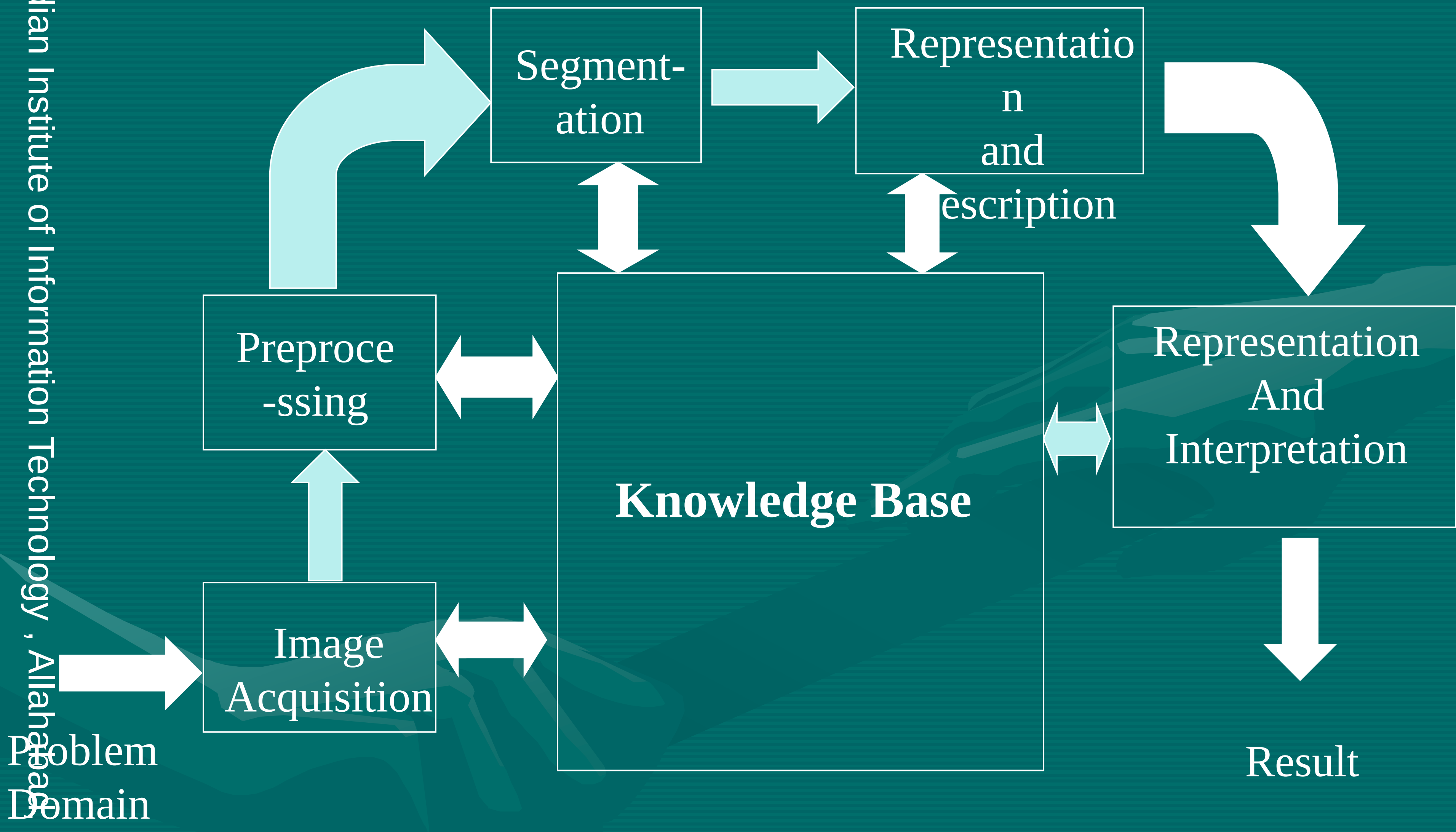
Indian Institute of Information Technology, Allahabad, India

