Introduction to Digital Image Processing

Dr. Kundan Kumar
Ph.D. (IIT Kharagpur)
Assistant Professor
ECE Department (Cabin - E139)

Institute of Technical Education & Research (ITER)
Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India-751030

Digital Image Fundamentals

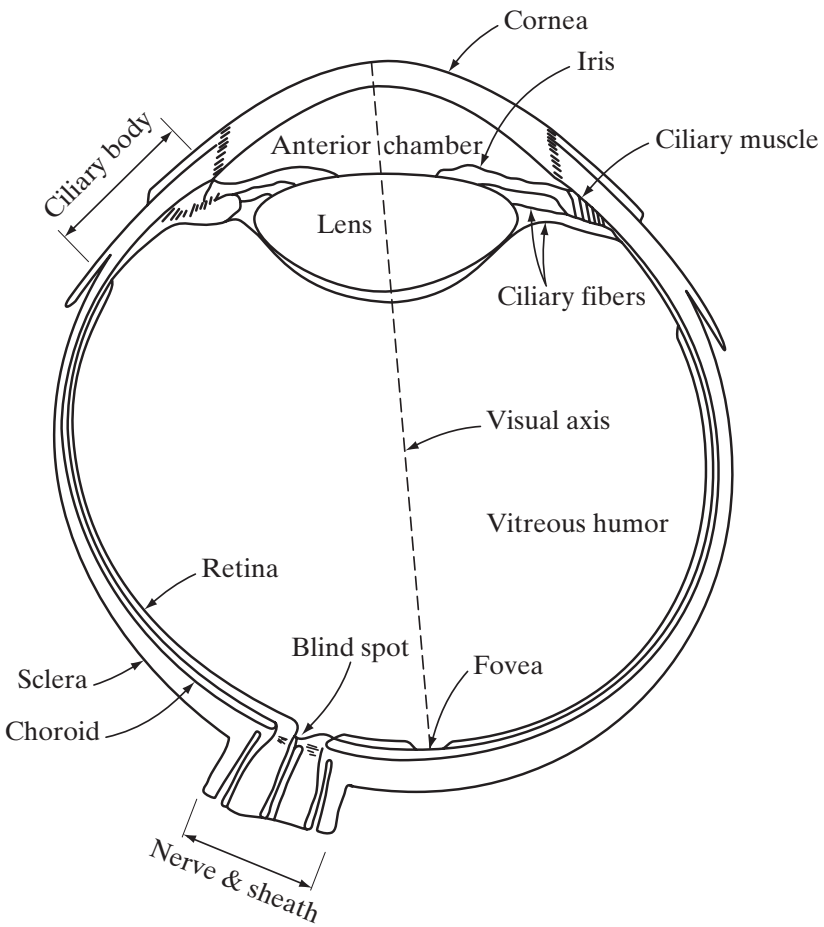
Human and computer vision

- We can't think of image processing without considering the human vision system. We observe and evaluate the images that we process with our visual system.
- Without taking this elementary fact into consideration, we may be much misled in the interpretation of images.

Simple Questions

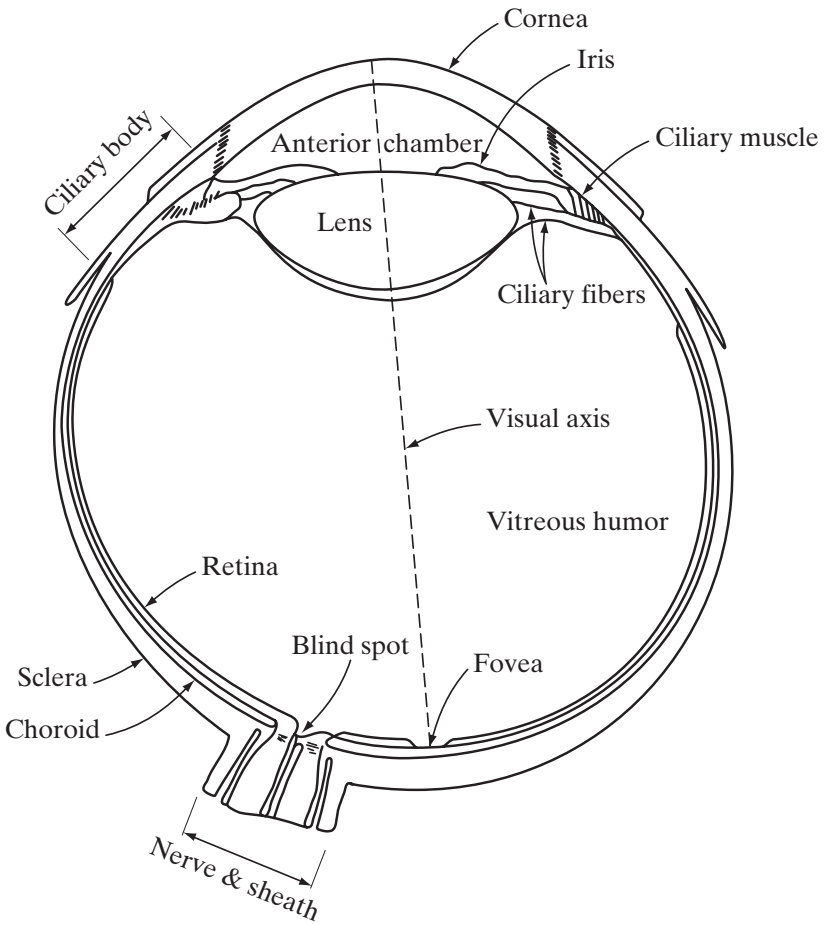
- What intensity differences can we distinguish?
- What is the spatial resolution of our eye?
- How accurately we estimate and compare distances and areas?
- How do we sense colors?
- By which features can we detect and distinguish objects?

