Chapter 2

DIGITAL IMAGING FUNDAMENTALS

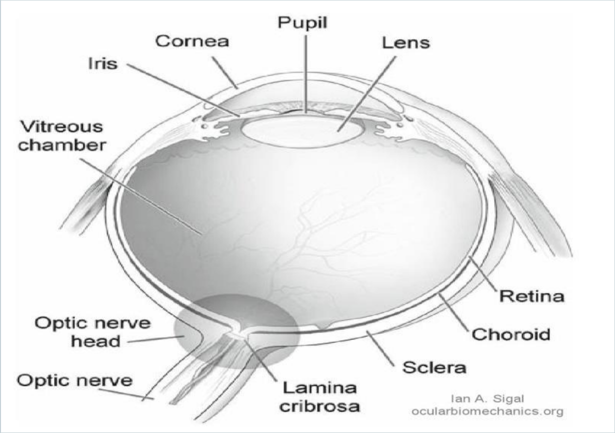
- This lecture will cover:
  - The human visual system.
  - Image formation in the eye.
  - Simultaneous contrast & optical illusion.
  - Image acquisition & formation.
  - Image sampling and quantization.
  - Image representation .
  - Spatial & intensity resolution.
  - Some basic relationships between pixels.
  - Mathematical tools used in image processing.
  - Image transforms.

Human Visual System

- The best vision model we have!
- Knowledge of how images form in the eye can help us with processing digital images
- We will see the structure of the human visual system

structure of human visual system

- The lens focuses light from objects onto the retina
- The retina is covered with light receptors called *cones* and *rods*.



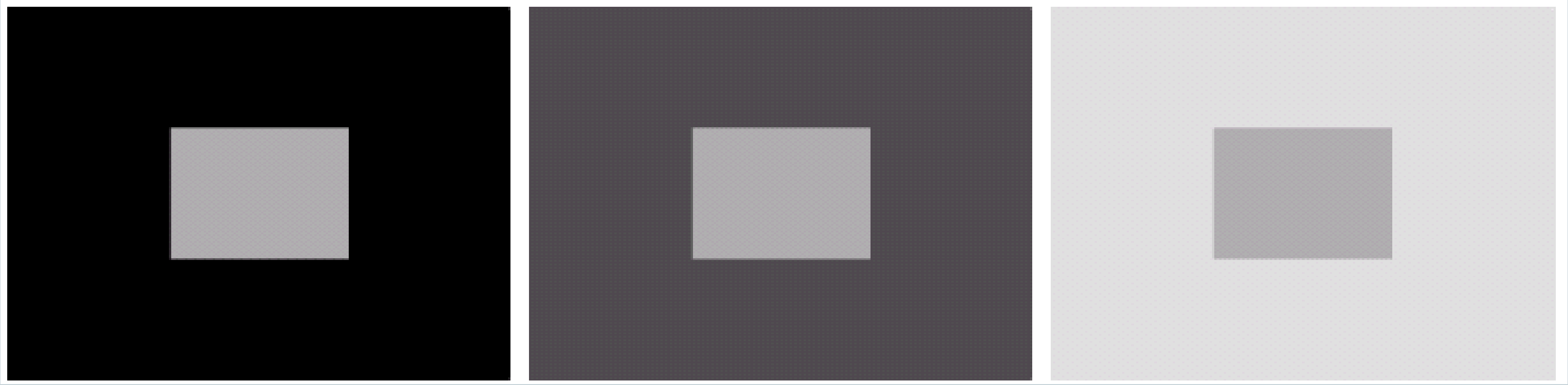
- Cones are (6-7 million) concentrated at the central portion of the retina and high sensetive to color.
- Rods are (75-150 million) distributed at retina surface and are sensitive to low levels of illumination.

Image Formation In The Eye

- In ordinary camera,the lens has a fixed focal length,and foucsing at various distances is achieved by varying the distance between the lens and the image palne.
- But in human eye, Muscles within the eye can be used to change the shape of the lens allowing us focus on objects that are near or far away

simultaneous contrast

- Is related to the fact that a region's perceived brightness does not depend on its intensity.
- Ex, All the inner squares have the same intensity,but they appear progressively darker as the background becomes lighter.



Optical Illusions

- In which the eye fills in nonexisting information or wrongly perceives geometrical properties of objects

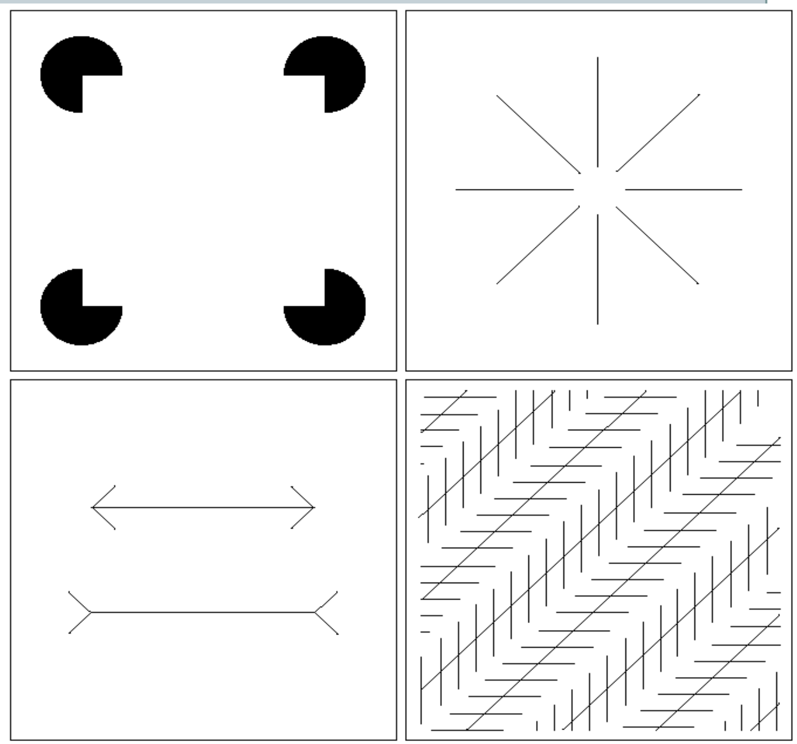
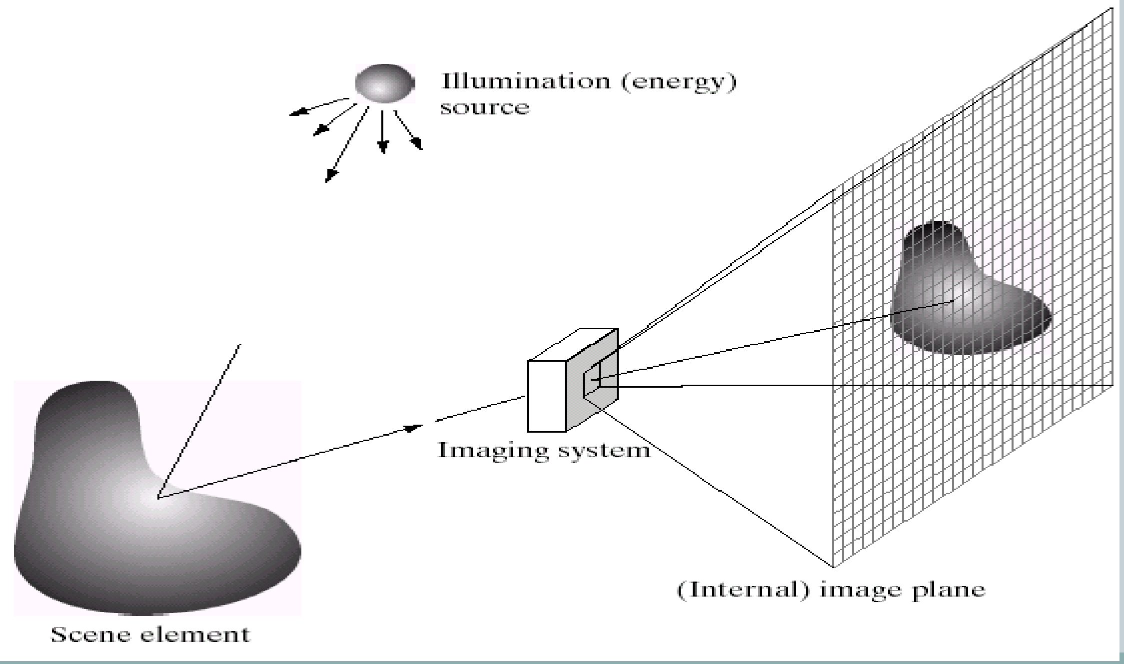


Image Acquisition

- to acquire a digital image
- Images are generated by the combination of *illumination* source and reflection of energy by the objects in that scene.



