UNIT-I

DIGITAL IMAGE FUNDAMENTALS & IMAGE TRANSFORMS

What Is Digital Image Processing?

- The field of digital image processing refers to processing digital images by means of a digital computer.

What is a 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.*

- When x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a *digital image*

Picture elements, Image elements, pels, and pixels

- A digital image is composed of a finite number of elements, each of which has a particular location and value.
- These elements are 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.

The Origins of Digital Image Processing

- **One of the first applications of digital images was in the newspaper industry, when pictures were first sent by submarine cable between London and New York.**

- **Specialized printing equipment coded pictures for cable transmission and then reconstructed them at the receiving end.**

- Figure was transmitted in this way and reproduced on a telegraph printer fitted with typefaces simulating a halftone pattern.

- The initial 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

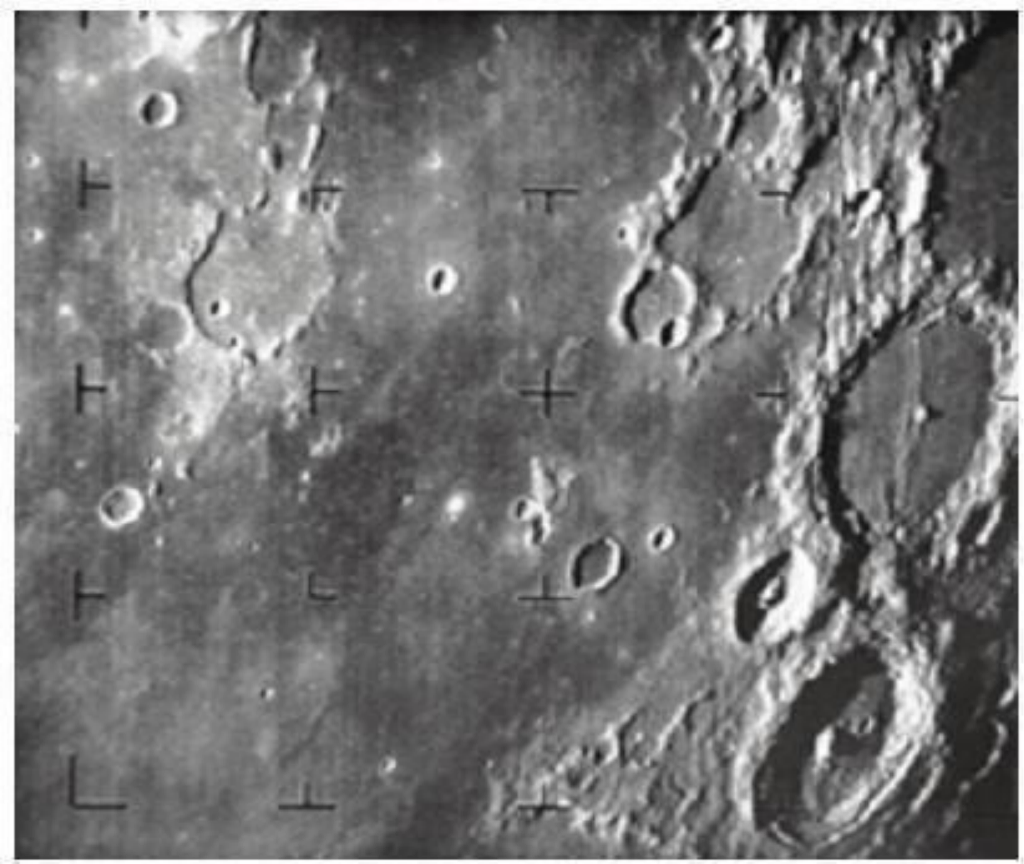
- The printing technique based on photographic reproduction made from tapes perforated at the telegraph receiving terminal from 1921.

- Figure shows an image obtained using this method.
- The improvements are tonal quality and in resolution.

- The early Bartlane systems were capable of coding images in five distinct levels of gray.
- This capability was increased to 15 levels in 1929.

- Figure is typical of the type of images thatcould be obtained using the 15-tone equipment.

- Figure shows the first image of the moon taken by *Ranger*



Applications of DIP

- The field of image processing has applications in medicine and the space program.

