is defined as

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

In this case, the pixels with D8 distance from (x, y) less than or equal to some value r form a square centered at (x, y). For example, the pixels with D8 distance <= 2 from (x, y) (the center point) form the following contours of constant distance:

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

The pixels with D8=1 distance are the 8-neighbors of D (x, y).

Question Bank:

Part- A

1. Justify the need for Image Processing
2. Illustrate the concept of 4-neighbor pixel relationship
3. List the components of Image processing system
4. Comment on the need for image sampling
5. Illustrate how 2D image are acquired
6. List the components of human visual system

Part- B

1. Explain the Image sampling and Quantization process in detail
2. Enumerate how image is being sensed by Human visual perception system
3. Appraise on the various Image sensing procedures
4. Summarize on the concepts of relationships available between the pixels present in an image

Citation: All images shown in this unit were adopted from “Rafael C Gonzalez, Richard E Woods, "Digital Image Processing", 3th Edition, Pearson Education.

TEXT / REFERENCES BOOKS

1. Rafael C Gonzalez, Richard E Woods, "Digital Image Processing", 4th Edition, Pearson Education, 2017.
2. Milan Sonka, Vaclav Hlavac, Roger Boyle, "Image Processing Analysis and Machine Vision", 4th Edition, Cengage Learning, 4th Edition, 2014.
3. B. Chanda ,D.Dutta Majumdar , "Digital Image Processing and Applications", Prentice Hall of India, 2011.
4. William K Pratt, "Digital Image Processing", 4th Edition, John Willey 2007.
5. Rangaraj M. Rangayyan, “Biomedical Image Analysis”, CRC Press LLC, Boca Raton, FL, 2005.
6. Jain A.K., "Fundamentals of Digital Image Processing", Pearson Education, 1989.

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 ELECTRICAL AND ELECTRONICS ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIT –II – IMAGE ENHANCEMENT AND RESTORATION – SECA 3009

UNIT II

IMAGE ENHANCEMENT AND RESTORATION

Enhancement by Point Processing, Histogram Processing, Arithmetic/Logic Operations, Image Averaging, Spatial Filters for Smoothing and Sharpening, Frequency domain filters for Smoothing and Sharpening; Image Degradation & Restoration Model, Noise Models, Inverse Filtering, Geometric Mean Filter.

2.1 POINT PROCESSING

Image enhancement is to process the given image such that the result is more suitable to process than the original image. It sharpens the image features such as edges, boundaries or contrast the image for better clarity. It does not increase the inherent information content of the data, but increase the dynamic range of feature chosen. The main drawback of image enhancement is quantifying the criterion for enhancement and therefore large number of image enhancement techniques is empirical and require interactive procedure to obtain satisfactory results. Point Processing is the image enhancement at any point in an image depends only on the gray level at that point. Some of the basic intensity transformation functions are

i) Linear Functions:

- Identity Transformation
- Negative Transformation

ii) Logarithmic Functions:

- Log Transformation
- Inverse-log Transformation

iii) Power-Law Functions:

- nth power transformation
- nth root transformation

iv) Piecewise Transformation function

- Contrast Stretching
- Gray-level Slicing
- Bit-plane slicing

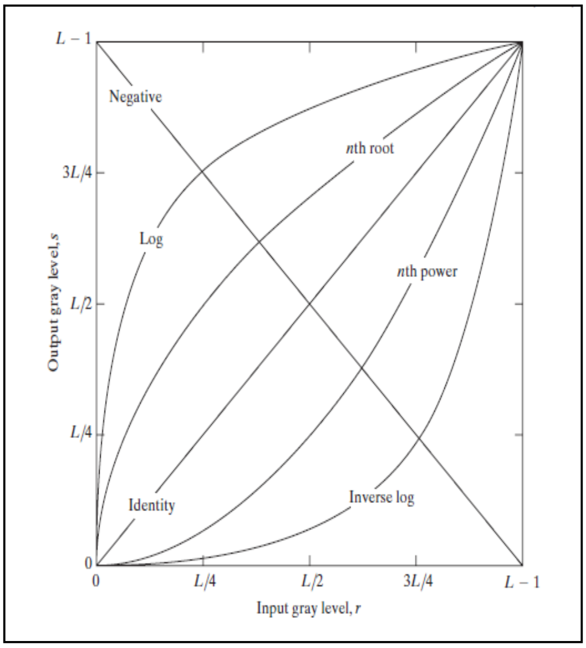


Figure 2.1 Transformation Functions

Linear transformation: Linear transformation includes

- simple identity and
- negative transformation
- Identity transition is shown by a straight line
- In this transition, each value of the input image is directly mapped to each other value of output image. That results in the same input image and output image. Hence is called identity transformation output gray

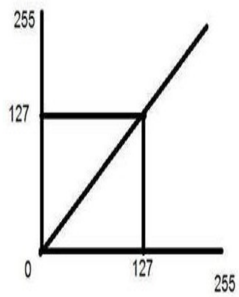
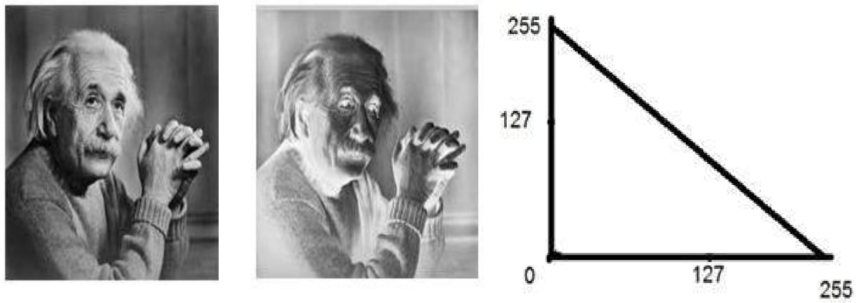


Figure 2.2 Identity Transformation Function

Negative transformation

The negative of an image with gray level in the range **[0, L-1]** is obtained by using the negative transformation, the expression is given by **s = L-1-r**

- Reversing the intensity level of an image produces the equivalent of photographic negative
- Suitable for enhancing white or gray detail embedded in dark regions of an image, when the black areas are dominant in size



Output gray level Input image Output image

Fig. 2.3 Negative Transformation

Input gray level: Logarithmic transformations

The log transformations can be defined by $s = c \log(r + 1)$ where c is a constant and $r \geq 0$. During log transformation, the dark pixels in an image are expanded and the higher pixel values are compressed. The inverse log transform is opposite to log transform. Log transforms has the important characteristics: it compresses the dynamic range of images with large variation in pixel values.

Input Gray Level Input image

Fig. 2.4 Logarithmic Transformation

Output image: Power – Law transformations

This includes nth power and nth root transformations. It is given by the expression: $s=c\, r^{\gamma}$ (or) $s=c(r+\varepsilon)\gamma$ where γ is called gamma, due to which this transformation is also known as gamma transformation. The exponent in the power law equation is referred to as gamma, the process used to correct this power-law response phenomenon is called gamma correction.

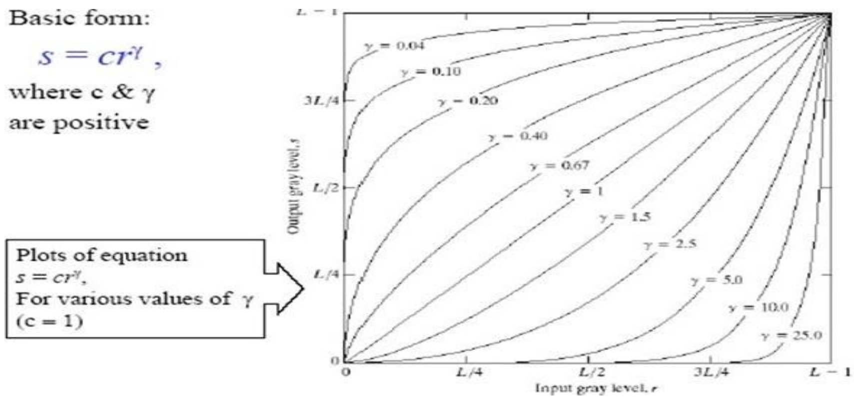


Fig. 2.5 Power Law Transformation

Piecewise- Linear Transformation

One of the simplest piecewise linear functions is a contrast-stretching transformation, which is used to enhance the low contrast images. Low contrast images may result from poor illumination and wrong setting of lens aperture during image acquisition.

Contrast stretching

Figure 2.5 shows a typical transformation used for contrast stretching. The locations of points (r1, s1) and (r2, s2) control the shape of the transformation function.

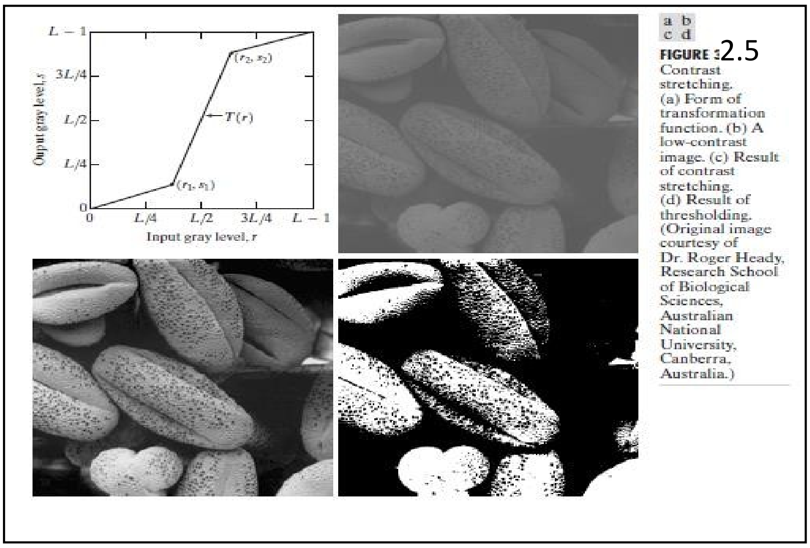


Fig. 2.6 Contrast Stretching

If r1 = s1 and r2 = s2, the transformation is a linear function that produces no changes in gray levels. If r1 = r2, s1 = 0 and s2 = L-1, the transformation becomes a thresholding function that creates a binary image. Intermediate values of (r1, s1) and (r2, s2) produce various degrees of spread in the gray levels of the output image, thus affecting its contrast. In general, r1 ≤ r2 and s1 ≤ s2 is assumed, so the function is always increasing. Figure (b) shows an 8-bit image with low contrast. Fig. (c) shows the result of contrast stretching, obtained by setting (r1, s1) = (r_{min}, 0) and (r2, s2) = (r_{max},L-1) where r_{min} and r_{max} denote the minimum and maximum gray levels in the image, respectively. Thus, the transformation function stretched the levels linearly from their original range to the full range [0, L-1]. Finally, Fig. (d) shows the result of using the thresholding function defined previously, with r1=r2=m, the mean gray level in the image.

Gray-level Slicing: This technique is used to highlight a specific range of gray levels. It can be implemented in several ways, but the two basic themes are:One approach is to display a high value for all gray levels in the range of interest and a low value for all other gray levels. This transformation, shown in Fig.2.6 (a), produces a binary image. The second approach, based on the transformation shown in Fig. (b), this brightens the desired range of gray levels but preserves gray levels unchanged Fig.(c) shows a gray scale image, and fig.(d) shows the result of using the transformation in Fig.(a).

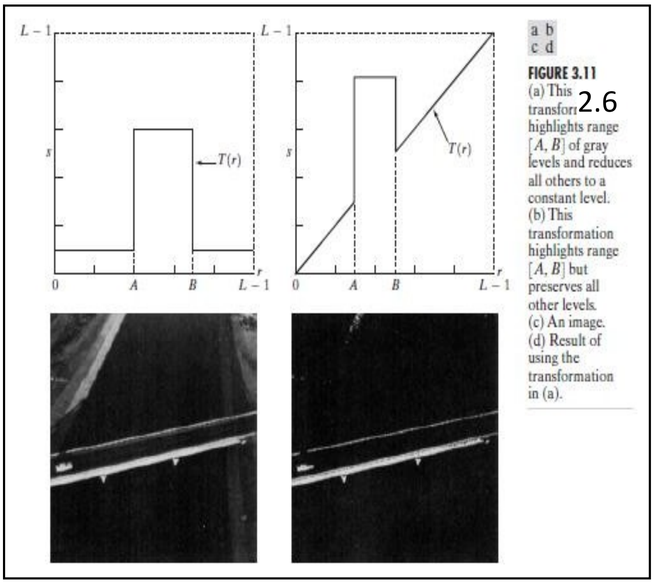


Fig. 2.7 Gray level slicing

Bit-plane Slicing

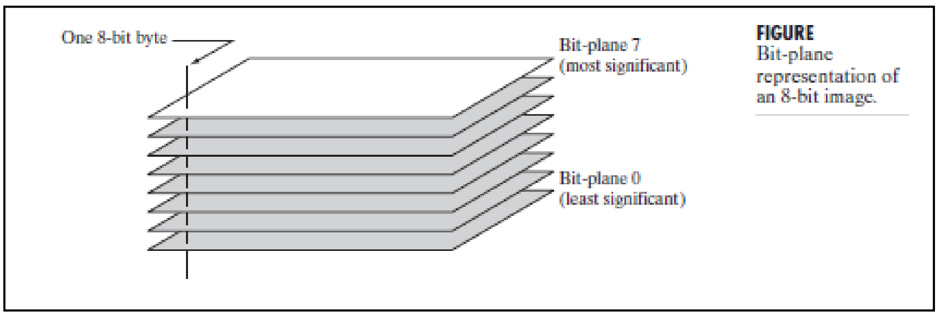


Fig. 2.8 Bit Plane Slicing

Pixels are digital numbers, each one composed of bits. Instead of highlighting gray-level range, we could highlight the contribution made by each bit. This method is useful and used in image compression. Most significant bits contain the majority of visually significant data.

Fig. 2.9 Example for bit plane slicing

2.2 ARITHMETIC/LOGIC OPERATIONS

Image arithmetic applies one of the standard arithmetic operations or a logical operator to two or more images. The operators are applied in a pixel-by-pixel way, i.e. the value of a pixel in the output image depends only on the values of the corresponding pixels in the input images. Hence, the images must be of the same size. Although image arithmetic is the most simple form of image processing, there is a wide range of applications.

Logical operators are often used to combine two (mostly binary) images. In the case of integer images, the logical operator is normally applied in a bitwise way.

$$s(x, y) = f(x, y) + g(x, y)$$
$$d(x, y) = f(x, y) - g(x, y)$$
$$p(x, y) = f(x, y) \times g(x, y)$$
$$v(x, y) = f(x, y) \div g(x, y)$$

Arithmetic/logical operations are performed on pixel-by-pixel basis based on two or more images. When dealing with logical operations on gray-scale images, pixel values are processed as strings of binary numbers. In AND and OR image masks, light represents a binary 1 and dark represents a binary 0. Masking refers to as Region of Interest (ROI) processing.

Image Subtraction

- Enhancement of differences between images
- Key usefulness of subtraction is the enhancement of differences between images.

