# UNIT-1

# Introduction and Image Transforms

## 1.1 Introduction

An image is nothing more than a two dimensional signal. It is defined by the mathematical function f(x,y) where x and y are the two co-ordinates horizontally and vertically and the amplitude of f at any pair of coordinate (x, y) is called the intensity or gray level of the image at that point. When x, y and the amplitude values of f are all finite discrete quantities, we call the image a digital image.

The field of digital image processing refers to the processing of digital image by means of a digital computer. A digital image is composed of a finite number of elements, each of which has a particular location and values of these elements are referred to as picture elements, image elements, pels and pixels.

**Applications**

Some of the major fields in which digital image processing is widely used are mentioned below

- Gamma Ray Imaging- Nuclear medicine and astronomical observations.
- X-Ray imaging – X-rays of body.
- Ultraviolet Band –Lithography, industrial inspection, microscopy, lasers.
- Visual and Infrared Band – Remote sensing.
- Microwave Band – Radar imaging.
- Radio Band – Medicine and Astronomy

## 1.2 Origins of Digital Image Processing

Initially, digital image processing was employed in the newspaper industry, when pictures were first sent by submarine cable between London and New York. Introduction of the Bartlane cable picture transmission system in the early 1920s reduced the time required to transport a picture across the Atlantic from more than a week to less than three hours. Specialized printing equipment coded pictures for cable transmission and then reconstructed them at the receiving end.

But there were two major problems in improving the visual quality of these early digital pictures were related to the selection of printing procedures and the distribution of intensity levels. To overcome these problems photographic reproduction made from tapes perforated at the telegraph receiving terminal. This method improves both in tonal quality and in resolution.

In 1929 the image coding capability of Bartlane systems was increased to 15 levels from five distinct levels of gray. The history of digital image processing is intimately tied to the development of the digital computer. In fact, digital images require so much storage and computational power that progress in the field of digital image processing has been dependent on the development of digital computers and of supporting technologies that include data storage, display, and transmission.

In 1940's John von Neumann introduced of two key concepts. They are (1) a memory to hold a stored program and data, and (2) conditional branching. These two ideas are the foundation of a central processing unit (CPU), which is at the heart of computers. Starting with von Neumann, there were a series of key advances that led to computers powerful enough to be used for digital image processing. Briefly, these advances may be summarized as follows:

- The invention of the transistor by Bell Laboratories in 1948
- The development in the 1950s and 1960s of the high-level programming languages COBOL (Common Business-Oriented Language) and FORTRAN (Formula Translator)
- The invention of the integrated circuit (IC) at Texas Instruments in 1958
- The development of operating systems in 1960
- The development of the microprocessor (a single chip consisting of the central processing unit, memory, and input and output controls) by Intel in the year 1970.
- Introduction by IBM of the personal computer in 1981
- Progressive miniaturization of components, starting with large scale integration (LI) in the late 1970s and then very large scale integration (VLSI) in the 1980s, to the present use of ultra large scale integration (ULSI).

In addition to these advances there were developments in the areas of mass storage and display systems, both of which are fundamental requirements for digital image processing. The first computers powerful enough to carry out meaningful image processing tasks appeared in the early 1960s. In 1964 pictures of the moon transmitted by *Ranger 7* were processed by a computer to correct various types of image distortion inherent in the on-board television camera.

### 1.3 Fundamental Steps in Digital Image Processing

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

- Methods whose input and outputs are images.
- Methods whose inputs may be images but outputs are attributes extracted from those images.

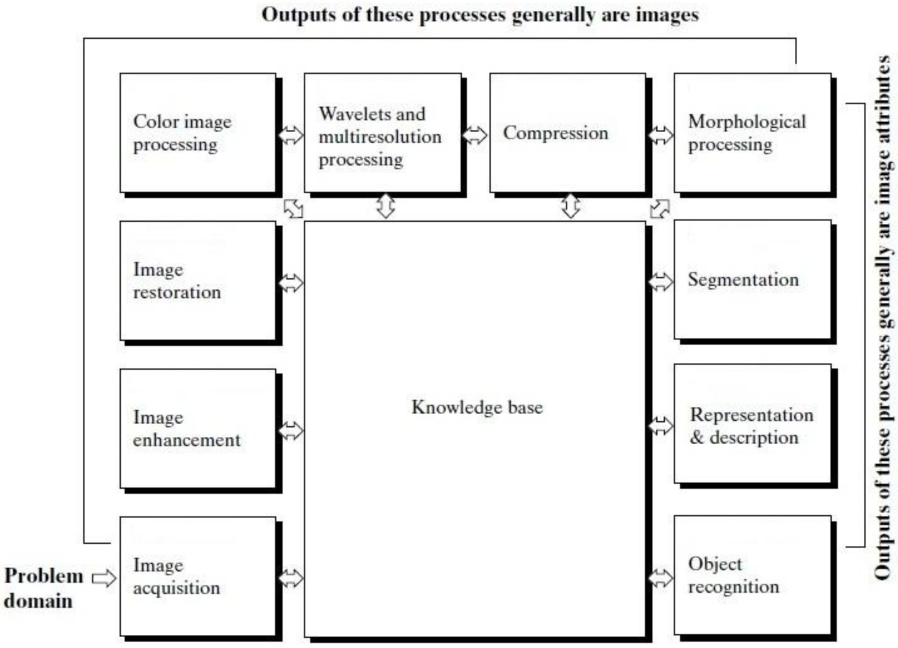



Fig: Fundamental Steps in Digital Image Processing

***Image Acquisition*** could be as simple as being given an image that is already in digital form. Generally, this stage involves preprocessing, such as scaling.

***Image Enhancement*** is used to bring out detail that is obscured, or simply to highlight certain features of interest in an image. It is a very subjective area of image processing. Enhancement is based on human subjective preferences regarding what constitutes a "good" enhancement result. A familiar example of enhancement is when we increase the contrast of an image because to looks better."

***Image Restoration*** is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

***Color Image Processing*** is an area that has been gaining in importance because of the significant increase in the use of digital images over the Internet.

***Wavelets and Multi-resolution Processing*** is the foundation for representing images in various degrees of resolution. It is employed for image data compression and for pyramidal representation where images are subdivided into successively into smaller regions.

***Compression*** deals with techniques for reducing the storage required saving an image, or the bandwidth required transmitting it. Image compression is familiar to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

***Morphological Processing*** deals with tools for extracting image components that are useful in the representation and description of shape and boundary of objects. It is majorly used in automated inspection applications.

***Segmentation*** is defined as portioning an input image into its constituent parts or objects. It is very important to distinguish between different objects in an image as in the case of systems employed for traffic control, or crowd control. In character recognition, the key role of segmentation is to extract individual characters and words from the background.

***Representation and Description*** always follow the output of a segmentation stage. It is a process which transforms raw data into a form suitable for subsequent computer processing. The first decision is to choose between boundary representation and regional representation. Boundary representation is used when the details of external shape characteristics is important where as the regional representation is used when the internal properties are important.

***Recognition*** is the process that assigns a label to an object based on its descriptors. The information provided by its descriptors and the recognized object is interpreted by assigning a meaning to it.