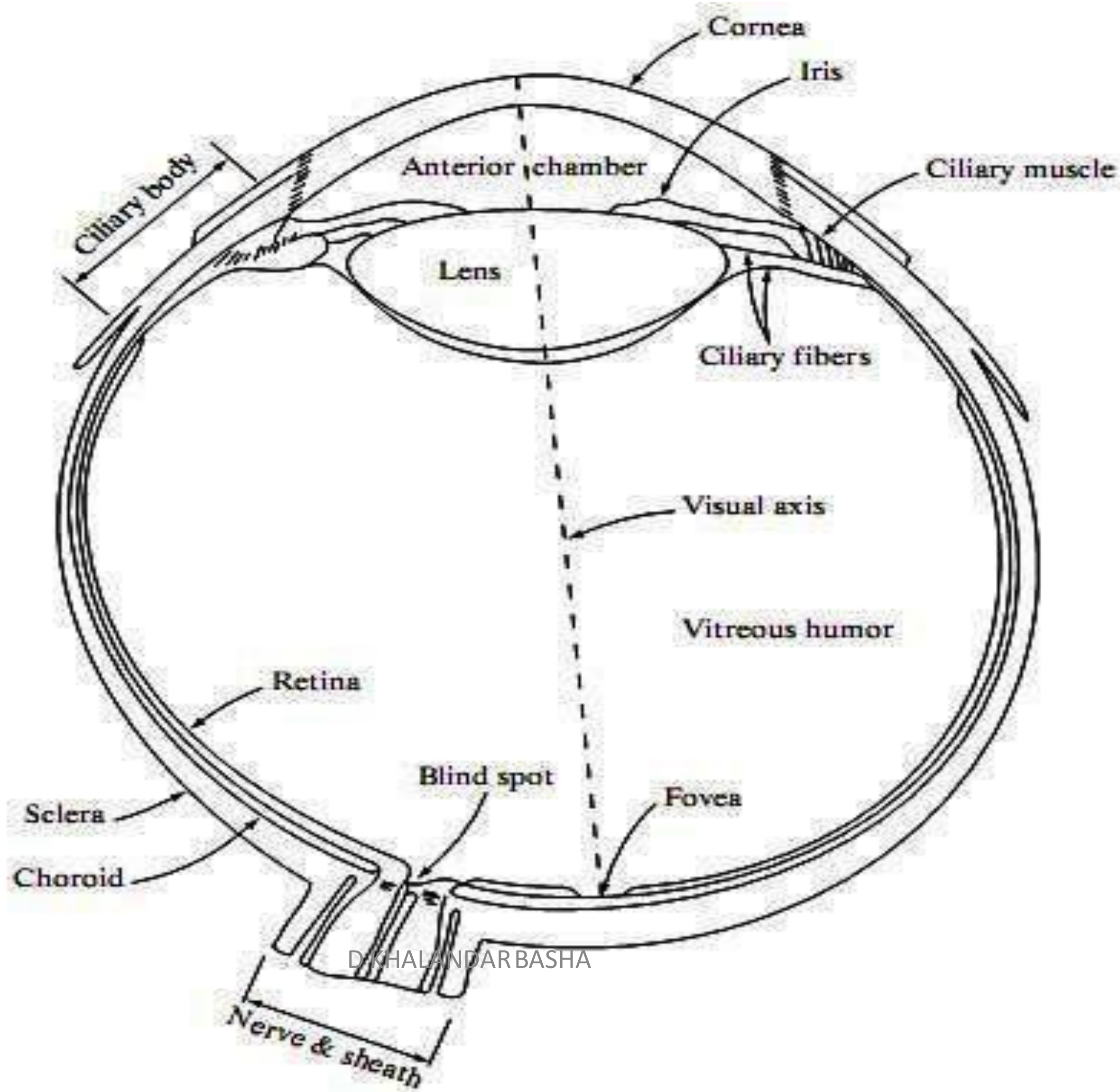
- Computer procedures are used to enhance the contrast or code the intensity levels into color for easier interpretation of X-rays and other images used in industry, medicine, and the biological sciences.

- Geographers use the same or similar techniques to study pollution patterns from aerial and satellite imagery

- **Image enhancement and restoration procedures are used to process degraded images of unrecoverable objects or**

- **Experimental results too expensive to duplicate.**

Structure of the Human Eye



- The eye is nearly a sphere, with an average diameter of approximately 20mm.

- Three membranes enclose the eye:

- The cornea and sclera outer cover the choroid the retina.

Cornea

- The cornea is a tough, transparent tissue that covers the anterior surface of the eye.

- Continuous with the cornea,the sclera is an opaque membrane that encloses the remainder of the optic globe.

Choroid

- The choroid lies directly below the sclera.
- This membrane contains a net- work of blood vessels that serve as the major source of nutrition to the eye.
- The choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe.