- If the difference in the pixel value is small, then the image appears black when displayed in 8-bit display.
- To bring more detail contrast stretching can be performed

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

Image Averaging

- Noisy image g(x,y) formed by the addition of noise

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

- Averaging K different noisy images η(x,y) to an original image f(x,y)
- Objective is to reduce the noise content by adding a set of noisy images $\{g_i(x,y)\}$

$$\boxed{\bar{g}(x, y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x, y)} \qquad \boxed{E\{\bar{g}(x, y)\} = f(x, y)}$$

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

$E\{\bar{g}(x, y)\}$ → Expected value of $\bar{g}$

2.3 HISTOGRAM PROCESSING

- Histogram is a graphical representation showing a visual impression ofthe distribution of data
- An Image Histogram is a type of histogram that acts as a graphical representation of the lightness/color distribution in a digital image
- It plots the number of pixels for each value
- The histogram of a digital image with gray levels in the range [0, L-1] is a discrete

  function $h(r_k) = n_k$ where r_k is the kth gray level and n_k is the number of pixels in the image having gray level r_k

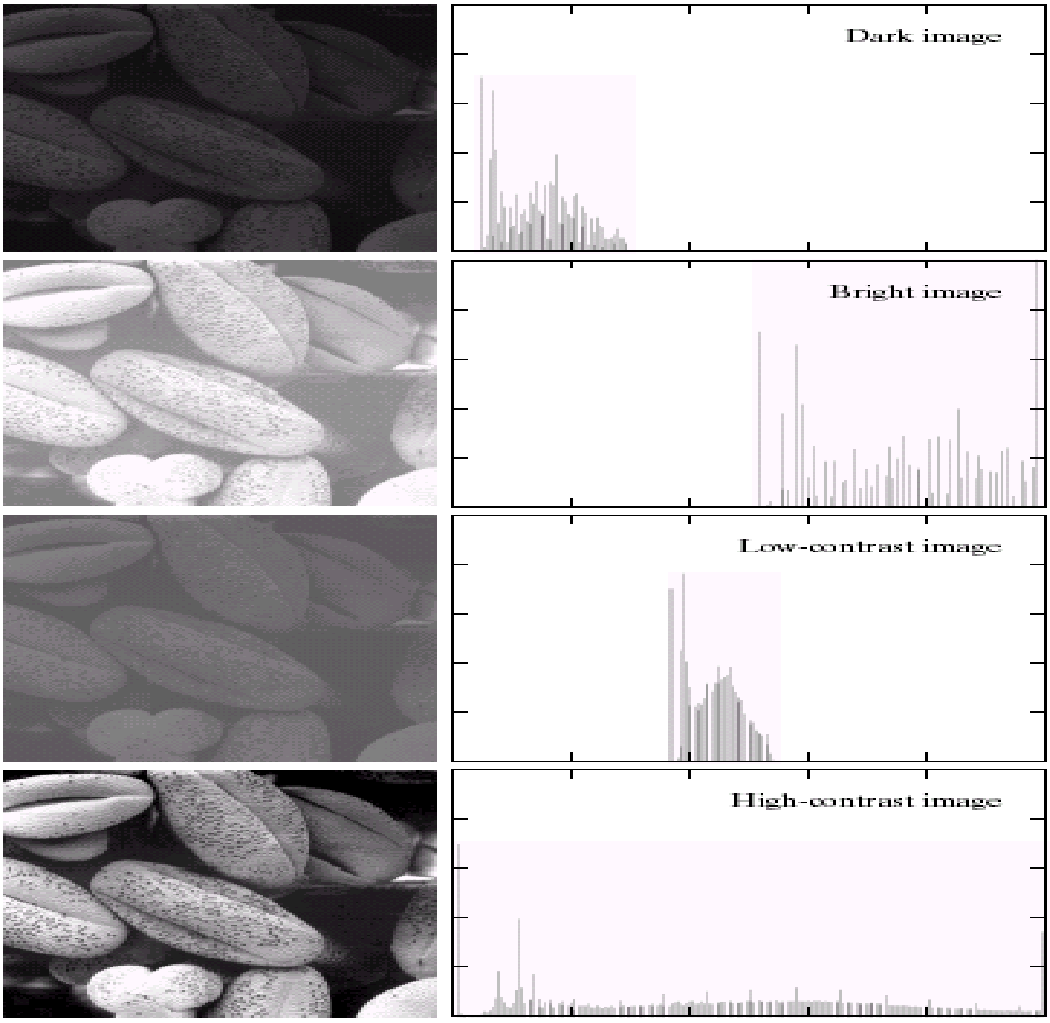


Fig. 2.10 Histogram of different types of images

It is common practice to normalize a histogram by dividing each of its values by the total number of pixels in the image, denoted by n.

A normalized histogram is given by

$$p(r_k) = n_k / n \text{ for } k = 0, 1, \ldots, L - 1$$

Thus, $p(r_k)$ gives an estimate of the probability of occurrence of gray level r_k

<u>Note:</u> The sum of all components of a normalized histogram is equal to 1

Histogram Equalisation

The gray levels in an image is viewed as random variables in the interval[0,1] Fundamental descriptors of a random variables is its probability density function $p_s(s)$ and $p_s(r)$ are PDF of s and r

$$\pi\sigma(\sigma) = \pi\rho(\rho)$$

Transformation has the particular importance in image processing

$$\sigma = T(\rho) = \int_{0}^{\rho}$$

$$\pi\rho\,(\omega)\delta\omega$$

Discrete version of transformation- histogram equalization or histogram linearization

$$\sigma\kappa \frac{n_j}{n} = T(\rho\kappa) = \sum_{j=0}^{k}$$

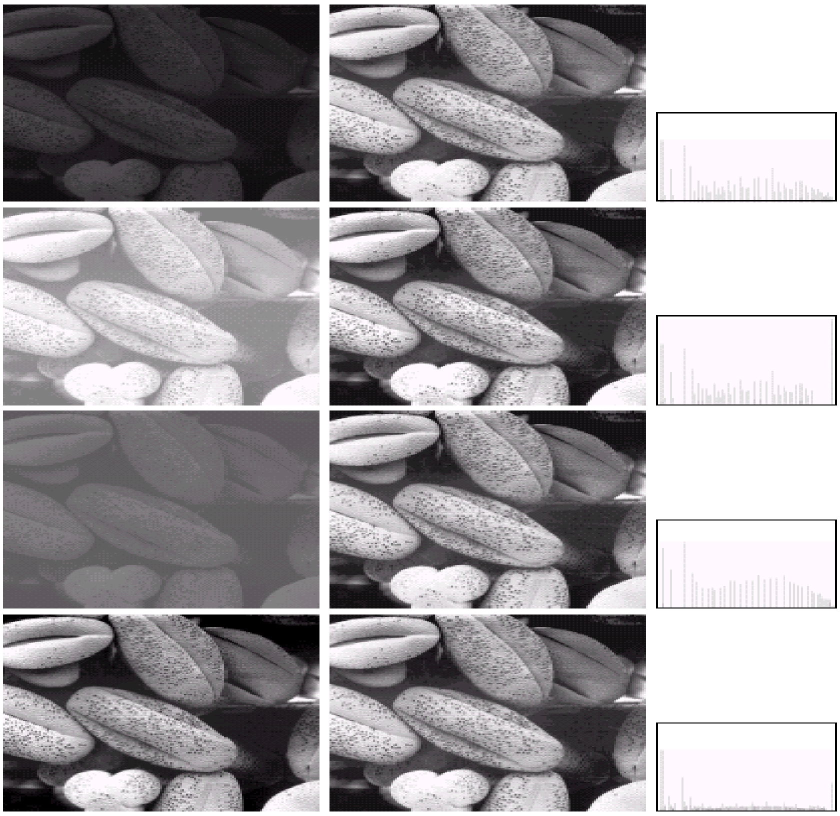


Fig. 2.11 Images after Histogram equalization

Assume the images have 64 x 64 = 4096 pixels in 8 gray levels. The following table shows the equalization process

Table 2. 1: sample Equalization procedure

Original Image Gray Level	No. of pixels (frequency)	Probability	Cumulative Probability	Multiply by Max. Gray Level	Rounding
0	790	0.19	0.19	1.33	1
1	1023	0.25	0.44	3.08	3
Original Image Gray Level	**No.of pixels (frequency)**	**Probability**	**Cumulative Probability**	**Multiply by Max. Gray Level**	**Rounding**
2	850	0.21	0.65	4.55	4
3	656	0.16	0.81	5.67	5
4	329	0.08	0.89	6.23	6
5	245	0.06	0.95	6.65	6
6	122	0.03	0.98	6.86	6
7	81	0.02	1.00	7	7

- r is in the range with representing black and representing white

- For r satisfying these conditions, we focus attention on transformations (intensity mappings) of the form

Table 2.2: Intensity Distribution table

r_k	n_k	$p_r(r_k) = n_k/MN$
$r_0 = 0$	790	0.19
$r_1 = 1$	1023	0.25
$r_2 = 2$	850	0.21
$r_3 = 3$	656	0.16
$r_4 = 4$	329	0.08
$r_5 = 5$	245	0.06
$r_6 = 6$	122	0.03
$r_7 = 7$	81	0.02

$$s = T(r) \quad 0 \le r \le L - 1$$

that produce an output intensity level s for every pixel in the input image having intensity r. We assume that:

(a) $T(r)$ is a monotonically† increasing function in the interval $0 \le r \le L - 1$; and

(b) $0 \le T(r) \le L - 1$ for $0 \le r \le L - 1$.

In some formulations to be discussed later, we use the inverse

$$r = T^{-1}(s) \quad 0 \le s \le L - 1$$

in which case we change condition (a) to

(a′) $T(r)$ is a strictly monotonically increasing function in the interval $0 \le r \le L - 1$.

Histogram Matching (Specification)

Procedure for histogram matching:

- Obtain histogram of given image
- Use the equation to pre compute a mapped level s_k for each level r_k

$$sk = Tr(k) = \sum_{j=0}^{k} pr(rj) = \sum_{j=0}^{k} nj / n$$

- Obtain the transformation function G from the given $p_z(z)$ using the equation

$$(G(\hat{z}) - sk) \ge 0; k = 0,1,2 \qquad L - 1$$

- Precompute z_k for each value of s_k using the iterative scheme defined in connection with equation

$$(G(\hat{z}) - sk) \ge 0; k = 0,1,2 \qquad L - 1$$

- For each pixel in the original image, if the value of that pixel is r_k, map this value to the corresponding level s_k; then map levels s_k into the final level z_k

Processed image that has a specified histogram is called histogram matching or histogram specification.

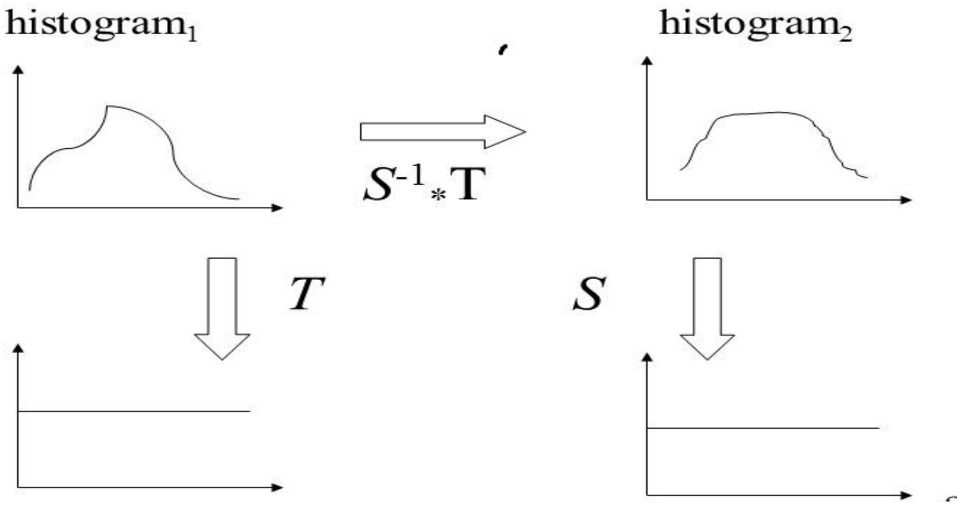


Fig. 2.12 Histogram Specification

2.4 SPATIAL FILTERING

- The output intensity value at (x,y) depends not only on the input intensity value at (x,y) but also on the specified number of neighboring intensity values around (x,y)
- Spatial masks (also called window, filter, kernel, template) are used and convolved over the entire image for local enhancement (spatial filtering)
- The size of the masks determines the number of neighboring pixels which influence the output value at (x,y)
- The values (coefficients) of the mask determine the nature and properties of enhancing technique.
- The mechanics of spatial filtering
- For an image of size *M* x *N* and a mask of size *m* x *n*
- The resulting output gray level for any coordinates x and y is given by

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

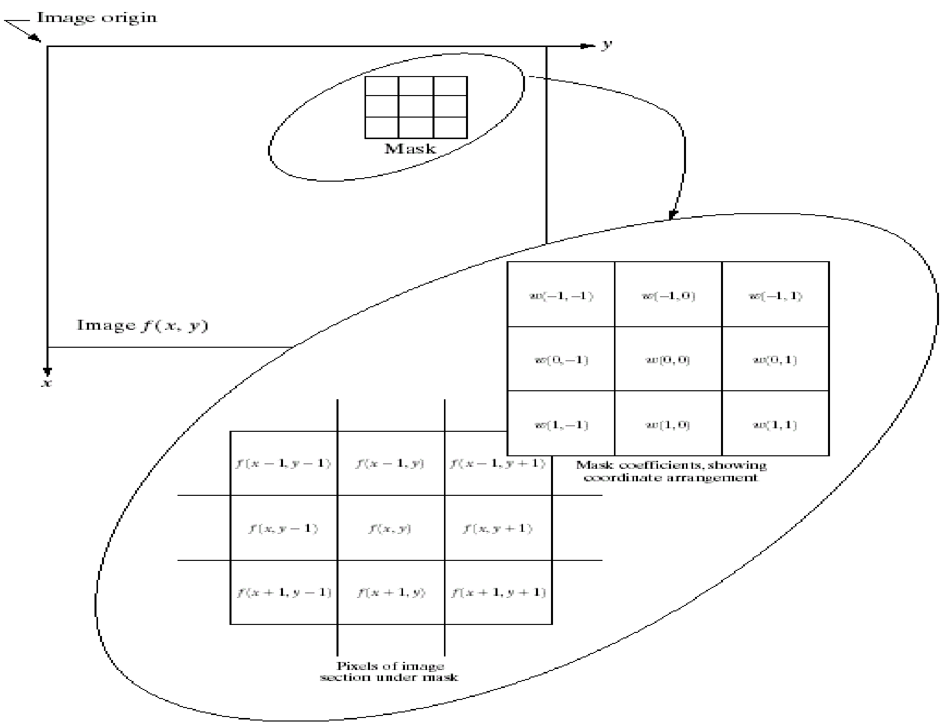
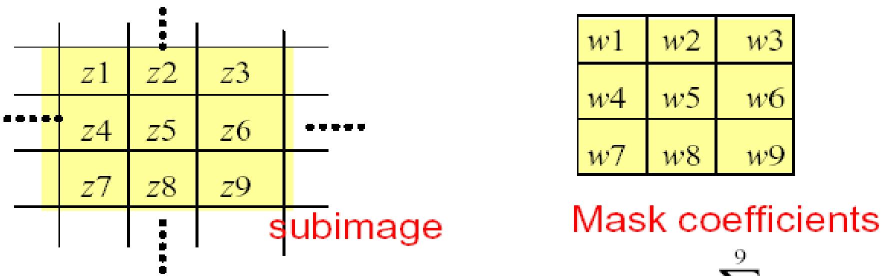


Fig. 2.13 Filter Window

Given the 3×3 mask with coefficients: $w_1, w_2, \ldots, w_9$

The mask cover the pixels with gray levels: $z_1, z_2, \ldots, z_9$



$$z_5 \leftarrow z = w_1 z_1 + w_2 z_2 + \cdots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

Fig. 2.14 Filter Window procedure

z gives the output intensity value for the processed image (to be stored in a new array) at the location of z_5 in the input image

Mask operation near the image border

Problem arises when part of the mask is located outside the image plane; to handle the problem:

- Discard the problem pixels (e.g. $512\text{x}512_{\text{input}}$ $510\text{x}510_{\text{output}}$ if mask size is 3x3)

- Zero padding: expand the input image by padding zeros ($512x512_{input}$ $514x514_{output}$)
- Zero padding is not good - create artificial lines or edges on the border
- We normally use the gray levels of border pixels to fill up the expanded region (for 3x3 mask). For larger masks a border region equal to half of the mask size is mirrored on the expanded region.

Spatial Filtering for Smoothing

- For blurring/noise reduction;
- Smoothing/Blurring is usually used in preprocessing steps,

e.g., to remove small details from an image prior to object extraction, or to bridge small gaps in lines or curves

- Equivalent to Low-pass spatial filtering in frequency domain because smaller (high frequency) details are removed based on neighborhood averaging (averaging filters)

  Implementation: The simplest form of the spatial filter for averaging is a square mask (assume $m \times m$ mask) with the same coefficients 1/m2 to preserve the gray levels (averaging).