***Knowledge Base*** is used to guide the operation of each processing module and control the interaction between them. It stores the information about a problem domain is coded into an image processing system. This knowledge may be as simple as detailing regions of an image where the information of the interest in known to be located. The knowledge base also can be quite complex such interrelated list of all major possible defects in a materials inspection problems or an image database containing high resolution satellite images of a region in connection with change detection application.

## 1.4 Components of Image Processing System

The basic components comprising a typical *general-purpose* system used for digital image processing is shown in the following figure

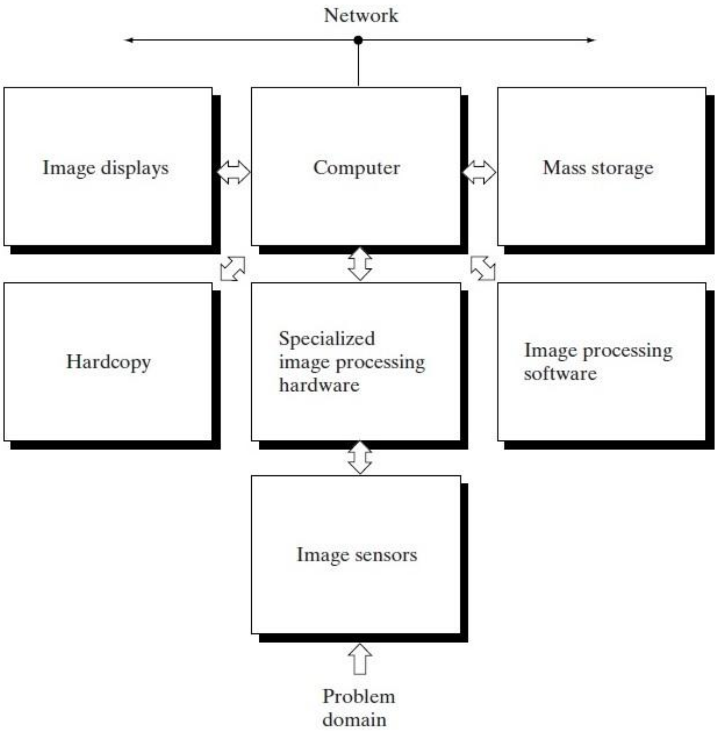



Fig: Components of a general-purpose image processing system

**Image Sensors**

With reference to sensing, two elements are required to acquire digital image. The first is a physical device that is sensitive to the energy radiated by the object we wish to image. The second is a digitizer which converts the output of the physical sensing device into digital form.

**Specialized Image Processing Hardware**

It consists of the digitizer plus hardware that performs other primitive operations such as an arithmetic logic unit, which performs arithmetic such addition and subtraction and logical operations in parallel on images.

**Computer**

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

**Software**

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

**Mass storage**

The storage capability is must in image processing applications. An image of size 1024 x 1024 pixels, in which the intensity of each pixel is an 8- bit quantity requires one megabytes of storage space if the image is not compressed. Digital Image processing applications falls into three principal categories of storage

- Short term storage for use during processing is computer memory
- On line storage for relatively fast recall are magnetic disks or optical media storage
- Archival storage such as magnetic tapes and disks

**Image Displays**

Image displays in use today are mainly color TV monitors. These monitors are driven by the outputs of image and graphics displays cards that are an integral part of computer system. In some cases, it is necessary to have stereo displays, and these are implemented in the form of headgear containing two small displays embedded in goggles worn by the user.

**Hardcopy Devices**

The devices for recording image includes laser printers, film cameras, heat sensitive devices inkjet units and digital units such as optical and CD ROM disk. Films provide the highest possible resolution, but paper is the obvious medium of choice for written applications. For presentations, images are displayed on film transparencies or in a digital medium if image projection equipment is used.

**Networking**

It is almost a default function in any computer system in use today because of the large amount of data inherent in image processing applications. The key consideration in image transmission is bandwidth. But communications with remote sites via the Internet are not always as efficient. This situation is overcome by improving optical fiber and other broadband technologies.

### 1.5 Elements of Visual Perception

The digital image processing field is built on a foundation of mathematical and probabilistic formulations but human visual perception is also has equal importance.

**Structure of the Human Eye**

The horizontal cross section of the human eye is nearly a sphere, with an average diameter of approximately 20 mm. The eye is enclosed with three membranes

a) The cornea and sclera: The cornea is a tough, transparent tissue that covers the anterior surface of the eye. Continuous with the cornea, the sclera encloses the reminder of the optic globe.

b) The choroid: It lies directly below the sclera and contains a network of blood vessels that serve as the major source of nutrition to the eyes. It helps to reduce extraneous light entering in the eye. It consists of two parts the Iris and Ciliary Body. The front of the iris contains the visible pigment of the eye, whereas the back contains a black pigment. The Iris contracts or expands to control the amount of light that enters the eyes. The Ciliary Body contains 60 t0 70% of water and more protein than any other tissue in the eye.

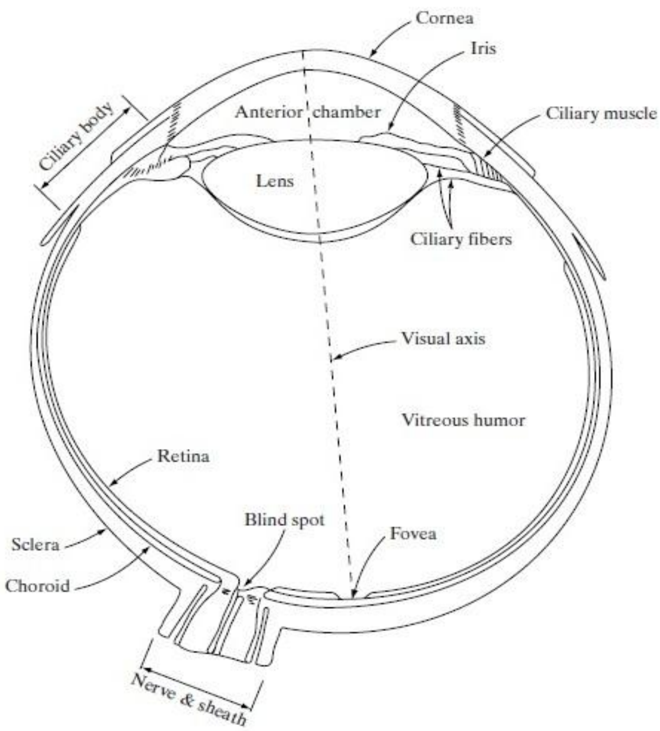



Fig: Cross Section of Human Eye

(c) Retina: It is innermost membrane of the eye. When the eye is properly focused, light from an object outside the eye is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina. There are two major classes of the receptors are,

- Cones: It is in the number about 6 to 7 million for each eye. These are located in the central portion of the retina called the fovea are highly sensitive to color. Human can resolve fine details with these cones because each one is connected to its own nerve end. Cone vision is called photopic or bright light vision.
- Rods: These are very much in number from 75 to 150 million and are distributed over the entire retinal surface. The large area of distribution and the fact that several roads are connected to a single nerve give a general overall picture of the field of view. They are not involved in the color vision and are sensitive to low level of illumination. Rod vision is called is scotopic or dim light vision. The absent of reciprocators is called blind spot.

