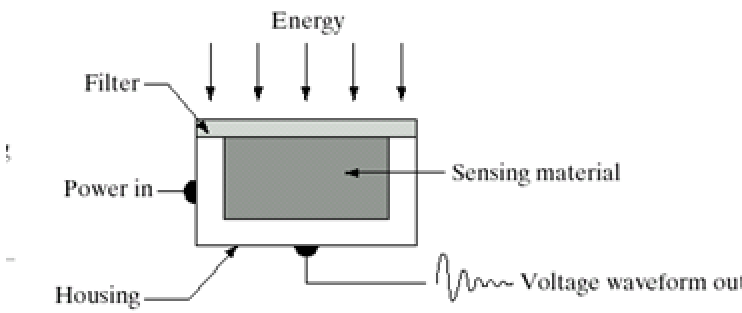
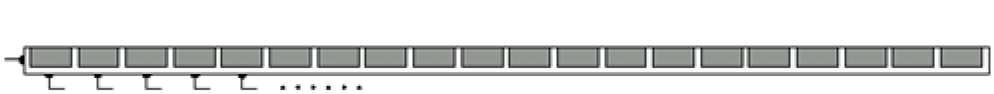
UNIT - 2
Image Sensing and Acquisition

Image Sensing and Acquisition

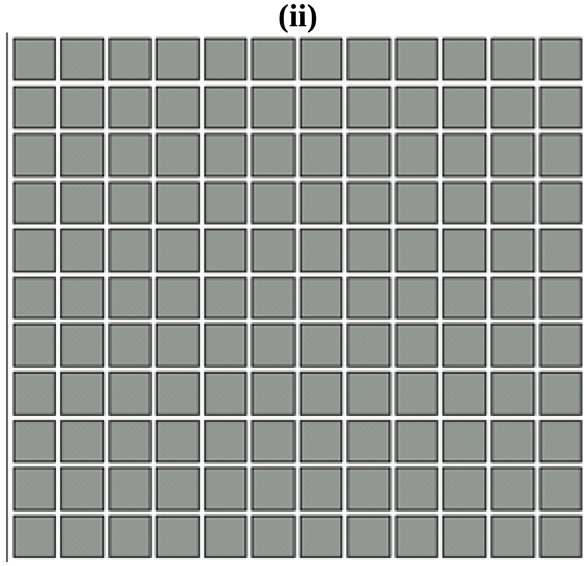
There are 3 principal sensor arrangements (produce an electrical output proportional to light intensity).
(i)Single imaging Sensor (ii)Line sensor (iii)Array sensor



(i)



(ii)



(iii)

Fig: (i)Single image (ii)Sensor line sensor (iii)Array sensor

Image Acquisition using a single sensor

The most common 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 favours 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.

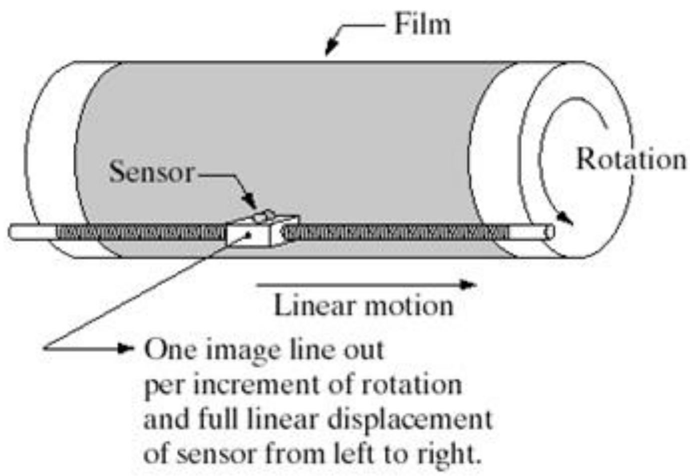


Fig: Combining a single sensor with motion to generate a 2-D image

In order to generate a 2-D image using a single sensor, there have to be relative displacements in both the x- and y-directions between the sensor and the area to be imaged. 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.

Image Acquisition using Sensor Strips

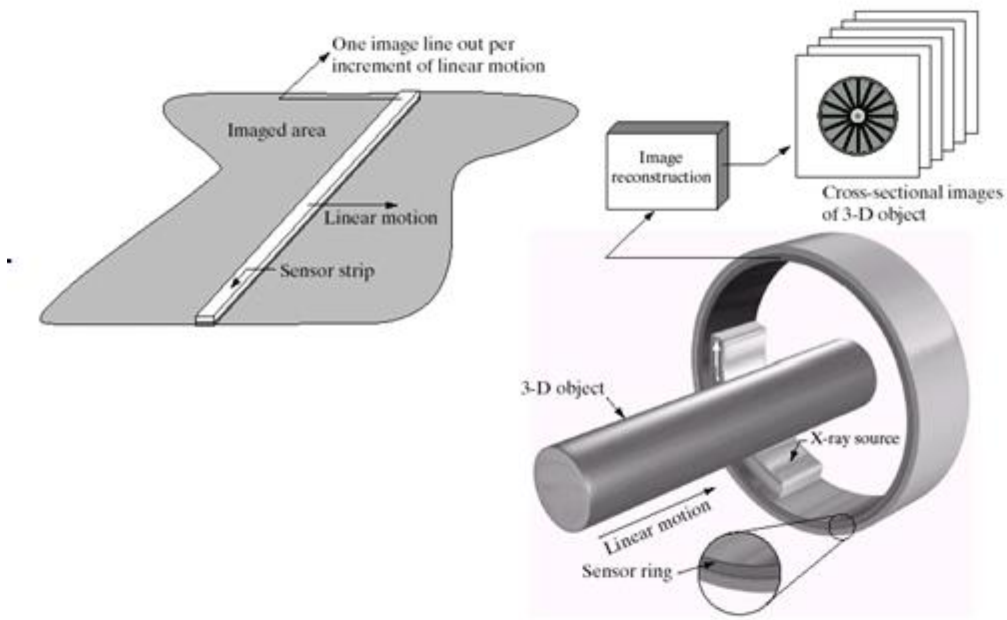


Fig: (a) Image acquisition using linear sensor strip (b) Image acquisition using circular sensor strip.