  Applications: Reduce noise; smooth false contours Side effect: Edge blurring

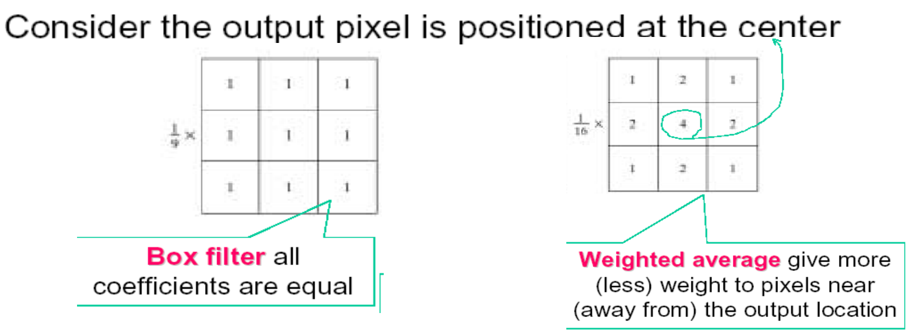


Fig. 2.15 Edge Blurring

Spatial Filtering for Sharpening

Background: to highlight fine detail in an image or to enhance blurred detail

Applications: electronic printing, medical imaging, industrial inspection, autonomous target detection (smart weapons)

Foundation:

- Blurring/smoothing is performed by spatial averaging (equivalent to integration)
- Sharpening is performed by noting only the gray level changes in the image that is the differentiation

Operation of Image Differentiation

- Enhance edges and discontinuities (magnitude of output gray level >>0)
- De-emphasize areas with slowly varying gray-level values (output gray level: 0)

  Mathematical Basis of Filtering for Image Sharpening

- First-order and second-order derivatives
- Approximation in discrete-space domain
- Implementation by mask filtering

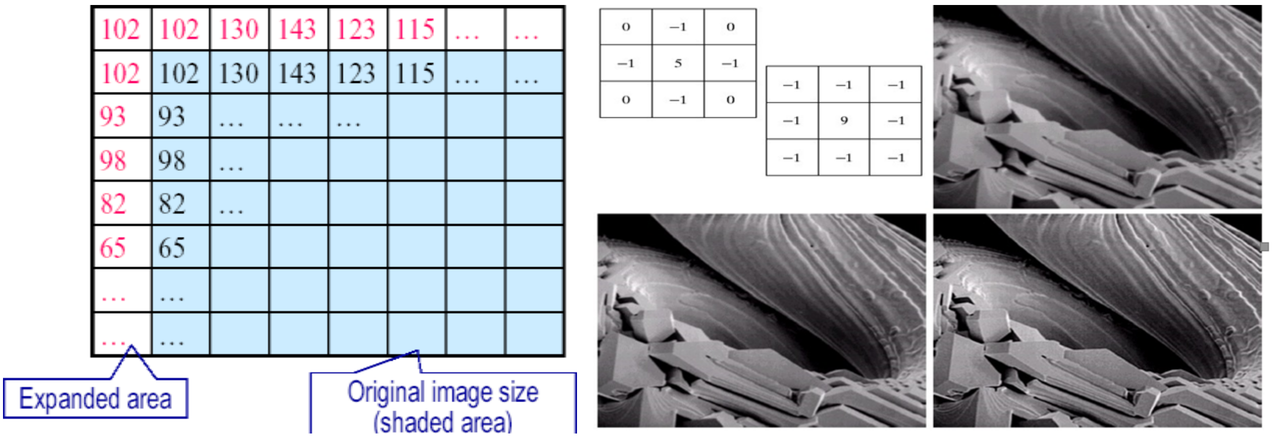


Fig. 2.16 Spatial Sharpening

2.5 FREQUENCY DOMAIN FILTERS

- Any function that periodically repeats itself can be expressed as the sum of sines and/or cosines of different frequencies, each multiplied by a different coefficient (Fourier series).
- Even functions that are not periodic (but whose area under the curve is finite) can be expressed as the integral of sines and/or cosines multiplied by a weighting function (Fourier transform).
- The **frequency domain** refers to the plane of the two dimensional discrete Fourier transform of an image.

- The purpose of the Fourier transform is to represent a signal as a linear combination of sinusoidal signals of various frequencies.

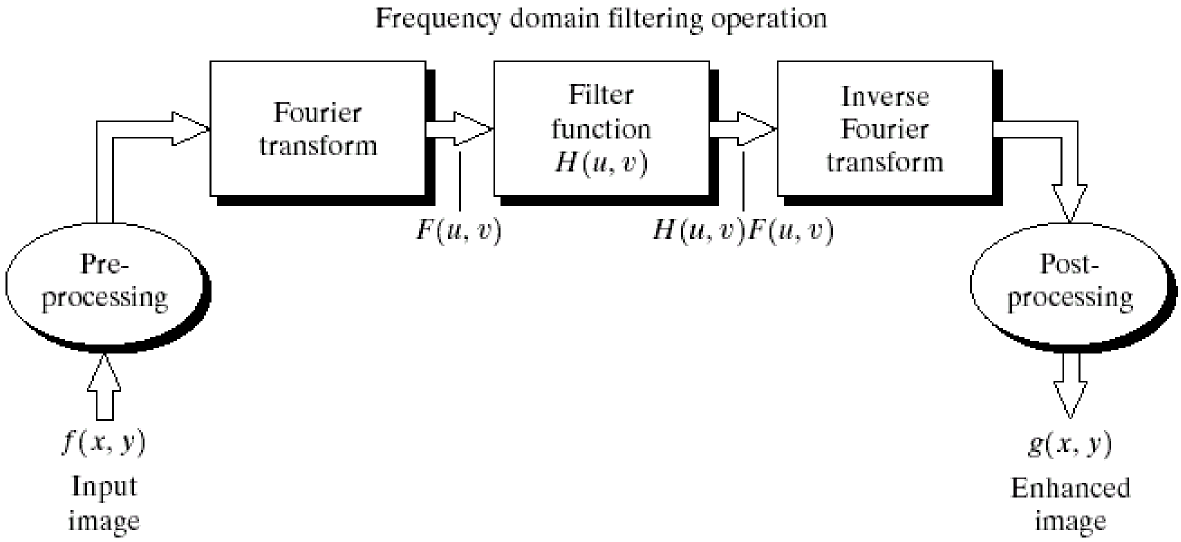


Fig. 2.17 Frequency domain operations

Frequency Domain Filters - Smoothing Ideal Low Pass Filter

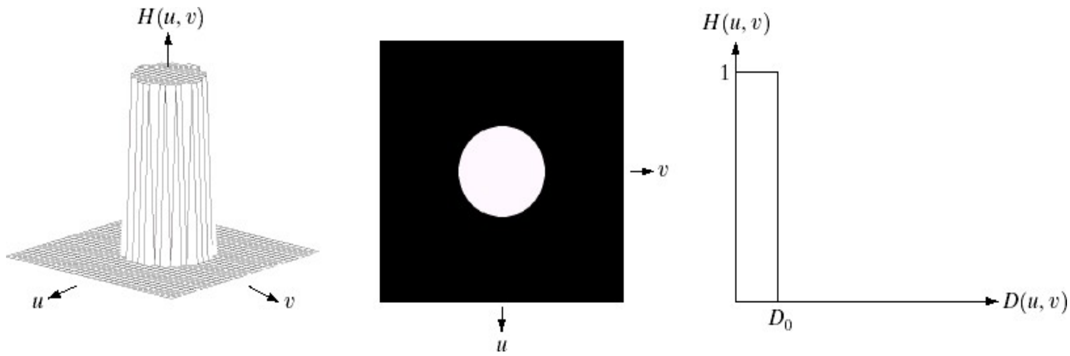


Fig. 2.18 Ideal Low Pass filter

Butterworth Low Pass Filter

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

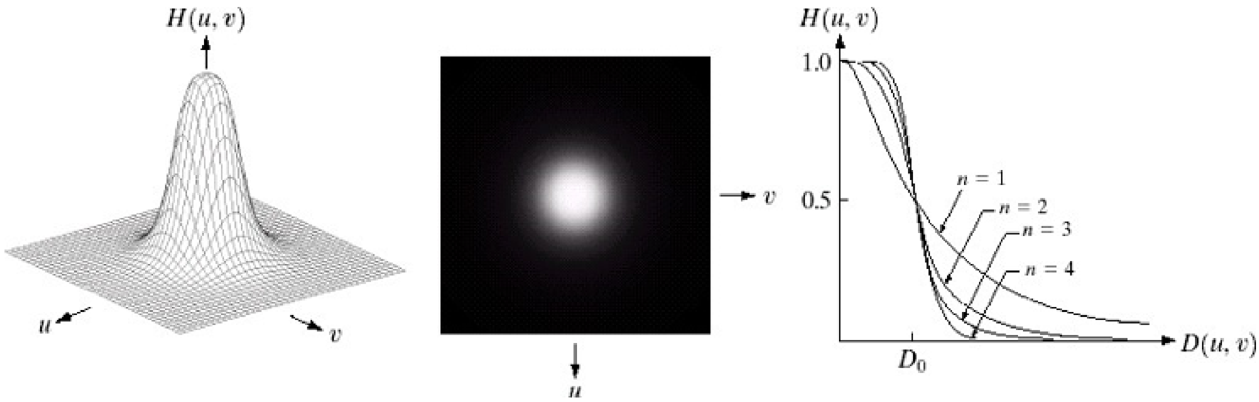


Fig.2.19 Butterworth Low pass filter

Gaussian Low Pass Filter

$H\,(u,v) = e^{-\,D^2\,(u,v)/\,2\,D^2_0}$

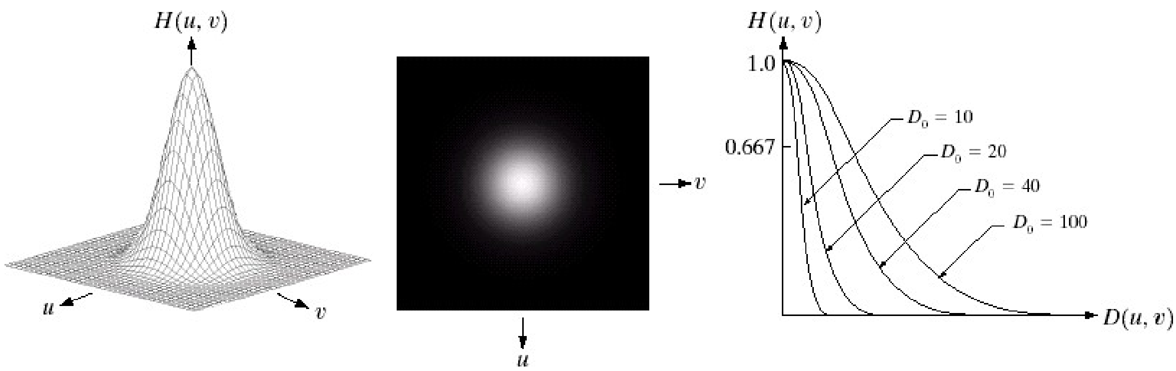


Fig. 2.20 Gaussian Low Pass filter

Frequency Domain Filters - Sharpening

- Image details corresponds to high-frequency
    - Sharpening: high-pass filters
    - $H_{hp}(u,v)=1-H_{lp}(u,v)$

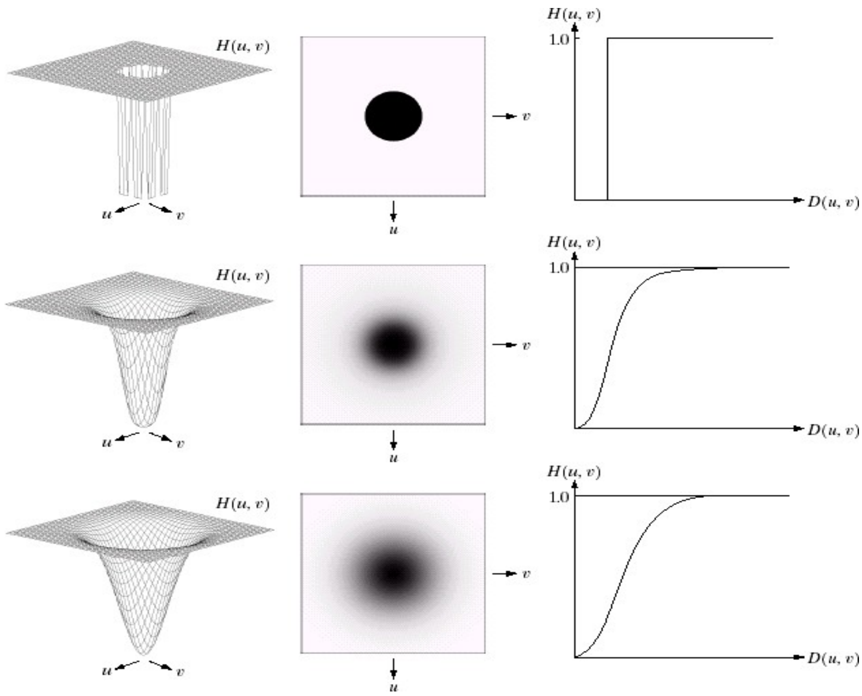


Fig. 2.21 High Pass filter a) Ideal b) Butterworth c) Gaussian

2.6 IMAGE DEGRADATION MODEL

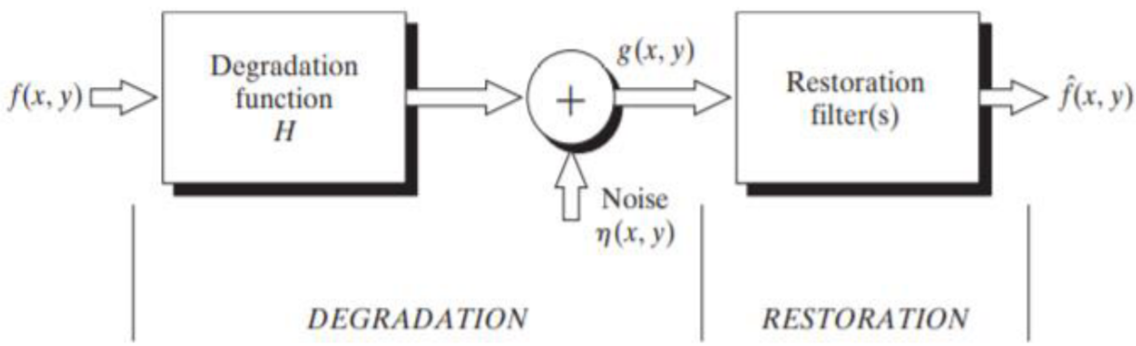


Fig. 2.22 Image Degradation Function

Degradation function along with some additive noise operates on f(x, y) to produce degraded image g(x, y) . Given g(x, y), some knowledge about the degradation function H and additive noise η(x, y), objective of restoration is to obtain estimate f ’(x, y) of the original image. If H is linear, position invariant process then degraded image in spatial domain is given by:

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

Since convolution in Spatial domain = multiplication in Frequency Domain

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

2.7 NOISE MODELS

The sources of noise in digital images arise during image acquisition (digitization) and transmission. Imaging sensors can be affected by ambient conditions. Interference can be added to an image during transmission.