Human Eye



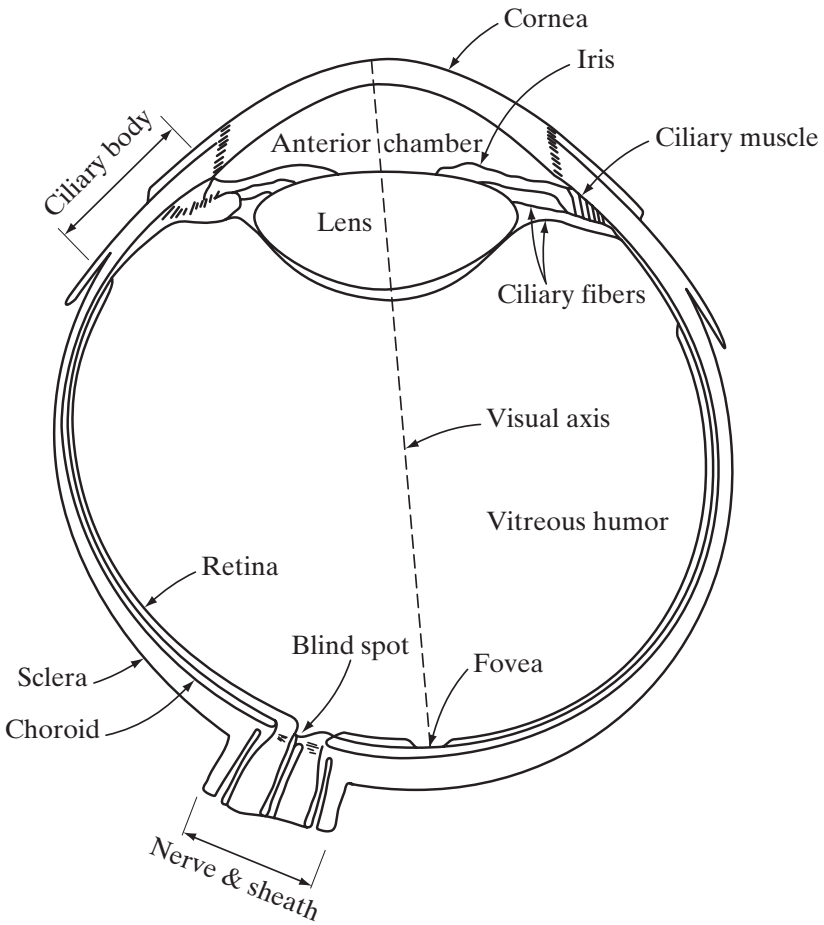
- Shape is nearly a sphere.
- Average diameter = 20 mm.
- 3 membranes:
  - Cornea and Sclera - outer cover
  - Choroid
  - Retina -enclose the eye

Human Eye



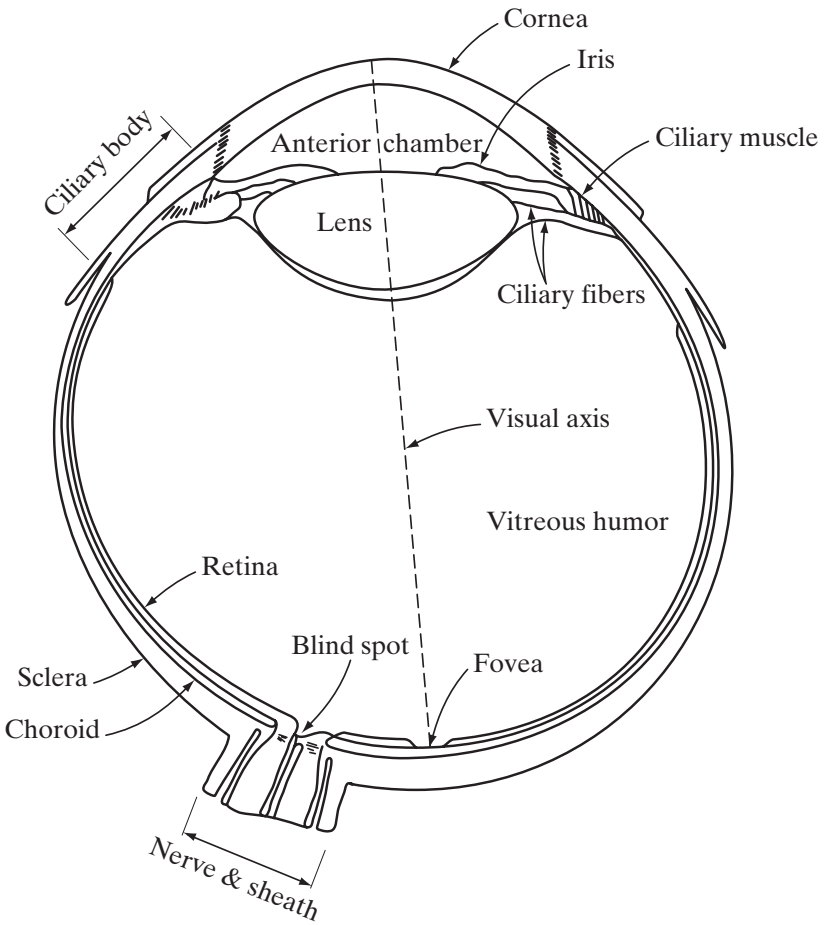
- Cornea:
  - tough, transparent tissue, covers the anterior surface of the eye.
- Sclera
  - Opaque membrane, encloses the remainder of the optic globe