The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction. 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. Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional ("slice") images of 3-D objects. 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) imaging.

Image Acquisition using Sensor Arrays

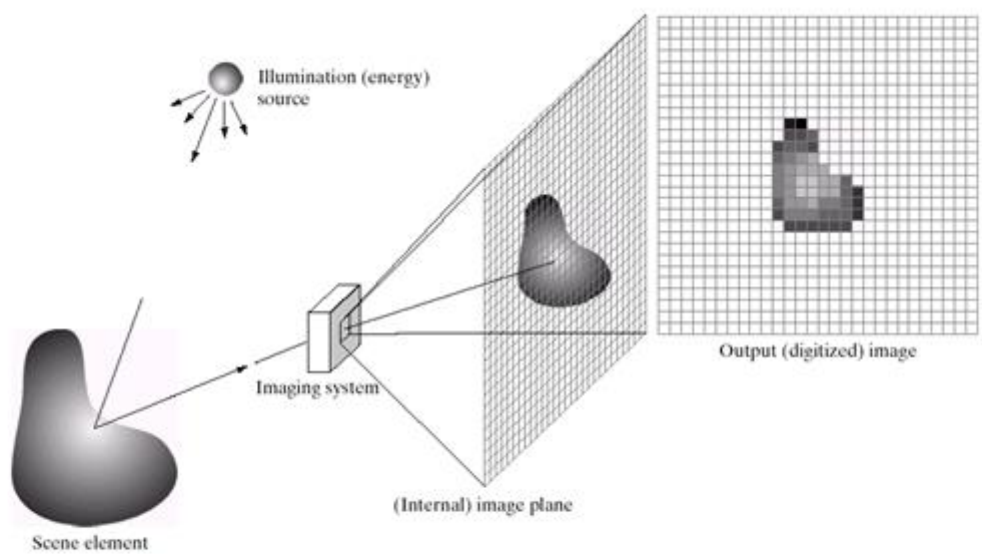


Fig: An example of the digital image acquisition process (a) energy source (b) An element of a scene (d) Projection of the scene into the image (e) digitized image

This type of arrangement is 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.

The first function performed by the imaging system 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. The sensor array, which is coincident with the focal plane, produces outputs proportional to the integral of the light received at each sensor.

A Simple Image Formation Model

An image is defined by two dimensional function $f(x,y)$. The value or amplitude of f at spatial coordinates (x,y) is a positive scalar quantity. When an image is generated from a physical process, its value are proportional to energy radiated by physical source. As a consequence, $f(x,y)$ must be nonzero and finite; that is,

$$0 < f(x,y) < \infty$$

The function $f(x,y)$ may be characterized by two components: (1) the amount of source illumination incident on the scene being viewed and (2) the amount of illumination reflected by the objects in the scene. These are called illumination and reflectance components denoted by $i(x,y)$ and $r(x,y)$ respectively. The two function combine as product to form $f(x,y)$:

$$f(x,y)=i(x,y)\, r(x,y)$$

Where $0 < i(x,y) < \infty$ and $0 < r(x,y) < 1$ $r(x,y)=0$ means total absorption $r(x,y)=1$ means total reflectance

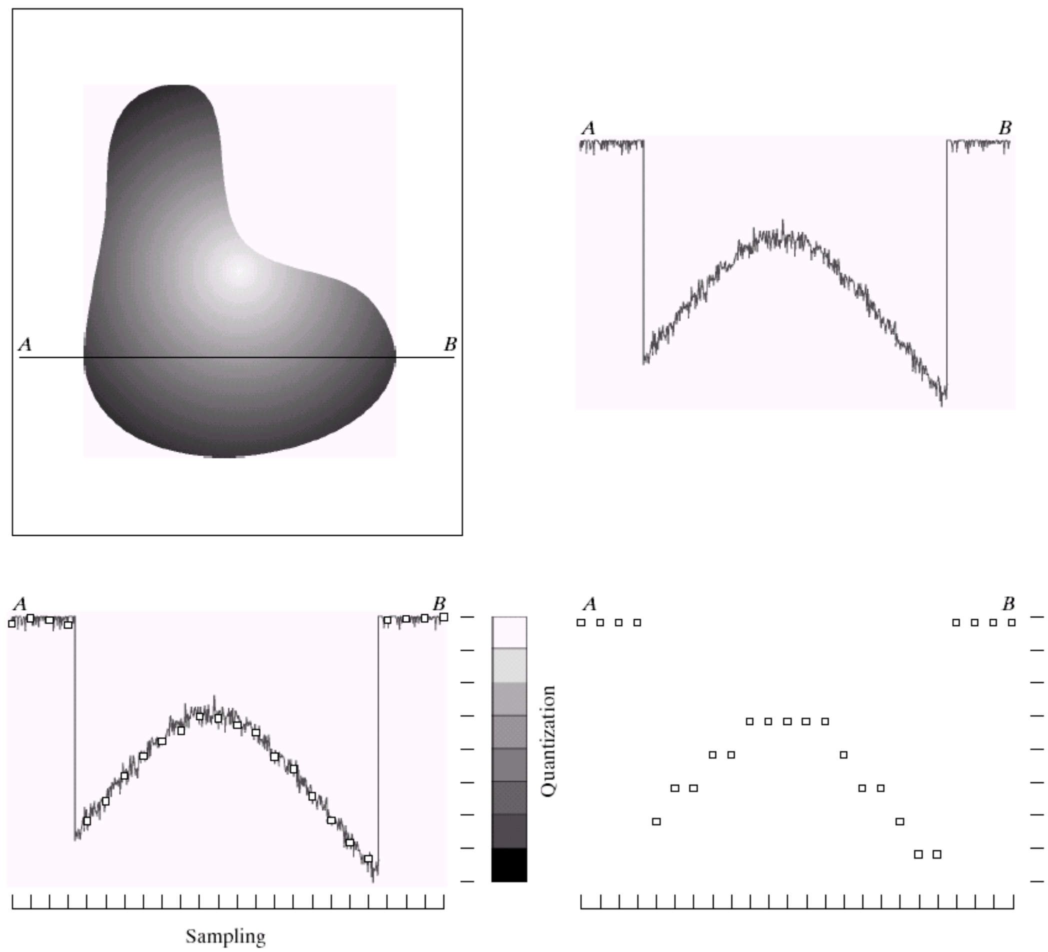
We call the intensity of a monochrome image at any coordinates (x,y) the gray level (l) of the image at that point. That is $l=f(x,y)$.