We can consider a noisy image to be modelled as follows:

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

Where f(x, y) is the original image pixel, η(x, y) is the noise term and g(x, y) is the resulting noisy pixel. If we can estimate the noise model we can figure out how to restore the image. There are many different models for the image noise term η(x, y):

Gaussian Noise:

- Most frequently used noise model
- PDF of Gaussian random variable z is given by:
  $$p(z) = \frac{1}{\sqrt{(2\pi)}\sigma} e^{-(z-\mu)^2/2\sigma^2}$$
- z is Gray level
- μ is Mean of average value of z
- σ is Standard Deviation of z
- σ2 is Variance of z When z is defined by this equation then
  - About 70% of its values will be in the range [(μ - σ),(μ + σ)] and
  - About 95% of its values will be in the range [(μ - 2σ),(μ + 2σ)]

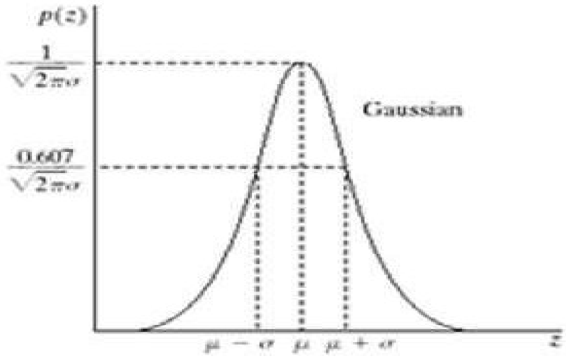


Fig. 2.23 Gaussian Noise

Rayleigh Noise

PDF of Rayleigh Nois $p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & \text{for } z \geq a \\ 0 & \text{for } z \prec a \end{cases}$

- z is Gray level
- μ is Mean of average value ofz
- σ2 is Variance of z

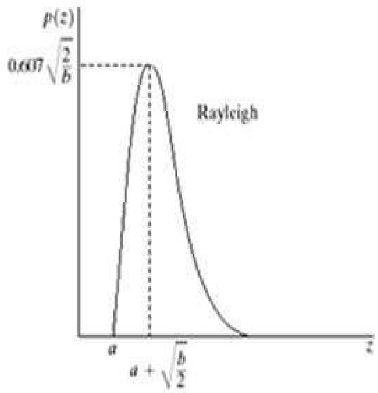


Fig 2.24 Rayleigh Noise

Erlang (Gamma) Noise:

PDF of Erlang Noise is $g_{p(z)} = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-z} & \text{for } z \geq 0 \\ 0 & \text{for } z \prec 0 \end{cases}$

- z is Gray level
- μ is Mean of average value ofz

- σ2 is Variance of z

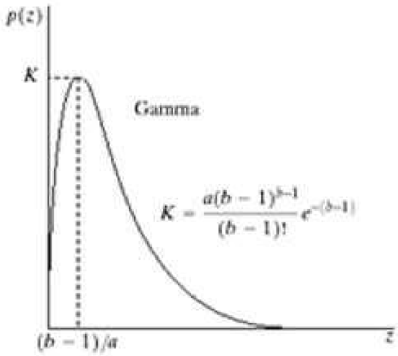


Fig 2.25 Erlang Noise

Exponential Noise

PDF of Exponential Noise is given by: $p(z) = \begin{cases} a e^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases}$

- z is Gray level
- μ is Mean of average value of z

  σ2 is Variance of z

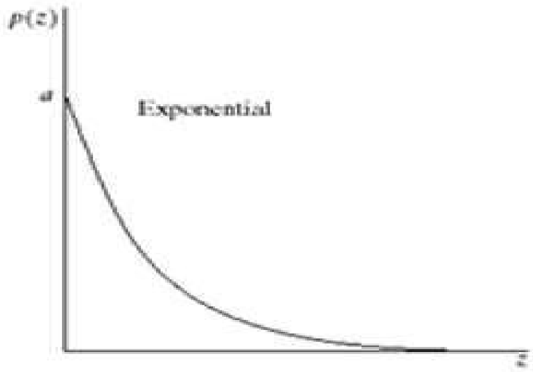


Fig 2.26 Exponential Noise

2.8 INVERSE FILTERING

- Simplest approach to restore an image
- we compute an estimate, of the transform of the original image simply by dividing the transform of the degraded image, , by the degradation function
- Compute an estimate F'(u, v) of the transform of the original image by:

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

Divisions are made between individual elements of the functions

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

- Even if we know degradation function, we can not recover the undegraded image [Inverse Fourier Transform of F(u, v)] exactly because
- N(u, v) is random function whose Fourier Transform is not known
- If degradation has ZERO or less value then N(u, v) / H(u, v) dominates the estimated F’(u, v)
- No explicit provision for handling Noise with α and β being positive, real constants

2.9 GEOMETRIC MEAN FILTER

$$\hat{F}(u, v) = \left[\frac{H^*(u, v)}{|H(u, v)|^2}\right]^{\alpha}\left[\frac{H^*(u, v)}{|H(u, v)|^2 + \beta\left[\frac{S_\eta(u, v)}{S_f(u, v)}\right]}\right]^{1-\alpha} G(u, v)$$

The geometric mean filter consists of the two expressions in brackets raised to the powers α and 1-α , respectively

$$\hat{F}(u, v) = \left[\frac{H^*(u, v)}{|H(u, v)|^2}\right]^{\alpha}\left[\frac{H^*(u, v)}{|H(u, v)|^2 + \beta\left[\frac{S_\eta(u, v)}{S_f(u, v)}\right]}\right]^{1-\alpha} G(u, v)$$

When α=1 this filter reduces to the inverse filter. With α=0 the filter becomes the so-called parametric Wiener filter, which reduces to the standard Wiener filter when β=1.If α = ½ the filter becomes a product of the two quantities raised to the same power, which is the definition of the geometric mean, thus giving the filter its name. With β =1 as α decreases below 1/2, the filter performance will tend more toward the inverse filter Similarly, when α increases above 1/2, the filter will behave more like the Wiener filter.

Citation: All images shown in this unit were adopted from “Rafael C Gonzalez, Richard E Woods, "Digital Image Processing", 3th Edition, Pearson Education

Question Bank:

Part- A

1. Justify the need for Image Enhancement
2. Illustrate the concept of image subtraction
3. Interpret the concept of Contrast Enhancement
4. Comment on spectral high pass filter
5. Illustrate Gaussian noise with its PDF
6. Infer how Geometric mean filter is implemented to remove noises present in a image

Part- B

1. Explain Image Enhancement based on pixel point processing
2. Enumerate how histogram matching is used to increase the enhancement level of a given image
3. Appraise on the various spatial filtering procedures used in DIP system
4. Discuss in detail the different noise models available in DIP systems
5. Enumerate in detail the Image degradation model with necessary diagrams

TEXT / REFERENCE BOOKS

1. Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, 2nd Edition, Pearson Education, Inc., 2004.
2. Anil K. Jain, “Fundamentals of Digital Image Processing”, PHI Learning Private Limited, New Delhi, 2002.
3. William K. Pratt, “Digital Image Processing”, 3rd Edition, John Wiley & Sons, Inc., 2001.
4. Rafael C.Gonzalez, Richard E.Woods and Steven L. Eddins, “Digital Image Processing using Matlab”, Pearson Education, Inc., 2004.
5. Bernd Jähne, “Digital Image Processing”, 5th Revised and Extended Edition, Springer, 2002.

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 ELECTRICAL AND ELECTRONICS ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIT – III – DIGITAL IMAGE PROCESSING – SECA 3009

UNIT III

IMAGE SEGMENTATION & COMPRESSION

Image Segmentation - Detection of Discontinuities, Edge Linking and Boundary Detection, Thresholding, Region based Segmentation, Coding Redundancy, Inter pixel Redundancy, Image Compression model, Error Free Compression, Variable Length Coding, Lossy Compression.

3.1 IMAGE SEGMENTATION

Image segmentation divides an image into regions that are connected and have some similarity within the region and some difference between adjacent regions. The goal is usually to find individual objects in an image. For the most part there are fundamentally two kinds of approaches to segmentation: discontinuity and similarity.

- Similarity may be due to pixel intensity, color or texture.
- Differences are sudden changes (discontinuities) in any of these, but especially sudden changes in intensity along a boundary line, which is called an edge.

There are three kinds of discontinuities of intensity: points, lines and edges. The most common way to look for discontinuities is to scan a small mask over the image. The mask determines which kind of discontinuity to look for. Only slightly more common than point detection is to find one-pixel wide line in an image. For digital images the only three-point straight lines are only horizontal, vertical, or diagonal (+ or –450).

3.2 EDGE LINKING & BOUNDARY DETECTION

Two properties of edge points are useful for edge linking:

  – the strength (or magnitude) of the detected edge points
  – their directions (determined from gradient directions)

- This is usually done in local neighborhoods.
- Adjacent edge points with similar magnitude and direction are linked.
- For example, an edge pixel with coordinates (x_0,y_0) in a predefined neighborhood of (x,y) is similar to the pixel at (x,y) if

$$|\nabla f(x,y) - \nabla(x_0, y_0)| \leq E, \quad \underline{E:} \text{ a nonnegative threshold}$$

$$|\alpha(x,y) - \alpha(x_0, y_0)| < A, \quad \underline{A:} \text{ a nonegative angle threshold}$$

Hough transform: a way of finding edge points in an image that lie along a straight line. Example: *xy*-plane v.s. *ab*-plane (parameter space)

$$yi = axi + b$$

- The Hough transform consists of finding all pairs of values of θ and ρ which satisfy the equations that pass through (*x*,*y*).
- These are accumulated in what is basically a 2-dimensional histogram.
- When plotted these pairs of θ and ρ will look like a sine wave. The process is repeated for all appropriate (*x*,*y*) locations.

3.3 THRESHOLDING

Global – T depends only on gray level values

Local – T depends on both gray level values and local property Dynamic or Adaptive – T depends on spatial coordinates

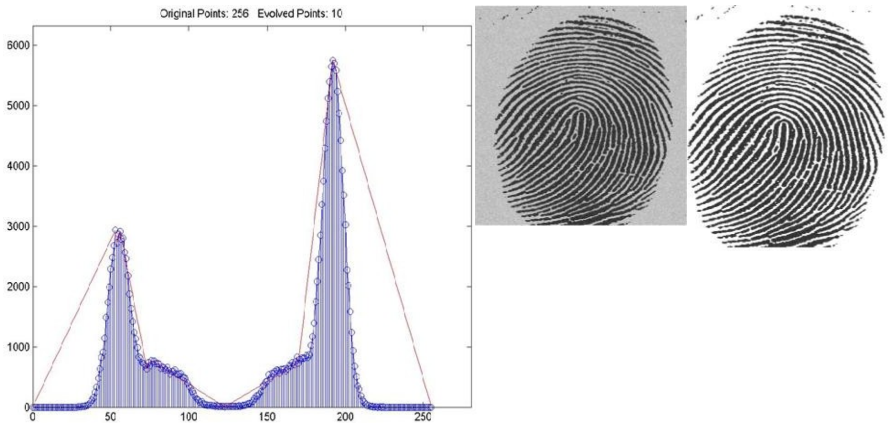


Figure 3.1 Gray level thresholding

Different approaches possible in Gray level threshold are

(1) Interactive threshold

(2) Adaptive threshold

(3) Minimization method

3.4 REGION BASED SEGMENTATION

- Edges and thresholds sometimes do not give good results for segmentation.
- Region-based segmentation is based on the connectivity of similar pixels in a region.
  - Each region must be uniform.
  - Connectivity of the pixels within the region is very important.
- There are two main approaches to region-based segmentation: region growing and region splitting.

Basic Formulation

- Let R represent the entire image region.

For example: $P(R_k)$=TRUE if all pixels in R_k have the same gray level. Region splitting is the opposite of region growing.

  - First there is a large region (possible the entire image).
  - Then a predicate (measurement) is used to determine if the region is uniform.
  - If not, then the method requires that the region be split into two regions.
  - Then each of these two regions is independently tested by the predicate (measurement).
  - This procedure continues until all resulting regions are uniform.

The main problem with region splitting is determining where to split a region. One method to divide a region is to use a quad tree structure. Quadtree: a tree in which nodes have exactly four descendants. The split and merge procedure:

  - Split into four disjoint quadrants any region R_i for which $P(R_i)$ = FALSE.

  Merge any adjacent regions R_j and R_k for which $P(R_j \text{U} R_k)$ = TRUE. (the quadtree structure may not be preserved)

    - Stop when no further merging or splitting is possible.

3.5 REGION SPLIT AND MERGE

If only splitting is used, the final partition normally contains adjacent regions with identical properties. This drawback can be remedied by allowing merging as well as splitting. The steps involved in split and merging is given as follows

1. Split into four disjoint quadrants any region R_i for which $Q(R_i)$ = FALSE.
2. When no further splitting is possible, merge any adjacent regions R_j and R_k for which $Q(R_j \cup R_k)$ = TRUE.
3. Stop when no further merging is possible.

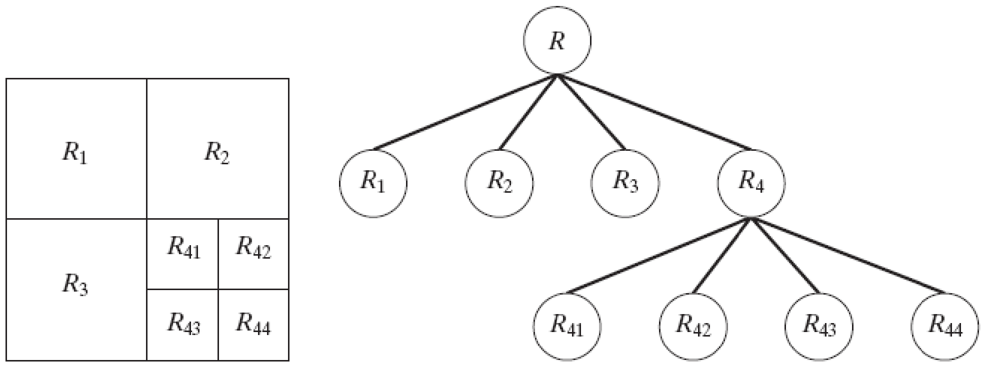


Figure 3.2 Region splitting procedure

3.6 IMAGE COMPRESSION

The term data compression refers to the process of reducing the amount of data required to represent a given quantity of information. Data are the means by which information is conveyed. Various amounts of data may be used to represent the same amount of information. Such might be the case, for example, if a long-winded individual and someone who is short and to the point were to relate the same story. Here, the information of interest is the story; words are the data used to relate the information. If the two individuals use a different number of words to tell the same basic story, two different versions of the story are created, and at least one includes nonessential data. That is, it contains data (or words) that either provide no relevant information or simply restate that which is already known. It is thus said to contain data redundancy.

Data redundancy is a central issue in digital image compression. It is not an abstract concept but a mathematically quantifiable entity. If n1 and n2 denote the number of information-carrying units in two data sets that represent the same information, the relative data redundancy RD of the first data set (the one characterized by n1) can be defined as

$$R = 1 - \frac{1}{C}$$

And

$$C = \frac{b}{b'}$$

where b & b’ are the number of pixels in the original image and the compressed images respectively.

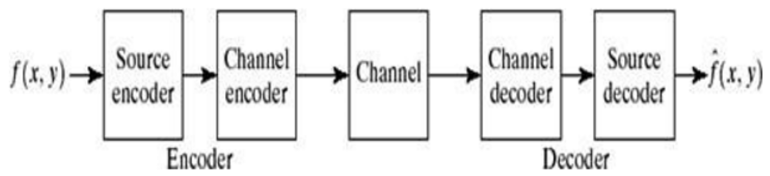
In digital image compression, three basic data redundancies can be identified and exploited: coding redundancy, interpixel redundancy, and psychovisual redundancy. Data compression is achieved when one or more of these redundancies are reduced or eliminated.

1. Coding redundancy. A code is a system of symbols (letters, numbers, bits, and the like) used to represent a body of information or set of events. Each piece of information or event is assigned a sequence of code symbols, called a code word. The number of symbols in each code word is its length. The 8-bit codes that are used to represent the intensities in most 2-D intensity arrays contain more bits than are needed to represent the intensities.

2. Spatial and temporal redundancy. Because the pixels of most 2-D intensity arrays are correlated spatially (i.e., each pixel is similar to or dependent on neighboring pixels), information is unnecessarily replicated in the representations of the correlated pixels. In a video sequence, temporally correlated pixels (i.e., those similar to or dependent on pixels in nearby frames) also duplicate information.

3. Irrelevant information. Most 2-D intensity arrays contain information that is ignored by the human visual system and/or extraneous to the intended use of the image. It is redundant in the sense that it is not used.

3.7 GENERAL COMPRESSION MODEL



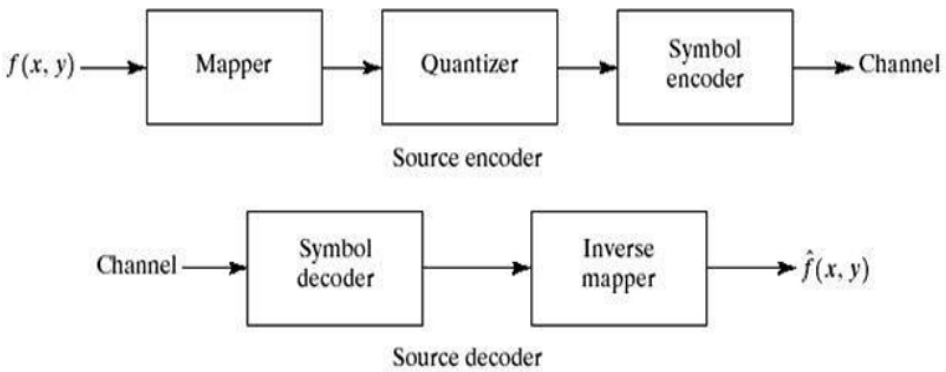


Figure 3.3 Image compression Model

The encoder of Fig. 3.3 is designed to remove the redundancies through a series of three independent operations. In the first stage of the encoding process, a mapper transforms into a (usually nonvisual) format designed to reduce spatial and temporal redundancy. This operation generally is reversible and may or may not reduce directly the amount of data required to represent the image.