Human Eye



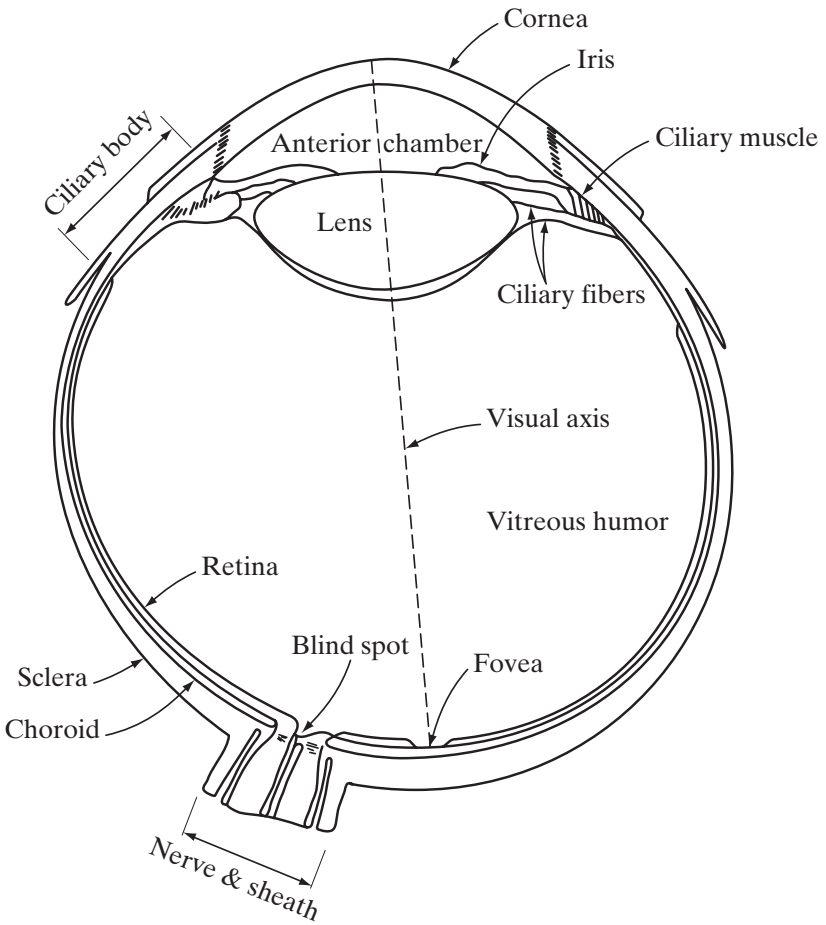
- Choroid:
  - Lies below the sclera, contains network of blood vessels that serve as the *major source of nutrition* to the eye.
  - Choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe.

Human Eye



- Lens:
  - both infrared and ultraviolet light are absorbed appreciably by proteins within the lens structure and, in excessive amounts, can cause damage to the eye.
- Retina:
  - Innermost membrane of the eye which lines inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina.

Human Eye



Receptors:

- Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina.
- Receptors are divided into 2 classes:
  - Cones
  - Rods