The interval of l ranges from [0,L-1]. Where $l=0$ indicates black and $l=1$ indicates white. All the intermediate values are shades of gray varying from black to white.

Image Sampling and Quantization

Sampling and quantization are the two important processes used to convert continuous analog image into digital image. Image sampling refers to discretization of spatial coordinates (along x axis) whereas quantization refers to discretization of gray level values (amplitude (along y axis)).

(Given a continuous image, f(x,y), digitizing the coordinate values is called sampling and digitizing the amplitude (intensity) values is called quantization.)



a b
c d

FIGURE 2.16 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.

The one dimensional function shown in fig 2.16(b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in fig 2.16(a). The random variation is due to the image noise. To sample this function, we take equally spaced samples along line AB as shown in fig 2.16 (c).In order to form a digital function, the gray level values also must be converted(quantizes) into discrete quantities. The right side of fig 2.16 (c) shows the gray level scale divided into eight discrete levels, ranging from black to white. The result of both sampling and quantization are shown in fig 2.16 (d).

Representing Digital Image

An image may be defined as a two-dimensional function, f(x, y), where x and y are *spatial* (plane) 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.

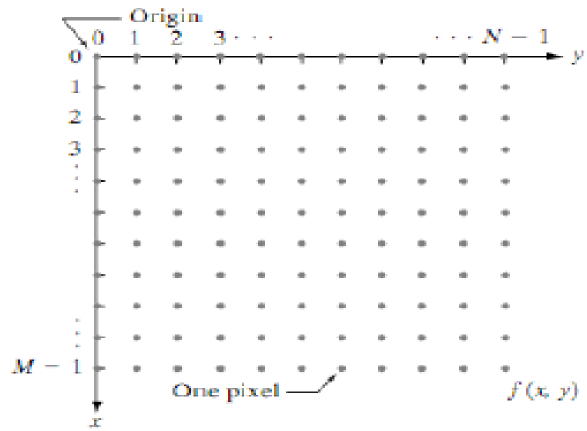


Fig: Coordinate convention used to represent digital images

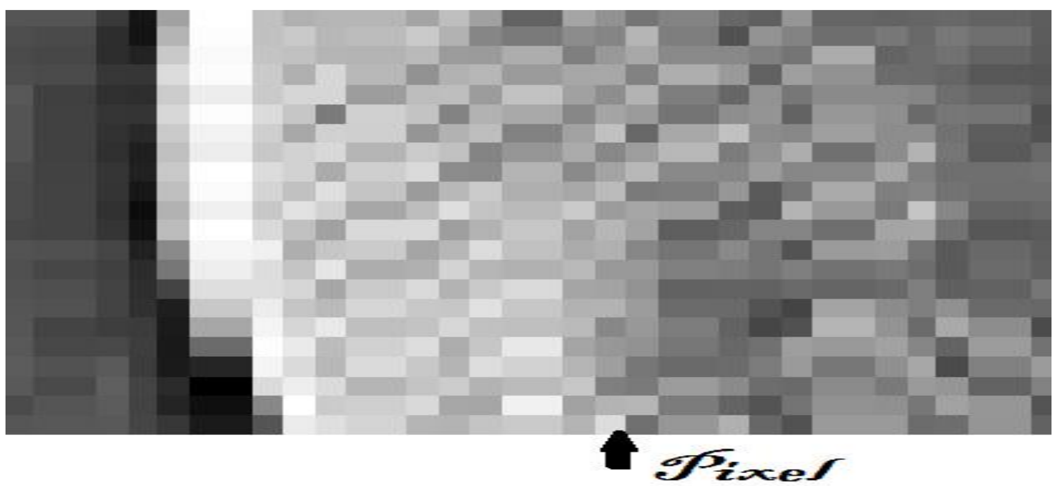


Fig: Zoomed image, where small white boxes inside the image represent pixels

Digital image is composed of a finite number of elements referred to as *picture elements, image elements, pels,* and *pixels*. *Pixel* is the term most widely used to denote the elements of a digital image.

We can represent M*N digital image as compact matrix as shown in fig below

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

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 processing digital images by means of a digital computer.

If k is the number of bits per pixel, then the number of gray levels, L, is an integer power of 2.

$$L = 2^k$$

When an image can have 2^k gray levels, it is common practice to refer to the image as a "k-bit image". For example, an image with 256 possible grey level values is called an 8 bit image.