A simple image formation model

- Images are two dimensional functions of the form f(x,y) which mean the value or amplitude of ‘f’ at spatial coordinates (x,y) .
- F(x,y) characterized by i(x,y) and r(x,y) as following

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

Where,

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

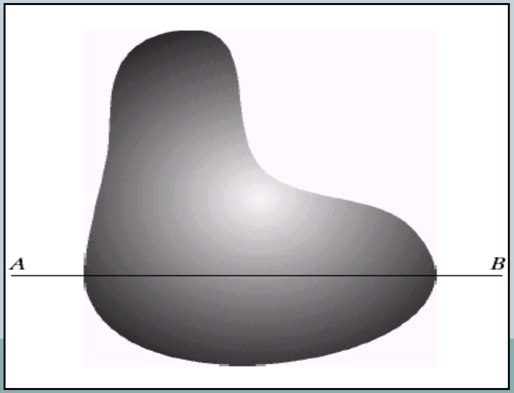
and $$0<r(x,y) <1$$

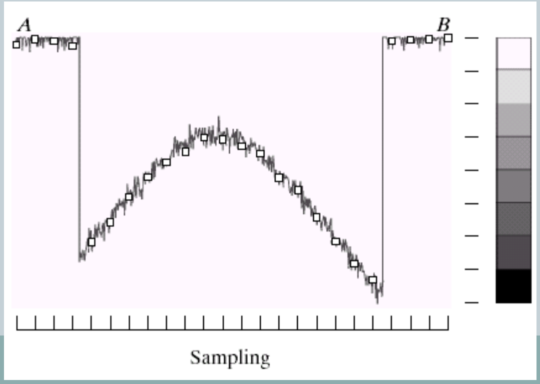
- Reflectance is bounded by 0 (total absorption) and 1(total reflectance).

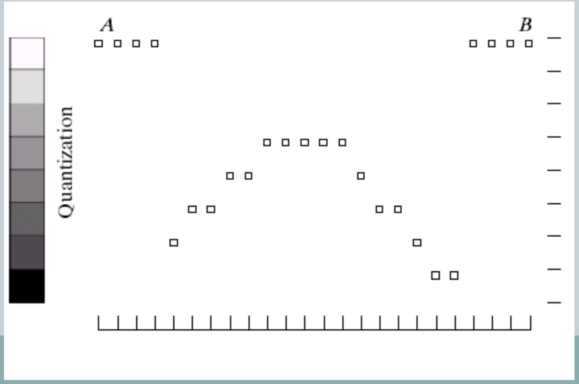
Image Sampling And Quantisation

- The ouput of most sensors is a continuous waveform so to create a digital image we need to do two processes :

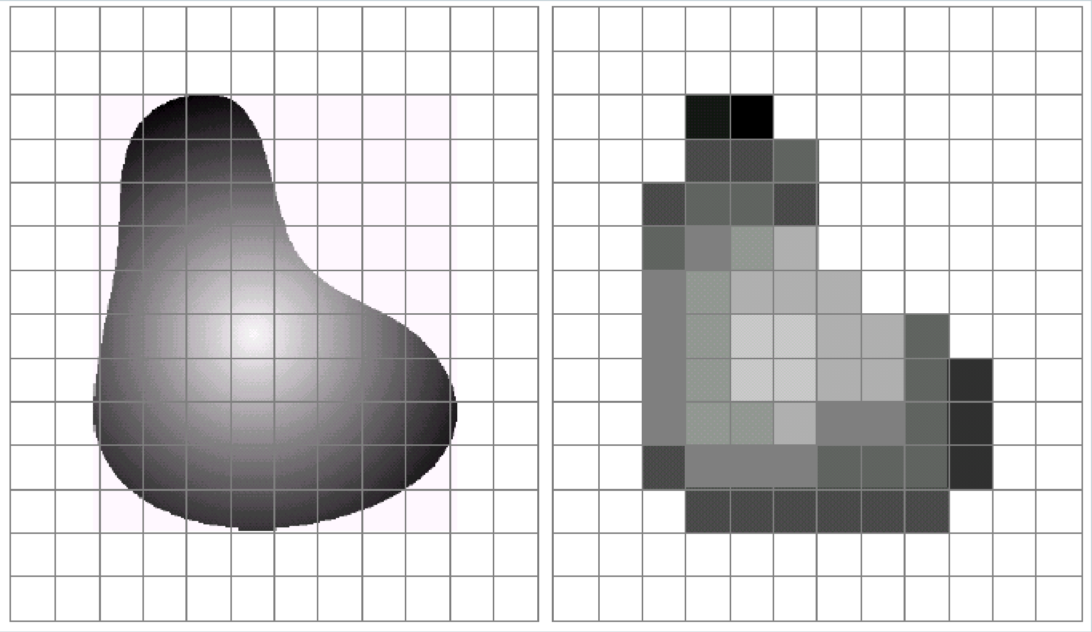
1. Image sampling : is to digitizing the image coordinates.
2. Image quantization: is to digitizing the image amplitude(intensity level).







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



(a) (b)

Representing digital images

- Image representation: to convert the input data to a form suitable for computer processing.

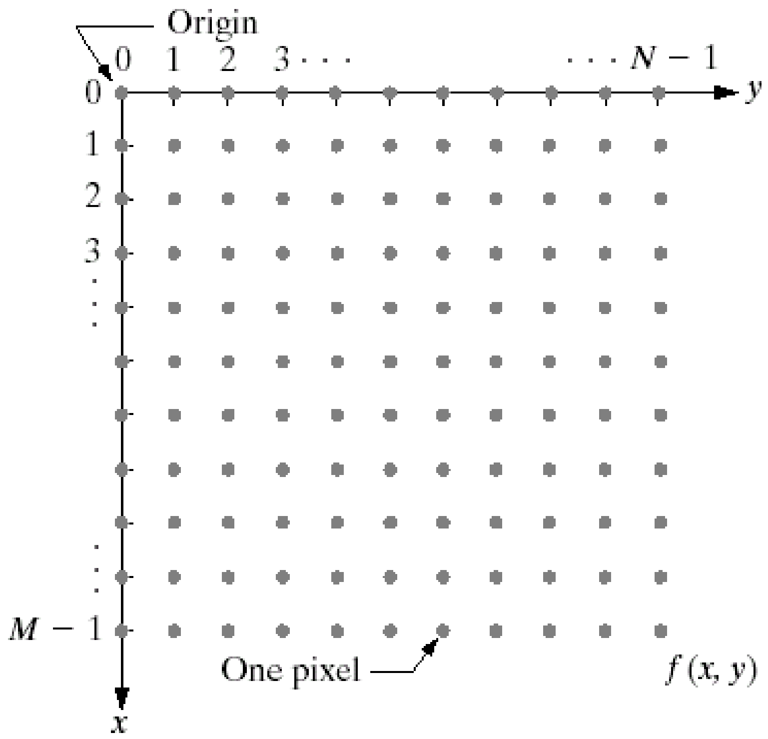
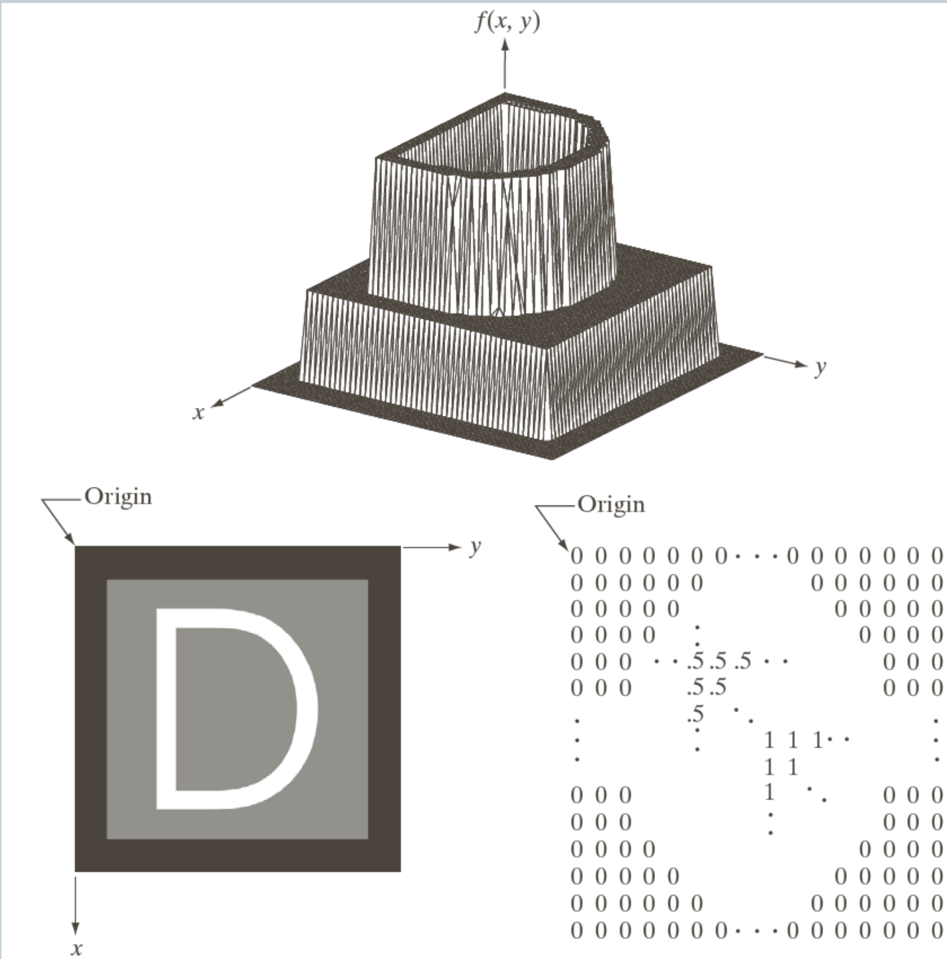