Introduction to Signal and Image Processing

Steps in Image Processing



Steps in Image Processing

- **Image Acquisition**
  - Acquire digital images
  - Requires an image sensor and the capability to digitize the signal produced by the sensor.
- **Preprocessing**
  - Improve the image in ways that increase the chances of success of other processes. Examples:
    - Enhancing contrast
    - Removing noise
- **Segmentation**
  - Partition an input image into its constituent parts or objects.
  - **(Autonomous segmentation - Most difficult!!)**

Steps in Image Processing

- **Representation and Description**
  - Segmented image parts needs to be represented as either
    - Boundaries - This is appropriate when the focus is on external shape characteristics such as corners and inflections
    - Regions - This is appropriate when the focus is on internal properties such as texture
  - Once the appropriate representation scheme is chosen, then a method must be used to describe the data.
  - **Description (feature selection)** deals with extracting features that results in some quantitative information of interest or features that are basic for differentiating one class of objects from another.

Steps in Image Processing

- **Recognition and Interpretation**
  - Recognition - Process that assigns a label to an object based ion the information provided by its descriptors.
  - Interpretation - Assigning meaning to an ensemble of recognized objects.

Knowledge Base : Knowledge about a problem is coded into image processing systems in the form of knowledge database

Elements of Digital Image Processing Systems

- **Image Acquisition**
  - - ***Physical device t***hat is sensitive to a band in the electro-magnetic energy spectrum and that produces an electrical signal output proportional to the level of energy sensed.
  - - A ***Digitizer*** to convert the electrical output of the physical device into digital form.
- **Storage** - The amount of storage required for an image depends on the resolution of the sensing device as well as the number of bits required to code each pixel. Storage falls into 3 categories
  - 1. Short term storage for use during pre-processing
  - 2. on-line storage for relatively fast-recall.
  - 3. archival storage, characterized by infrequent access

Elements ….

- **Methods of storage**
  - - Computer Memory
  - - Frame Buffers
  - - Magnetic disks, Magneto-Optical Disks (online storage)
  - - Magnetic Tapes and Optical Disks - Archival Storage
- **Processing** - involves procedures that are usually expressed in algorithmic form.
  - - Usually done in software
  - - Use of hardware for real time applications

Elements ...

- **Communication**
  - Involves local communication (handled by the LAN communication protocols) and remote communication.
  - - Remote Communication - presents lots of challenges as images consume lot of bandwidth
- **Display** - This is one of the most important elements of the entire system as it enables the end user to see the results at each step of image processing.

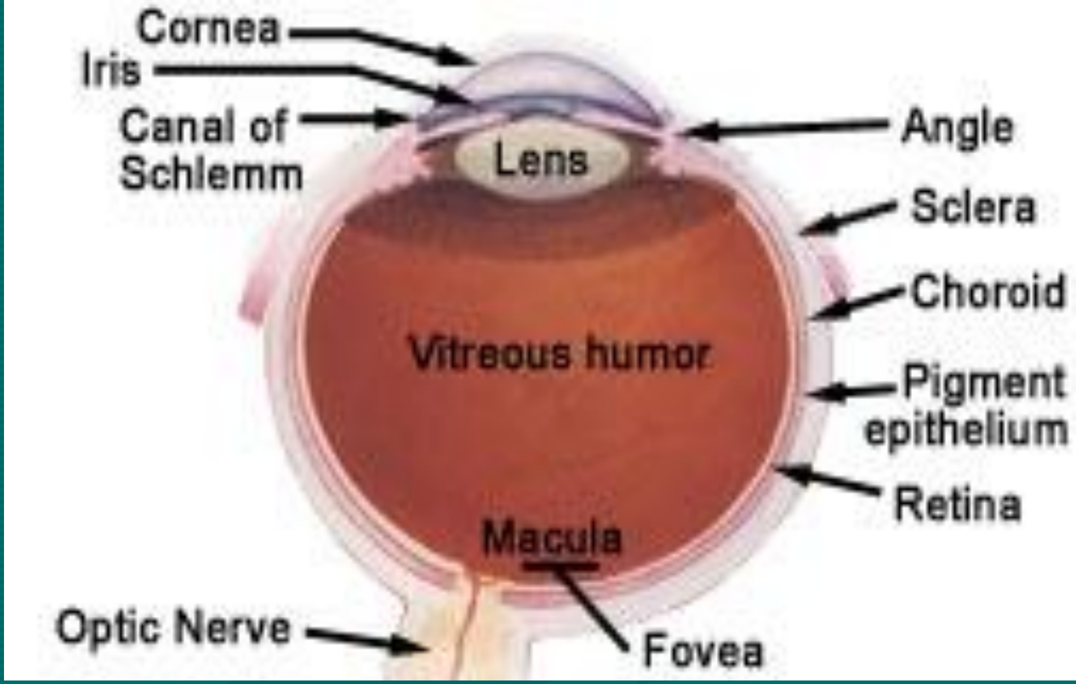
Image Processing is characterized by specific solutions