The quantizer in Fig. 3.3 reduces the accuracy of the mapper's output in accordance with a pre-established fidelity criterion. The goal is to keep irrelevant information out of the compressed representation. This operation is irreversible. It must be omitted when error-free compression is desired. In the third and final stage of the encoding process, the symbol coder of Fig. 3.3 generates a fixed- or variable-length code to represent the quantizer output and maps the output in accordance with the code. In many cases, a variable-length code is used. The shortest code words are assigned to the most frequently occurring quantizer output values—thus minimizing coding redundancy. This operation is reversible.

3.8 ERROR FREE CODING

When image compression is done the data are not lost. Without any error the data are compressed. One such code is the Huffman Coding. The procedure for Huffman Coding is given below.

Huffman Coding:

One of the most popular techniques for removing coding redundancy is due to Huffman (Huffman [1952]). When coding the symbols of an information source individually, Huffman coding yields the smallest possible number of code symbols per source symbol.

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

Figure 3.4 Huffman Source reduction

Original source			Source reduction							
Symbol	Probability	Code	1		2		3		4	
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

Figure 3.5 Huffman Code Assignment Procedure

The average length of this code is

$$L_{\text{avg}} = (0.4)(1) + (0.3)(2) + (0.1)(3) + (0.1)(4) + (0.06)(5) + (0.04)(5)$$
$$= 2.2 \text{ bits/pixel}$$

Acknowledgement /Citations: All Images, Pictures, Equations used in this material were adopted from “Rafael C Gonzalez, Richard E Woods, "Digital Image Processing", 3th Edition, Pearson Education.

Question Bank:

Part- A

1. Justify the need for Image Segmentation
2. Illustrate the concept of thresholding
3. Interpret the concept of Image Compression
4. Comment on Image Fidelity Criteria
5. Illustrate Edge detection.
6. Infer the concept of Region splitting algorithm

Part- B

1. Explain Image Region Growing based Segmentation algorithm
2. Enumerate how Region Split & Merge Algorithm is used for image object segmentation
3. Appraise on how Huffman Code is used to reduce the image storage by assuming suitable example
4. Discuss in detail the lossy Predictive Coding Model
5. Enumerate in detail the Image compression based on Wavelet Transforms.

TEXT / REFERENCE BOOKS

1. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2 nd Edition, Pearson Education, Inc., 2004.
2. Anil K. Jain, "Fundamentals of Digital Image Processing", PHI Learning Private Limited, New Delhi, 2002.
3. William K. Pratt, "Digital Image Processing", 3rd Edition, John Wiley & Sons, Inc., 2001.
4. Rafael C.Gonzalez, Richard E.Woods and Steven L. Eddins, "Digital Image Processing using Matlab", Pearson Education, Inc., 2004.
5. Bernd Jähne, "Digital Image Processing", 5th Revised and Extended Edition, Springer, 2002.

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 ELECTRICAL AND ELECTRONICS ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIT – IV – MORPHOLOGICAL & COLOUR IMAGE PROCESSING – SECA 3009

UNIT IV

MORPHOLOGICAL & COLOR IMAGE PROCESSING

Dilation and Erosion, Opening and Closing, Basic Morphological Algorithms: Boundary Extraction, Region Filling, Thickening and Thinning; Colour Image Representation, Colour Models, Pseudo Colour Image Processing, Colour Transformations, Smoothing and Sharpening, Segmentation based on Colour.

4.1 LOGIC OPERATIONS INVOLVING BINARY IMAGES

Mathematical Morphology is based on the algebra of non-linear operators operating on object shape and in many respects supersedes the linear algebraic system of convolution. It performs in many tasks – pre-processing, segmentation using object shape, and object quantification – better and more quickly than the standard approach. Mathematical morphology tool is different from the usual standard algebra and calculus. Morphology tools are implemented in most advanced image analysis.

Mathematical morphology is very often used in applications where shape of objects and speed is an issue—example: analysis of microscopic images, industrial inspection, optical character recognition, and document analysis. The non-morphological approach to image processing is close to calculus, being based on the point spread function concept and linear transformations such as convolution. Mathematical morphology uses tools of non-linear algebra and operates with point sets, their connectivity and shape. Morphology operations simplify images, and quantify and preserve the main shape characteristics of objects. Morphological operations are used for the following purpose:

- Image pre-processing(noise filtering, shape simplification)
- Enhancing object structure (skeleton zing, thinning, thickening, convex hull, object marking)
- Segmenting objects from the background
- Quantitative description of objects (area, perimeter, projections, Euler-Poincare characteristics)

Mathematical morphology exploits point set properties, results of integral geometry, and topology. The real image can be modelled using point sets of any dimension; the Euclidean 2D space and its system of subsets is a natural domain for planar shape description.

Computer vision uses the digital counterpart of Euclidean space – sets of integer pairs (∈) f or binary image morphology or sets of integer triples (∈) for gray-scale morphology or binary 3D morphology. Discrete grid can be defined if the neighbourhood relation between points is well defined. This representation is suitable for both rectangular and hexagonal grids. A morphological transformation is given by the relation of the image with another small point set B called structuring element. B is expressed with respect to a local origin. Structuring element is a small image-used as a moving window-- whose support delineates pixel neighbourhoods in the image plane. It can be of any shape, size, or connectivity (more than 1 piece, have holes). To apply the morphologic transformation () to the image means that the structuring element B is moved systematically across the entire image. Assume that B is positioned at some point in the image; the pixel in the image corresponding to the representative point O of the structuring element is called the current pixel. The result of the relation between the image X and the structuring element B in the current position is stored in the output image in the current image pixel position.

- **Reflection**

$\hat{B} = \{w | w = -b, \text{for } b \in B\}$

- **Translation**

$(B)_Z = \{c | c = b + z, \text{for } b \in B\}$

- **Dilation and Erosion Dilation**

With and as sets in 2, the dilation of by , denoted ⊕ , is defined as

A⊕B= { |(ˆ) ∩ ≠∅}

$$A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\}$$

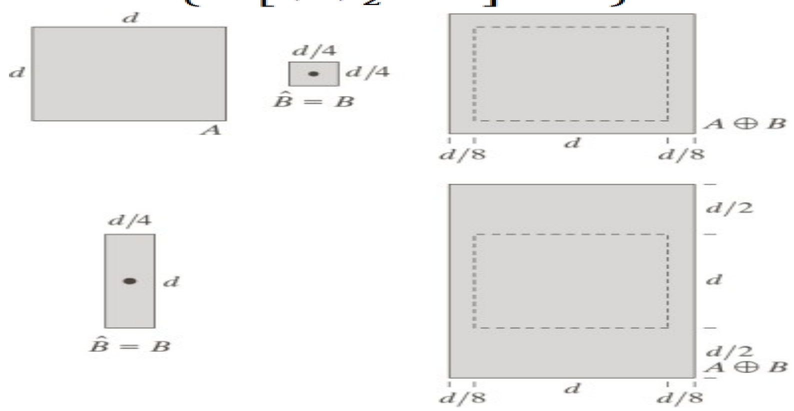
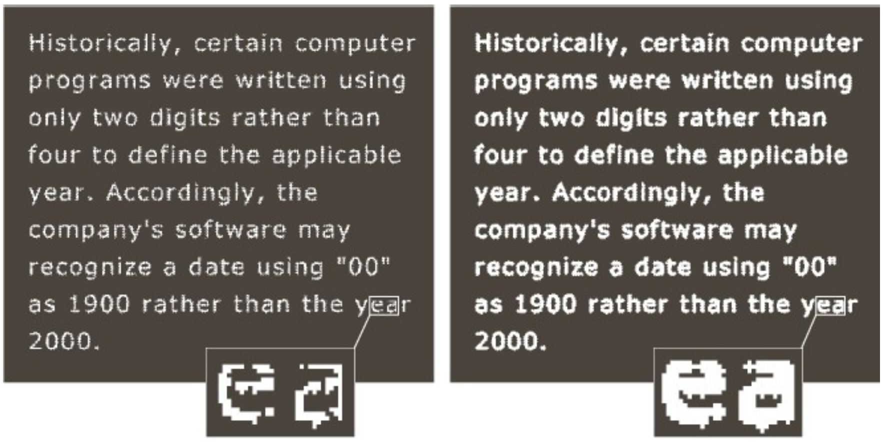


Figure 4.1.(a) SET A (b) Structuring Element (c) Dilation of A by B (d) Elongated structuring element and (e) Dilation of A using the element given in (d)



0	1	0
1	1	1
0	1	0

Figure 4.2 (a) Sample text of poor resolution with broken characters (b) Structuring element (c) Dilation of (a) by (b). Broken segment were joined

Erosion

With A and B as sets in Z^2, the erosion of A by B, denoted $A \quad B$,
defined

$$A \quad B = \{z | (B)_Z \subseteq A\}$$

The set of all points z such that B, translated by z, is contained by A.

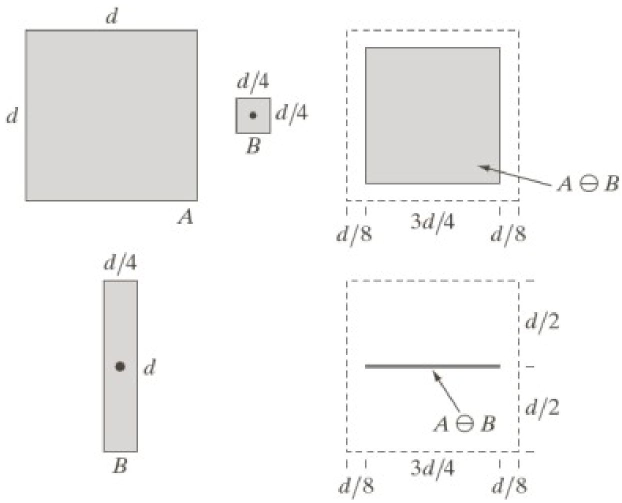


Figure 4.3 (a) Set A. (b) Square structuring element, B (c) Erosion of A by B, shown shaded (d) Elongated structuring element (e) Erosion of A by B using this element. The dotted border in (c) and (e) is the boundary of Set A, shown only for reference

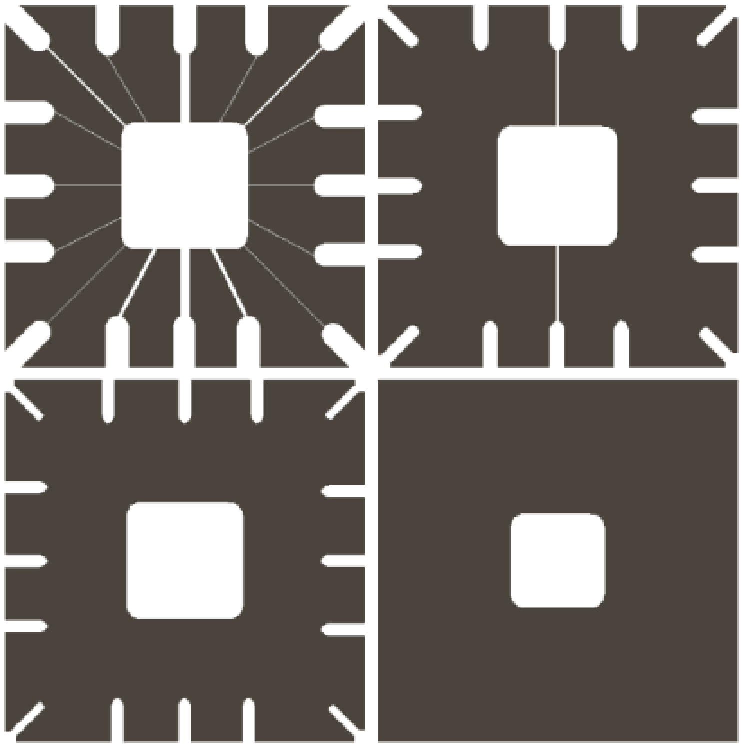


Figure 4.4 Using erosion to remove image components (a) A 486 × 486 binary image of a wire bond mask (b) – (d) image eroded using square structuring elements of size 11 × 11, 15 × 15 and 45 × 45 respectively. The elements of the SEs were all 1s.

- **Opening and Closing**

  Opening generally smooths the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions. Closing tends to smooth sections of contours but it generates f uses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour.

The opening of set A by structuring element B,denoted $A \circ B$, is defined as

$$A \circ B = (A - B) \oplus B$$

The closing of set A by structuring element B, denoted A
$\bullet B$, is defined as

$$A \bullet B = (A \oplus B) - B$$

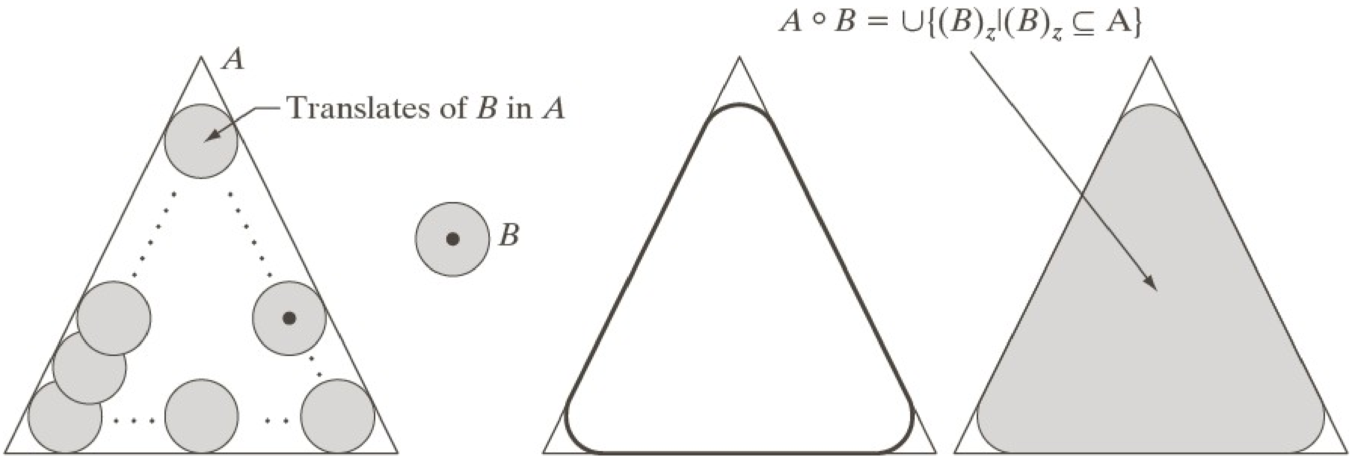


Figure 4.5 (a) Structuring element B "rolling" along the inner boundary of A (the dot indicates the origin of B), (b) structuring element (c) The heavy line is the outer boundary of the opening (d) Complete opening (shaded). We did not shade Shade A in (a) for clarity

- **Boundary Extraction**

The boundary of a set A, can be obtained by first eroding A by B and then performing the set difference between A and its erosion.

$$\beta(A) = A - (A - B)$$

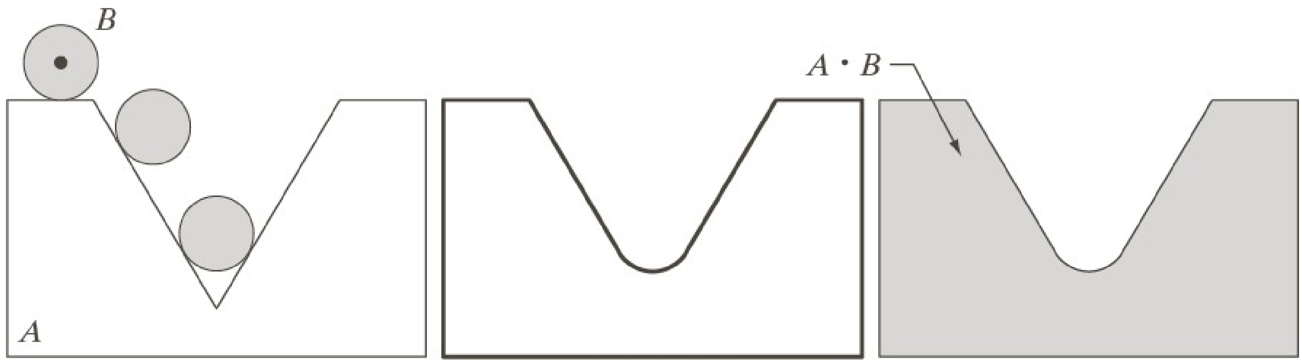


Figure 4.6 (a) Structuring element B "rolling" on the outer boundary of set A. (b) The heavy line is the outer boundary of the closing. (c) Complete closing (shaded). We did not shade A in (a) for clarity

