# Samrat Ashok Technological Institute
# Vidisha (MP)

**B.E. 7th Semester EI**
**Digital Image Processing (EI 1872)**
**Assignment-I** 20-09-2021
**Date of Submission: 30 September 2021**

---

## Q. 1. Explain the different steps involved in digital image processing.

### Fundamental Steps in Digital Image Processing

There are 11 fundamental steps in digital image processing (DIP) , all these steps may have sub-steps. The fundamental steps in DIP are described below with a neat block diagram.

### 1. Image Acquisition

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

### 2. Image Enhancement

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

### 3. Image Restoration

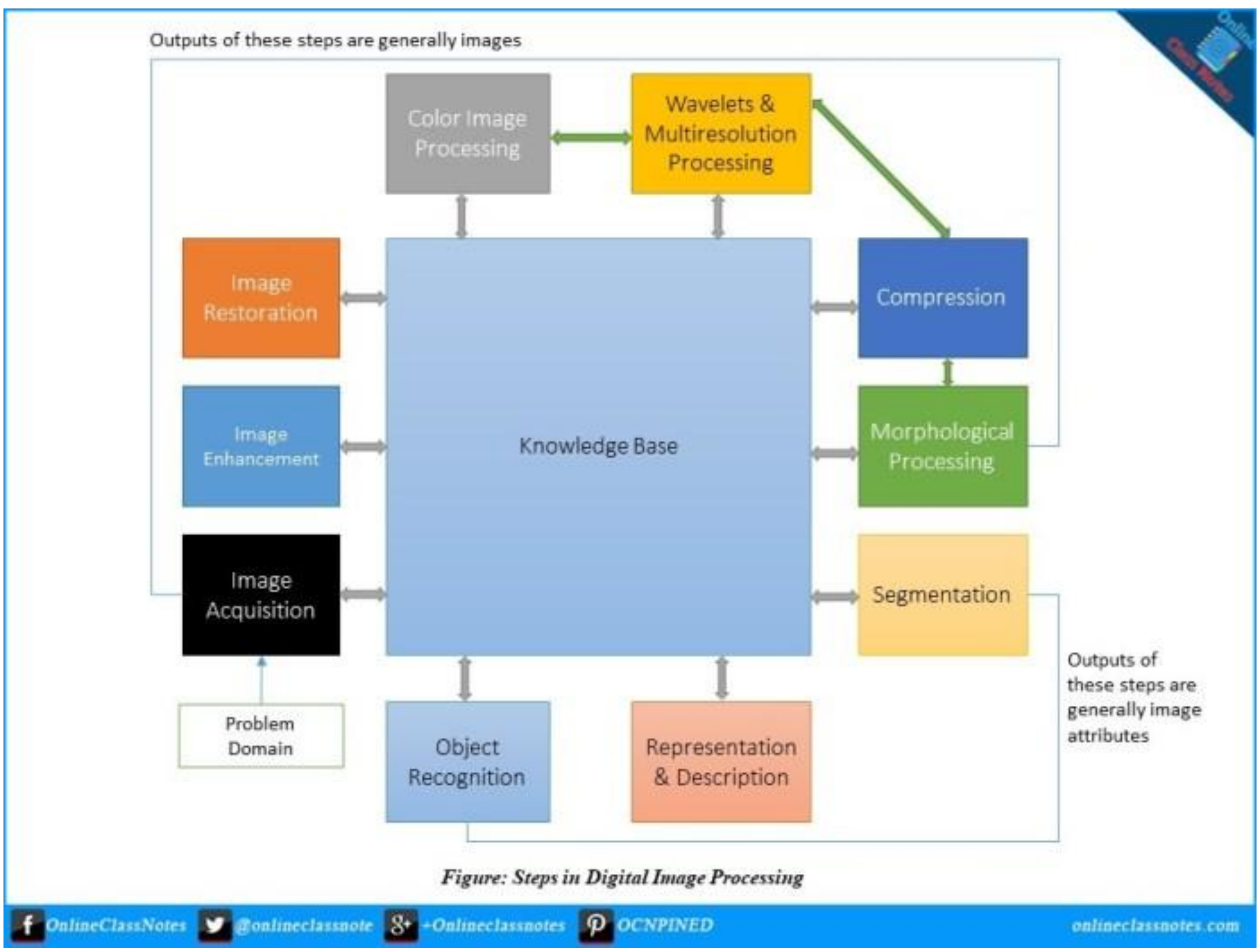
Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

### 4. Color Image Processing

Color image processing is an area that has been gaining its importance because of the significant increase in the use of digital images over the Internet. This may include color modeling and processing in a digital domain etc.