Therefore the number of bits required to store a digitalized image of size M*N is

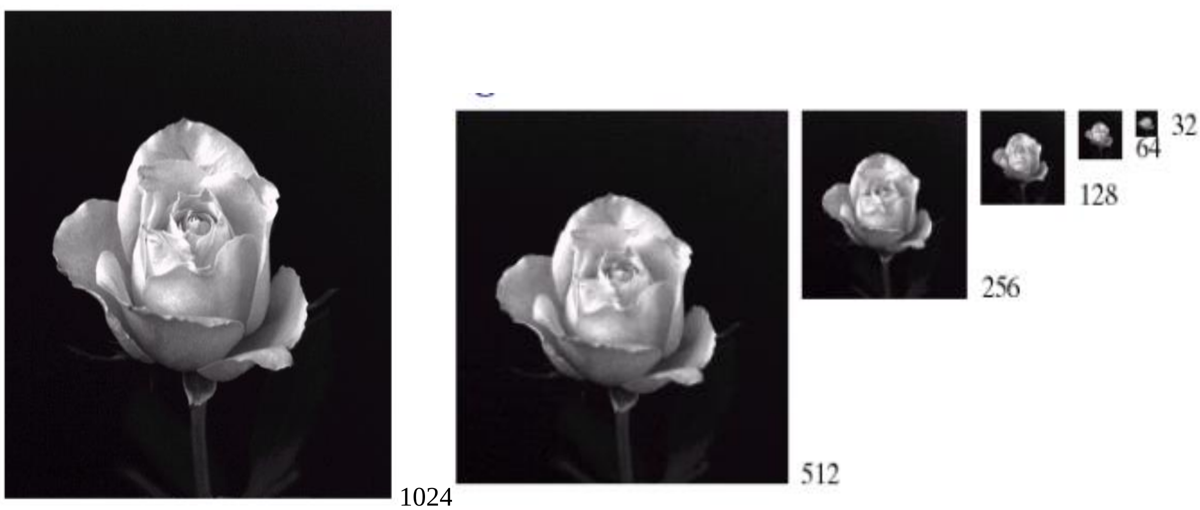
$$b = M * N * k$$

When M=N then $b = N^2 * k$

Spatial and Gray-Level Resolution

Spatial resolution is the smallest discernable (detect with difficulty)change in an image. Gray-level resolution is the smallest discernable (detect with difficulty) change in gray level.

Image resolution quantifies how much close two lines (say one dark and one light) can be to each other and still be visibly resolved. The resolution can be specified as number of lines per unit distance, say 10 lines per mm or 5 line pairs per mm. Another measure of image resolution is dots per inch, i.e. the number of discernible dots per inch.



1024

Fig: A 1024 x 2014, 8 bit image subsampled down to size 32 x 32 pixels. The number of allowable kept at 256.

Image of size 1024*1024 pixels whose gray levels are represented by 8 bits is as shown in fig above. The results of subsampling the 1024*1024 image. The subsampling was accomplished by deleting the appropriate number of rows and columns from the original image. For example, the 512*512 image was obtained by deleting every other row and column from the 1024*1024 image. The 256*256 image was generated by deleting every other row and column in the

512*512 image, and so on. The number of allowed gray levels was kept at 256. These images show dimensional proportions between various sampling densities, but their size differences make it difficult to see the effects resulting from a reduction in the number of samples. The simplest way to compare these effects is to bring all the subsampled images up to size 1024 x 1024.