**Image Formation in the Eye**

In an ordinary photographic camera, the lens has a fixed local length and focusing at various distances achieved by varying the distance between the lens and the image plane, where the film is located. In the human eye, the distance between the lens and imaging region 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 flattering or thickening the lens for distant or near objects, respectively. The refractive power of the lens is increased from its minimum to maximum value, when the distance between the centre of the lens and the retina along the visual axis is varies from 17mm to 14mm approximately. The lens exhibits lowest refractive power, when the eye is focused on objects at distances greater than about 3m. Similarly, the lens exhibits highest refractive power, when the eye is focused on near object.

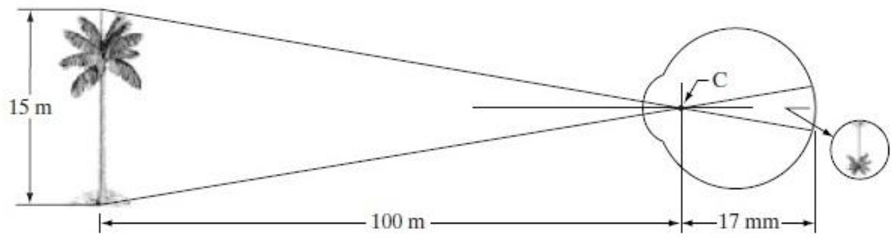



**Fig: Graphical representation of the eye looking at a tree. Point C is the optical center of the lens**

The above geometry illustrates how to obtain the dimensions of image formed on the retina. For example, the observer is looking at a tree 15 m high at a distance of 100 m. If h is the height in mm of that object in the retinal image, the geometry of Fig. yields 15/100=h/17 or h=2.55 mm. The retinal image is reflected primarily in the area of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

**Brightness Adaptation and Discrimination**

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

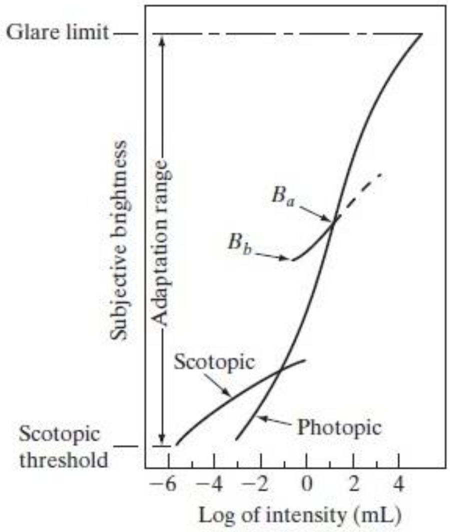
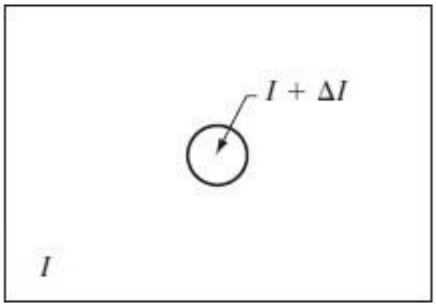



Fig: Range of subjective brightness sensations showing a particular adaptation level.

The curve represents the range of intensities to which the visual system can adopt. But the visual system cannot operate over such a dynamic range simultaneously. Rather, it is accomplished by change in its overcall sensitivity called brightness adaptation. The total range of distinct intensity levels it can discriminate simultaneously is rather small when compared with the total adaptation range. For any given set of conditions, the current sensitivity level to which of the visual system is called brightness adoption level , $B_a$ in the curve. The small intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level. It is restricted at level $B_b$, at and below which all stimuli are perceived as indistinguishable blacks. The upper portion of the curve is not actually restricted. Whole simply raise the adaptation level higher than $B_a$. The ability of the eye to discriminate between changes in light intensity at any specific adaptation level is also of considerable interest.

Take a flat, uniformly illuminated area large enough to occupy the entire field of view of the subject. It may be a diffuser such as an opaque glass, that is illuminated from behind by a light source whose intensity, I can be varied. To this field is added an increment of illumination ΔI in the form of a short duration flash that appears as circle in the center of the uniformly illuminated field.




If ΔI is not bright enough, the subject cannot see any perceivable changes. As ΔI gets stronger the subject may indicate of a perceived change. $\Delta I_c$ is the increment of illumination discernible 50% of the time with background illumination I. Now, $\Delta I_c$ /I is called the Weber ratio. Small value means that small percentage change in intensity is discernible representing "Good" brightness discrimination. Large value of Weber ratio means large percentage change in intensity is required representing "poor brightness discrimination".

### 1.6 Light and the Electromagnetic Spectrum

In 1666, Sir Isaac Newton discovered that when a beam of sunlight is passed through a glass prism, the emerging beam of light is not white but consists instead of a continuous spectrum of colors ranging from violet to red.

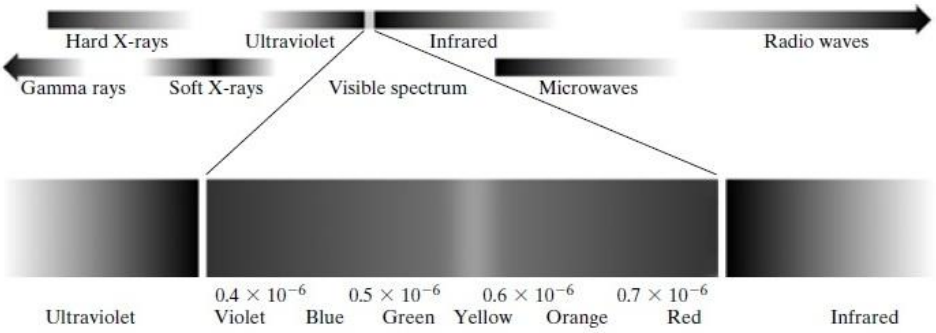



Fig: The electromagnetic spectrum

The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy. Wavelength (λ) and frequency ($v$) are related by the expression

$$\lambda = \frac{c}{v}$$

Where c is the speed of light ($2.998\times10^8$ m/s).The energy of the various components of the electromagnetic spectrum is given by the expression

$$E=hv$$

Where *h* is Planck's constant. The units of wavelength are meters, with the terms *microns* and *nanometers*. Frequency is measured in Hertz (Hz), with one Hertz being equal to one cycle of a sinusoidal wave per second. A commonly used unit of energy is the electron-volt.

Electromagnetic wave can be defined as a stream of massless particles, each traveling in a wavelike pattern and moving at the speed of light. Each massless particle contains a certain amount of energy. Each bundle of energy is called a *photon*. This energy is proportional to frequency, so the higher-frequency electromagnetic wave carries more energy per photon.

Light is a one type of electromagnetic radiation that can be seen and sensed by the human eye. The visible band of the electromagnetic spectrum spans the range from approximately 0.43 µm (violet) to about 0.79 µm (red). The color spectrum is divided into six broad regions: violet, blue, green, yellow, orange, and red. There are two types of lights,

- **Monochromatic Light** is a void of color that defines in terms of intensity of gray level. It ranges from black to grays and finally to white. The monochromatic images are called as Gray-Scale Images.
- **Chromatic Light** contains colors from violet to red. The quality of a chromatic light can be described by three basic quantities,
    - **Radiance** is the total amount of energy that flows from the light source, and it is usually measured in watts (W).
    - **Luminance** is a measure of the amount of energy an observer perceives from a light source, and it is measured in terms of lumens (lm).
    - **Brightness** is a subjective descriptor of light perception and it cannot be measured practically.