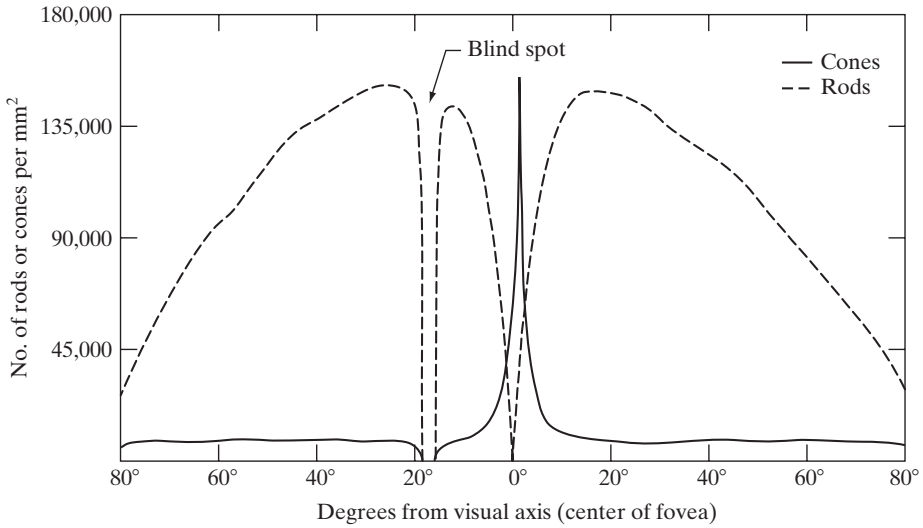
Cones

- 6-7 million, located primarily in the central portion of the retina (the *fovea*, muscles controlling the eye rotate the eyeball until the image falls on the fovea).
- Highly sensitive to color.
- Each is connected to its own nerve end thus human can resolve fine details.
- Cone vision is called *photopic* or *bright-light vision*.

Rods

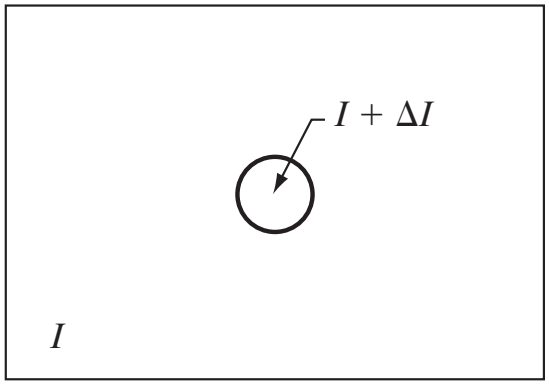
- 75-150 million, distributed over the retina surface.
- Several rods are connected to a single nerve end reduce the amount of detail discernible.
- Serve to give a general, overall picture of the field of view.
- Sensitive to low levels of illumination.
- Rod vision is called *scotopic* or *dim-light vision*.

Cross section of the right eye



- *Blind spot* $\Rightarrow$ the absence of receptors area.
- Receptor density is measured in degrees from the fovea.
- Cones are most dense in the center of the retina (in the area of the fovea).
- Rods increase in density from the center out to approx. 20^o off axis and then decrease in density out to the extreme periphery of the retina.