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

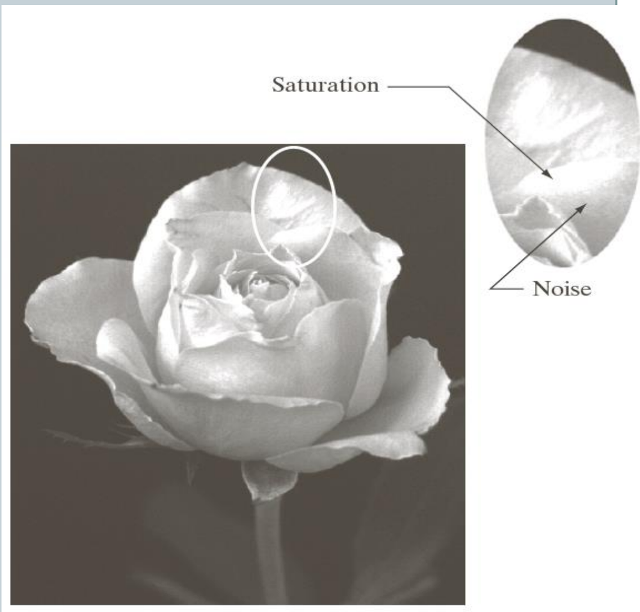


a
b c

FIGURE 2.18
(a) Image plotted as a surface.
(b) Image displayed as a visual intensity array.
(c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

Dynamic range

- It can be defined as the ratio between maximum measurable intensity level and minimum detectable intensity level in an image.
- As a rule, the upper limit is determined by the saturation and the lower limit by noise.
- Image contrast : is the difference in intensity between the highest and lowest intensity level in image.

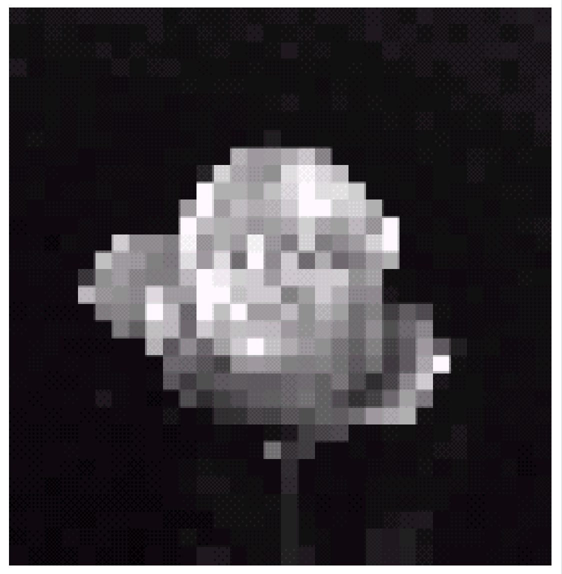


Spatial and intensity resolution

- Spatial resolution is:
  - a measure of the smallest detail in an image.
  - line pairs or dots per unit distance(dpi).
  - Note: to discriminate between two images we need to know dpi and size of the image.

1024

256



64

- intensity resolution:
  - is the smallest discernible change in intensity level.
  - the more intensity levels used, the finer the level of detail discernable in an image
  - Intensity level resolution is usually given in terms of the number of bits used to store each intensity level

Number of Bits	Number of Intensity Levels	Examples
1	2	0, 1
2	4	00, 01, 10, 11
4	16	0000, 0101, 1111
8	256	00110011, 01010101
16	65,536	1010101010101010

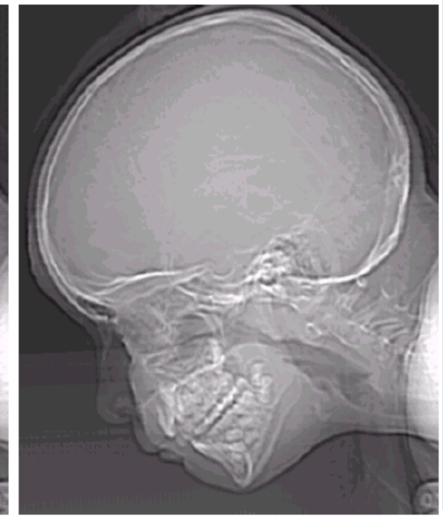
Ex,

256 grey levels (8 bits per pixel)

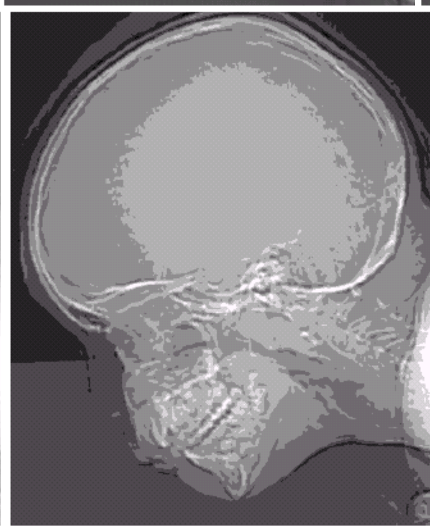
128 grey levels (7 bpp)

64 grey levels (6 bpp)

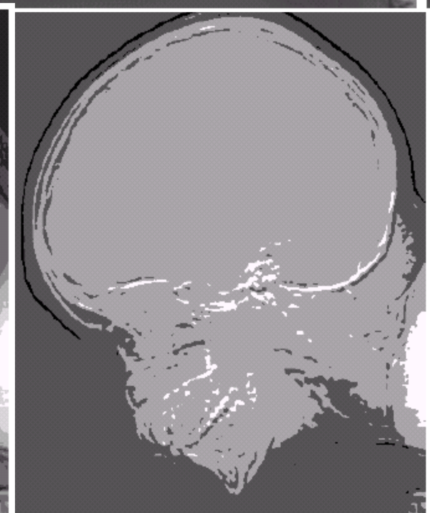
32 grey levels (5 bpp)



16 grey levels (4 bpp)



8 grey levels (3 bpp)



4 grey levels (2 bpp)

2 grey levels (1 bpp)

Some basic relationships between pixels

- Neighbors of pixel

  A pixel 'p' at coordinates(x,y) has

1. four horizontal and vertical neighbors called 4-neighbors N4(p):

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

1. Four diagonal neighbors of 'p' called ND(p) :

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

1. N4(p) and Nd(p) together are constructing N8(p).

Adjacency

•Let v be the set of intensity values used to define adjacency.
•In binary image V={1} if we refer to adjacency of pixels with values 1.
•But in grey image ,the adjacency of pixels with values from 0 till 255 ,so set V could be any subset from 0 till 255.
•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

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

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

$$D_e(p, q) = \left[(x - s)^2 + (y - t)^2\right]^{\frac{1}{2}} \tag{2.5-1}$$

The D_4 *distance* (called the *city-block distance*) between p and q is defined as

$$D_4(p, q) = |x - s| + |y - t| \tag{2.5-2}$$

The D_8 *distance* (called the *chessboard distance*) between p and q is defined as

$$D_8(p, q) = \max(|x - s|, |y - t|) \tag{2.5-3}$$

Mathematical tools used in digital image processing

- Consider the following 2 * 2 images

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

The *array product* of these two images is

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

On the other hand, the *matrix product* is given by

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

- In this book we use array product,meaning that the division is between corresponding pixels.