### 5. Wavelets and Multi-Resolution Processing

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




Fundamental steps in digital image processing

### 6. Compression

Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data.

### 7. Morphological Processing

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

### 8. Segmentation

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

### 9. Representation and Description

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

### 10. Object recognition

Recognition is the process that assigns a label, such as, "vehicle" to an object based on its descriptors.

### 11. Knowledge Base

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

---

## Q. 2. Illustrate the process of image acquisition.

**Image Sensing and Acquisition:**

The types of 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 energy.
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. We could even image a source, such as acquiring
images of the sun. 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 photo converter (e.g., a phosphor screen), which converts the energy
into visible light. Electron microscopy and some applications of gamma imaging use this
approach.

Figure 5.1 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.

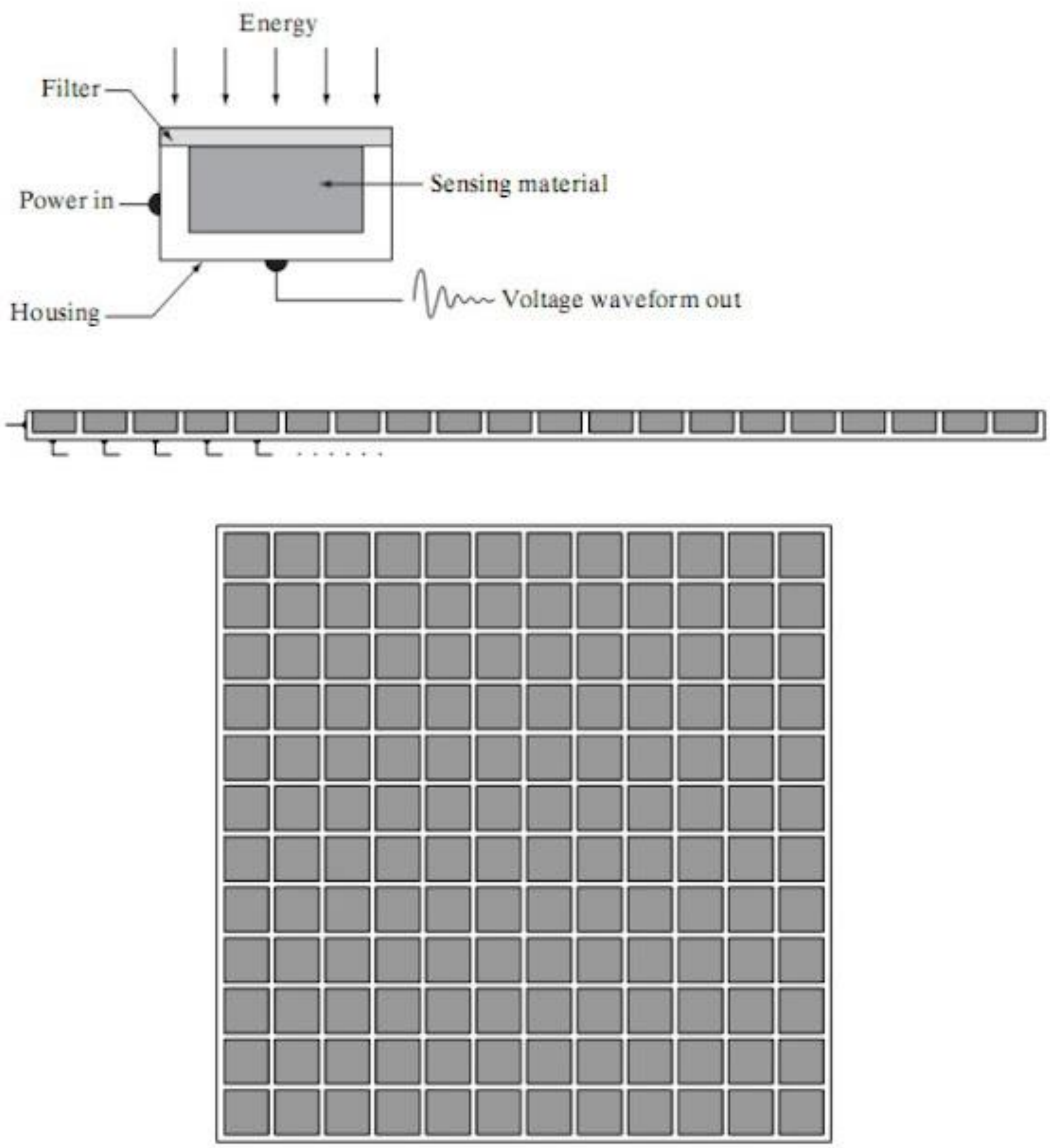


Fig.5.1 (a) Single imaging Sensor (b) Line sensor (c) Array sensor

## (1)Image Acquisition Using a Single Sensor:

Figure 5.1 (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 for 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 5.2 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. 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 sometimes are
referred to as microdensitometers.

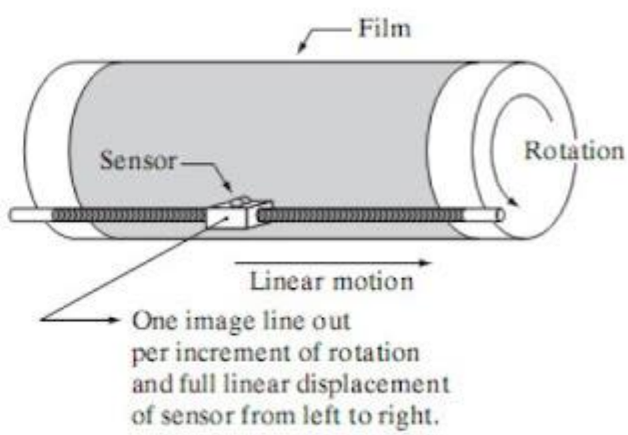


Fig.5.2. Combining a single sensor with motion to generate a 2-D image

## (2) 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 Fig. 5.1 (b) shows. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction, as shown in Fig. 5.3 (a).This is the type of arrangement used in most flat bed 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. Onedimensional 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 Fig. 5.3 (b) shows. 