### 1.7 Image Sensing and Acquisition

Images 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. The illumination energy is reflected from the object depending upon the nature of the source. For example the light reflected from a planar surface, X-rays pass through a patient's body for the purpose of generating a diagnostic X-ray film and the reflected or transmitted energy is focused onto a photo converter, which converts the energy into visible light. Electron microscopy and some applications of gamma imaging use this approach.

There are three principal sensor arrangements used to transform illumination energy into digital images. 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.

**(a) Image Acquisition Using a Single Sensor:**

The components of a single sensor are shown in the following figure. 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.

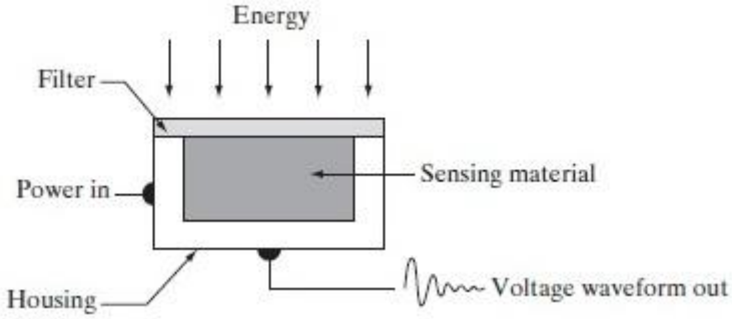



Fig: Single imaging sensor.

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. The following figure shows an arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension.

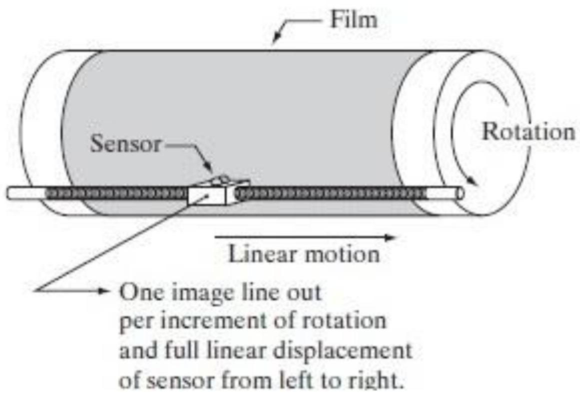



Fig: Combining a single sensor with motion to generate a 2-D image.

The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Since 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 are referred as "Microdensitometer's."

**(b) Image Acquisition Using Sensor Strips:**

A geometry that is used much more frequently than single sensors consists of an in-line arrangement of sensors in the form of a sensor strip, as shown in figure.

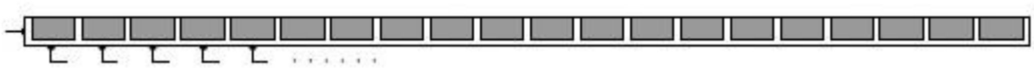


Fig: Line sensor

The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction, as shown in the following figure. This type of arrangement used in most flat bed scanners. Sensing devices with 4000 or more in-line sensors are possible.

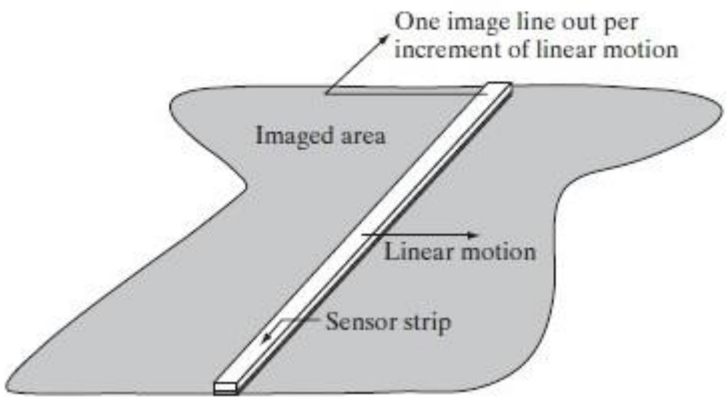



Fig: Image acquisition using a linear sensor strip

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 shown in the following figure.

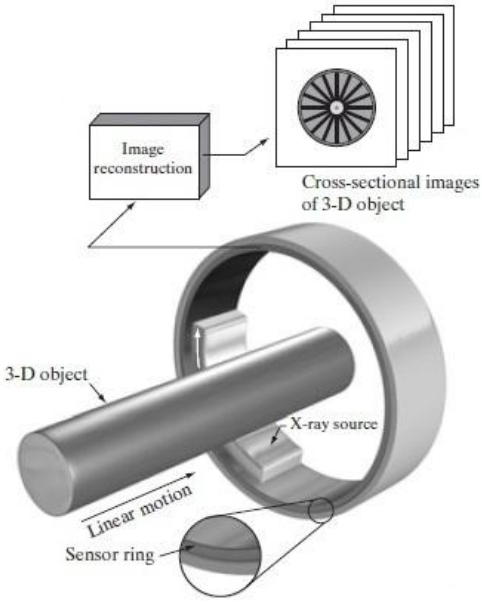



Fig: Image acquisition using a circular sensor strip.

A rotating X-ray source provides illumination and the portion of the sensors opposite the source collect the X-ray energy that pass through the object. This is the basis for medical and industrial computerized axial tomography (CAT) imaging. Hence the output of the sensors must be processed by reconstruction algorithms whose objective is to transform the sensed data into meaningful cross-sectional images. However, images are not obtained directly from the sensors by motion alone; they require extensive processing. A 3-D digital volume consisting of stacked images is generated as the object is moved in a direction perpendicular to the sensor ring. Other modalities of imaging based on the CAT principle include magnetic resonance imaging (MRI) and positron emission tomography (PET).The illumination sources, sensors, and types of images are different, but conceptually they are very similar to the basic imaging.

**(c) Image Acquisition Using Sensor Arrays:**

Numerous electromagnetic and some ultrasonic sensing devices are used the sensors arranged in 2-D 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 4000 × 4000 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.

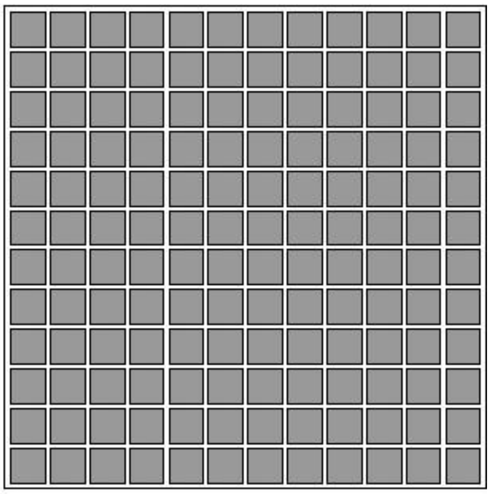


Fig: Array sensor

Since the sensor array shown in figure is two dimensional, which useful to obtain complete image by focusing the energy pattern onto the surface of the array. The following figure shows that the energy from an illumination source being reflected from a scene element and also transmitted through the scene elements. The first function performed by the imaging system shown in figure (c) is to collect the incoming energy and focus it onto an image plane.