Contrast Sensitivity



Weber's ratio: $\nabla I_c/I$

Good brightness discrimination

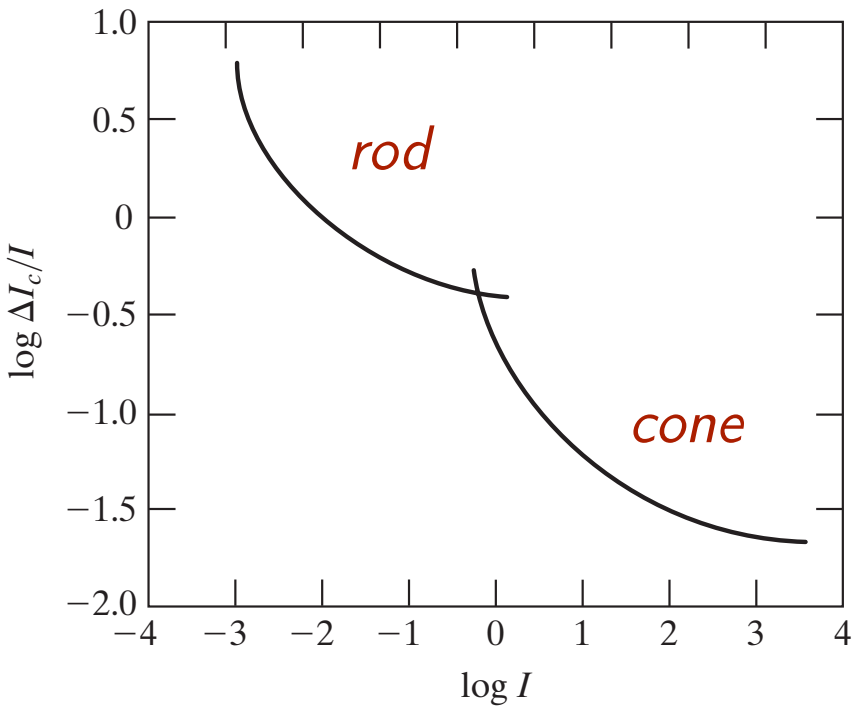
$\Rightarrow \nabla I_c/I$ is small

Bad brightness discrimination

$\Rightarrow \nabla I_c/I$ is large

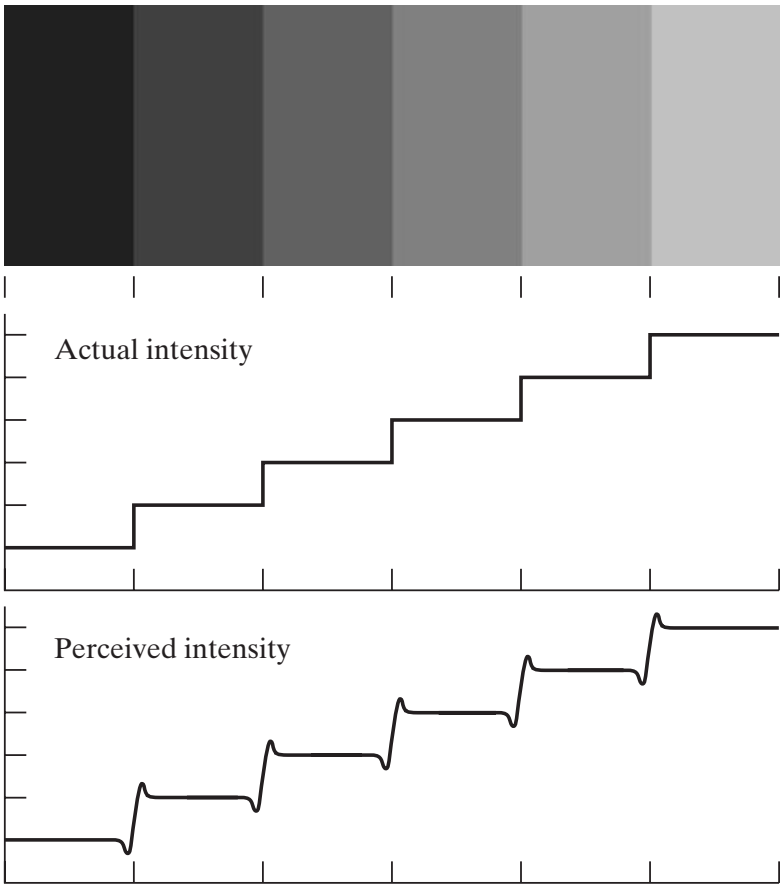
- where I is uniformly illumination on a flat area and ∇I_c is the change in the object brightness required to just distinguish the object from the background
- The ability of the eye to discrimination b/w changes in brightness at any specific adaptation level is of considerable interest.

Weber ratio



- brightness discrimination is poor (the Weber ratio is large) at low levels of illumination and improves significantly (the ratio decreases) as background illumination increases.
- hard to distinguish the discrimination when it is bright area but easier when the discrimination is on a dark area.

Brightness vs. Function of intensity

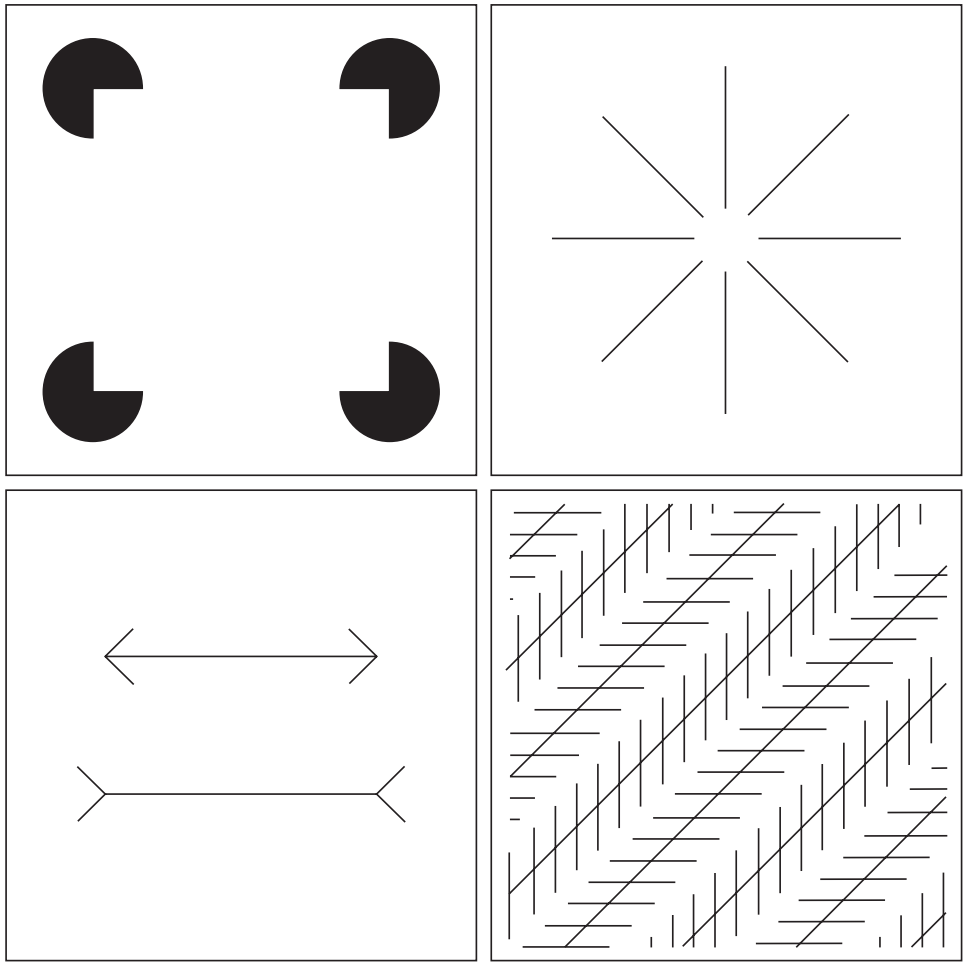


- Brightness is not a simple function of intensity.
- visual system tends to undershoot or overshoot around the boundary of regions of different intensities.
- the intensity of the stripes is constant but we actually perceive a brightness pattern is strongly scalloped near the boundaries.