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 (the sensors obviously have to be sensitive to X-ray energy).This is the basis for medical and industrial computerized axial tomography (CAT). It is important to note that the output of the sensors must be processed by reconstruction algorithms whose objective is to transform the sensed data into meaningful cross-sectional images.

In other words, 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 approach shown in Fig. 5.3 (b).

Fig.5.3 (a) Image acquisition using a linear sensor strip (b) Image acquisition using a circular sensor strip.

## (3) Image Acquisition Using Sensor Arrays:

Figure 5.1 (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 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. Since
the sensor array shown in Fig. 5.4 (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. The principal
manner in which array sensors are used is shown in Fig.5.4. This figure shows the energy from
an illumination source being reflected from a scene element, but, 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 shown in Fig.5.4 (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 a lens, which projects the viewed scene onto the lens focal plane, as Fig. 2.15(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
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. 5.4 (e).

Fig.5.4 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

---

## Q. 3. Describe the sampling and quantization process used to obtain a digital image.

### Sampling & Quantization in Digital Image Processing

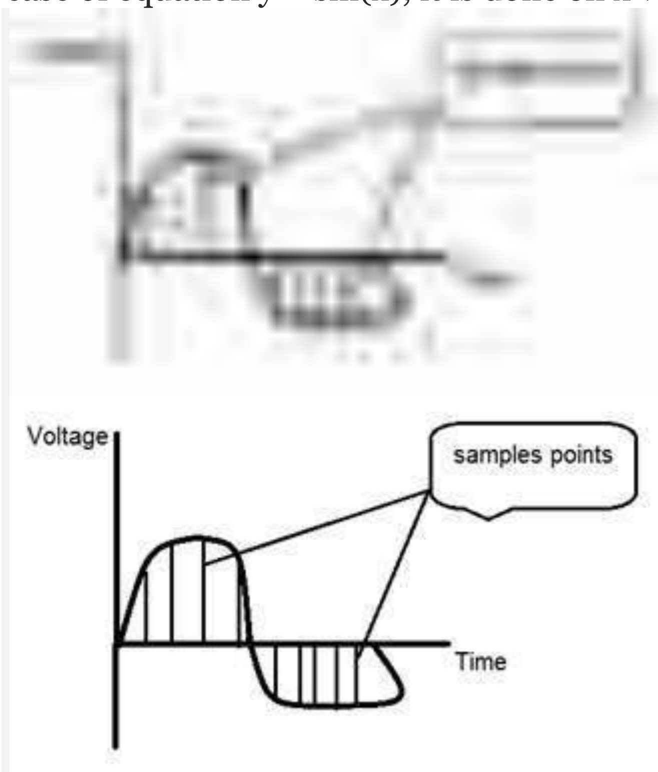
Process. In order to become suitable for digital processing, an image function f(x,y) must be digitized both spatially and in In world need to be translated into digital form by “Digitization” Digital Image Processing, signals captured from the physical amplitude. This digitization process involves two main processes called

1. **Sampling:** Digitizing the co-ordinate value is called sampling.
2. **Quantization:** Digitizing the amplitude value is called quantization

Typically, a frame grabber or digitizer is used to sample and quantize the analogue video signal.