Linear versus non linear operations

- Consider a general operator H that produce an output g(x,y) from an input image f(x,y)

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

- So H is said to be linear operator if

$$\begin{aligned} H[a_i f_i(x, y) + a_j f_j(x, y)] &= a_i H[f_i(x, y)] + a_j H[f_j(x, y)] \\ &= a_i g_i(x, y) + a_j g_j(x, y) \end{aligned} \qquad (2.6\text{-}2)$$

- Where a_i, a_j are constants and $f_i(x, y)$, and $f_j(x, y)$ are images of same size.

Linear example

As a simple example, suppose that H is the sum operator, Σ; that is, the function of this operator is simply to sum its inputs. To test for linearity, we start with the left side of Eq. (2.6-2) and attempt to prove that it is equal to the right side:

$$\begin{aligned}\sum\left[a_i f_i(x, y) + a_j f_j(x, y)\right] &= \sum a_i f_i(x, y) + \sum a_j f_j(x, y) \\ &= a_i \sum f_i(x, y) + a_j \sum f_j(x, y) \\ &= a_i g_i(x, y) + a_j g_j(x, y)\end{aligned}$$

- The first step come from the fact that summtion is distributive.so the summtion is alinear operator.

Non linear example

- Consider the following two images

$$f_1 = \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad f_2 = \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}$$

and suppose that we let $a_1 = 1$ and $a_2 = -1$. To test for linearity, we again start with the left side of Eq. (2.6-2):

$$\max\left\{(1)\begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1)\begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}\right\} = \max\left\{\begin{bmatrix} -6 & -3 \\ -2 & -4 \end{bmatrix}\right\}$$
$$= -2$$

Working next with the right side, we obtain

$$(1)\max\left\{\begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix}\right\} + (-1)\max\left\{\begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}\right\} = 3 + (-1)7$$
$$= -4$$

The left and right sides of Eq. (2.6-2) are not equal in this case, so we have proved that in general the max operator is nonlinear.

Arithmetic operations

- Four arithmetic operation are denoted as:

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

- Images in arithmetic operations must be of the same size.

Averaging noisy images for noise reduction

- as

■ Let $g(x, y)$ denote a corrupted image formed by the addition of noise, $\eta(x, y)$, to a noiseless image $f(x, y)$; that is,

$$g(x, y) = f(x, y) + \eta(x, y) \tag{2.6-4}$$

if an image $\bar{g}(x, y)$ is formed by averaging K different noisy images,

$$\bar{g}(x, y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x, y) \tag{2.6-5}$$

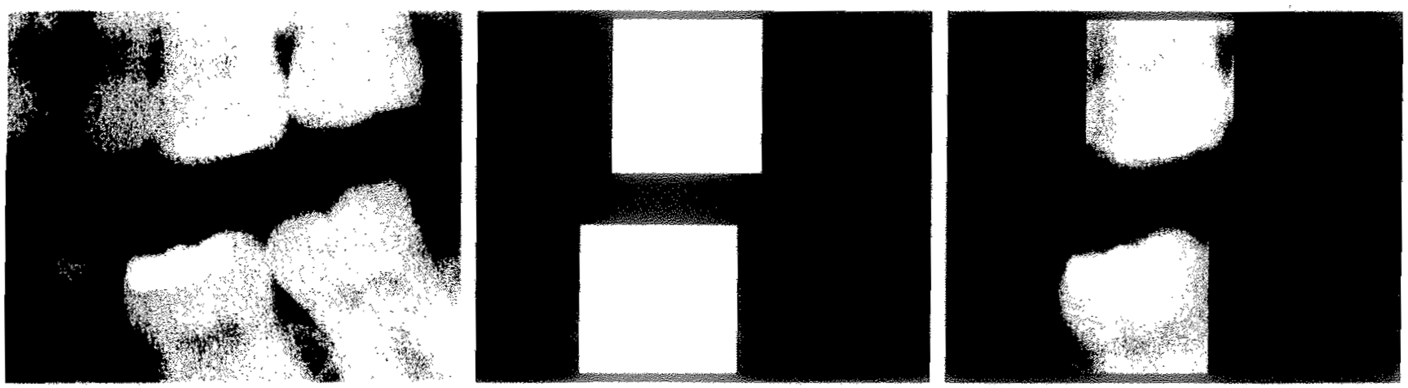
then it follows that

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

Using multiplication in masking(ROI)

- Region of interest operation simply consists of :

Multiplying agiven image by a mask that has 1s in the ROI and 0s elsewhere.



a b c

FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

- Note:

In arthmetic operations such as difference between two 8-bit images can range from a min. Of -255 and a max. Of 255 ,and the values of sum two images can range from 0 to 510.

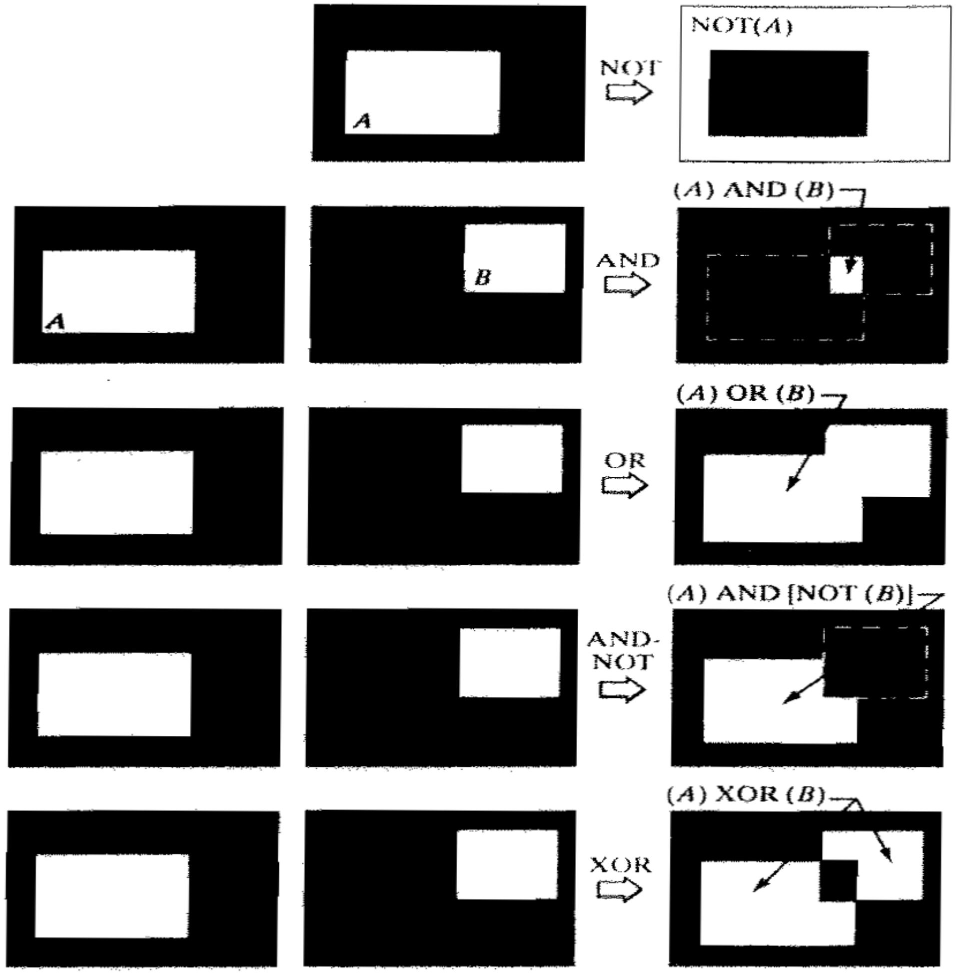
So we need to set all negative values to 0 and set to 255 all values exceed this limit.

Set and logical operations

- Basic set operations
  - If a is an element of A ,we write $a \in A$
  - If a is not an element of A , we write $a \notin A$
  - The set with no elemnts called null set or empty set and is denoted by the symbol $\varnothing$
  - If every elemnt in set A is an element in set B then A is said to be subset of B $A \subseteq B$
  - The union of two sets A and B denoted as $C = A \cup B$
  - The set of elements belongs to either A , B or both $D = A \cap B$
  - Two sets A and B are said to be disjoint or mutually exclusive if they have no common elemnts $A \cap B = \varnothing$
  - Complement set of A is all elemnts not in A $A^c = \{w | w \notin A\}$
  - Diffrence of two sets denoted as A-B $A - B = \{w | w \in A, w \notin B\} = A \cap B^c$

Logical operations

FIGURE 2.33
Illustration of logical operations involving foreground (white) pixels. Black represents binary 0s and white binary 1s. The dashed lines are shown for reference only. They are not part of the result.