Techniques that work well in one area can be totally inadequate in another.

Digital Image Fundamentals

1. Elements of Visual Perception

Structure of the Human Eye



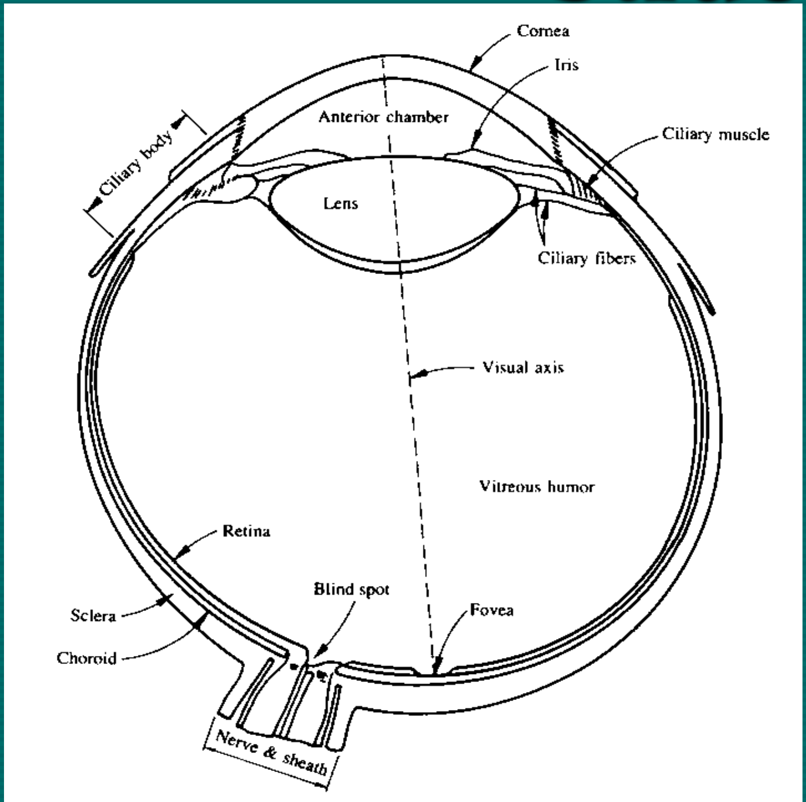
- Eye: Sphere, diameter of 20 mm
- Consists of 3 membranes:
  1. Cornea and sclera
  2. Choroid
  3. Retina
- Cornea: transparent
- Sclera: opaque, connected to cornea
- Choroid: network of blood vessels