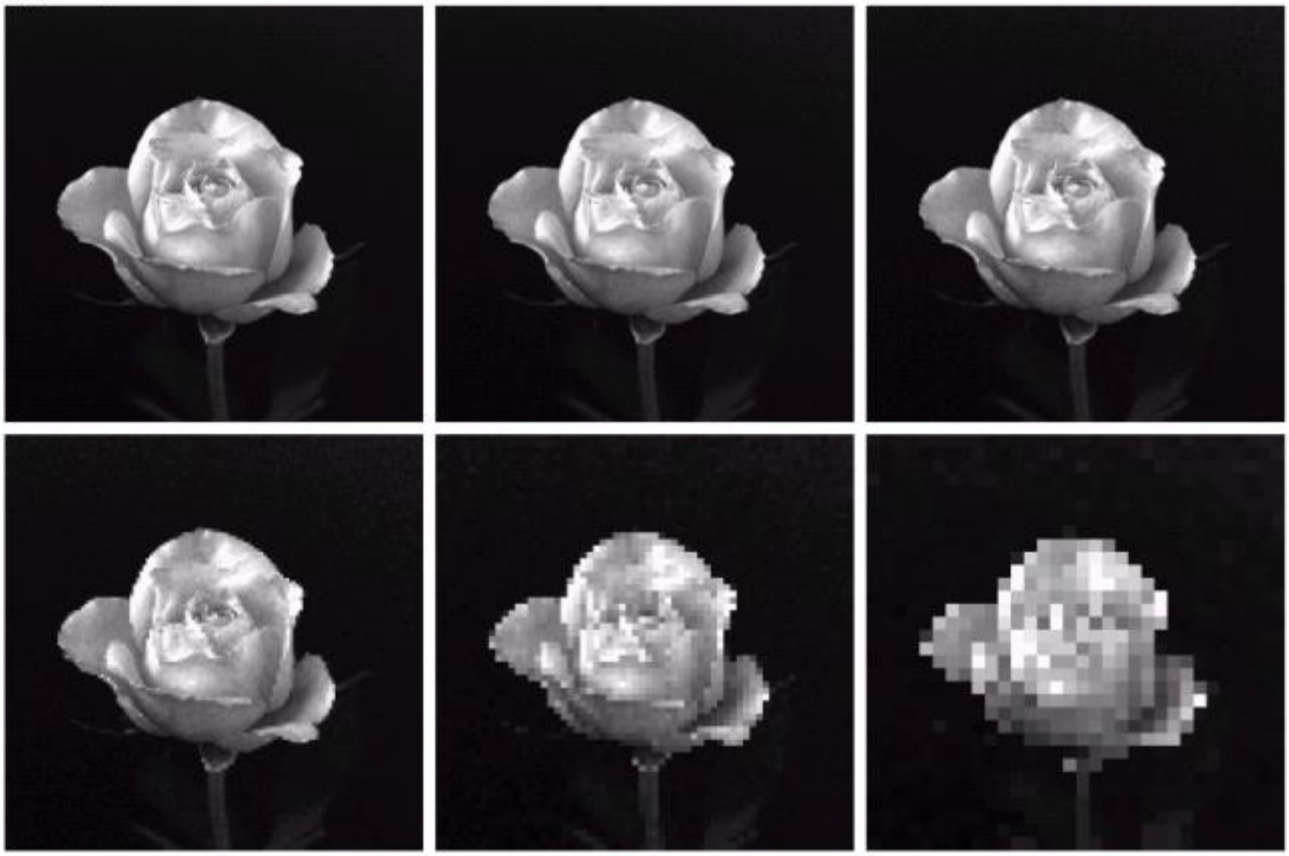


Fig: (a) 1024 x 1024, 8 bit image (b) 512 x 512 image resampled into 1024 x 1024 pixels by row and column duplication. (c) through (f) 256 x 256, 128 x 128, 64 x 64, and 32 x 32 images resampled into 1024 x 1024 pixels.

Gray-Level Resolution

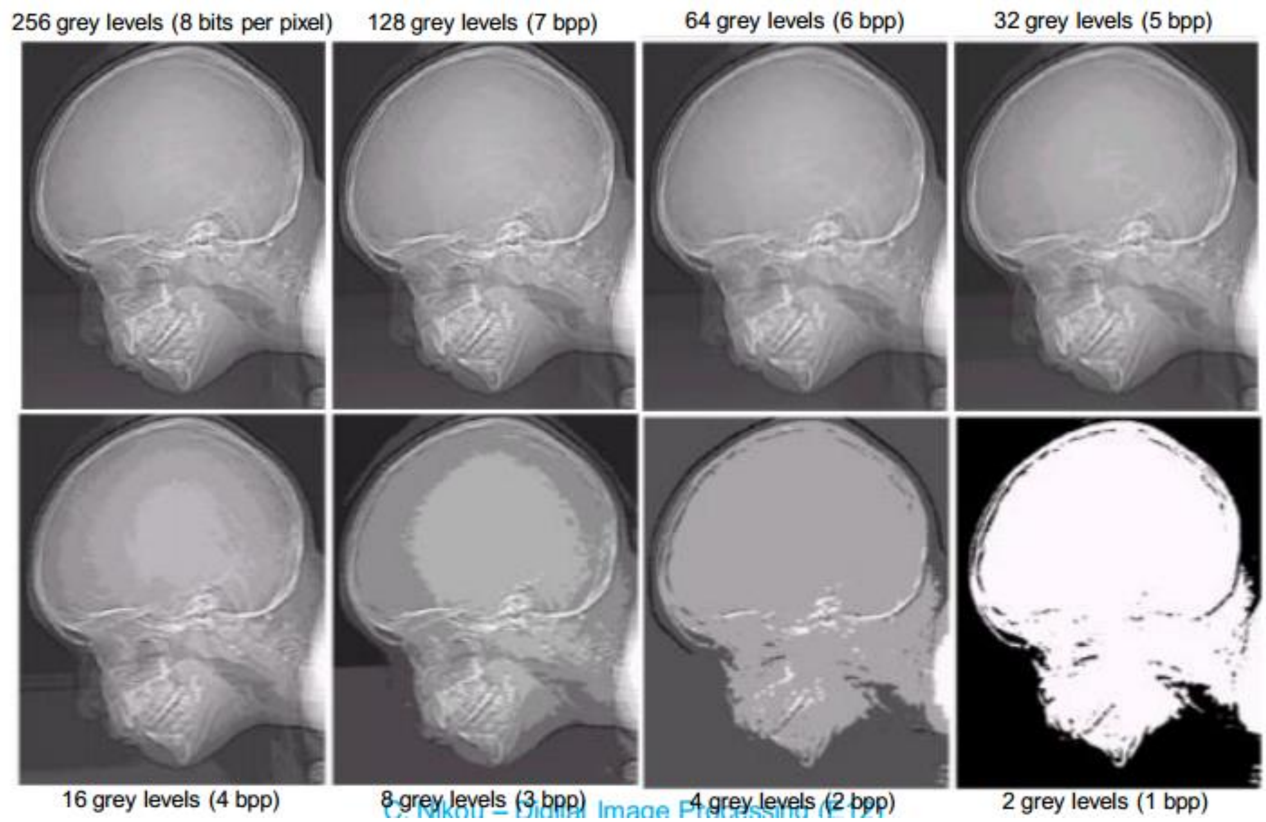


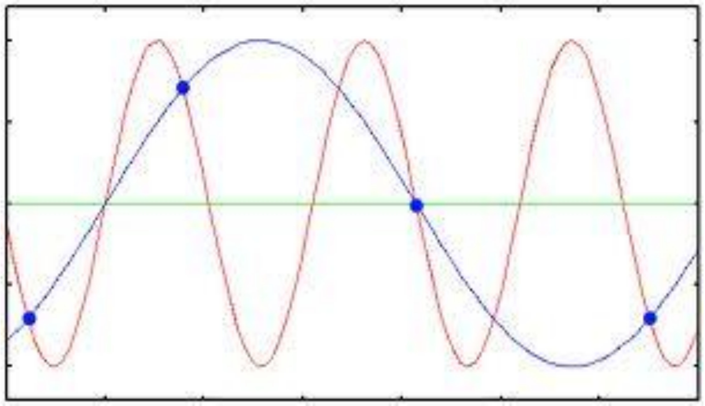
Fig: 452 x 374, 256 level image. (b)- (d) Image displayed in 128, 64, 32,16,8,4 and 2 grey level, while keeping spatial resolution constant.