Spatial operations

- Spatial operations are performed directly on the pixels of a given image.
  - Point : the output value at a specific coordinate is dependent only on the input value at that same coordinate.
  - Local : the output value at a specific coordinate is dependent on the input values in the neighborhood of that same coordinate
  - Global : the output value at a specific coordinate is dependent on all the values in the input image

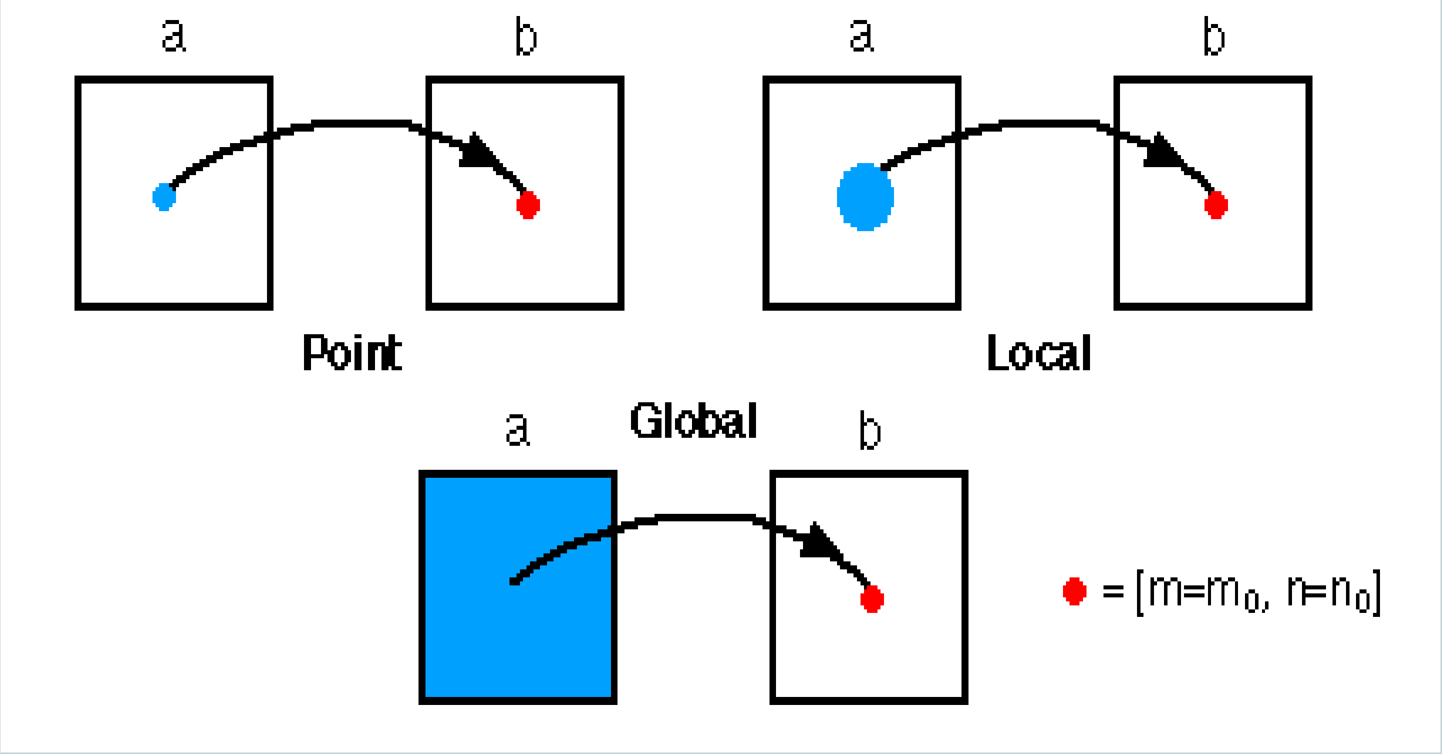

a
b
a
b
Point
Local
a
Global
b
= [m=m0, n=n0]

Image transforms

- Most of image processing approches works directly on spatial domain .but in some cases it is better to work on transform domain and applying the inverse transform to return to spatial domain .
- A 2-D linear transform can be expressed as:

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) r(x, y, u, v)$$

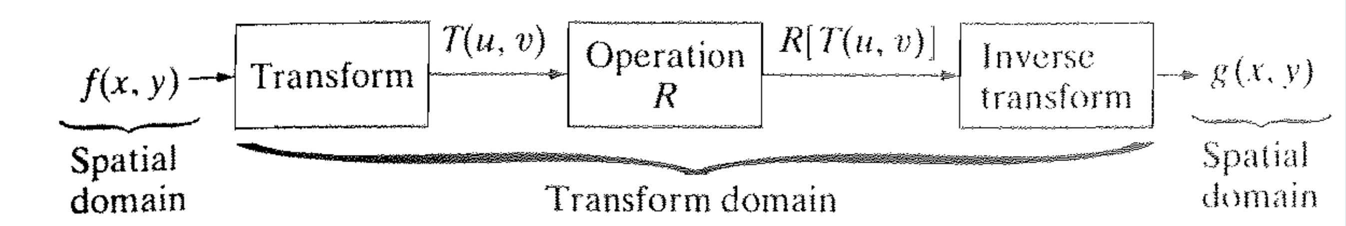
- Where f(x,y) is the input image and r(x,y,u,v) is the forward transformation kernel.

$$r(x, y, u, v) = e^{-j2\pi(ux/M + vy/N)}$$

- We can recover f(x,y) using the inverse transform T(u,v).

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) s(x, y, u, v)$$

- Where s(x,y,u,v) is called inverse transformation kernel. $s(x, y, u, v) = \frac{1}{MN} e^{j2\pi(ux/M + vy/N)}$



General approach for operating in the linear transform domain.

Probabilistic methods

let z_i, $i = 0, 1, 2, \ldots, L - 1$, denote the values of all possible intensities in an $M \times N$ digital image. The probability, $p(z_k)$, of intensity level z_k occurring in a given image is estimated as

$$p(z_k) = \frac{n_k}{MN}$$

where n_k is the number of times that intensity z_k occurs in the image and MN is the total number of pixels. Clearly,

$$\sum_{k=0}^{L-1} p(z_k) = 1$$

Once we have $p(z_k)$, we can determine a number of important image characteristics. For example, the mean (average) intensity is given by

$$m = \sum_{k=0}^{L-1} z_k \, p(z_k) \tag{2.6-44}$$

Similarly, the variance of the intensities is

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 \, p(z_k)$$

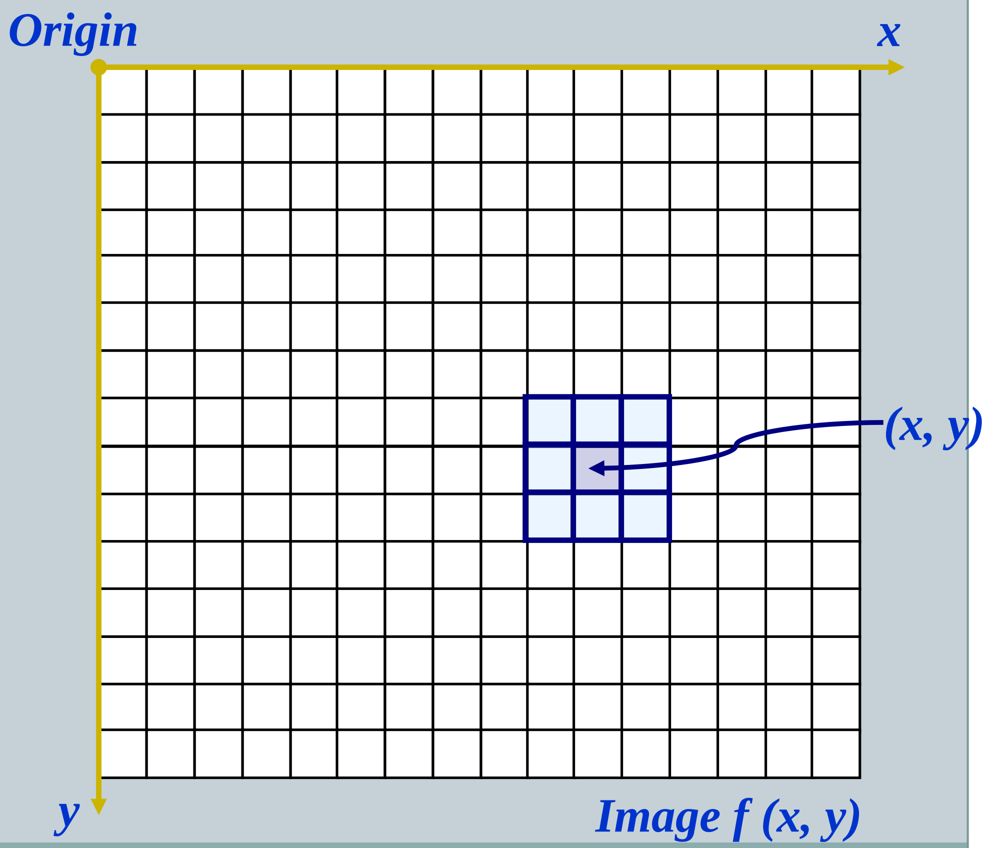
Chapter 3

Image Enhancement

- is the process of manipulating images so that the result is more suitable than the original for aspecific application. As following:
    - Highlighting interesting detail in images.
    - Removing noise from images.
    - Making images more visually appealing.
- Enhancement are problem oriented techniques.
- Ex, the method for enhancing X-ray images may be not suitable for enhancing satellite images .