- At its anterior extreme, the choroid is divided into the ciliary body and the iris diaphragm.

- The latter contracts or expands to control the amount of light that enters the eye

- The front of the iris contains the visible pigment of the eye, whereas the back contains a black pigment.

- The lens is made up of concentric layers of fibrous cells and is suspended by fibers that attach to the ciliary body.

- It contains 60 to 70% water, about 6% fat, and more protein than any other tissue in the eye.

Retina

- The innermost membrane of the eye is the retina, which lines the Inside of the ‖all’s entire posterior portion.

- When the eye is properly focused, light from an object outside the eye is imaged on the retina.

- Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina.

- There are two classes of receptors: cones and rods.

- The cones in each eye number between 6 and 7 million.

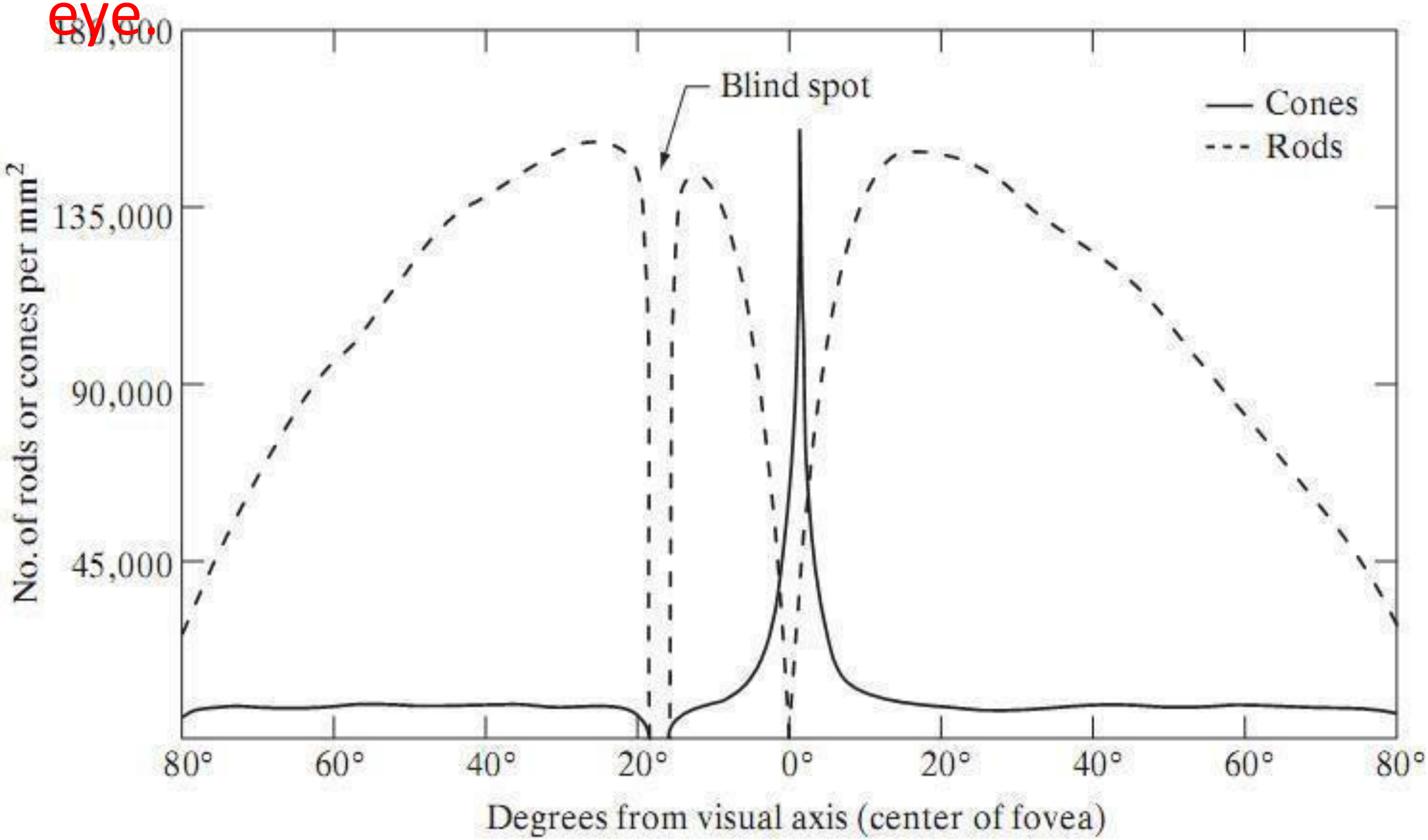
- They are located primarily in the central portion of the retina, called the fovea, and are highly sensitive to color.

- Muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea.

- Cone vision is called photopic or bright-light vision.

- The number of rods is much larger: Some 75 to 150 million are distributed over the retinal surface.

- Figure shows the density of rods and cones for a cross section of the right eye passing through the region of emergence of the optic nerve from the eye.



- The absence of receptors in this area results in the so-called blind spot.

- Fig. shows that cones are most dense in the center of the retina (in the center area of the fovea)

Image Formation in the Eye

- The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible.

- The shape of the lens is controlled by tension in the fibers of the ciliary body.

- To focus on distant objects, the controlling muscles cause the lens to be relatively flattened.

- Similarly, these muscles allow the lens to become thicker in order to focus on objects near the eye.

- The distance between the center of the lens and the retina called the *focal length varies from approximately 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum.*

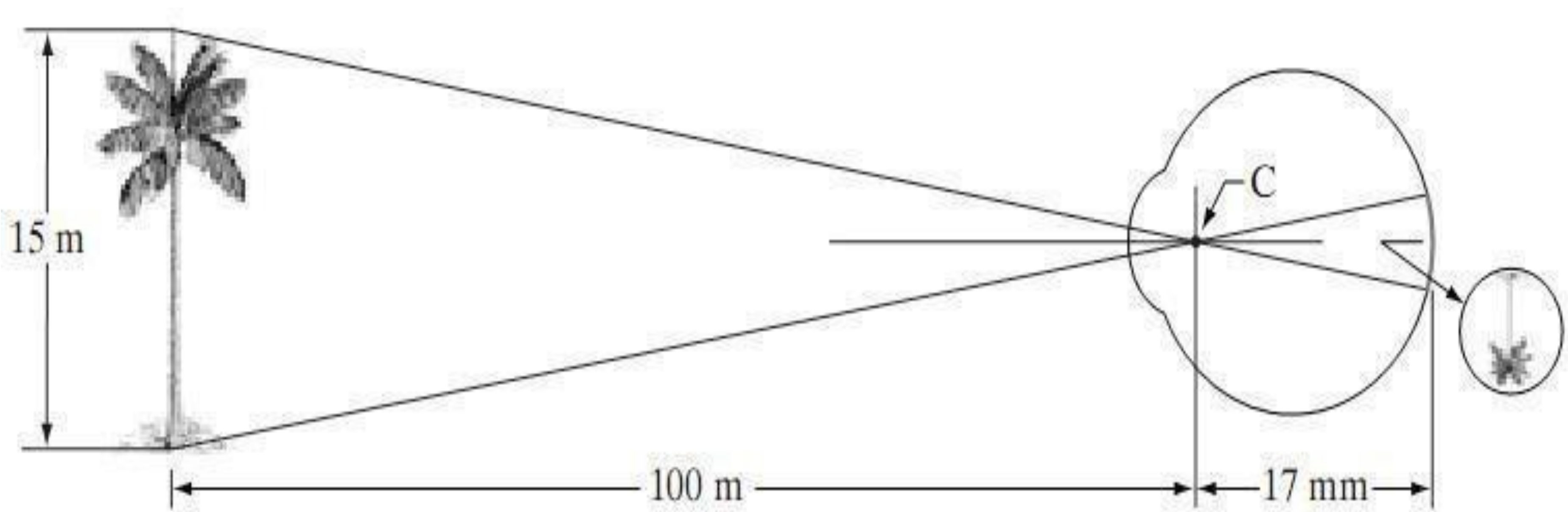
- *When the eye* focuses on an object farther away the lens exhibits its lowest refractive power.

- When the eye focuses on a nearby object, the lens is most strongly refractive.