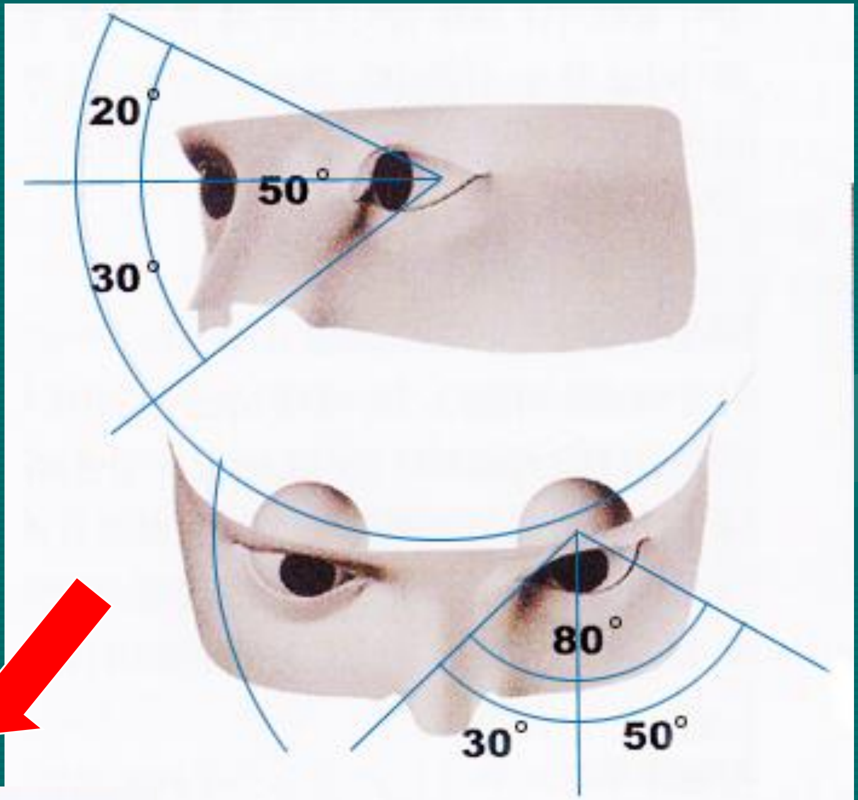
In front choroid is connected to iris diaphragm

- Iris: contracts or expands to control amount of light
- Pupil: central opening of iris, 2 to 8 mm in diameter

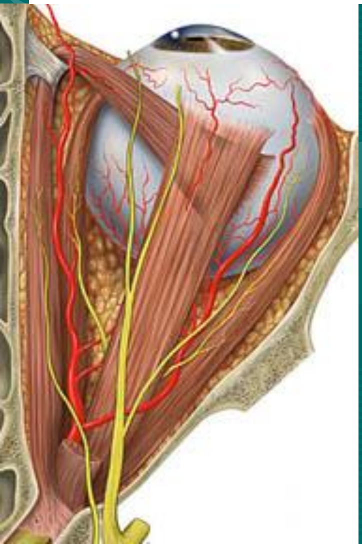
Indian Institute of Information Technology , Allahabad, India

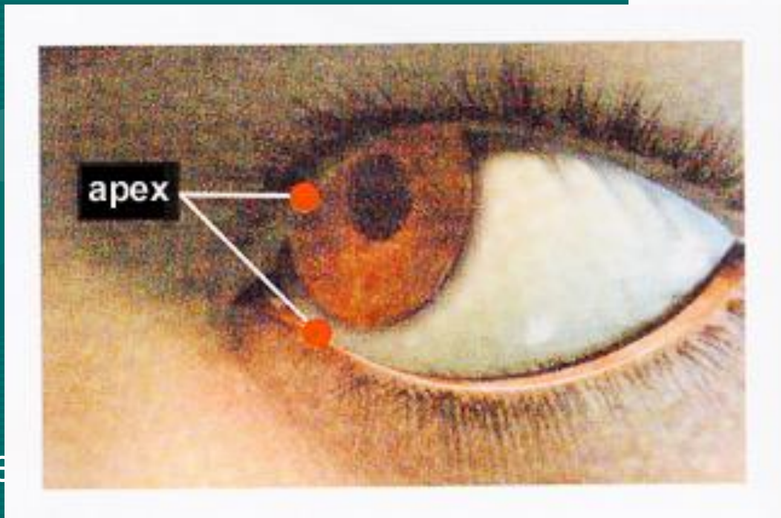
Structure of Human Eye ..