Spatial Domain Image Enhancement

- Most spatial domain enhancement operations can be reduced to the form
  - $g(x, y) = T[f(x, y)]$
- where $f(x, y)$ is the input image, $g(x, y)$ is the processed image and T is some operator defined over some neighbourhood of (x, y)



Point Processing

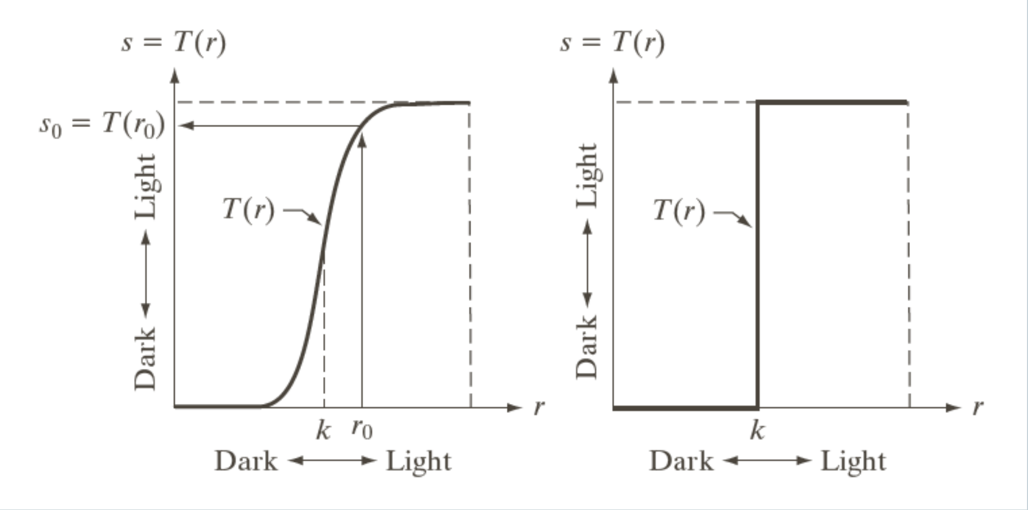
- The simplest spatial domain operations occur when the neighbourhood is simply the pixel itself
- Point processing operations take the form
    - $s = T(r)$
- where s refers to the processed image pixel value and r refers to the original image pixel value .

Intensity Transformations

- Values of r lower than k are compresed by the transformation function into a narrow range of s, toward black. The opposite is true for values greater than k .this is called contrast stretching.
- This transformation

Function is called

Thresholding function.



Basic intensity transformation functions

- There are three basic Types of functions:
  - Linear(negative and Identity transformation).
  - Logarithmic (Log and invers-log tranformation).
  - Power-low(nth-power and nth-root transformation).

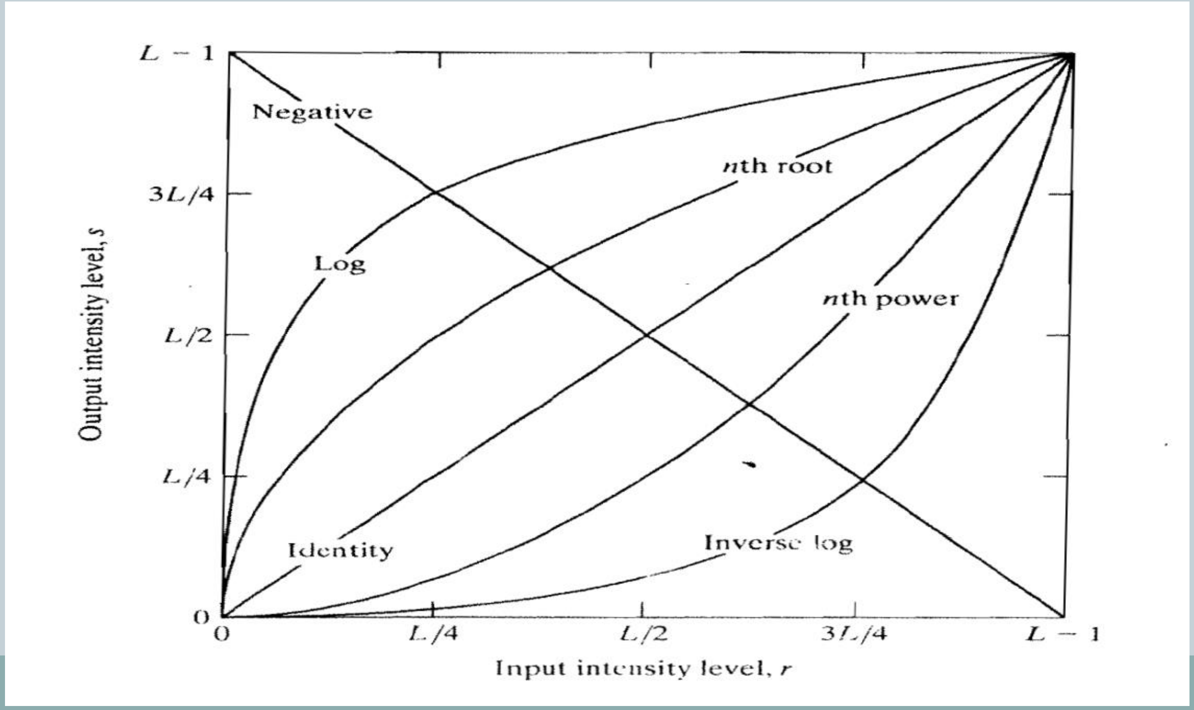
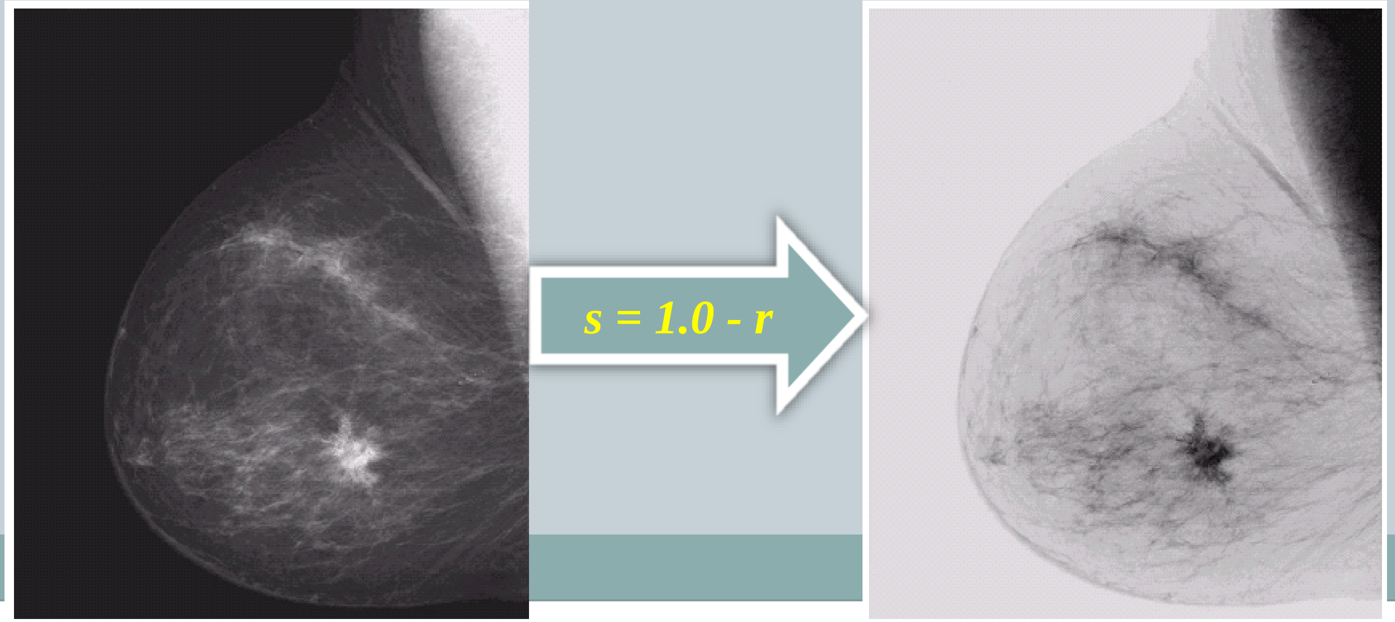


image negatives

- The negative of the image with intensity level in the range [0,L-1] is given by the expression:
  - $S=L-1-r$
- it is suitable for enhancing white or gray details embeded in black regions of images.