## Sampling

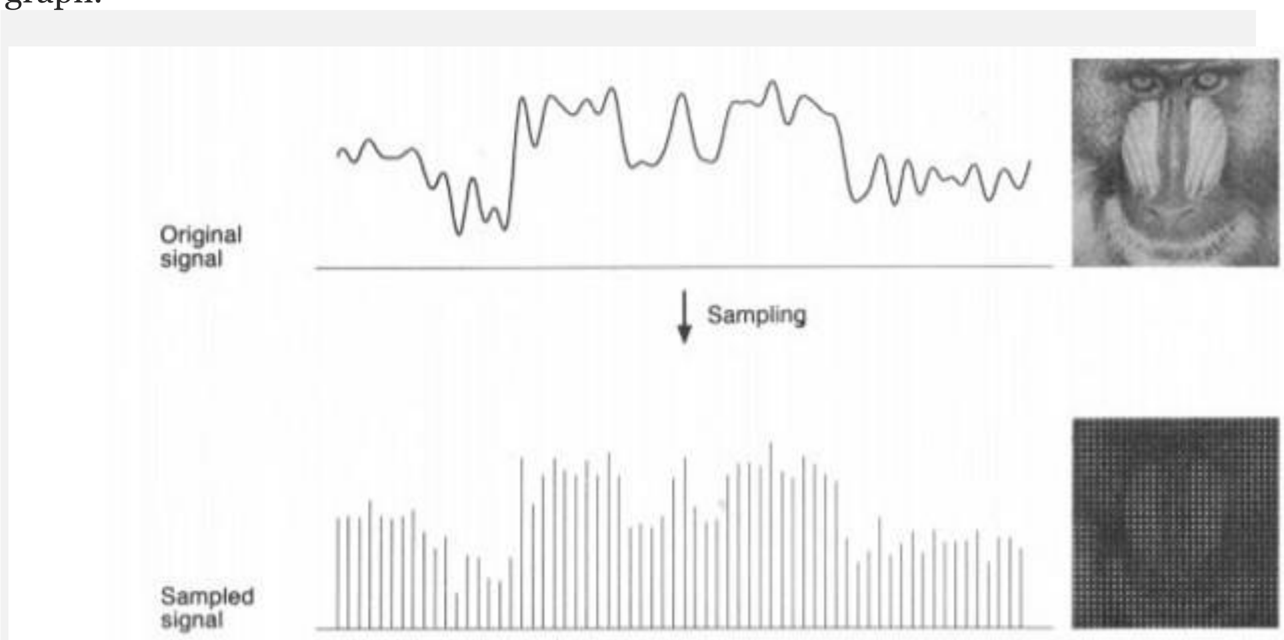
Since an analogue image is continuous not just in its co-ordinates (x axis), but also in its amplitude (y axis), so the part that deals with the digitizing of co-ordinates is known as sampling. In digitizing sampling is done on independent variable. In case of equation y = sin(x), it is done on x variable.




When looking at this image, we can see there are some random variations in the signal caused by noise. In sampling we reduce this noise by taking samples. It is obvious that more samples we take, the quality of the image would be more better, the noise would be more removed and same happens vice versa. However, if you take sampling on the x axis, the signal is not converted to digital format, unless you take sampling of the y-axis too which is known as quantization.

Sampling has a relationship with image pixels. The total number of pixels in an image can be calculated as Pixels = total no of rows * total no of columns. For example, let's say we have total of 36 pixels, that means we have a square image of 6X 6. As we know in sampling, that more samples eventually result in more pixels. So it means that of our continuous signal, we have taken 36 samples on x axis. That refers to 36 pixels of this image. Also the number sample is directly equal to the number of sensors on CCD array.