Simultaneous contrast

- All the small squares have exactly the same intensity, but they appear to the eye progressively darker as the background becomes brighter.
- Region's perceived brightness does not depend simply on its intensity.

Human Perception Phenomena



Signals

- A signal is a function that carries information.
- Usually content of the signal changes over some set of *spatiotemporal* dimensions.

  Vocabulary: existing in both space and time; having both spatial extension and temporal duration.
- Some signals vary over time only called *time-varying signals*.

$$f(t)$$

for example: audio signal, speech signal, acoustic signal

Spatially-varying signals

- Signals can vary over space as well.
- An image can be thought of as being a function of 2 spatial dimensions:

$$f(x, y)$$

- For monochromatic images, the value of the function is the amount of light at that point.
- Computerized Axial Tomography (CAT) and Magnetic Resonance Imaging (MRI) scanners produce images that are functions of 3 spatial dimensions:

$$f(x, y, z)$$

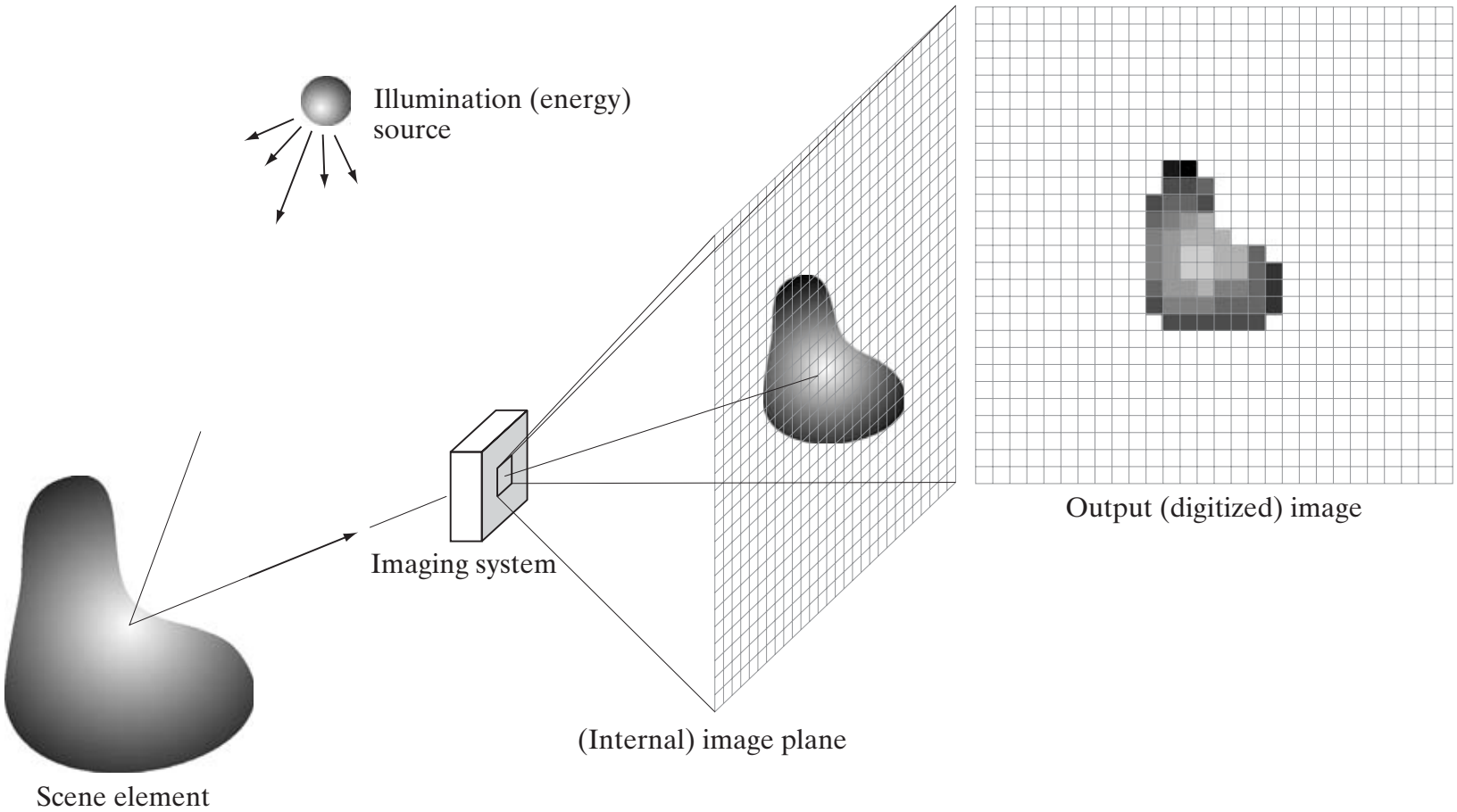
Spatiotemporal Signals

- What do you think a signal of this form is?