Origina l Image



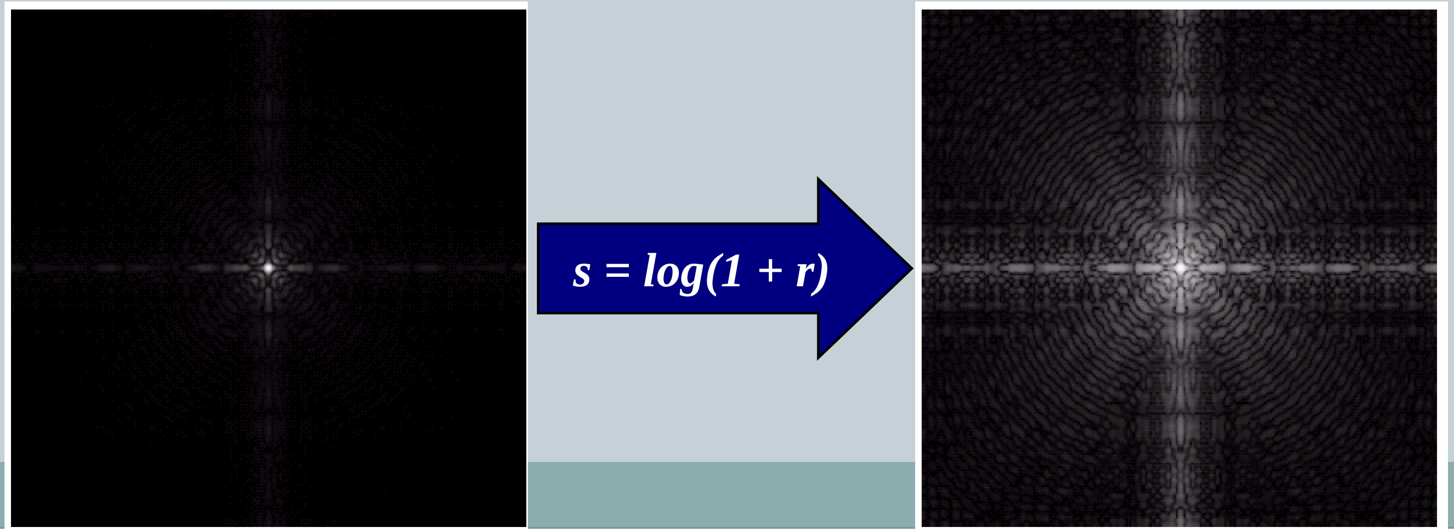
Negative Image

Logarithmic transformation

- General form:
  - $s = c * log(1 + r)$
- The log transformation maps a narrow range of low input grey level values into a wider range of output values .(brights images).
- The inverse log transformation performs the opposite transformation .

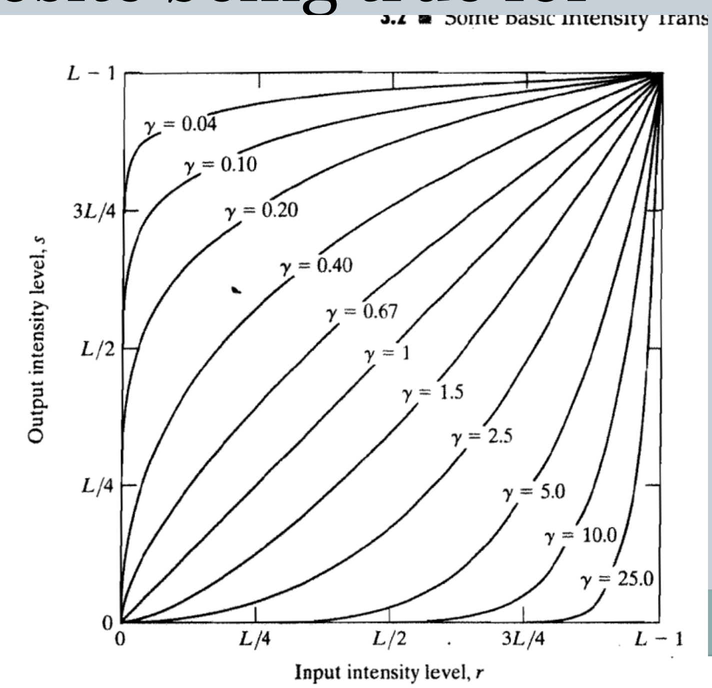
Logarithmic Transformations (cont…)

- Log functions are particularly useful when the input grey level values may have an extremely large range of values
- In the following example the Fourier transform of an image is put through a log transform to reveal more detail



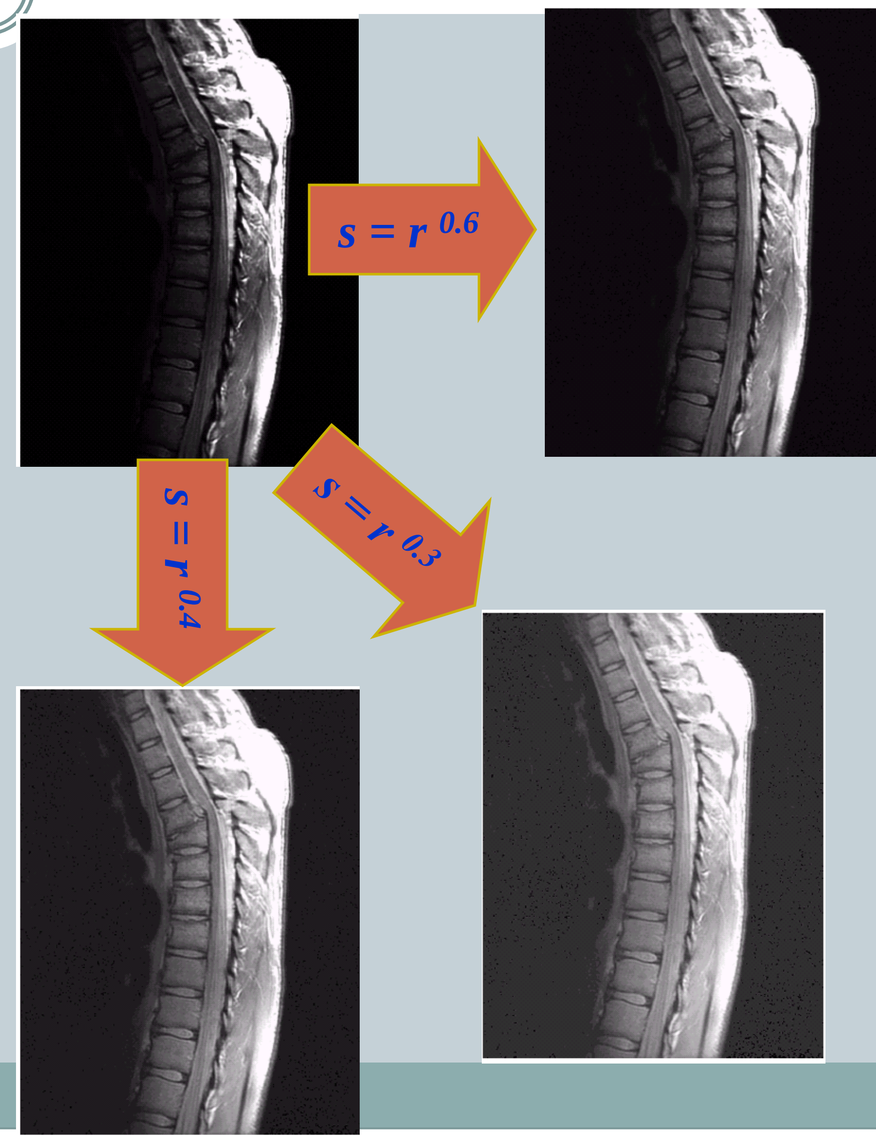
Power-law (gamma) transformation

- It has the following form:
  - $s = c * r^{\gamma}$
- Power-law curves with fracional values of γ map a narrow range of dark input values into a wider range of output values, with the opposite being true for higher values off input levels.