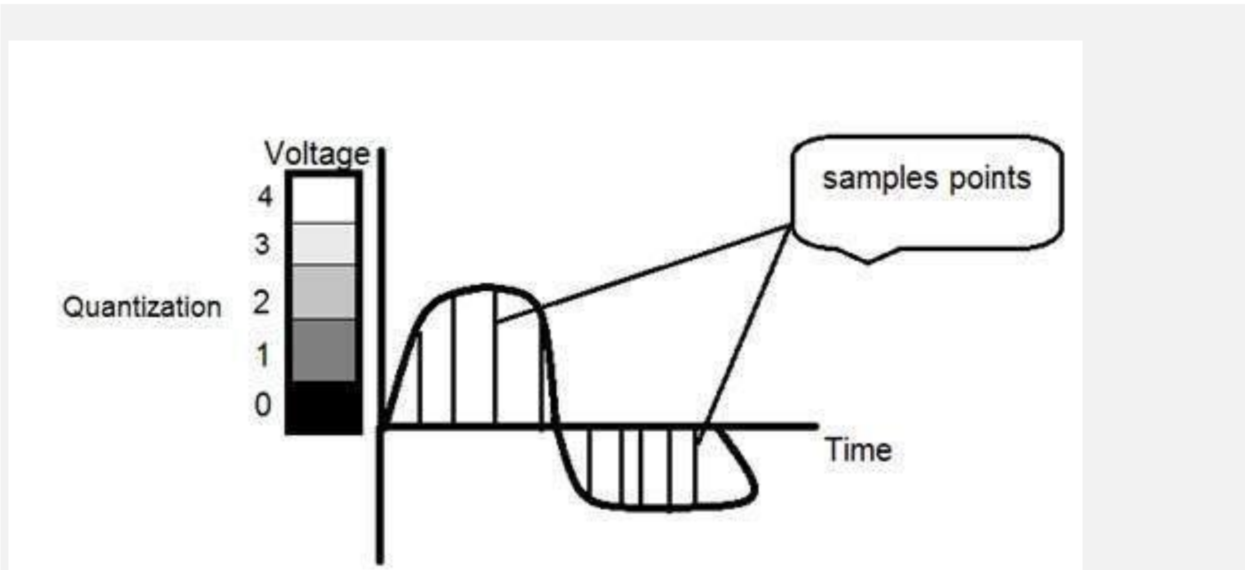
Here is an example for image sampling and how it can be represented using a graph.




## Quantization

Quantization is opposite to sampling because it is done on "y axis" while sampling is done on "x axis". Quantization is a process of transforming a real valued sampled image to one taking only a finite number of distinct values. Under quantization process the amplitude values of the image are digitized. In simple words, when you are quantizing an image, you are actually dividing a signal into quanta(partitions).

Now let's see how quantization is done. Here we assign levels to the values generated by sampling process. In the image showed in sampling explanation, although the samples has been taken, but they were still spanning vertically to a continuous range of gray level values. In the image shown below, these vertically ranging values have been quantized into 5 different levels or partitions. Ranging from 0 black to 4 white. This level could vary according to the type of image you want.




There is a relationship between Quantization with gray level resolution. The above quantized image represents 5 different levels of gray and that means the image formed from this signal, would only have 5 different colors. It would be a black and white image more or less with some colors of gray.

When we want to improve the quality of image, we can increase the levels assign to the sampled image. If we increase this level to 256, it means we have a gray scale image. Whatever the level which we assign is called as the gray level. Most digital IP devices uses quantization into k equal intervals. If b-bits per pixel are used,

No. of quantization levels = $k = 2^b$

The number of quantization levels should be high enough for human perception of fine shading details in the image. The occurrence of false contours is the main problem in image which has been quantized with insufficient brightness levels.