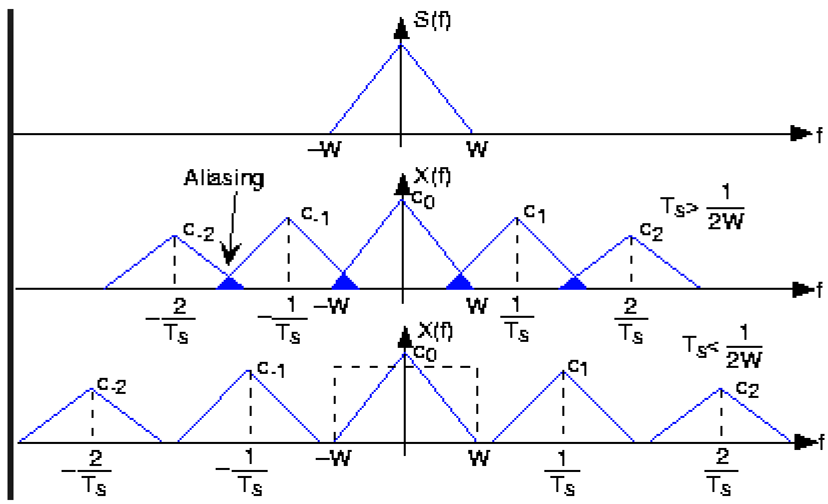
Inthis example, we keep the number of samples constant and reduce the number of grey levels from 256 to 2.

Aliasing and Moire Patterns

Shannon sampling theorem tells us that, if the function is sampled at a rate equal to or greater than twice its higest frequency($f_s \geq f_m$)it is possible to recover completely the original function from the samples. If the function is undersampled, then a phenonmemnon called **alising(distortion) (If two pattern or spectrum overlap, the overlapped portion is called aliased)** 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 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*.





A **moiré pattern** is a secondary and visually evident superimposed pattern created, for example, when two identical (usually transparent) patterns on a flat or curved surface (such as closely spaced straight lines drawn radiating from a point or taking the form of a grid) are overlaid while displaced or rotated a small amount from one another.

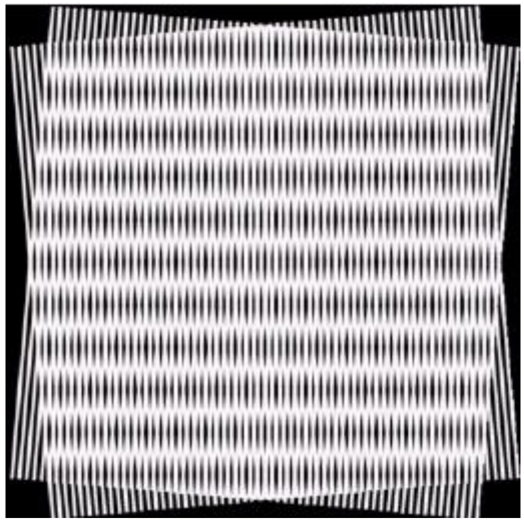


Fig: Illustration of the moire effect

Zooming and Shrinking Digital Image

Zooming may be viwed as oversampling and shrinking many be viewed as undersampling.

Zooming is a method of increasing the size of a given image. Zooming requires two steps: creation of new pixel locations, and the assigning new grey level values to those new locations

Nearest neighbor interpolation: Nearest neighbor interpolation is the simplest method and basically makes the pixels bigger. The intensity of a pixel in the new image is the intensity of the nearest pixel of the original image. If you enlarge 200%, one pixel will be enlarged to a 2 x 2 area of 4 pixels with the same color as the original pixel.

Bilinear interpolation: Bilinear interpolation considers the closest 2x2 neighborhood of known pixel values surrounding the unknown pixel. It then takes a weighted average of these 4 pixels to arrive at its final interpolated value. This results in much smoother looking images than nearest neighbour.

Fig: Images zoomed from 128 x 128, 64 x 64 and 32 x 32 pixels to 1024 x 1024 pixels using nearest neighbor grey-level interpolation.

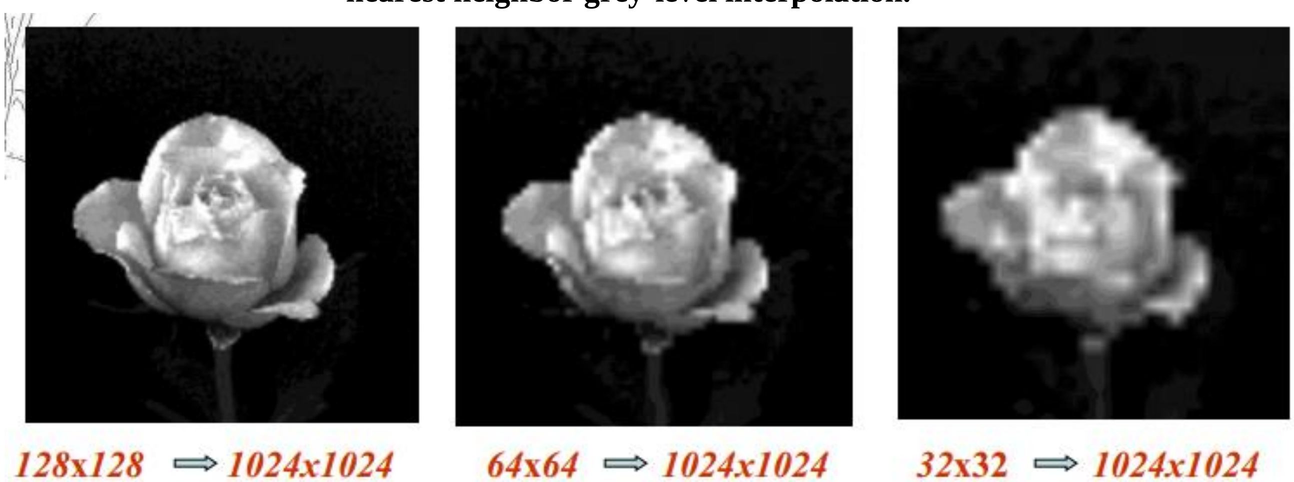


Fig: Images zoomed from 128 x 128, 64 x 64 and 32 x 32 pixels to 1024 x 1024 pixels using bilinear interpolation.