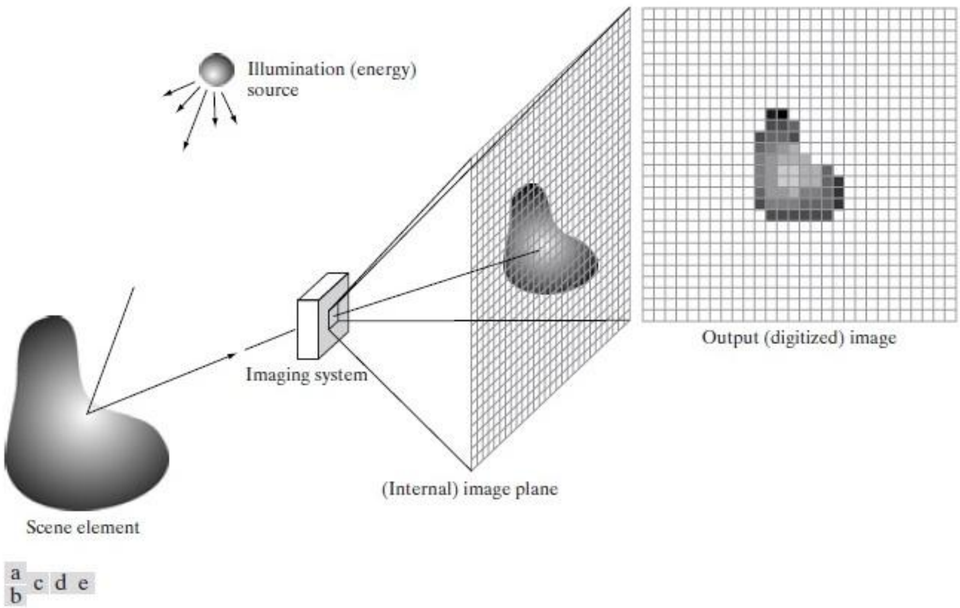



**Fig:** An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

If the illumination is light, the front end of the imaging system is a lens, which projects the viewed scene onto the lens focal plane. 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 sweeps these outputs and converts them to a video signal, which is then digitized by another section of the imaging system. The output is a digital image, as shown diagrammatically in Fig. (e).

**1.9 Image Sampling and Quantization**

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

**Generating a Digital Image**

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.

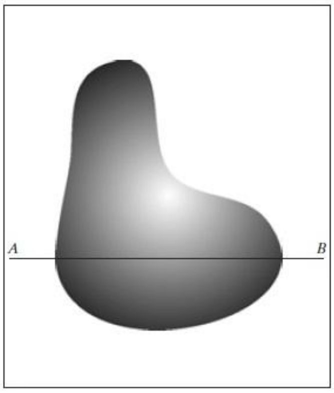



Fig: (a) Continuous Image

Let us take the plot of amplitude (gray level) values of the continuous image along the line segment AB in Fig. (a).

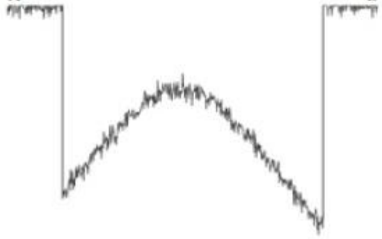



Fig. (b) A scan line from A to Bin the continuous image, used to illustrate the concepts of sampling and quantization

The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in Fig. (c). the location of each sample is given by a vertical tick mark in the bottom part of the figure. The samples are shown as small white squares superimposed on the function.

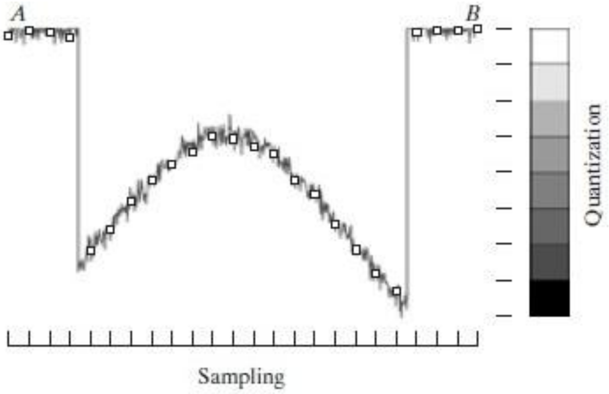



Fig: (c) Sampling and quantization

The set of these discrete locations gives the sampled function. However, the values of the samples still span (vertically) a continuous range of gray-level values. In order to form a digital function, the gray-level values also must be converted (quantized) into discrete quantities. The right side of Fig.(c) shows the gray-level scale divided into eight discrete levels, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight gray levels. The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown in Fig. (d).

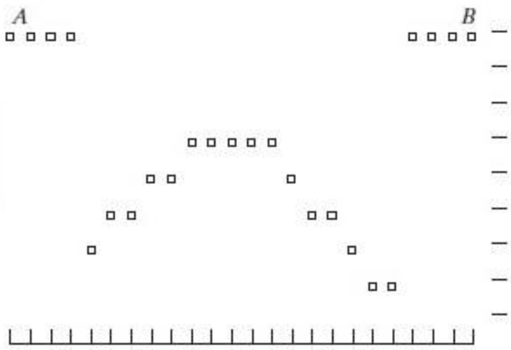



Fig: (d) Digital scan line

Starting at the top of the image and carrying out this procedure line by line produces a two-dimensional digital image.

**1.8 Representing Digital Images**

An image is represented 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) represents the intensity or gray level of image at that point. In a digital image both the coordinates and the amplitude values of *f* are all finite and discrete quantities. Hence, a digital image is composed of a finite number of elements, each of which has a particular location value. These elements are called image elements, picture elements or pixels.

An image f(x, y) is sampled so that the resulting digital image has M rows and N columns. Hence a digital image can be considered as a matrix whose rows and columns indices identify a point on the image and the corresponding matrix element value identifies the gray level at that point. An image can be represented as

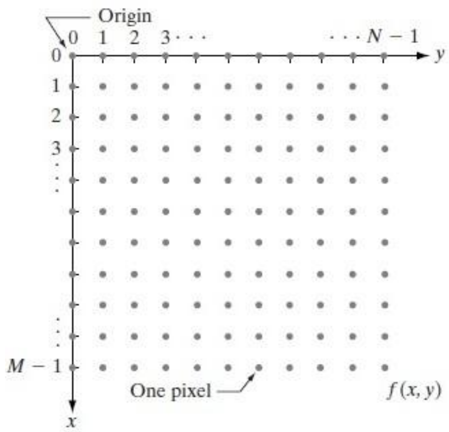



Fig: Coordinate convention used to represent digital images

The complete M×N digital image represented in the following compact matrix form:

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

Each element of this matrix array is called an image element, picture element, pixel, or pel. In general, the matrix notation to denote a digital image and its elements as

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

Where $a_{ij}$=f(x=i, y=j)=f(i, j)

The digitization process requires decisions about values for M, N, and for the number, L, of discrete gray levels allowed for each pixel. However, due to processing, storage, and sampling hardware considerations, the number of gray levels is an integer power of 2.

$$L = 2^k$$

We assume that the discrete levels are equally spaced and that they are integers in the interval [0, L-1]. Sometimes the range of values spanned by the gray scale is called the *dynamic range* of an image. The images having a high dynamic range will exhibit high contrast. Conversely, an image with low dynamic range tends to have a dull, washed out gray look. The number, b, of bits required to store a digitized image is

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

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

When an image can have $2^k$ gray levels, it is referred to as "k- bit". An image with 256 possible gray levels is called an "8- bit image" ($256=2^8$).