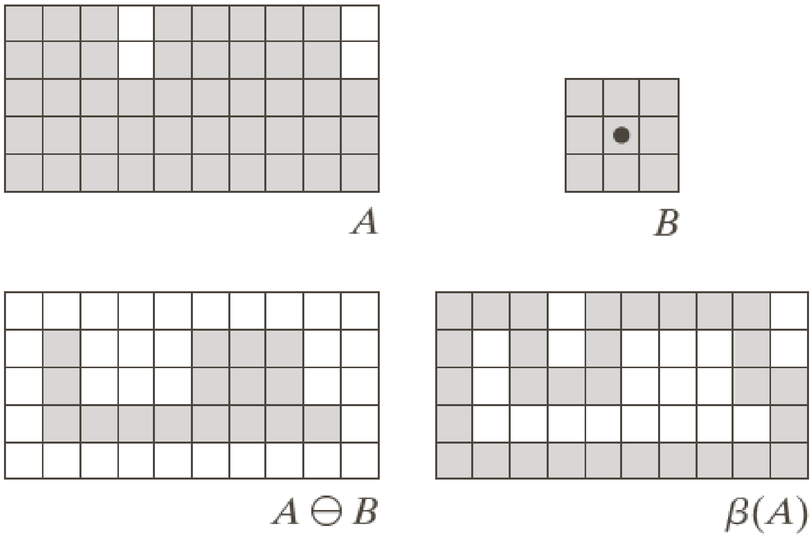


Figure 4.7 (a) Set A. (b) Structuring element B (c) A eroded by B. (d) Boundary given by the set difference between A and its erosion

Figure 4.8(a) A simple binary image, with 1s represented in white (b) Result of using equations with the structuring element

- **Region Filling**

A hole may be defined as a background region surrounded by a connected border of foreground pixels. Let A denote a set whose elements are 8-connected boundaries, each boundary enclosing a background region (i.e., a hole). Given a point in each hole, the objective is to fill all the holes with 1s.

1. Forming an array X_0 of 0s (the same size as the array containing A), except the locations in X_0 corresponding to the given point in each hole, which we set to 1.

2. $X_k = (X_{k-1} + B)$ Ac   k=1,2,3,…

Stop the iteration if $X_k = X_{k-1}$

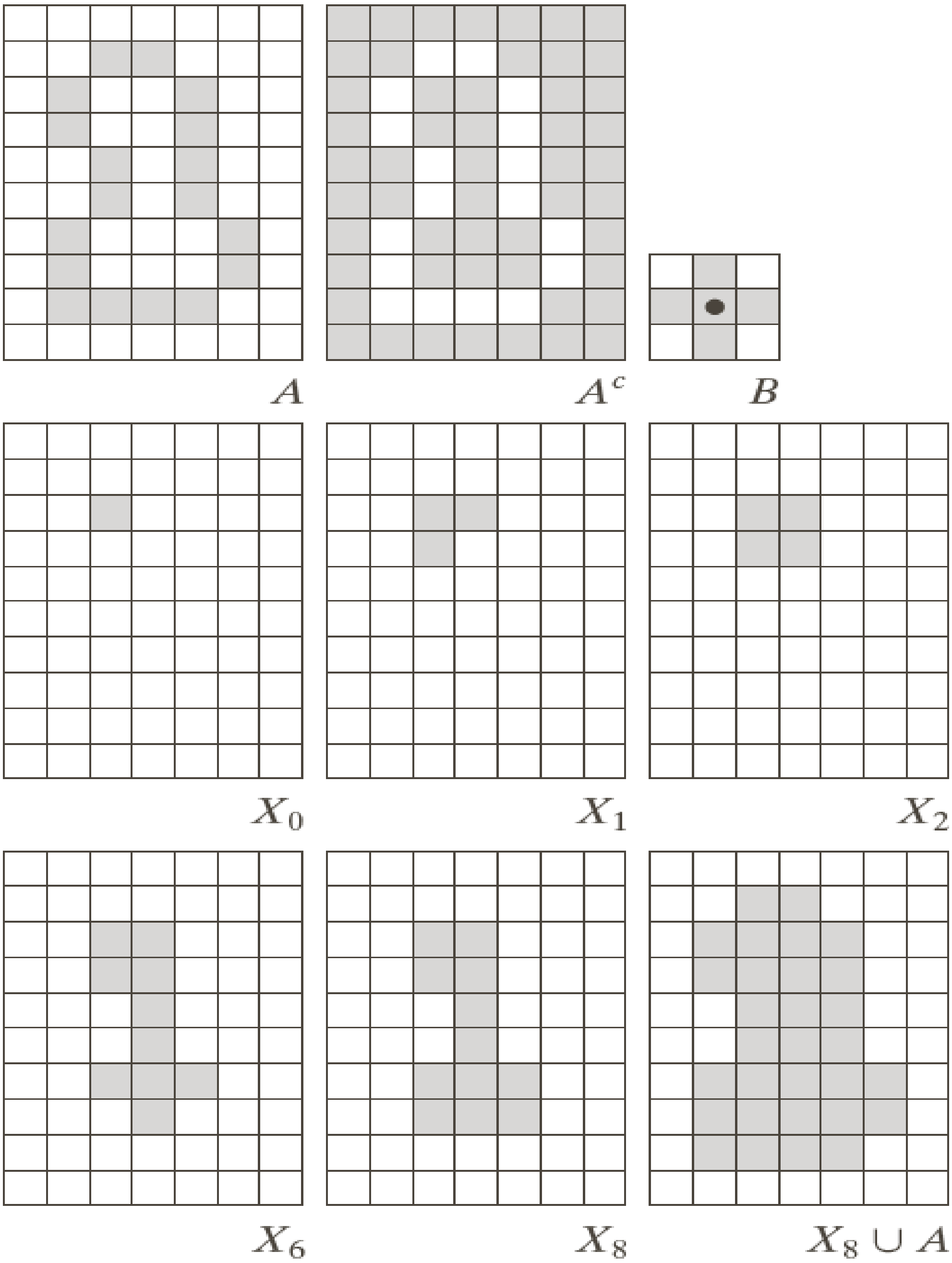


Figure 4.9 Hole filling. (a) Set A (shown shaded). (b) Complement of A. (c) Structuring element B. (d) Initial point inside the boundary (e) – (h) various steps of (i) Final result

- **Thickening**

The thickening is defined by the expression

$$\odot \quad = \quad \cup(\quad * \quad)$$

The thickening of by a sequence of structuring element { }

⊙{ }=((...((⊙ ^1)⊙ ^2)...)⊙ ^)

In practice, the usual procedure is to thin the background of the set and then complement the result.

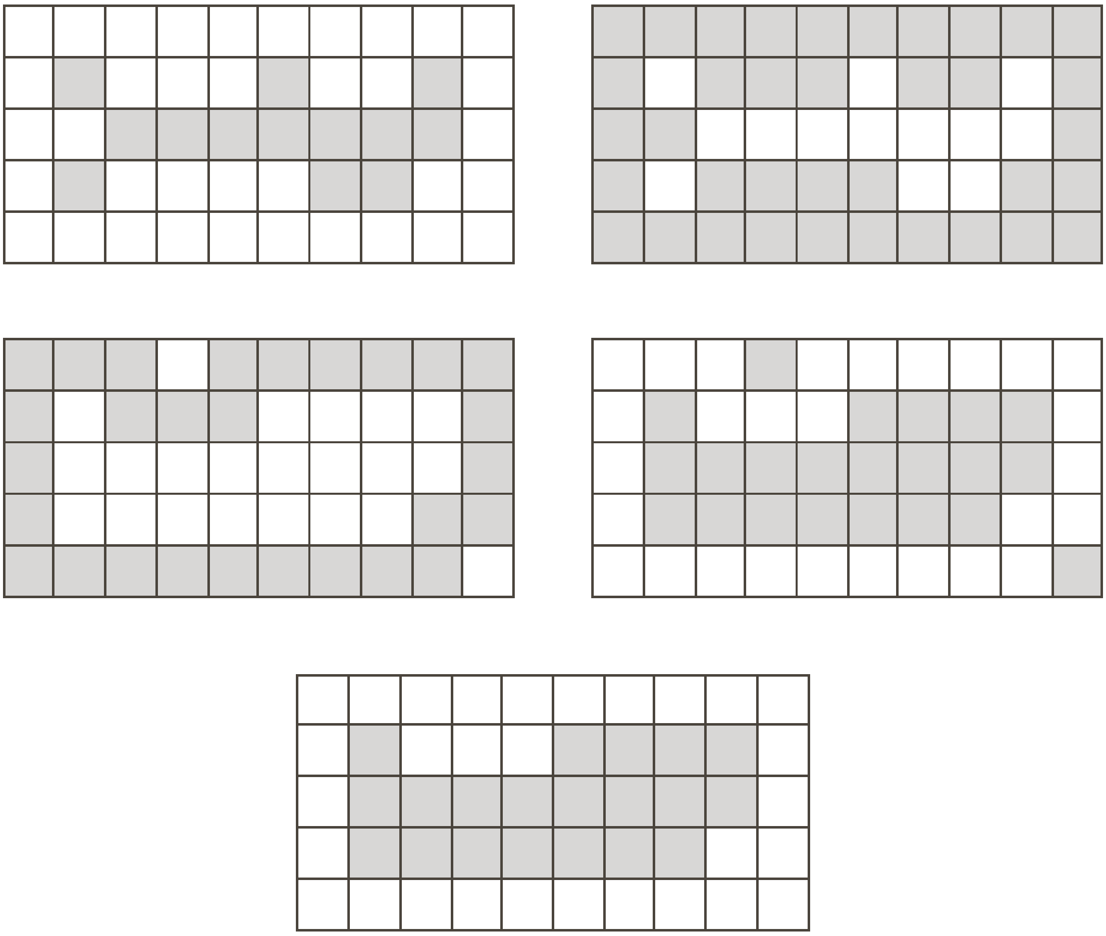


Figure 4.10 Set A. (b) Complement of A. (c) Result of thinning the complement of A. (d) Thickened set obtained by complementing (c). (e) Final result, with no disconnected points.

- **Thinning**

The thinning of a set A by a structuring element B, defined

$$A \otimes B = A - (A * B)$$

$$= A \cap (A * B)^c$$

$$\{B\} = \{B^1, B^2, B^3, \ldots, B^n\}$$

where B^i is a rotated version of B^{i-1}

The thinning of A by a sequence of structuring element $\{B\}$

$$A \otimes \{B\} = ((...((A \otimes B^1) \otimes B^2)...) \otimes B^n)$$

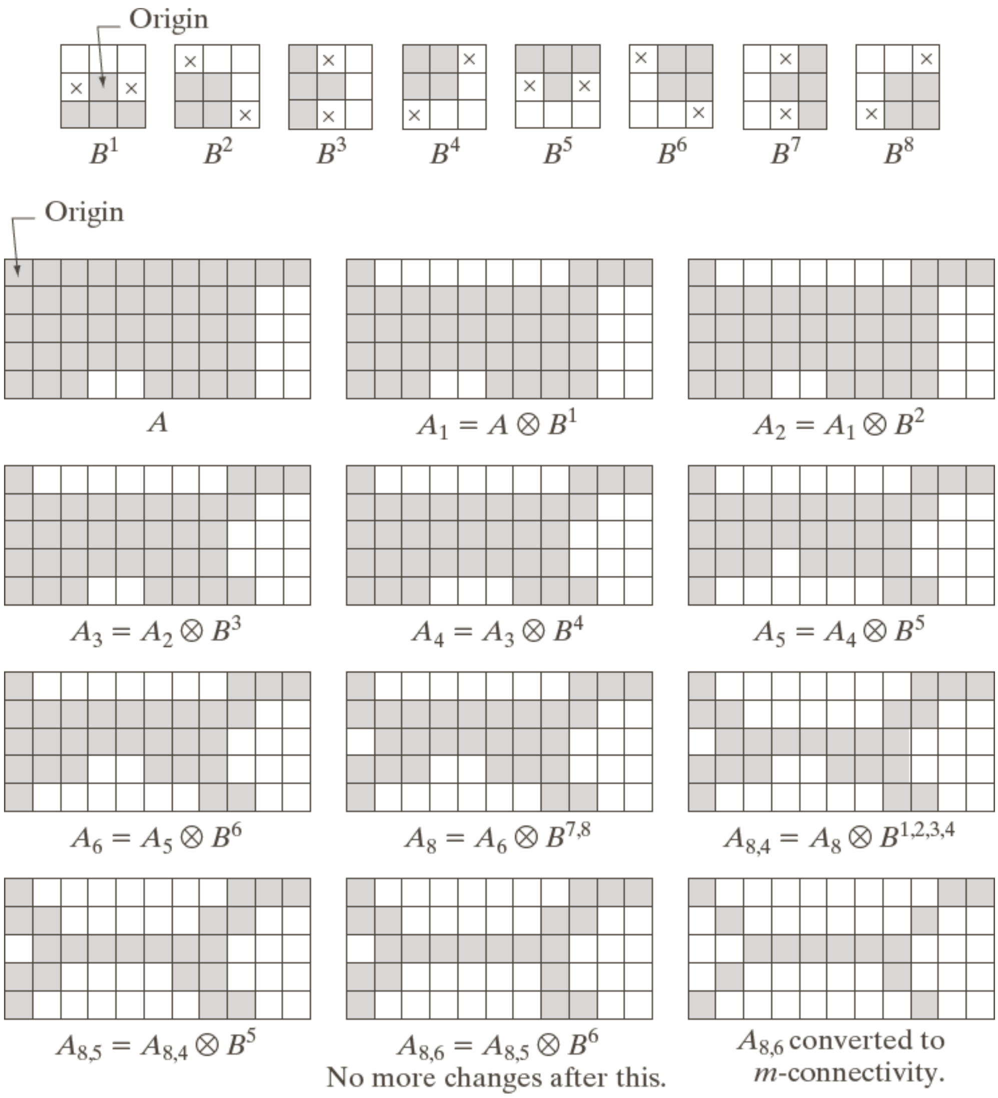


Figure 4.11 (a) Sequence of rotated structuring elements used for thinning. (b) Set A. (c) Result of thinning with the first element (d) – (i) Results of thinning with the next seven elements (there was no change between the seventh and eighth elements). (j) Result of using the first four elements again (l) Result after convergence (m) Conversion to m-connectivity.