- 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

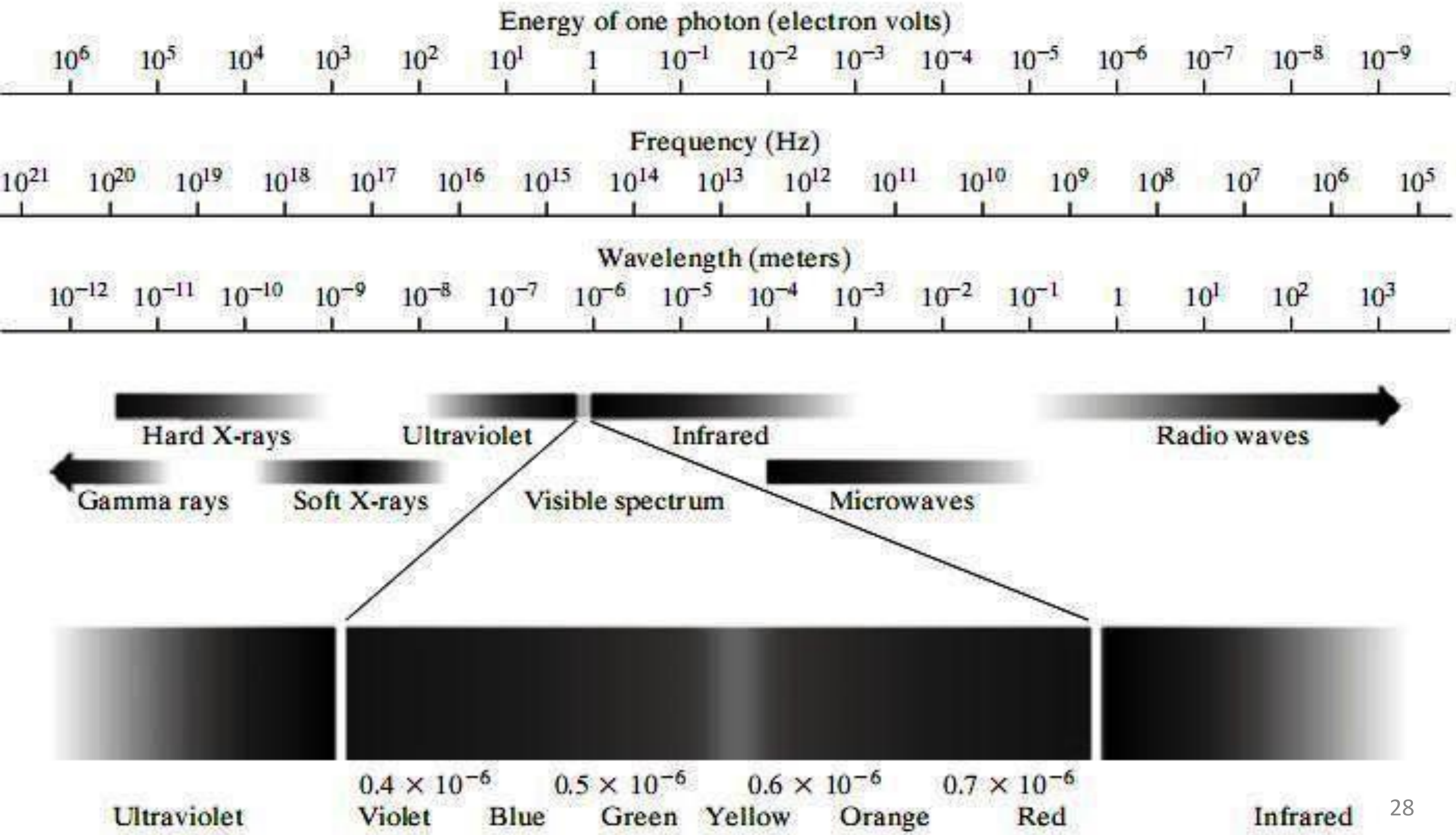
15/100 = h/17 or h=2.55mm.



Light and the Electromagnetic Spectrum

- 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 at one end to red at the other.

The electromagnetic spectrum



- The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy.

- Wavelength (l)and frequency (n)are related by the expression

- where c is the speed of light (2.998*10^8 m s)

- The energy of the electromagnetic spectrum is given by the expression E = hv

- where *h is Plank"s constant*

A Simple Image Formation Model

- Images by two-dimensional functions of the form f(x, y).

- The value or amplitude of f at spatial coordinates (x, y) gives the intensity (brightness) of the image at that point.

- As light is a form of energy, f(x,y) must be non zero and finite.

- The function f(x, y) may be characterized by two components:

  (1)the amount of source illumination incident on the scene being viewed

  (2)the amount of illumination reflected by the objects in the scene.