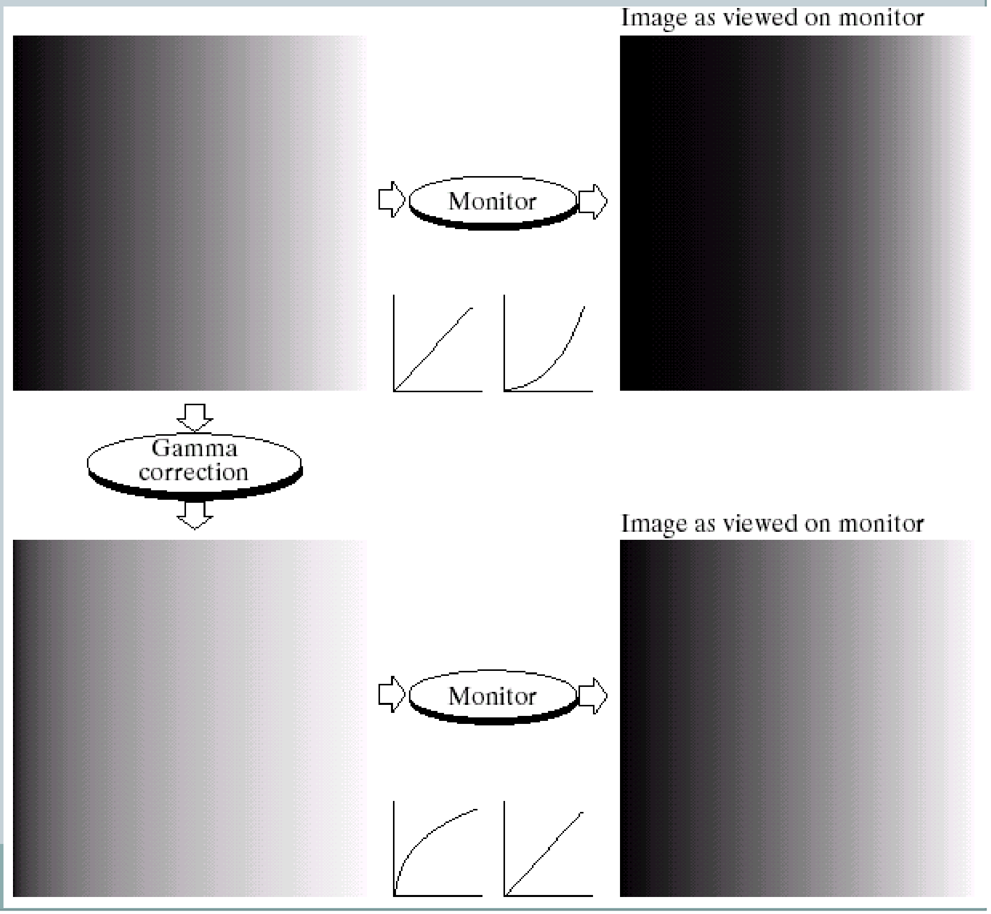
Power Law Example (cont…)

- The images to the right show a magnetic resonance (MR) image of a fractured human spine
- Different curves highlight different detail



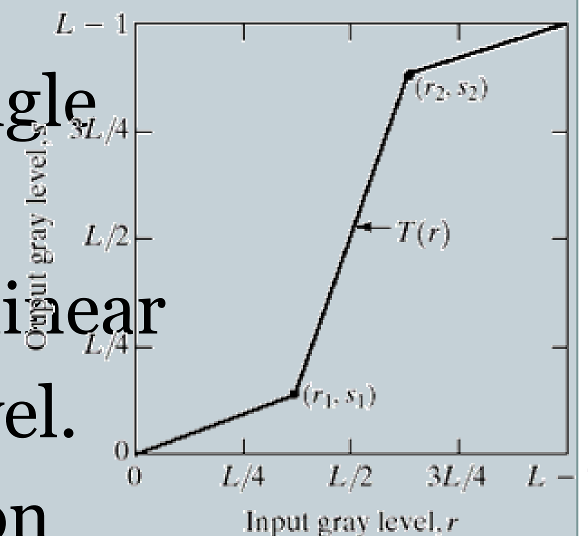
Gamma Correction

- Many of you might be familiar with gamma correction of computer monitors
- Problem is that display devices do not respond linearly to different intensities
- Can be corrected using a log transform



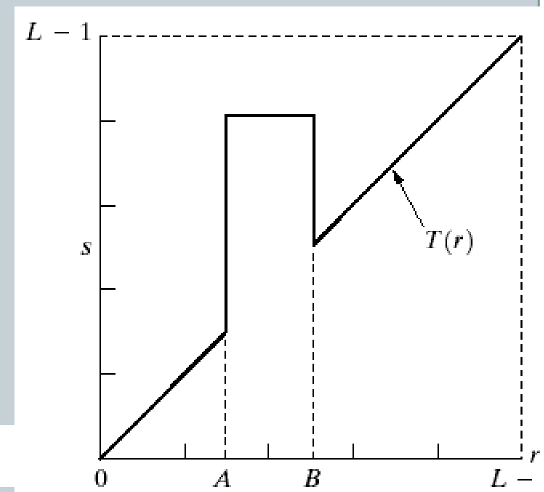
Piecewise Linear Transformation Functions

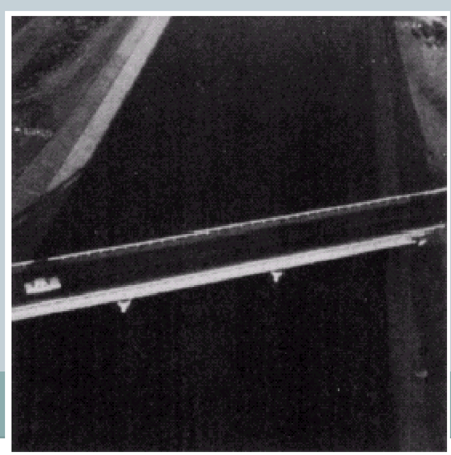
- Contrast stretching: is a process that expands the range of intensity levels in an image so that it spans the full intensity range of the device .
- $r_1 \le r_2$ and $s_1 \le s_2$ so the function is single valued and monotonically increasing..
- If $r_1 = s_1$ and $r_2 = s_2$,so the function is linear that produce no change in intensity level.
- If $r_1 = r_2$, $s_1 = 0$ and $s_2 = L - 1$,the function is thresholding that creates a binary image.

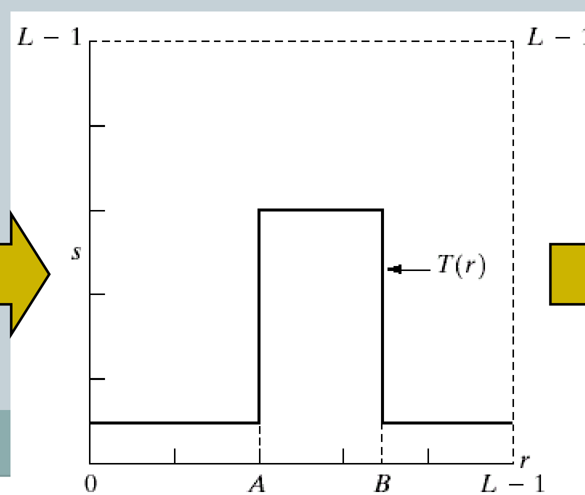


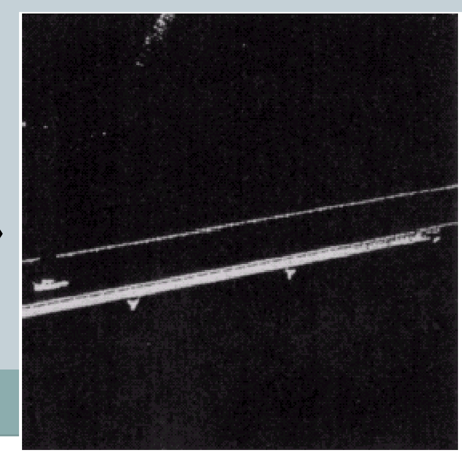
Intensity level slicing

- Highlights a specific range of grey levels.
  - To display in one value(white) all the values in the range of interest and in another(black) all other intensities.
  - Brightens or darkens the desired range of intenities but leaves all other intensity levels in the image unchanged.









Bit Plane Slicing

- Often by isolating particular bits of the pixel values in an image we can highlight interesting aspects of that image
  - Higher-order bits usually contain most of the significant visual information
  - Lower-order bits contain subtle details