### 1.9 Spatial and Gray Level Resolution

**Spatial resolution** is the smallest discernible details are an image. Spatial resolution can be defined with line pairs per unit distance and pixels per unit distance. Consider a 1024× 1024 pixel image with 256 gray level digital image. Now reduce the number of pixels from 1024 to 512 pixels in each row and column of the image matrix without any change in the gray level. Hence the image will nearly be same and it is very difficult to find any differences between the two.

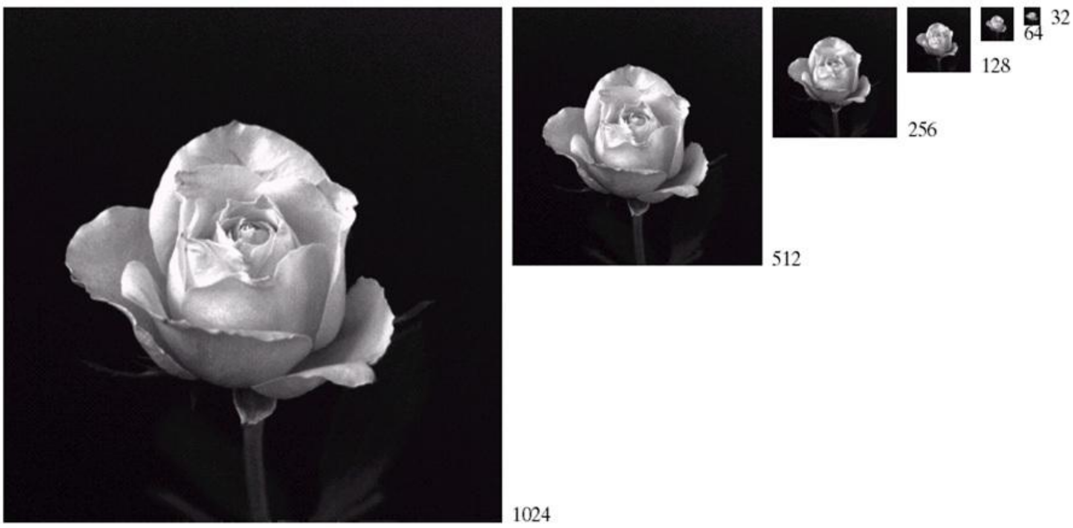



Fig: A 1024×1024, 8-bit image sub sampled down to size 32×32 pixels. The number of allowable gray levels was kept at 256.

The numbers of pixels are further reduced to 256, 128, 64 and 32 respectively by keeping the gray level constant and display area is kept the same as for 1024× 1024 display field. So, very slight check board pattern appear through the image begins to appear in the 256×256 image which begins to increase in the 128×128 image which becomes more in the 64×64 and the 32×32 image respectively.

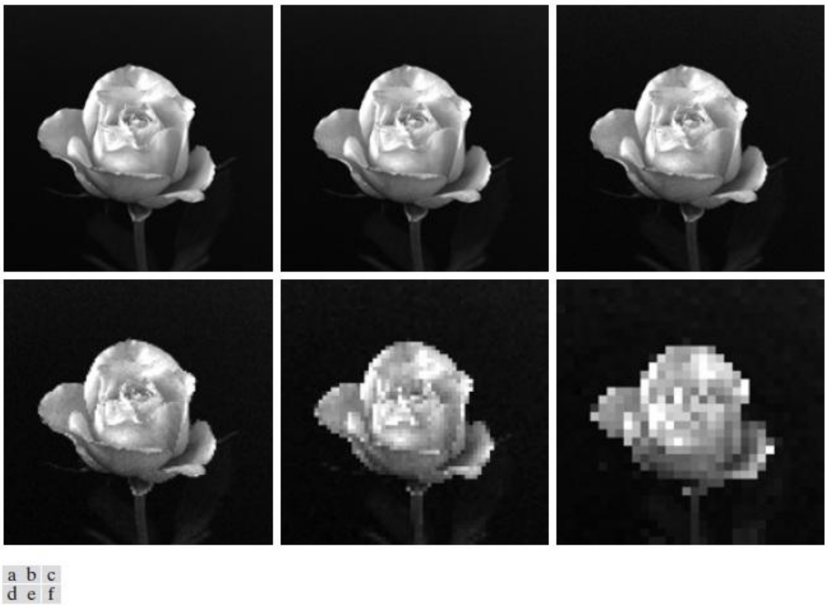



Fig: (a) 1024×1024, 8-bit image. (b) 512×512 image resampled into 1024×1024 pixels by row and column duplication. (c) Through (f) 256×256, 128×128, 64×64, and 32×32 images resampled into 1024×1024 pixels

**Gray level** resolution refers to smallest discernible change in gray levels. Consider an image of 1024× 1024 pixel and 256 gray levels are reduced from 256 to 2 by varying the value of k from 8 to 1 while keeping the spatial resolution constant.

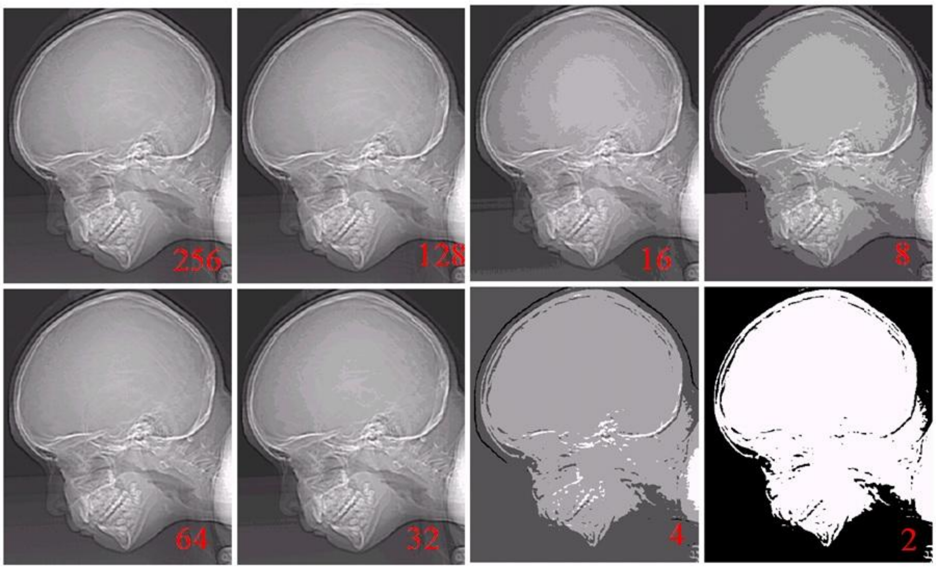



Fig: A 1024×1024, image displayed in 256,128 to 2 Gray levels while keeping the spatial resolution constant.