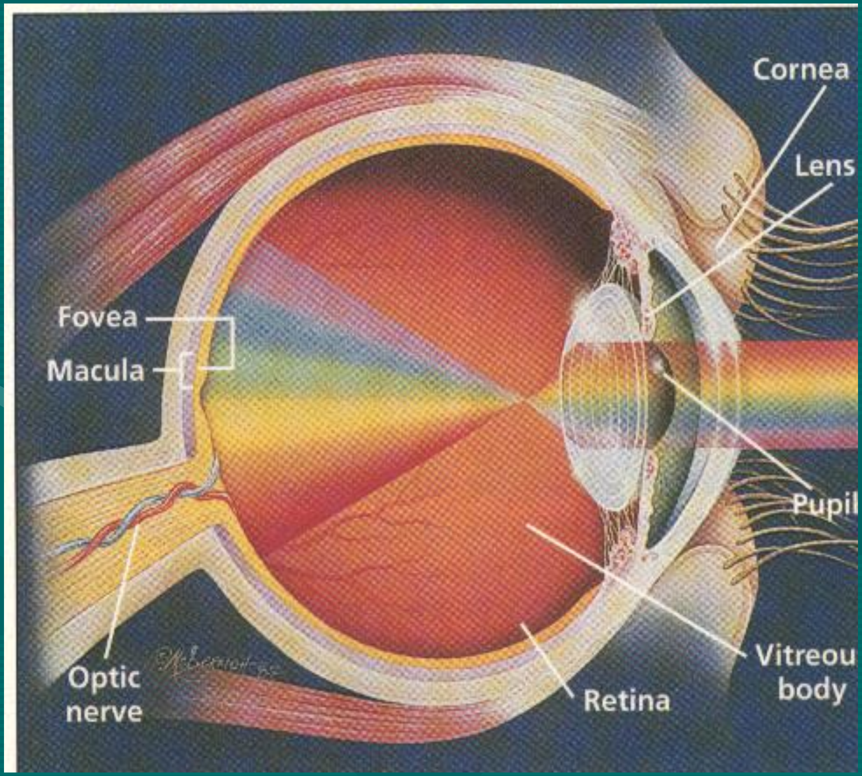
Movable range of an eye

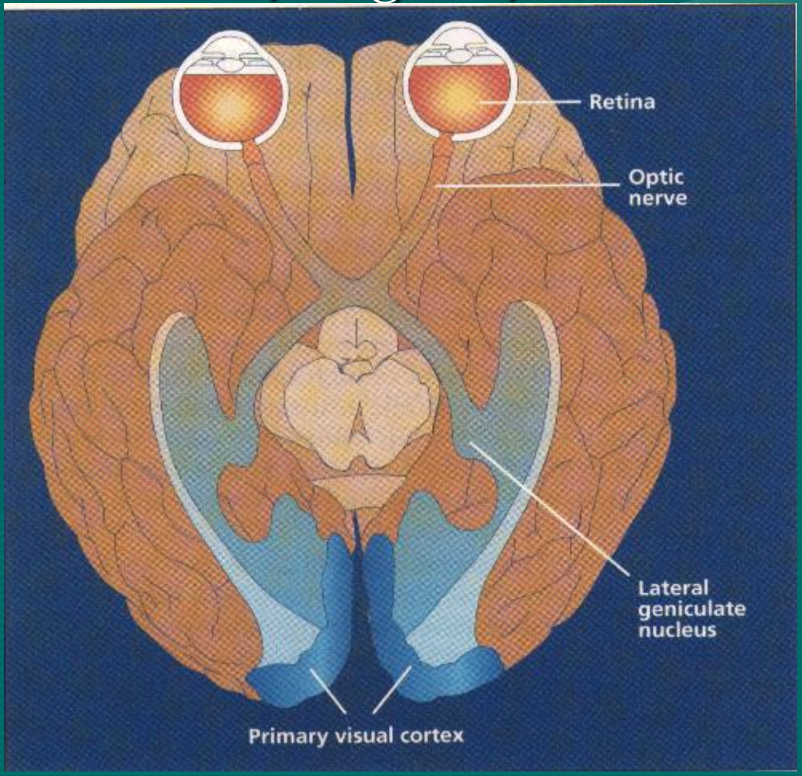




The Eye - What can we get from it ?