---

## Q. 4. What is image formation model?

### *Fundamentals:*
### *Models of Image Formation*

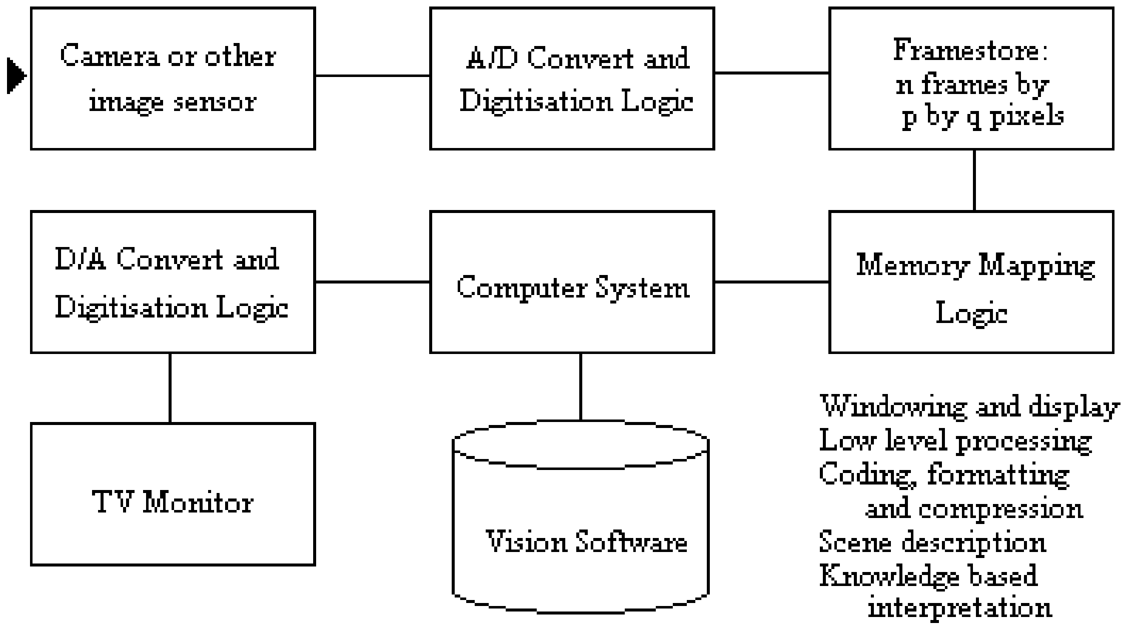
interacted with physical objects. For example, an electronic camera converts reflected In the context of image formation, a sensor registers information about radiation that has light of varying intensity and hue into a two-dimensional matrix of luminance and chrominance values, a laser rangefinder converts received laser radiation reflected from the scene when a transmitter is scanned across it into a ``depth map'' constructed from the receiver's viewpoint. A model of the imaging process has several different components :-

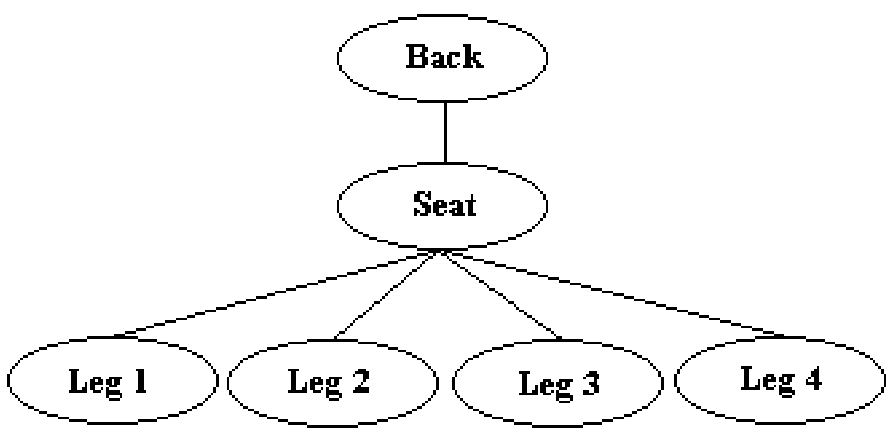
(a)

The **image function** is a mathematical representation of the image. In particular we are concerned primarily with a discrete ( digitised ) image function, by nature of the electronic processing involved. Most image functions are expressed in terms of two spatial variables, $f(x,y)$, where **x** and **y** are the two spatial coordinates. $f(x,y)$ might be intensity in a range from 0 ( black ) to 255 ( white ), or colour, where

$$f(x,y) = \{f(x,y)green, f(x,y)red, f(x,y)blue\} \qquad (1)$$

or depth, where $f(x, y)$ refers to the **z** coordinate, or distance to an imaged point from the sensor. In **Figure 3:** Intensity and depth images, the depth data is encoded by intensity, the brighter the point the nearer the viewer.




Colour spaces are a way of organising the colours perceived by humans in the range 400nm.(blue) to 700nm.(red) approx. The colour signal perceived by an electronic system may be a weighted combination of three signal channels, i.e. red, green, and blue, but this does not give any direct correspondence to the human capability to see things in black-and-white, effectively deriving intensity information from colour receptors. There are various 3-variable colour combinations in use throughout the world, e.g. IHS (intensity-saturation-hue) and YIQ respectively red-cyan, magenta-green, white-black.