Image shrinking is done in a similar manner as just described for zooming. 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.

Perspective Projection	Some of the important characteristics of Perspective Projection : • Objects which are of the same size but at different distances from the image plane would appear to be of different sizes. The farther ones would appear smaller. • Parallel lines in the scene would appear to converge at some finite point. E.g. if you stand on the railway tracks and look them up, the pair of rails would appear to intersect at some far-off but finite point.

	• The expanse of an object along the line of sight appears much smaller compared to the expanse along a direction perpendicular to the line of sight. • Consider a plane which is parallel to the image plane and passing through the center of projections. All the world points lying on this plane will not get projected on to the image plane.

Some Basic Relationship between Pixels

Neighbors of a Pixel

In a 2-D coordinate system each pixel p in an image can be identified by a pair of spatial coordinates (x, y).
Referring to the Fig below, a pixel p has two horizontal neighbors (x−1, y), (x+1, y) and two vertical neighbors (x, y−1), (x, y+1). These 4 pixels together constitute the 4-neighbors of pixel p, denoted as $N_4(p)$.

The pixel p also has 4 diagonal neighbors which are: (x+1, y+1), (x+1, y−1), (x−1, y+ 1), (x−1, y−1). The set of 4 diagonal neighbors forms the diagonal neighborhood denoted as $N_D(p)$. The set of 8 pixels surrounding the pixel p forms the 8-neighborhood denoted as $N_8(p)$. We have $N_8(p) = N_4(p) \cup N_D(p)$.

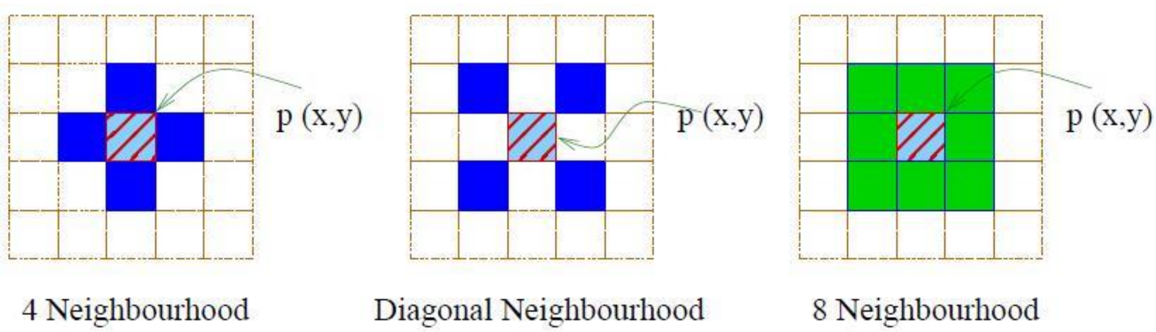


Figure : Pixel Neighborhood: The center pixel p is shown in a dashed pattern and the pixels inthe defined neighborhood are shown in filled color.

Adjacency
The concept of adjacency has a slightly different meaning from neighborhood. Adjacency takes into account not just spatial neighborhood but also intensity groups. Suppose we define a set S={0,L-1} of intensities which are considered to belong to the same group. Two pixels *p* and *q* will be termed adjacent if both of them have intensities from set S and both also conform to some definition of neighborhood.

4 Adjacency: Two pixels *p* and *q* are termed as 4-adjacent if they have intensities from set S and *q* belongs to $N_4(p)$.
8 Adjacency Two pixels *p* and *q* are termed as 4-adjacent if they have intensities from set S and *q* belongs to $N_8(p)$.

Mixed adjacency or m-adjacency : (there shouldn't be closed path)

Two pixels with intensity values from set S are m-adjacent if

- $q \in N_4(p)$,

OR

- $q \in N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose intensity values are from set S.

Mixed adjacency is a modification of 8-adjacency and is used to eliminate the multiple path connections that often arise when 8-adjacency is used.

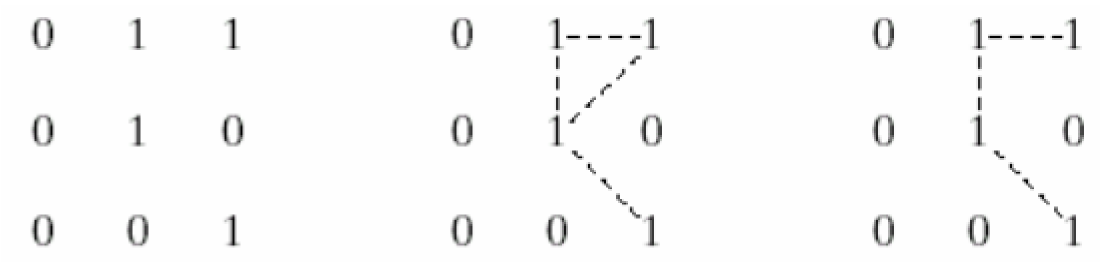
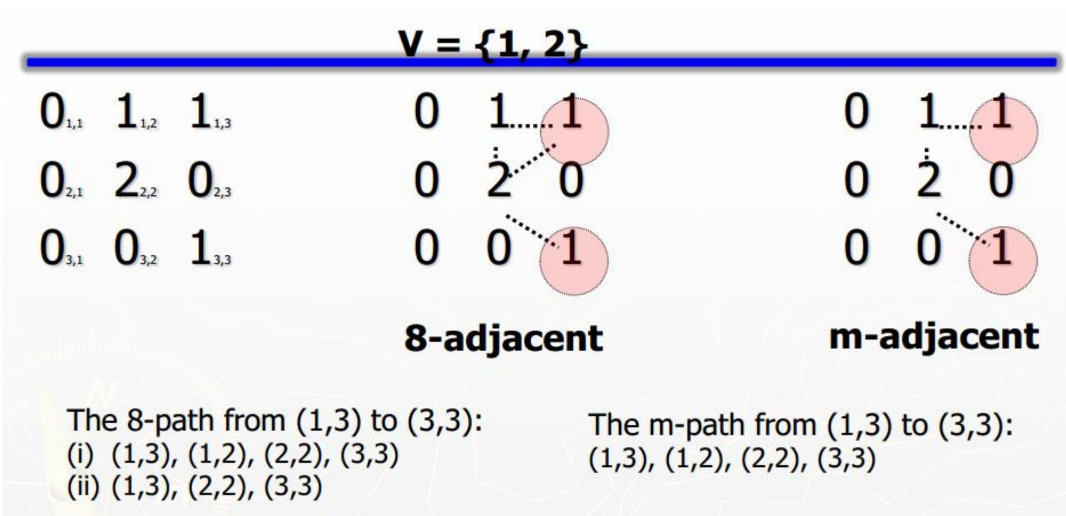