- **Human Eye**
- **Definition of terms** - Smooth Pursuit, Vergence,





The Eye

- **Visual Front End Processor**
  - Simplifies the complexity of input image to a great extent before passing it to upper layer (i.e. to Brain)

Structure of the Human Eye

- **Lens:**

– focuses light on retina

– Contains 60% to 70% water

– Absorbs 8% of visible light

– High absorption in infrared and ultraviolet (can cause damage to eye)

- Retina: the inner most layer, covers the posteriori portion of eye
- When eye is properly focused, light of an object is imaged on the retina
- Light receptors are distributed over the surface of retina

Structure of Human Eye

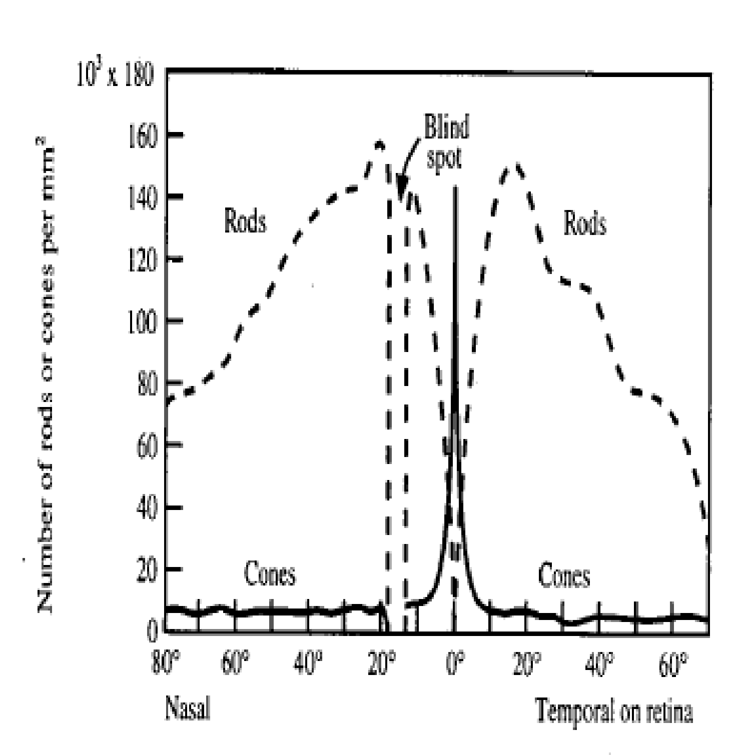
- **Retina contains light receptors: Cones & rods**

– **Cones**:

- 6 to 7 million,
- located mainly in central part of retina fovea)
- Sensitive to color,
- Can resolve fine details because each one is connected to its nerve
- Cone vision: photopic or bright-light

– Rods:

- 75 to 150 million,
- No color vision, responsible for lowlight vision,
- Distributed a wide region on the retina
- Rod vision: scotopic or dim-light



Structure of Human Eye

- Blind spot: a region of retina without receptors, optic nerves go through this part
- Fovea: a circular area of about 1.5 mm in diameter
- A comparison between eye (fovea) and a CCD camera:

– Density of cones in fovea: 150,000 /mm2

– Number of cones: 337,000

– A medium resolution CCD chip has the same number of elements in a 7mm x 7mm area.