- These are called the *illumination and reflectance components and are denoted by i(x, y) and r(x, y), respectively.*

- The two functions combine as a product to form f(x, y):

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

$$r(x, y) = 0 \quad \text{---} \quad \text{total absorption}$$

$$1 \quad \text{---} \quad \text{total reflection}$$

- The intensity of a monochrome image f at any coordinates (x, y) the *gray level (l) of the image at that point.*

That is, $l = f(x_0, y_0)$

L lies in the range $L_{\min} \leq \ell \leq L_{\max}$

In practice, $L_{\min} = i_{\min} r_{\min}$ and $L_{\max} = i_{\max} r_{\max}$.

GRAY SCALE

- The interval [L_{min} , L_{max}] is called the*gray scale.*

- *Common practice is to shift this interval numerically to the interval [0, L-1],*

- *where L = 0 is considered black and*

L = L-1 is considered white on the gray scale.

All intermediate values are shades of gray varying from black to white.

Basic Relationships Between Pixels

- <u>1. **Neighbors of a Pixel :-**</u>

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.

N_D(p) and N_8(p)

- 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, arecalled the 8-*neighbors of p, denoted by* N_8(p).

- If some of the points in N_D(p) and N_8(p) fall outside the image if (x, y) is on the border of theimage.

Adjacency, Connectivity, Regions, and Boundaries

- To establish whether two pixels are connected, it must be determined if they are neighbors and
- if their gray levels satisfy a specified criterion of similarity (say, if their gray levels are equal).

- For instance, in a binary image with values 0 and 1, two pixels may be 4-neighbors,

- but they are said to be connected only if they have the same value

- Let *V be the set of gray-level values used todefine connectivity. In a binary image, V={1} for the connectivity of pixels with value 1.*

- *In a grayscale image, for connectivity of pixels with a range of intensity values of say 32, 64 V typically contains more elements.*

- *For example,*

- *In the adjacency of pixels with a range of possible gray-level values 0 to 255,*
- *set V could be any subset of these 256 values.We consider three types of adjacency:*

- **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 ***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).

$N_4(p) \cap N_4(q)$

(d) 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 theset whose values are from V.*

- A *path from pixel p with coordinates (x, y) to pixel q with coordinates (s, t) is a sequence of distinct pixels with coordinates*

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

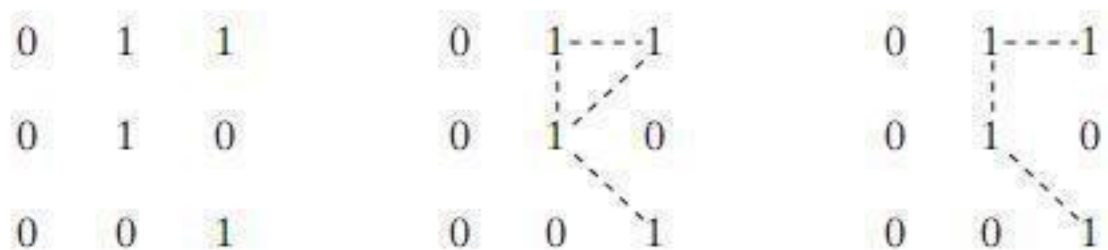
- where $(x_{\text{б’}}\ \ y_0) = (x, y)$ and $(x_{\text{Ŷ’}}\ \ y_n) = (s, t)$,

$(x_{i’}\ \ y_i)$ and $(x_{i\text{-б’}}\ \ y_{i\text{-}1})$ pixels and are adjacent for б ч i ч Ŷ. IŶ this Đase, Ŷ is the *length of the path.*

- *If* $(x_{\text{б’}}\ \ y_0) = (x_{\text{Ŷ’}}\ \ y_n)$ the path is a *closed path*.

- Two pixels p and q are said to be *connected in S if there exists a path between them consisting entirely of pixels in S.*

- *For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S.*



a b c

FIGURE (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) m-adjacency.

Relations, equivalence

- A binary relation R on a set A is a set of pairs of elements from A. If the pair (a, b) is in R, the notation used is aRb (ie a is related to b)

- Ex:- the set of points A = { p1,p2,p3,p4} arranged as

P1p2

P3

p4

- In this case R is set of pairs of points from A that are 4-connected that is R = {(p1,p2), (p2,p1), (p1,p3), (p3,p1)} .

thus p1 is related to p2 and p1 is related to p3 and vice versa but p4 is not related to any other point under the relation .

Reflective - Symmetric - Transitive

- Reflective

if for each a in A, aRb

- Symmetric

if for each a and b in A, aRb implies bRa

- Transitive

if for a, b and c in A, aRb and bRc implies aRc

A relation satisfying the three properties is called an equivalence relation.