Now reduce the gray level of the image from 256 to 128 but it looks identical. But, the number of gray levels are further decreased a set of fine bridge structures are developed in areas of smooth gray levels. This effect caused by the use of insufficient number of gray levels is called false contouring. It is generally quite visible in image displayed using 16 or less uniformly spaced gray levels.

### 1.10 Isopreference Curves

An image with M×N pixels and 'L' gray levels and we can obtain the different images by varying the values of N and k ($L=2^k$) are visualized to the observers, who can rank those images using their subjective quality. The results obtained from these rankings are produced in a curve form. These curves are called "Isopreference curves".

Let us consider three images. The woman's face is representative of an image with relatively little detail; the picture of the cameraman contains an intermediate amount of detail; and the crowd picture contains, by comparison, a large amount of detail.

a b c

Fig: (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail.

Sets of these three types of images were generated by varying *N* and k, and observers were then asked to rank them according to their subjective quality. Results were summarized in the form of so-called *isopreference curves* in the Nk-plane.

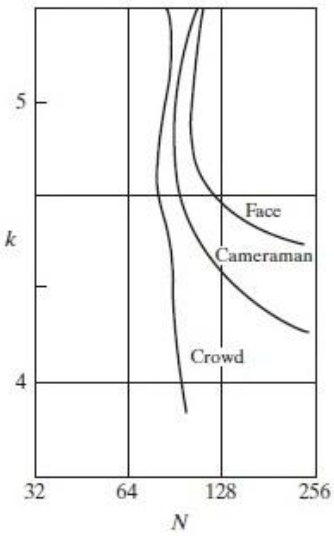



Fig: Representative Isopreference curves for the three types of images.

From the above figure we can observe that Isopreference curve for the crowd is nearly vertical. This indicates that, for a fixed value of *N*, the perceived quality for this type of image is nearly independent of the number of gray levels used. It is also note that perceived quality in the other two image categories remained the same in some intervals in which the spatial resolution was increased, but the number of gray levels actually decreased. Hence this result is that a decrease in k tends to increase the apparent contrast of an image, a visual effect that humans often perceive as improved quality in an image.

**1.11 Zooming and Shrinking Digital Images**

Zooming may be defined as oversampling and shirking as under sampling these techniques are applied to a digital image. There are two steps of zooming,

- Creation of new pixel locations
- Assignment of gray level to those new locations

In order to perform gray –level assignment for any point in the overly, we look for the closet pixel in the original image and assign its gray level to the new pixel in the grid. When we are done with all points in the overlay grid, we simply expand it to the original specified size to obtain the zoomed image. This method of gray-level assignment is called *nearest neighbor interpolation*.

*Pixel replication* is a method used to generate a special case of nearest neighbor interpolation. Pixel replication is applicable when we want to increase the size of an image an integer number of times. If we want to double the size of an image, we can duplicate each column and row. The same procedure is used to enlarge the image by any integer number of times (triple, quadruple, and so on). Duplication is just done the required number of times to achieve the desired size. The gray-level assignment of each pixel is predetermined by the fact that new locations are exact duplicates of old locations. Although nearest neighbor interpolation is fast but it has a undesirable feature that it produces a check board effect at high factors of magnification. To overcome this problem, the gray-level assignment is done by *bilinear interpolation* using the four nearest neighbors of a point. Let (x', y') denote the coordinates of a point in the zoomed image and let $v$(x', y') denote the gray level assigned to it. For bilinear interpolation, the assigned gray level is given by

$$v\,(x', y') = ax' + by' + cx'y' + d$$

Image shrinking is similar to zooming but the equivalent process of pixel replication is row-column deletion. For example, to shrink an image by one-half, we delete every other row and column. It may be leads to aliasing effect. To reduce possible aliasing effect by blur an image slightly before shrinking it.

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

**Connectivity:**

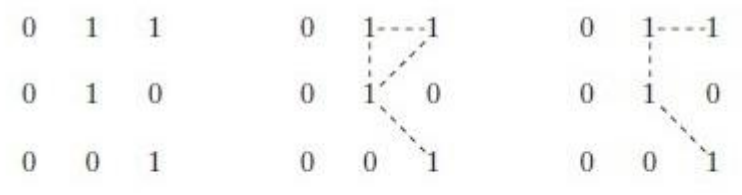
The relationship between two or more pixels is defined by pixel connectivity. Connectivity information is used to establish the boundaries of the objects. The pixels p and q are said to be connected if certain conditions on pixels brightness specified by the set *V* and spatial adjacency are satisfied. For a binary image, this set *V* will be {0,1} and for grey scale images might be a range of possible gray-level values 0 to 255, set *V* could be any subset of these 256 values.

**4-*Connectivity:*** The pixels p and q are said to be in 4-connectivity when both have the same values as specified by the set *V* and if q is said to be in the set $N_4$ (p). This implies any path from *p* to *q* on which every other pixel is 4-connected to the next pixel.

***8-Connectivity:*** It is assumed that the pixels *p* and *q* share a common grey scale value. The pixel p and q are said to be 8-connectivity if q is in the set $N_8$ (p).

**M-Connectivity*:*** It is also known as "Mixed Connectivity". The two pixels p and q are said to be in m-connectivity when

- q is in $N_4$(p), *or*
- q is in $N_D$(p) *and* the set has no pixels whose values are from *V*.



a b c

(a) Arrangement of pixels (b) 8-connectivty to the center pixel (c) m- connectivity

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

**Adjacency:**

Let *V* be the set of gray-level values used to define adjacency. In a binary image V={1} if we are referring to adjacency of pixels with value 1. There are 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 $N_4$(p).

(b) 8-*adjacency*. Two pixels p and q with values from *V* are 8-adjacent if q is in the set $N_8$(p).

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

- q is in $N_4$(p), *or*
- q is in $N_D$(p) *and* the set has no pixels whose values are from *V*.

Let *R* be a connected subset of pixels in an image, then R is called as *region* of the image. The *boundary* (also called *border* or *contour*) of a region *R* is the set of pixels in the region that have one or more neighbors that are not in *R*. If *R* happens to be an entire image then its boundary is defined as the set of pixels in the first and last rows and columns of the image.

**Distance Measures:**

Let us consider the pixels p, q, and z, with coordinates (x, y), (s, t), and (v, w), respectively. D is a distance function if

**(a)** D (p, q) ≥0 (D (p, q) =0 if p=q)

**(b)** D (p, q) =D (q, p), and

**(c)** D (p, z) ≤ D (p, q) +D (q, z).

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

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

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).

The ***$D_4$ distance*** **or** ***city-block distance*** between p and q is defined as

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

The ***$D_8$ distance*** **or** ***chessboard distance*** between p and q is defined as

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

**1.13 Aliasing and Moiré Patterns**

The function is sampled at a rate equal to or greater than twice its highest frequency, it is possible to recover completely the original function from its samples. If the function is *under sampled*, then *aliasing* corrupts the sampled image. The corruption is in the form of additional frequency components being introduced into the sampled function. These are called *aliased frequencies*. The *sampling rate* in images is the number of samples taken (in both spatial directions) per unit distance.