Fig: (a) Arrangement of pixels; (b) pixels that are 8-adjacent to the center pixel (c) m-adjacency



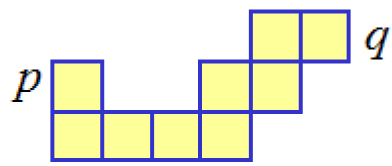
Path

A (digital) path (or curve) from pixel p with coordinates (x0,y0) to pixel q with coordinates (xn, yn) is a sequence of distinct pixels with coordinates

$$(x0, y0), (x1, y1), \ldots, (xn, yn)$$

Where (xi yi) and (xi-1, yi-1) are adjacent for $1 \le i \le n$.

- Here n is the length of the path.
- If (x0, y0) = (xn, yn), the path is closed path.
- We can define 4-, 8-, and m-paths based on the type of adjacency used.



Connected components

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 connected to it in S is called a connected component of S. If it has only one connected component, then set S is called connected set.

A **Region** R is a subset of pixels in an image such that all pixels in R form a connected component.

A **Boundary** 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 Measure

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

(a) $D(p,q) \ge 0$ $(D(p,q)=0 \text{ iff } p=q)$
(b) $D(p,q)= D(q,p)$
(c) $D(p,z) \le D(p,q) + D(q,z)$

The distance between two pixels p and q with coordinates (x_1, y_1), (x_2, y_2) respectively can be formulated in several ways:

Euclidean Distance= $[(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$

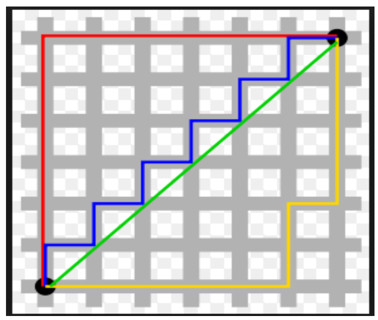
Pixels having a Euclidean distance r from a given pixel will lie on a circle of radius *r* centered at it and *r* = distance

D_4 distance(also called City-block Distance) $D_4(p,q) = |x - s| + |y - t|$

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

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

Pixels having a city-block distance 1 from a given pixel are its 4-neighbors.



Green line is **Euclidean distance.** Red, blue and yellow are **City Block Distance**

It's called city-block distance, because it is calculated as if on each pixel between your two coordinates stood a block (house) which you have to go around. That means, you can only go along the vertical or horizontal lines between the pixels but not diagonal. It's the same like the movement of the rook on a chess field.

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

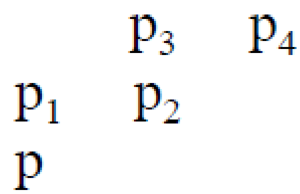
In this case, the pixels having a D_8 distance from (x,y) less than or equal to some value r from a square centered at (x,y). For example, the pixels with D_8 distance ≤ 2 from (x,y) (the center point) form the following contours of constant distance

To measure D_8 distance, you can only go along the vertical or horizontal or diagonal lines between the pixels.

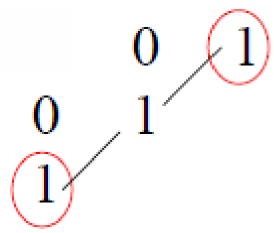
2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Pixels having a chess-board distance 1 from a given pixel are its 8-neighbors

D_m distance between two points is defined as the shortest m path between the points.

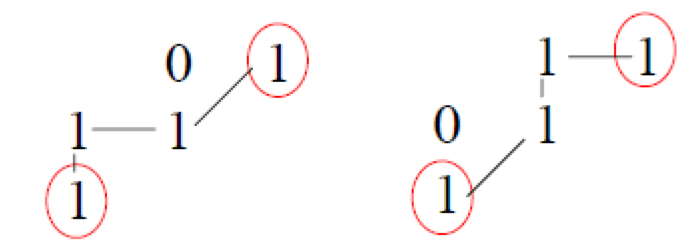


When V={1} p, p_2 and p_4 have value 1

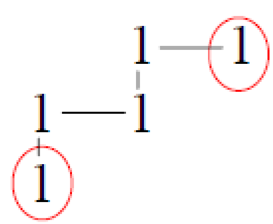


D_m shortest distance between p and p_4 is 2

When V={1} p, p_1, p_2 and p_4 have value 1 or p, p_2, p_3 and p_4 have value 1



D_m shortest distance between p and p_4 is 3



D_m distance between p and p_4 is 4

Linear and Nonlinear Operations

$$H(af+bh)=aH(f)+bH(g)$$

- Where H is an operator whose input and output are images.
- f and g are two images
- a and b are constants.

H is said to be linear operation if it satisfies the above equation or else H is a nonlinear operator.