Distance Measures

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

  (a) D(p, q)D;p, if p = q),

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

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

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

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

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

  •For example, the pixels with D_4 distance ч Ͳ 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		

- The pixels with D_4=1 are the 4-neighbors of (x, y).

- The D_8 *distance (also called chess board distance)between p and q is defined as*

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

- *For example, the pixels with* D_8 *distance ≤ 2 froŵ (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).

- The D_m 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, p2 and p4 have value 1 and that p1 and p3 can have a value of 0 or 1:

	p3	p4
P1	p2	
p		

- If only connectivity of pixels valued 1 is allowed, and p1 and p3 are 0 then the m distance between p and p4 is 2.

- If either p1 or p3 is 1, the distance is 3

- If both pi and p2 are 1, the distance is 4

DIGITAL IMAGE PROCESSING

UNIT 2: IMAGE ENHANCEMENT

Process an image so that the result will be more suitable than the original image for a application. specific

Highlighting interesting detail in images Removing noise from images

Making images more visually appealing

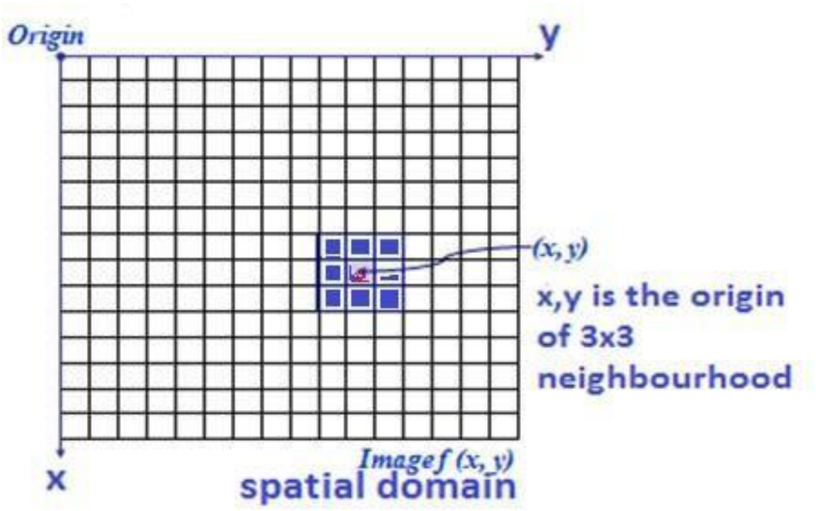
So, a technique for enhancement of x-ray image may not be the best for enhancement of microscopic images.

These spatial domain processes are expressed by:

G (x,y) = T (f(x,y)) depends only on the value of f at (x,y)

f(x,y) is the input image, G (x,y) is the output image

T is called a gray-level or intensity transformation operator which can apply to single image or binned images.



Window origin is moved from image origin along the 1st row and then second row etc.

At each location, the image pixel value is replaced by the value obtained after applying T operation on the window at the origin.

the neighborhood size may be different. We can have a neighborhood size of 5 by 5, 7 by 7 and so on depending upon the type of the image and the type of operation that we want to have.

Spatial domain techniques

Point Processing: Contrast stretching Thresholding

Intensity transformations / gray level transformations

> Image Negatives
> Log Transformations
> Power Law Transformations

Piecewise-Linear Transformation Functions Contrast stretching
Gray-level slicing Bit-plane slicing

Spatial filters

Smoothening filters Low pass filters Median filters
Sharpening filters High boost filters
Derivative filters

Suppose we have a digital image which can be represented by a two dimensional random field $f(x, y)$.

An image processing operator in the spatial domain may be expressed as a mathematical function $T[\cdot]$ applied to the input image $f(x, y)$ to produce a new image $g(x, y) = T[f(x, y)]$ as follows.

The operator T applied on $f(x, y)$ may be defined over some neighborhood of (x, y)

(i) A single pixel (x, y). In this case T is a gray level transformation (or mapping) function.

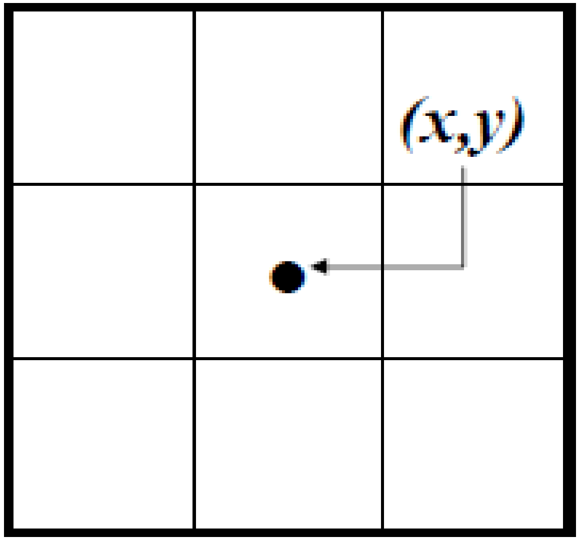
(ii) Some neighborhood of (x, y).