(b)

A **geometrical model** describes how the 3 world dimensions are translated into the dimensions of the sensor. In the context of a TV or still camera having a single 2D image plane, perspective projection is the fundamental mechanism whereby light is projected into a single **monocular** view. This type of projection does not yield direct information about the z-coordinate ( although several indirect ``clues'' are available ), and introduces some interesting distortions - 3D objects are more shrunken as they get further from the viewpoint, parallels converge etc. **Binocular** imaging uses a system with two viewpoints, in which the eyes do not normally converge, i.e. the eyes are aimed in parallel at an infinite point in the z direction. The depth information is encoded by it's different positions ( disparity ) in the two images.

(c)

A **radiometric model** illustrates the way in which the imaging geometry, the light sources and the reflectance properties of objects influence the light measured at the sensor. The brightness of the image at a point, the **image intensity** or **image irradiance** where irradiance defines the power per unit area falling on a surface $(Wm^{-2})$ depends on the following factors.

First, there is the **radiant intensity** of the source, i.e. the power per unit area emitted into a unit solid angle $(Wm^{-2}sr^{-1})$. Second there is the the **reflectance** of the objects in the scene, in terms of the proportion, spatial distribution and spectral variation of light reflected. The reflectance of a surface is generally somewhere between **specular,** i.e. mirror-like, and **Lambertian,** i.e. reflecting light ``evenly'' in all directions according to a cosine distribution. Of course, a distinction between sources and objects is over-simplistic, objects may radiate light and there will in general be multiple reflections. It is also worth noting that most electronic light transducers do not have a linear intensity response, or more markedly they have a very non-uniform spectral response ( like humans ! ).

(d)

The **digitising model** implies that the analogue scene data which varies continuously in intensity (say) and space, must be transformed into a discrete representation. Digitised images are **sampled**, i.e. only recorded at discrete locations, **quantised** , i.e. only recorded with respect to the nearest amplitude level, for example 256 levels of intensity, and **windowed**, i.e. only recorded to a finite extent in **x** and **y** etc. All these processes change fundamentally the world as seen by the camera or other sensor.

(e)

A **spatial frequency model** describes how the spatial variations of the image may be characterised in the spatial frequency domain, the more rapid the variations in the image, the higher the spatial frequency. An extension of Fourier analysis from a single dimension, e.g. time, to 2 or 3 spatial dimensions may be made, i.e. $(x, y)$ or $(x, y, z)$ instead of **t**. This type of analysis is fundamental to image processing, and since computer vision usually involves low level image processing operations such as convolutions it is dangerous to be unaware of the implications of the recording mechanism and the effects of this and subsequent processing on the space-spatial frequency duality.

There are many different mediums for image acquisition, including ultrasound, visible light, infra-red light, X-rays etc. Of these, visible light is the most widely used, and there are many acquisition systems based on TV cameras, spot rangers, laser scanners and solid state devices. To scan the image in two dimensions a raster scan system is commonly used, in which the scanning mechanism may be electrical, mechanical or a combination of both. Generally, access to the stored image data is on a random basis.

---

Q. 5. Discuss the following relationship between different pixels in an image.

a. Neighbours of pixels

b. Distance Measures

## Relationships between pixels(Neighbours and Connectivity)

An image is denoted by f(x,y) and p,q are used to represent individual pixels of the image.

## Neighbours of a pixel

A pixel p at (x,y) has 4-horizontal/vertical neighbours at (x+1,y), (x-1,y), (x,y+1) and (x,y-1). These are called the **4-neighbours of p : N4(p)**.

A pixel p at (x,y) has 4 diagonal neighbours at (x+1,y+1), (x+1,y-1), (x-1,y+1) and (x-1,y-1). These are called the **diagonal-neighbours of p : ND(p).**

The 4-neighbours and the diagonal neighbours of p are called **8-neighbours of p : N8(p).**

## Adjacency between pixels

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 N4(p).

**b) 8-adjacency:** Two pixels p and q with values from V are 8-adjacent if q is in the set N8(p).

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

1. q is in N4(p), or
2. 2) q is in ND(p) and the set N4(p)∩N4(q) has no pixels whose values are from V.

## Connectivity between pixels

It is an important concept in digital image processing.

It is used for establishing boundaries of objects and components of regions in an image.