The principal approach for reducing the aliasing effects on an image is to reduce its high-frequency components by blurring the image *prior* to sampling. However, aliasing is always present in a sampled image. The effect of aliased frequencies can be seen under the right conditions in the form of so called *Moiré patterns*. When a function is periodic, it may be sampled at a rate equal to or exceeding twice its highest frequency and it is possible to recover the function from its samples *provided* that the sampling captures *exactly* an integer number of periods of the function. This special case allows us to illustrate the Moiré effect.

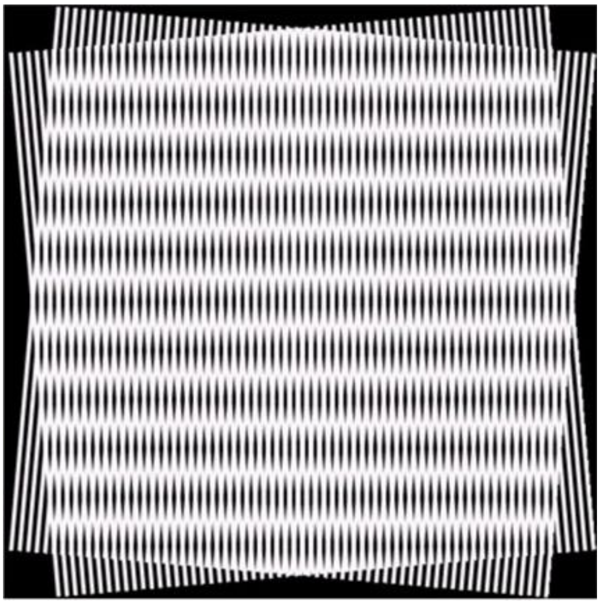


Fig: Illustration of the Moiré pattern effect

The above figure shows two identical periodic patterns of equally spaced vertical bars, rotated in opposite directions and then superimposed on each other by multiplying the two images. A Moiré pattern, caused by a breakup of the periodicity, is seen in figure as a 2-D sinusoidal waveform running in a vertical direction. A similar pattern can appear when images are digitized (e.g., scanned) from a printed page, which consists of periodic ink dots.

### 1.14 Mathematical Tools use in Digital Image Processing

Images are generally represented as in the form of matrices but operations can be performed are array operations. Array operations are different from the matrix operations. These operations are done on a pixel by pixel basis. There are different types of operations performed on images.

**Arithmetic Operations:**

Arithmetic Operations between images are array operations and these are carried out between images of equal size. Let us consider f $(x, y)$ and g$(x, y)$ are two images of equal size. The four arithmetic operations are denoted as,

$$s(x, y) = f(x, y) + g(x, y)$$

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

$$p(x, y) = f(x, y) \times g(x, y)$$

$$v(x, y) = f(x, y) \div g(x, y)$$

**Logical Operations:**

Logical operations includes AND, OR, NOT, XOR, NAND.

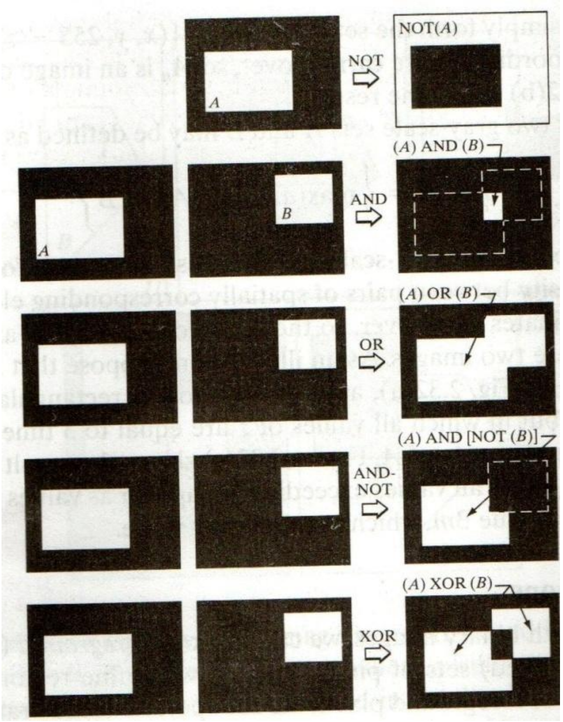
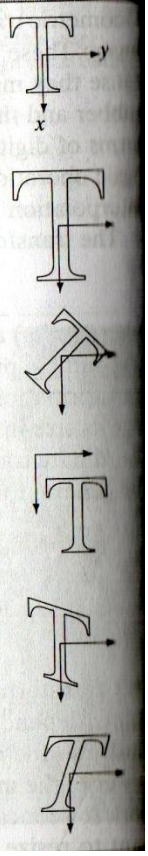



Fig: Illustration of logical operations. Black represents binary 0's and white binary 1's

**Geometrical Operations:**

| Transformation Name | Affine Matrix, T | Coordinate Equations | Example |
|---|---|---|---|
| Identity | $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v$<br>$y = w$ | |
| Scaling | $\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = c_x v$<br>$y = c_y w$ | |
| Rotation | $\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v\cos\theta - w\sin\theta$<br>$y = v\sin\theta + w\cos\theta$ | |
| Translation | $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$ | $x = v + t_x$<br>$y = w + t_y$ | |
| Shear (vertical) | $\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v + s_v w$<br>$y = w$ | |
| Shear (horizontal) | $\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v$<br>$y = s_h v + w$ | |



## PREVIOUS QUESTIONS

1. What are the fundamental steps involved in image processing? Explain
2. a) Assume that a 10m high structure is observed from a distance of 50m. What is the size of the retinal image?

   b) Discuss about the mathematical model of the human visual system.
3. a) An object is 15cm wide and is imaged with a sensor of size 8.8*6.6 mm from a distance of 0.7 m. What should be the required focal length?

   b) Discuss about digital imaging system and its classification.
4. a) Discuss about image quantization.

   b) A medical image has a size of 8*8 inches. The sampling resolution is 5 cycles/mm. How many pixels are required? Will an image of size 256*256 be enough?
5. a) With neat sketch, explain the components of image processing system.

   b) Describe the elements of visual perception.
6. Explain the following: a) Arithmetic operations on Images b) Logical operations on Images.
7. Define Fourier Transform. Explain the properties of two-dimensional Discrete Fourier transform.
8. a) Explain the concept of sampling and quantization of an image. Explain how images are digitally represented.

   b) Discuss about Discrete Cosine transform and write the properties?
9. How a digital image can be represented? Explain the effect of gray level resolution on digital images?
10. Explain the various fields that use Digital Image Processing?
11. Write an algorithm to generate haar basis and find haar basis for second order function.
12. Explain some of the basic relationships between pixels in a digital image?
13. Define the terms sampling and quantization. What is their role in image quality and size?
14. What is meant by Moire patterns? What are the reasons for Occurrence of such patterns?
15. a) Discuss a brief survey of image processing applications.

    b) Assume that a 15m high structure is observed from a distance of 75m. what is the size of the retinal image.