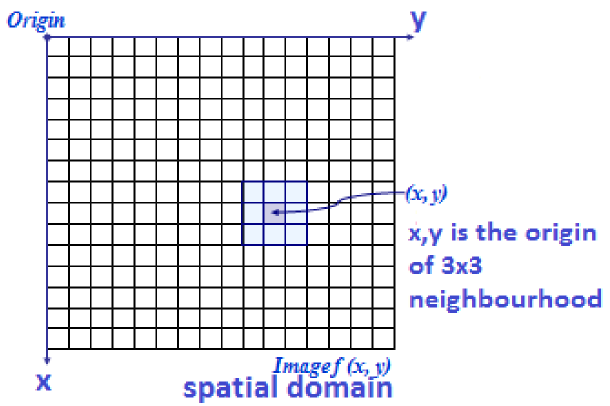
(iii) T may operate to a set of input images instead of a single image.

the neighborhood of a point (x, y) is usually a square sub image which is centered at point (x, y).

Mask/Filter

Neighborhood of a point (x,y) can be defined by using a square/rectangular (common used) or circular subimage area centered at (x,y)
The center of the subimage
is moved from pixel to pixel starting at the top of the corner





Spatial Processing :
intensity transformation -> works on single pi
contrast manipulation

image thresholding

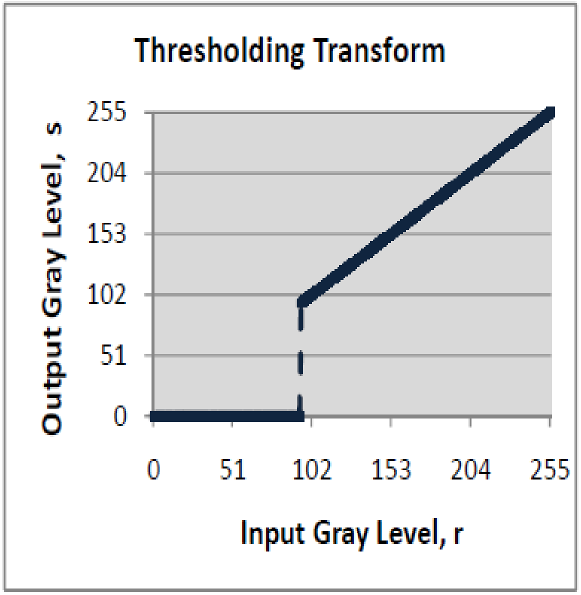
spatial filtering → Image sharpening (working on neighborhood of every pixel) or Neighborhood Processing:

Thresholding (piece wise linear transformation)

Produce a two-level (binary) image

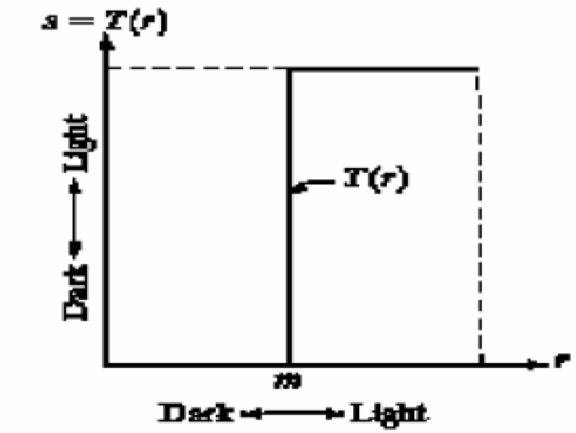
For any $0 < t < 255$ the threshold transform Thr_t can be defined as:

$$s = Thr_t(r) = \begin{cases} 0 & if\ r < t \\ r & otherwise \end{cases}$$



Thresholding has another form used to generate binary images from the gray-scale images, i.e.:

$$s = Thr_t(r) = \begin{cases} 0 & if\ r < t \\ 255 & otherwise \end{cases}$$



Thresholding transformations are particularl useful for segmentation in which we want to isolate an object of interest from a background