Two pixels are said to be connected:

- if they are adjacent in some sense(neighbour pixels,4/8/m-adjacency)
- if their gray levels satisfy a specified criterion of similarity(equal intensity level)

There are three types of connectivity on the basis of adjacency. They are:

**a) 4-connectivity:** Two or more pixels are said to be 4-connected if they are 4-adjacent with each others.

**b) 8-connectivity:** Two or more pixels are said to be 8-connected if they are 8-adjacent with each others.

**c) m-connectivity:** Two or more pixels are said to be m-connected if they are m-adjacent with each others.

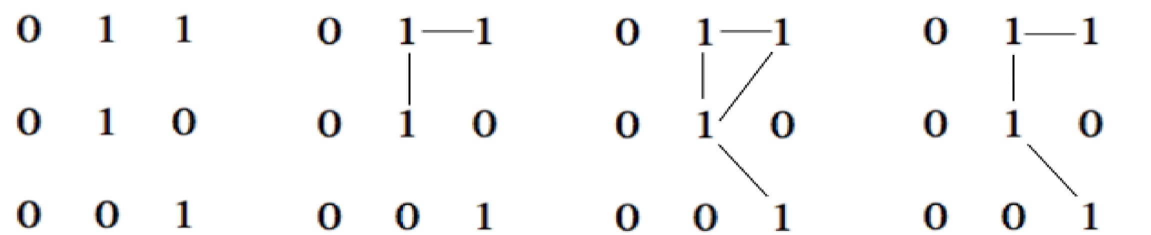



**Fig: An arrangement of pixels** **Fig: 4-connectivity of pixels** **Fig: 8-connectivity of pixels** **Fig: m-connectivity of pixels**

## Distance Measures:

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

**(a)** $D(p, q) \geq 0 \quad (D(p, q) = 0 \quad \text{iff} \quad p = q)$,
**(b)** $D(p, q) = D(q, p)$, and
**(c)** $D(p, z) \leq D(p, q) + D(q, z)$.

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).

The **D4 distance (also called city-block distance)** between p and q is defined as

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

In this case, the pixels having a $D_4$ distance from (x, y) less than or equal to some value r form a
diamond centered at (x, y). For example, the pixels with 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 $D_4 = 1$ are the 4-neighbors of (x, y).

The **D8 distance (also called 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 $D_8$=1 are the 8-neighbors of (x, y). Note that the $D_4$ and $D_8$ distances between p and q are independent of any paths that might exist between the points because these distances involve only the coordinates of the points. If we elect to consider m-adjacency, however, the Dm distance between two points is defined as the shortest m-path between the points. In this case, the distance between two pixels will depend on the values of the pixels along the path, as well as the values of their neighbors. For instance, consider the following arrangement of pixels and assume that p, $p_2$ , and $p_4$ have value 1 and that $p_1$ and $p_3$ can have a value of 0 or 1:

$$\begin{array}{ccc} & p_3 & p_4 \\ p_1 & p_2 & \\ p & & \end{array}$$

Suppose that we consider adjacency of pixels valued 1 (i.e. = {1}). If $p_1$ and $p_3$ are 0, the length of the shortest m-path (the Dm distance) between p and $p_4$ is 2. If $p_1$ is 1, then $p_2$ and p will no longer be m-adjacent (see the definition of m-adjacency) and the length of the shortest m-path becomes 3 (the path goes through the points $pp_1p_2p_4$). Similar comments apply if $p_3$ is 1 (and p1 is 0); in this case, the length of the shortest m-path also is 3. Finally, if both $p_1$ and $p_3$ are 1 the length of the shortest m-path between p and $p_4$ is 4. In this case, the path goes through the sequence of points $pp_1p_2p_3p_4$.

6. Evaluate the chessboard distance, city block distance and the Euclidean distance between the pixels whose coordinates are p(5,7) and q(1,4).

Distance= $\sqrt{(X2-X1)+(Y2-Y2)^2}$

$P=(5,7) \quad Q=(1,4)$

$X1=5,\ X2=1,\ Y1=7,\ Y2=4$

$=\sqrt{(1-5)^2 +(4-5)^2}$

$=\sqrt{(4)^2+(3)^2}$

$=\sqrt{16+9}$

$=\sqrt{25}$

$=5$

Distance = 5 Ans.

---


By- Ashutosh Awasthi

Roll No.- 0108ei181011

Sch. No.- 28498


Awasthi