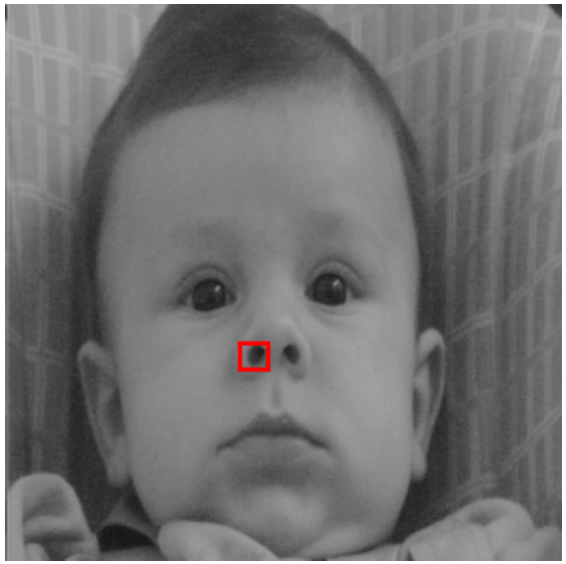
$$f(x, y, t)$$

- x and y are spatial dimensions; t is time.
- Perhaps, it is a video signal animation or other time-varying picture sequence.

Image Formation Model



Digital Image Representation



Pixel values in highlighted region

99	71	61	51	49	40	35	53	86	99
93	74	53	56	48	46	48	72	85	102
101	69	57	53	54	52	64	82	88	101
107	82	64	63	59	60	81	90	93	100
114	93	76	69	72	85	94	99	95	99
117	108	94	92	97	101	100	108	105	99
116	114	109	106	105	108	108	102	107	110
115	113	109	114	111	111	113	108	111	115
110	113	111	109	106	108	110	115	120	122
103	107	106	108	109	114	120	124	124	132

Camera → Digitizer → A set of number in 2D grid

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

Digital Image Representation

- Image is a two-dimensional signal denoted as $f(x, y)$.
- The amplitude value of f at spatial coordinates (x, y) is a positive scalar quantity whose physical meaning is determined by the source of light.
- The function $f(x, y)$ may be characterized by
  - the amount of source illumination incident on the scene being viewed
  - the amount of illumination reflected by the objects in the scene

Illumination and Reflectance

- $f(x, y)$ can be defined as

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

- Illumination can be determined by the nature of the light source.

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

- Reflectance is to be determined by the nature of the objects in a scene $\Rightarrow$ bounded from total absorption to total reflectance.

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

Pixels in image

- A pixel is the smallest element in an image.
- A digital image is a grid of pixels.
- Each pixel corresponds to any one value called pixel intensity.
- The intensity of an image varies with the location of a pixel.
- If I be an image and (x, y) is the location (or coordinate) of any pixel then the image is represented as a function of location : $I(x, y)$, where x and y are integers. Thus an image $I(x, y)$ is a matrix of pixels.

Gray levels

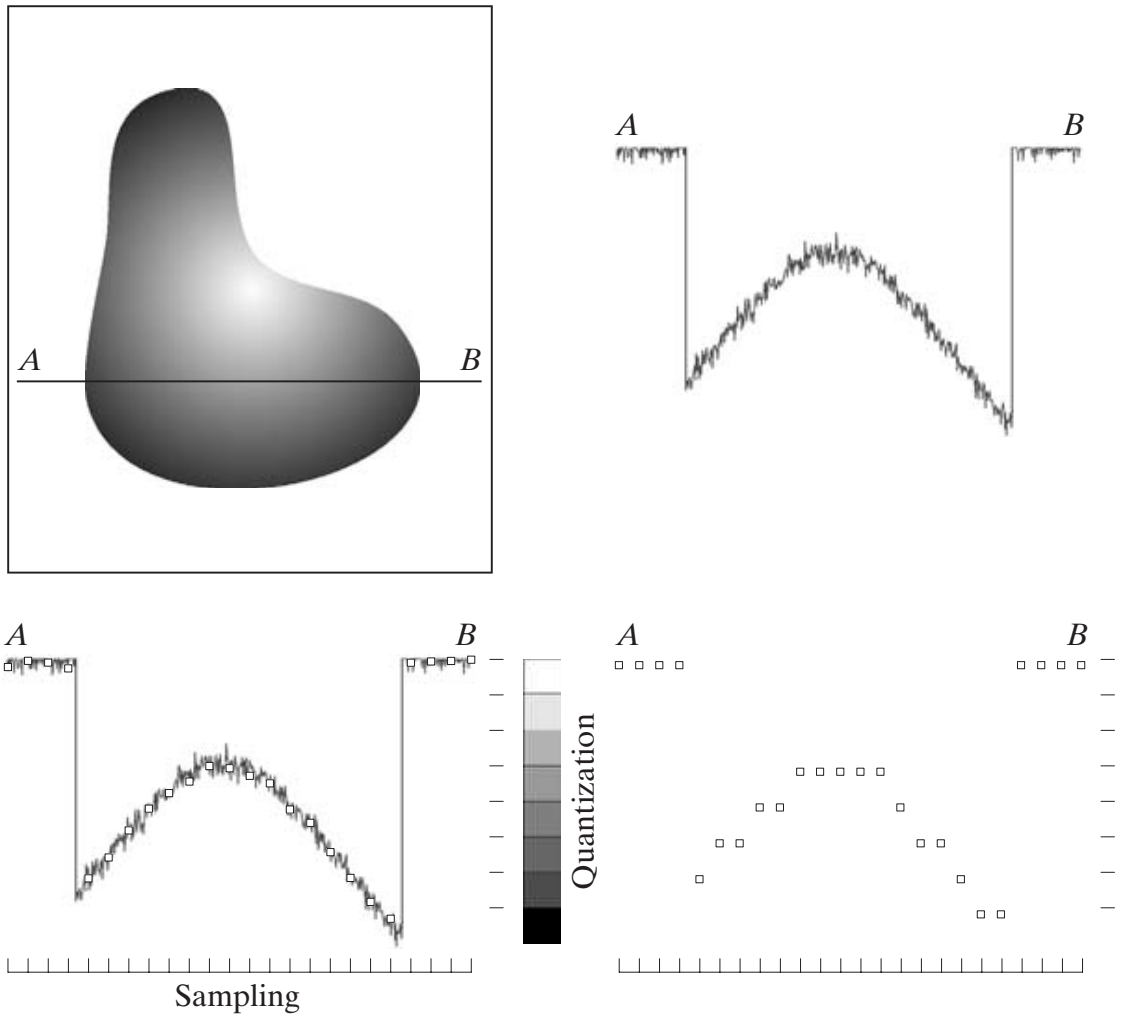
- Let the intensity of monochrome image at any coordinate (x_0, y_0) is denoted as