- **Color**

Color Vision Models

- The color image is represented by the **RN , GN , BN** coordinates at each pixel.
- The matrix A transforms the input into the three **cone responses αk (x, y, C),**

$k = 1, 2, 3$ where (x, y) are the spatial pixel coordinates and C refers to its color

In Fig., we have represented the **normalized cone responses**

- In analogy with the definition of tristimulus values, Tk are called the retinal cone tristimulus coordinates
- The cone responses undergo nonlinear point transformations to give three f ields Tk (x, y), k = 1, 2, 3
- The 3 x 3 matrix B transforms the {f(x, y)} into {Ck(x, y)} such that C1 (x, y) is the monochrome (achromatic) contrast field c(x, y), as in simplified model, and C2 (x, y) and C3(x, y) represent the corresponding chromatic fields

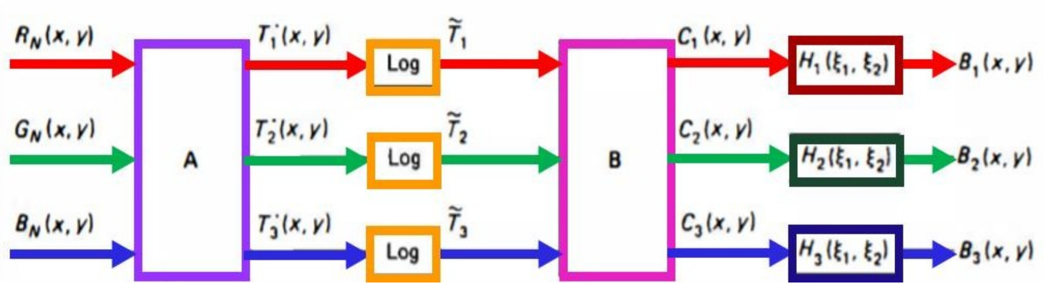


Figure 4.12 Colour Vision Model

- The spatial filters $H_k(\xi_1 , \xi_2)$, k = 1, 2, 3, represent the frequency response of the visual system to luminance and chrominance contrast signals
- Thus $H1(\xi_1 , \xi_2)$ is the same as H (ξ 1 , ξ 2) in simplified model and is a bandpass

$$T_k' \triangleq \frac{\alpha_k(x, y, C)}{\alpha_k(x, y, W)}, \qquad k = 1,2,3$$

filter that represents the lateral inhibition phenomenon

- The visual frequency response to chrominance signals are not well established but are believed to have their passbands in the lower frequency region, as shown in figure

$$\mathbf{A} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.127 & 0.724 & 0.175 \\ 0.000 & 0.066 & 1.117 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 21.5 & 0.0 & 0.00 \\ -41.0 & 41.0 & 0.00 \\ -6.27 & 0.0 & 6.27 \end{pmatrix}$$

The 3 x 3 matrices A and B are given as follows:

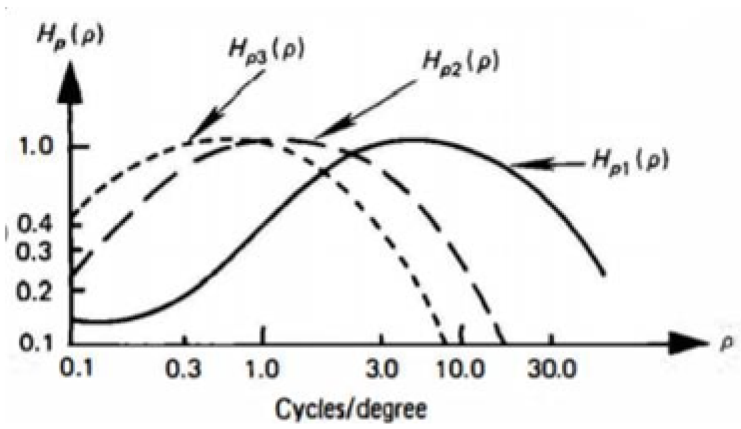


Figure 4.13 Color Vision Model

- From the model , a criterion for color image fidelity can be defined
- For example, for two color images {RN , GN , BN} and {R,,V, G,,V, BN}, their subjective mean square error could be defined by

$$e_{ls} = \frac{1}{A}\sum_{k=1}^{3}\iint_{\Re}(B_k(x,y) - B_k^{\cdot}(x,y))^2\,dx\,dy$$

Where r is the region which over the image is defined (or available), A is its area. And

$\{B_k(x,y)\}$ and $\{B_{kdot}(x,y)\}$ are the outputs of the model for the two colour images.

4.2 COLOR FUNDAMENTALS COLOR SPECTRUM

- In 1666, Sir Isaac Newton
- When a beam of sunlight is passed through a glass prism
- The emerging beam of light consists of a continuous spectrum of colors ranging from violet to red (not white)
- Divided into six broad regions
- Violet, blue, green, yellow, orange, and red
- Each colorblends smoothly into the next

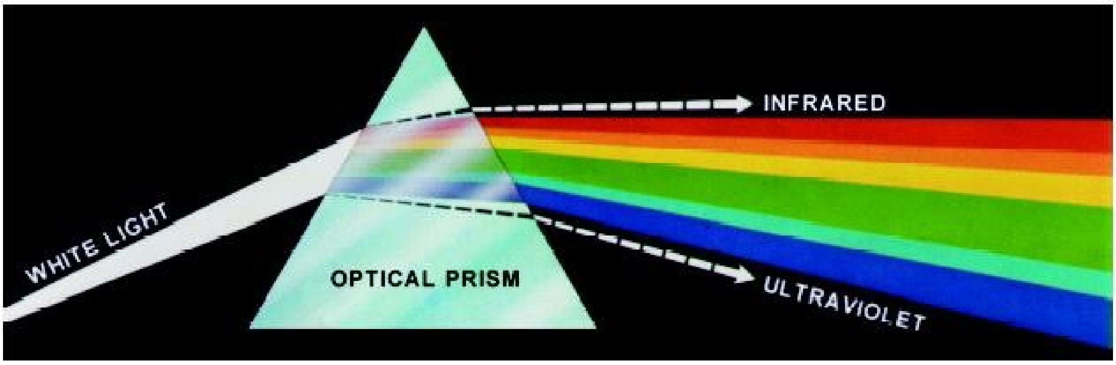


Figure 4.14 Colour Vision Model

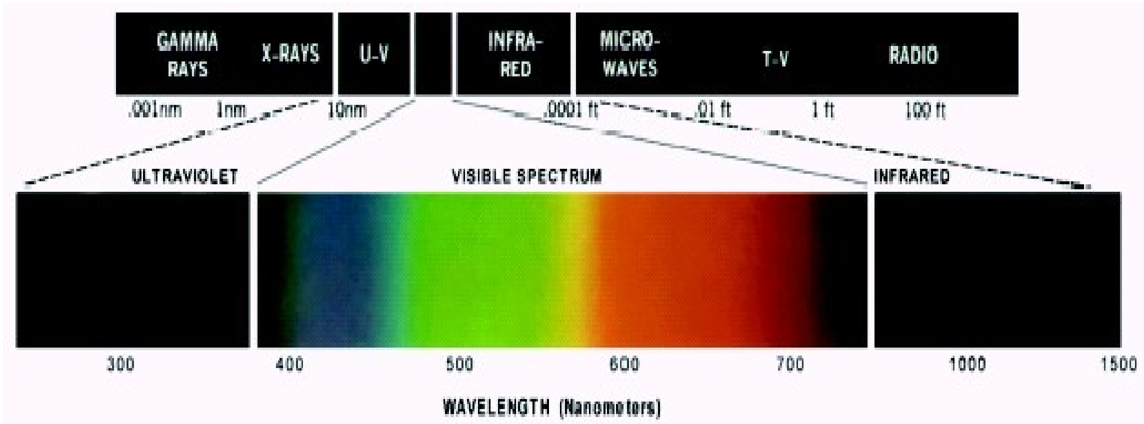


Figure 4.15 Color Vision Model

Primary colors of pigments or colorants

- cyan, magenta, yellow
- A primary color of pigments is defined as one that subtracts or absorbs a primary color of light and reflects or transmits the other two

Secondary colors of pigments or colorants

- red, green, blue
- Combination of the three pigment primaries, or a secondary with its opposite primary, produces black

Characteristics of colors

- Brightness:
- The chromatic notion of intensity
- Hue:
- An attribute associated with the dominant wavelength in a mixture of light waves
- Representing dominant color as perceived by an observer
- Saturation
- Referring to relative purity or the amount of white mixed with a hue
- Saturation is inversely proportional to the amount of white light Hue and saturation taken together are called chromaticity
- A color may be characterized by its brightness and chromaticity
- The amounts of red, green, and blue needed to form any particular color are called the tristimulus values (Denoted X(red), Y(green), and Z(blue)).

Color Models

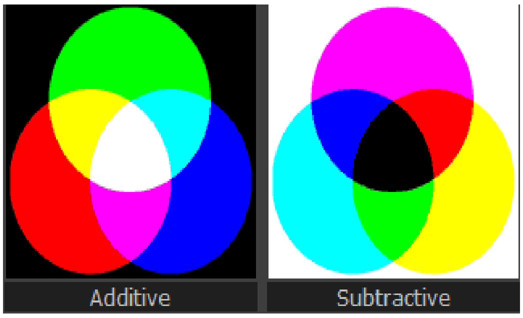
The purpose of a color model is to facilitate the specification of colors in some standard Color models are oriented either toward hardware or applications

- Hardware-oriented
    - Color monitor or Video camera : RGB
    - Color printer : CMY
    - Color TV broadcast : YIQ (I : inphase, q : quadrature)
- Color image manipulation : HSI, HSV
- Image processing : RGB, YIQ, HIS
    - Additive processes create color by adding light to a dark background (Monitors)
    - Subtractive processes use pigments or dyes to selectively
    - block white light (Printers)

RGB color Model

Images represented in the RGB color model consist of three independent image planes, one for each primary color. The number of bits used to represent each pixel in RGB space is called the pixel depth. The term full-color image is used often to denote 24-bit RGB color image

- RGB model is based on a Cartesian coordinate system
- The color subspace of interest is the cube
- RGB values are at three corners
- Colors are defined by vectors extending from the origin
- For convenience, all color values have been normalized
- All values of R, G, and B are in the range [0, 1]



Additive Color Mixing			Subtractive Color Mixing		
Red + Green	→	Yellow	Cyan + Magenta	→	Blue
Green + Blue	→	Cyan	Magenta + Yellow	→	Red
Blue + Red	→	Magenta	Yellow + Cyan	→	Green
Red + Green + Blue	→	White	Cyan + Magenta + Yellow	→	Black

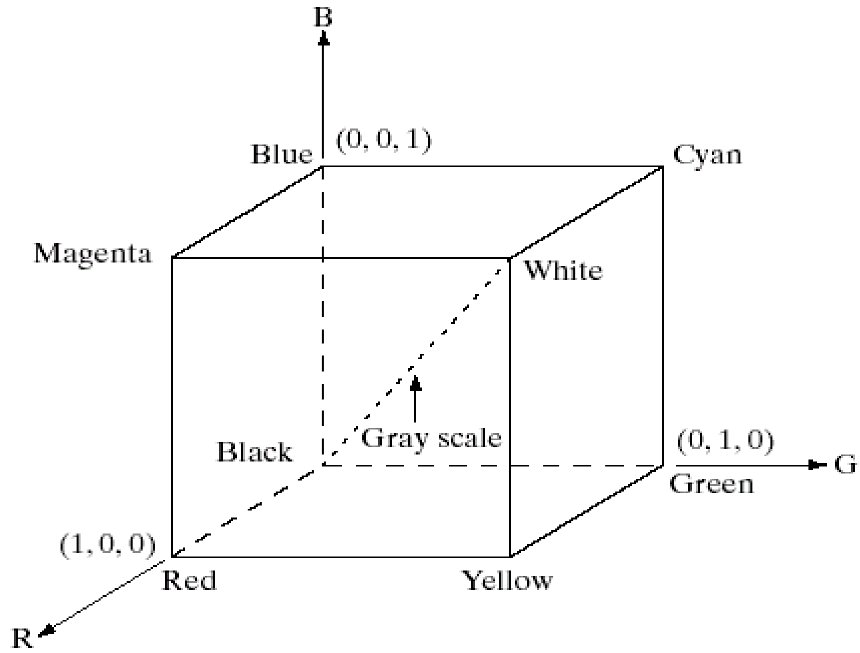


Figure 4.15 RGB color cube

CMY Colour Model

General purpose of CMY color model is to generate hardcopy output

- The primary colors of pigments
- Cyan, Magenta, and Yellow
- C = W - R, M = W - G, and Y = W - B
- Most devices that deposit colored pigments on paper require CMY data input
- Converting RGB to CMY
- The inverse operation from CMY to RGB is generally of no practical interest

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

Color Transformations

Color can be described by its red (R), green (G) and blue (B) coordinates (the well-known RGB system), or by some its linear transformation as XYZ, CMY, YUV, IQ, among others. The CIE adopted systems CIELAB and CIELUV, in which, to a good approximation, equal changes in the coordinates result in equal changes in perception of the color. Nevertheless, sometimes it is useful to describe the colors in an image by some type of cylindrical-like coordinate system, it means by its hue, saturation and some value representing brightness. If the RGB coordinates are in the interval from 0 to 1, each color can be represented by the point in the cube in the RGB space. Let us imagine the attitude of the cube, where the body diagonal linking ”black” vertex and ”white” vertex is vertical. Then the height of each point in the cube corresponds to the brightness of the color, the angle or azimuth corresponds to the hue and the relative distance from the vertical diagonal corresponds to the saturation of the color.

The present color models have some disadvantages in practical use. E.g. we convert an image in some image processing application into some brightness-hue-saturation model and we would like to work with individual components (coordinates) as with separate images. There is desirable regarding to the back conversion to have all combinations of the values. It means we need such model, where the range of values of saturation is identical f or all hues. From this point of view, the GLHS color model [2] is probably the best from the current ones, particularly for $w_{min} = w_{mid} = w_{max} = 1/3$. The good model should satisfy some demands as:

- The brightness should be a linear combination of all three RGB components. At least, it must be continuous growing function of all of them.
- The hue differences between the basic colors (red, green and blue) should be 120◦ and similarly between the complement colors (yellow, purple and cyan). The hue difference between a basic color and an adjacent complement one (e.g. red and yellow) should be 60◦.
- The saturation should be 1 for the colors on the surface of the RGB color cube, it means in case of one of the RGB components is 0 or 1 except black and white vertices and it is 0 in case of R=G=B.

In our opinion, the best brightness, hue and saturation system consists of the brightness as linear combination of the RGB values, the hue as actual angle in the color cube and saturation as relative distance from the body diagonal to the surface of the color cube. Such a system, called YHS, is presented in [1]. It satisfies all three demands and makes easier some color manipulations.

Color Image Segmentation

Color image segmentation that is based on the color feature of image pixels assumes that homogeneous colors in the image correspond to separate clusters and hence meaningful objects in the image. In other words, each cluster defines a class of pixels that share similar color properties. As the segmentation results depend on the used color space, there is no single-color space that can provide acceptable results for all kinds of images. Clustering is the process of partitioning a set of objects (pattern vectors) into subsets of similar objects called clusters. Pixel clustering in three-dimensional color space on the basis of their color similarity is oneof popular approaches in the field of color image segmentation. Clustering is often seen as an unsupervised classification of pixels. Generally, the a priori knowledge about the image is not used during a clustering process. Colors, dominated in the image, create dense clusters in the color space in natural way. Region-based techniques group pixels into homogeneous regions. In this family of techniques, we can find following techniques: region growing, region splitting, region merging and others. Particularly the region growing technique, proposed for grayscale images so long ago, is constantly popular in color image processing. The region growing is a typical bottom-up technique. Neighboring pixels are merged into regions, if their attributes, for example colors, are sufficiently similar. This similarity is often represented by a homogeneity criterion. If a pixel satisfied the homogeneity criterion, then the pixel can be included to the region and then the region attributes (a mean color, an area of region etc.) are updated. The region growing process, in its classical version, is starting from chosen pixels called seeds and is continued so long as all pixels will be assigned to regions. Each of these techniques varies in homogeneity criteria and methods of seeds location. The advantages of region growing techniques result from taking into consideration two important elements: the color similarity and the pixel proximity in the image.

Question Bank:

Part- A

1. Justify the need for Morphology based Segmentation
2. Illustrate the concept of Dilation and Erosion
3. Interpret the concept of Color Transformation
4. Comment on Opening and Closing Operators
5. Illustrate Thickening and Thinning operations
6. Infer the concept of Boundary detection

Part- B

1. Appraise on different Morphology based segmentation operators in detail
2. Enumerate how Color Models are created with the help of chromaticity diagram
3. Appraise on how Color Image Segmentation is used to identify a color image object
4. Discuss in detail how the Boundary regions are extracted from an image object
5. Enumerate in detail the concept of region filling

Acknowledgement/Citation: All images, equations used in this material were adopted from "Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2 nd Edition, Pearson Education, Inc."

TEXT / REFERENCE BOOKS

1. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2 nd Edition, Pearson Education, Inc., 2004.
2. Anil K. Jain, "Fundamentals of Digital Image Processing", PHI Learning Private Limited, New Delhi, 2002.
3. William K. Pratt, "Digital Image Processing", 3rd Edition, John Wiley & Sons, Inc., 2001.
4. Rafael C.Gonzalez, Richard E.Woods and Steven L. Eddins, "Digital Image Processing using Matlab", Pearson Education, Inc., 2004.
5. Bernd Jähne, "Digital Image Processing", 5th Revised and Extended Edition, Springer, 2002.