$$\ell = f(x_0, y_0)$$

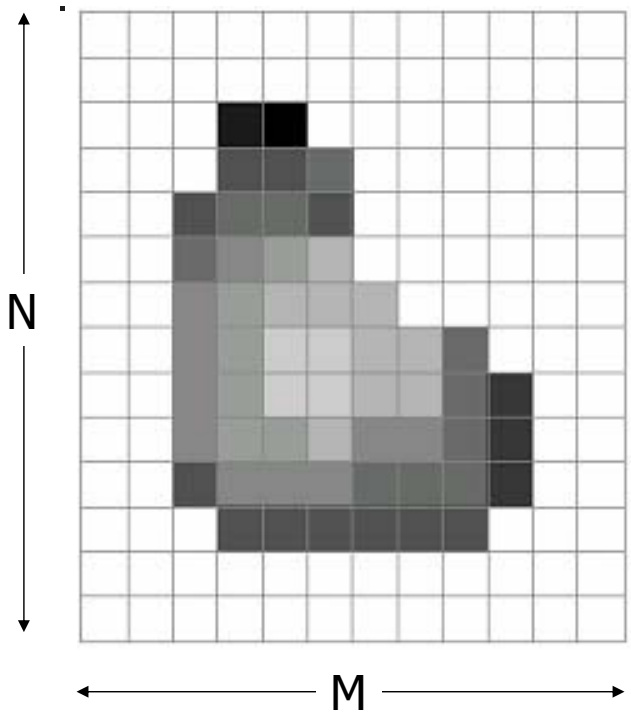
  where $L_{min} \leq \ell \leq L_{max}$

- The interval $[L_{min}, L_{max}]$ is called grayscale.
- Commonly, interval $[0, L-1]$ is used in practice. Where, $\ell = 0$ is considered black and $\ell = L-1$ is considered white on the gray scale.
- All intermediate values are shades of gray varying from black to white.

Image Sampling and Quantization



Number of bits



- The number of gray levels typically is an integer power of 2

$$L = 2^k$$

- Number of bits required to store a digitized image

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

Exercise Problems

Question 01:

(a) what is the range of pixel values for 16-bit images.

(b) what is the maximum intensity value in an 4-bit image.

(b) What would be the required memory to save a 256×256 8-bit grayscale image.

(c) What would be the required memory to save a 512×512 24-bit color image.

Types of Images

- 8 bit, 16 bit, 24 bit images, etc.
- Grayscale images, Color images, (RGB, HSV, etc.), Infrared images.

- Image format: png, bmp, jpg, jpeg, pgm, tiff, etc.
- Low contrast, high contrast, dark, light Images, etc.

Spatial Resolution

- *Spatial resolution* is a measure of the smallest identifiable detail in an image.
- It is related to the sampling interval.
- Higher is the spatial resolution, greater is the sampling rate i.e., lower is the image area represented by each sampled point.
- *Unit* PPI (pixel per inch) or DPI (dot per inch)
- News paper - 75 dpi, Magazine - 133 dpi, Glossy Paper - 175 dpi, Book pages printed at 2400 dpi

Intensity Resolution

- It is also known as *Gray-level Resolution*.
- It is related to the quantization interval.
- Higher is the graylevel interval/resolution, more is the quantization level i.e., lower is the quantization error.


Thank You
Queries?