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 ELECTRICAL AND ELECTRONICS ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIT – v – MEDICAL IMAGE PROCESSING – SECA 3009

UNIT V

MEDICAL IMAGE PROCESSING

Noise Reduction in Nuclear Medicine Imaging, Contrast enhancement of mammograms, Detection of Spinal Canal, Detection of calcifications by multi-tolerance region growing, Shape analysis of calcifications, Analysis of Ligament Healing.

5.1 NOISE REDUCTION IN NUCLEAR MEDICINE IMAGING

Nuclear medicine images are attained under low photon conditions which produces Poisson noises in the images. Counting the photons and countering them reduces the effect of noises present in a nuclear image. Thus, the quality of the image can be increased. When the patient is subjected to longer duration of image capturing, he or she is put into undue stress and motion artifices comes into picture. Also, various practical limitations stop us from going into longer image capturing mode. Figure 5.1 shows the SPECT images of one section of a resolution phantom acquired over 2,15 and 40s. Each image has been scaled such that its minimum and maximum values are mapped to the display range of [0-255]. Also shown is a schematic representation of the section: the circles represents cross-sections of cylindrical holes in a plexiglass block; the diameter of the entire phantom is 200 mm. The two large circles at the extremes of the two sides are of diameter 30 mm. The two inner arrays have circles of diameter 22,17,14,12,9,8 and 6 mm. The phantom was filled with a radiopharmaceutical such that the cylindrical holes would be filled with the radioactive Material. It is evident that the image quality improves as the photon counting time is increased. The circles of diameter 9 and8 mm are distinctly visible only in image (c). However, the circles of diameter 5mm are not visible in any of the images. This is due to the reason that lower dimension merges with the poison noise.

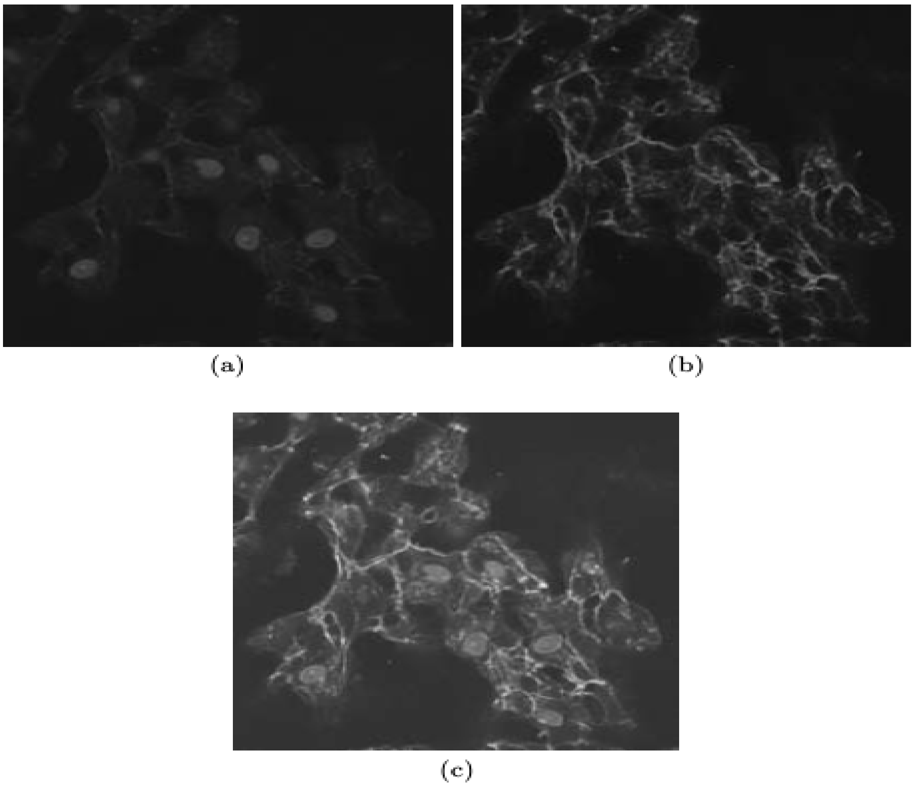


Figure 5.1 : SPECT image (a) The red dye cell nuclei component of the confocal microscope image of the nucleus pulposus of a dog (b) The green_dye (c) Combination of images a & b

Hence to reduce the noises in any nuclear imaging procedure we have to increase the timing of exposure thereby noise effect can be minimized. To reduce noise, we have to increase the examination period or the radioactive materials activity. Increasing the materials activity makes the patient with more exposure of radiation dose. Hence there should be a compromise between the inspection time and the option for the patient to be immobile during the testing period.

Another way is by to decreasing the pixel size of the image. Changing from 256x256 to 128x128 pixels, for example, reduces the noise level. However, the pixel size should not be too large so that it is not changing the spatial resolution of the image. We can also use image processing procedures like background subtraction, filtering to decrease the noise levels and enhance the level of contrast present in the image.

5.2 CONTRAST ENHANCEMENT OF MAMMOGRAMS

Intensity of X- rays absorbed by the abnormality region is different from that of the normal region. Hence intensity variation in mammograms can be used as a tool for identifying the abnormality. In digitally processed mammograms, lesions appear as high intensity region and the normal tissues appear in low intensity regions. Hence the identification of high intensity region is a measure of abnormality. Presence and hence severity of the abnormality is dependent on the size, shape and texture of the abnormality region. Severity of the abnormality is directly related to the size of the high intensity region. Also, non-uniform texture is used as a measure of the severity.

Though manual interpretation appears simple, it is not so due to the following reasons:

- As the contrast between the background and the affected region is less, manual interpretation tests the expertise of the individual.
- Also, certain tissues and blood vessels appear as high intensity regions in mammograms. Hence the probability of false detection also increases.

Hence the paradigm has shifted to computer aided interpretation where image processing is performed to denoise the mammograms, enhance the contrast of the mammograms, isolate and characterise the abnormality region from the mammograms.

The different Procedures used for Mammogram Image enhancement are

- CLACHE algorithm
- Histogram based processing
- Gabor filter-based enhancement procedure

Not limited to this any Enhancement algorithm which helps in the identification of Microcalcifications or highlights the tumour region (Sharpening based algorithms) can be used for increasing the contrast level of the mammogram. While taking the mammogram a very low dose X-ray is made to fall on the region of interest. Hence the reflected rays will also be of less. This is the main reason for mammograms having low contrast.

(Any on enhancement algorithm discussed in Unit -2 can be explained here with respect to mammograms)

5.3 DETECTION OF THE SPINAL CANAL

In an application to analyze CT images of neuroblastoma, the spinal canal was observed to interfere with the segmentation of the tumor using the fuzzy connectivity algorithm. In order to address this problem, a method was developed to detect the center of the spinal canal in each CT slice, grow the D region containing the spinal canal, and remove the structure. The initializing seeds for the region growing procedure were automatically obtained with the following procedure.

The outer region in the CT volume containing materials outside the patient, the skin, and peripheral fat was first segmented and removed. The CT volume was then thresholder at +800 HU to detect the high-density bone structures. All voxels not within mm from the inner boundary of the peripheral fat layer were rejected, Regions were grown using each remaining voxel and all of the resulting regions were merged to form the bone volume. The inclusion criteria were in terms of the CT values being within +800±2σHU with Ў±σ= 103HU being the standard deviation of bone, and spatial Connectivity. The resulting CT volume was cropped to limit the scope of further analysis, as follows. The width of the image was divided into three equal parts, and the outer thirds were rejected. The height of the image was divided into six equal parts and the lower fourth and, fifth parts were included in the cropped region. In the interstice direction, the first 13% of the slices were removed, and the subsequent 20%slices were included in the cropped volume.

The cropped, binarized bone volume was subjected to a D derivative operator to produce the edges of the bone structures. The vertebral column is not continuous but made up of interlocking elements; As a result, the bone edge map could be sparse. The Hough transform for the detection of circles. The radius in the Hough space was limited to the range 6 to10mm. Because of the possibility of partial structures and edges in a given image, the global maximum in the Hough space may not relate to the inner circular edge of the spinal canal, as desired. In order to obtain the center and radius of the ROI, the CT values of bone marrow μ≈142HU and σ≈48HU and the spinal canal μ≈+30HU and σ≈8HU, were used as constraints. If the center of the circle corresponding to the Hough space maximum was not within the specified HU range, the circle was rejected and the next maximum in the Hough space was evaluated. This process was continued until a suitable circle was detected. The best fitting circle, which was not given by the global maximum in the Hough space, was obtained by applying the constraints defined above,

The centers of the circles detected as above were used as the seed voxels in a fuzzy connectivity algorithm to segment the spinal canal. The mean and standard deviation required for this procedure were estimated using a7x7x2 neighborhood around each seed voxel. The spinal canal volume was then removed from the CT volume, resulting in improved segmentation of the tumor volume.

5.4 DETECTION OF CALCIFICATIONS BY MULTI-TOLERANCE REGION GROWING

Detection of Micro calcification is the major hurdle present in the tumour segmentation process in mammogram analysis. To identify the breast cancer early stage itself we go for mammogram Imaging. If the patient has calcium sediments present in the breast region or in the ductal nodes then it can be easily mis represented has carcinoma tissue (cancer tissue).To avoid this confusion we have to go for the exact detection and isolation of the microcalcifications. Calcifications or small calcium deposits which are shown as a minute dots in mammograms as marked as a small circle in Figure 5.2.

To identify these regions, we can use any segmentation procedure. In this example we will execute Region growing procedure.

The algorithm for region growing based Microcalcification detection process is listed as below

- Select any pixel inside the mammogram as seed pixel

- Compare the 4 neighbors of the seed pixel. If any of the neighbors are having same intensity values group them. For grouping / clustering we can use any clustering algorithm. To compute the difference of pixel value of the initial seed point pi and its neighboring points, if the difference is smaller than the threshold (criterion) we define, the neighboring point could be classified into group Ci.
- Recompute the boundary of Ci and set those boundary points as new seed points pi (s). In addition, the mean pixel values of Ci have to be recomputed
- Repeat Step2 and 3 until all pixels in image have been allocated to a suitable cluster
- Since the micro dots have more or less same intensity values it will be grouped into a single group Ci. Then by using thresholding concept we can separate that particular group. Thus calcifications can be easily identified.

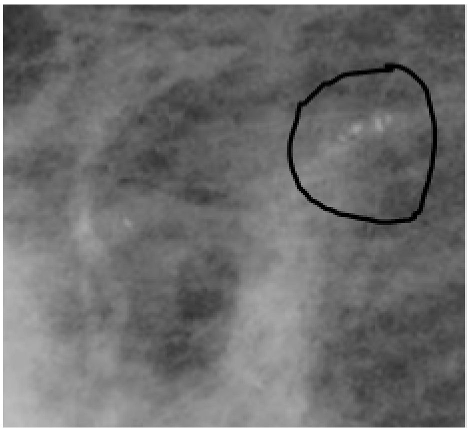


Figure 5.2: Calcium deposit present in mammograms

5.5 SHAPE ANALYSIS OF CALCIFICATIONS

Breast calcifications is the common findings done on a mammography and it can be of both benign and malignant type. Benign is not a life threatening one where as Malignant type is life threatening. Malignant tumor regions have to be identified at the early stage or they can spread into other regions of body. In Figure 5.3 three (3) micro calcifications are shown. It's hard to detect since they are overlapping with other breast muscles or ductal nodes. Effective and accurate identification of these dots increases the chance of early detection there by saving the patient. Figure 5.4 shows the morphology of micro calcifications.

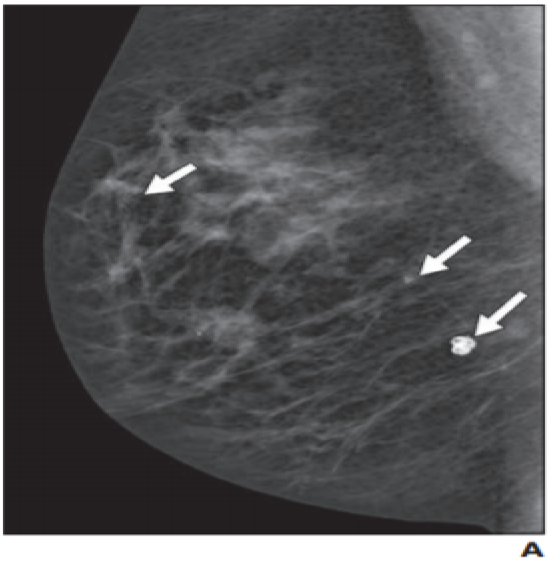


Figure 5.3: Micro Calcium clusters in a mammogram

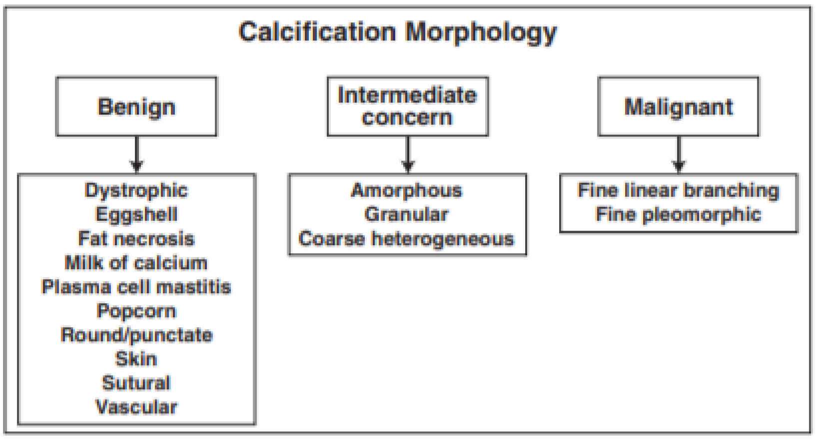


Figure 5.4: Flowchart for morphology of breast calcifications

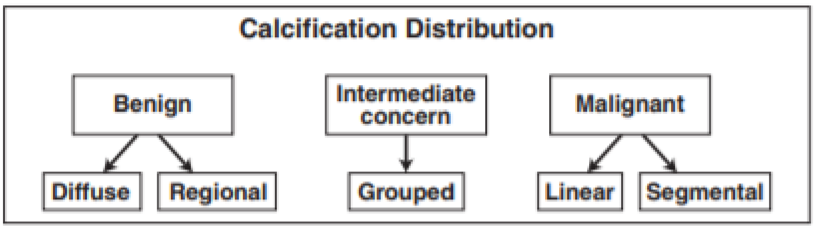


Figure 5.5: Flowchart for different types of calcifications

Figure 5.5 represents the different types of calcification's present in a mammogram. The benign type of calcium deposits will be more or less diffused / spread around the region while the malignant one's will have shar edges and of irregular shapes. With the help of the shape of these deposits we can easily classify them as benign or malignant tissues. For the identification of the shape different segmentation algorithms can be used as listed below

- Region based segmentation
- Contour (snake) based algorithms
- Thresholding based segmentation algorithm
- Morphological based segmentation procedures can be used

After segmentation The lesions has to be characterized by region descriptors or Teture features or shape features like Area, MajorAxisLength, MinorAxisLength, Solidity, Eccentricity, Centroid, Perimeter, ConvexArea etc.

Citations/Acknowledments: The images used in this sub section (5.3,5.4 &5.5) were adapted from https://www.ajronline.org/doi/pdf/10.2214/AJR.10.5732.

Question Bank:

Part- A

1. Justify the need for identification of microcalcification

2. Illustrate the concept of mammogram shape identification

3. Interpret the concept of spinal canal detection

4. Comment on any one procedure involved in enhancement of mammogram

5. Infer the effect of noises in medical images

Part- B

1. Appraise on different Morphology based segmentation operators used to isolate the cancer tissue present in mammogram.

2. Enumerate how Contrast enhancement increases the survivability of breast cancer patient

3. Appraise on how Micro calcifications are highlighted by using region based segmentation algorithm

4. Discuss in detail how the Boundary regions are extracted from spinal canal

Acknowledgement/Citation: some of the concepts , images used in this material were adopted from "*Biomedical Image Analysis,* Rangaraj M. Rangayyan".

TEXT / REFERENCE BOOKS

1. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2nd Edition, Pearson Education, Inc., 2004.

2. Anil K. Jain, "Fundamentals of Digital Image Processing", PHI Learning Private Limited, New Delhi, 2002.

3. William K. Pratt, "Digital Image Processing", 3rd Edition, John Wiley & Sons, Inc., 2001.

4. Bernd Jähne, "Digital Image Processing", 5th Revised and Extended Edition, Springer, 2002.