Image Formation in the Eye

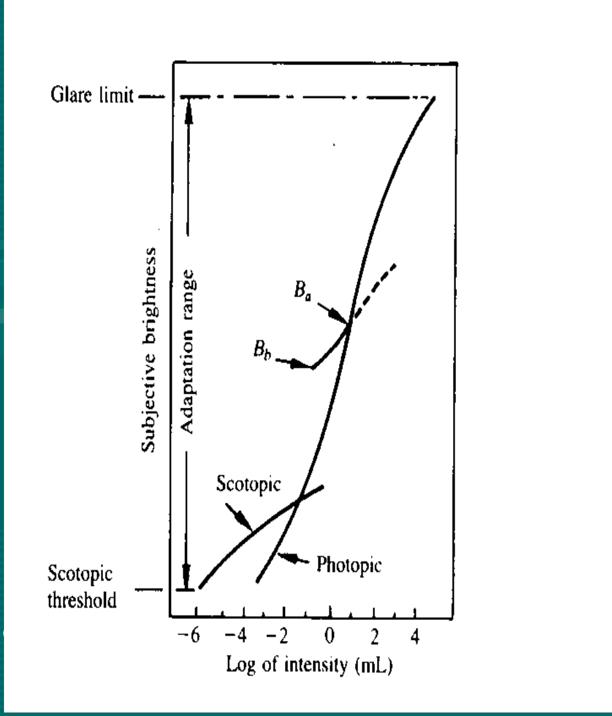
- ***Lens is flexible***
- Refraction of lens is controlled by its thickness
- Thickness is controlled by the tension of muscles connected to the lens
- ***Focus on distance objects***: lens is relatively flattened, ***refractive power is minimum***
- ***Focus on near objects***: lens is thicker, ***refractive power is maximum***

Brightness Adaptation and Discrimination

- The dynamic range of light intensity to which eye can adapt is enormous – (on the order of 10^{10})- from the ***scotopic threshold*** to the glare limit

- HVS can not operate over the entire range (see figure) simultaneously.

It accomplishes large variations due to ***brightness adaptation***



- Ability of the eye to ***discriminate between changes in brightness*** at any specific adaptation level.

- Brightness (light perceived by visual system) is a logarithmic function of light intensity.

Weber Experiments – ***brightness discrimination of HVS***

Image Model

Image - two dimensional light intensity function, denoted by *f(x,y)* , where the value or amplitude of f at spatial coordinates *(x,y)* gives the intensity (brightness) of the image at that point.

Perception of an object : Light reflected from that object

f(x,y) can be characterized by two components

1. The amount of source light incident on the scene being viewed

   Illumination component - *i(x,y)*

2. Amount of light reflected by the objects in the scene

   Reflectance component - *r(x,y)*

$$f(x,y) = i(x,y)r(x,y)$$
$$0 \le i(x,y) < \infty$$
$$0 \le r(x,y) \le 1$$

Image Model

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

The nature of ***i(x,y)*** is determined by the light source and ***r(x,y)*** is determined by the characteristics of the objects in a scene.

i(x,y):
Clear Day - 9000 foot-candles
Cloudy day - 1000 foot-candles
Full moon - 0.01 foot candles
commercial office - 100 foot-candles

Typical values of ***r(x,y)***
- Black Velvet - 0.01
- Stainless Steel - 0.65
- Flat-White wall paint - 0.80
- Silver plated metal - 0.90
- Snow - 0.93

Grey Level (l) at point (x,y)
- Intensity of monochrome image f at coordinates ***(x,y)***

Sampling & Quantization

For computer processing, an image function needs to be digitized both ***spatially and in amplitude***.

Image Sampling – Digitization of the spatial coordinates ***(x,y)***

Gray-level quantization – Digitization of Amplitude

Continuous image f(x,y) – Equally spaced samples arranged as ***(N x M)*** array (matrix)

Each element in the array is a discrete quantity.

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

Digital Image

Image Element (pixel)

Sampling and Quantization …

Need to decide the values for **N, M** and number of discrete gray levels allowed for each pixel **(G).**

In Digital Image Processing:

$$N = 2^n$$
$$M = 2^k$$
$$G = 2^m$$

G – No. of gray levels and it is assumed that these levels are equally spaced between ***0 and L*** in the gray scale.

Amount of storage required to store a digitized image

$$b = N \times M \times m \quad \text{(bytes)}$$

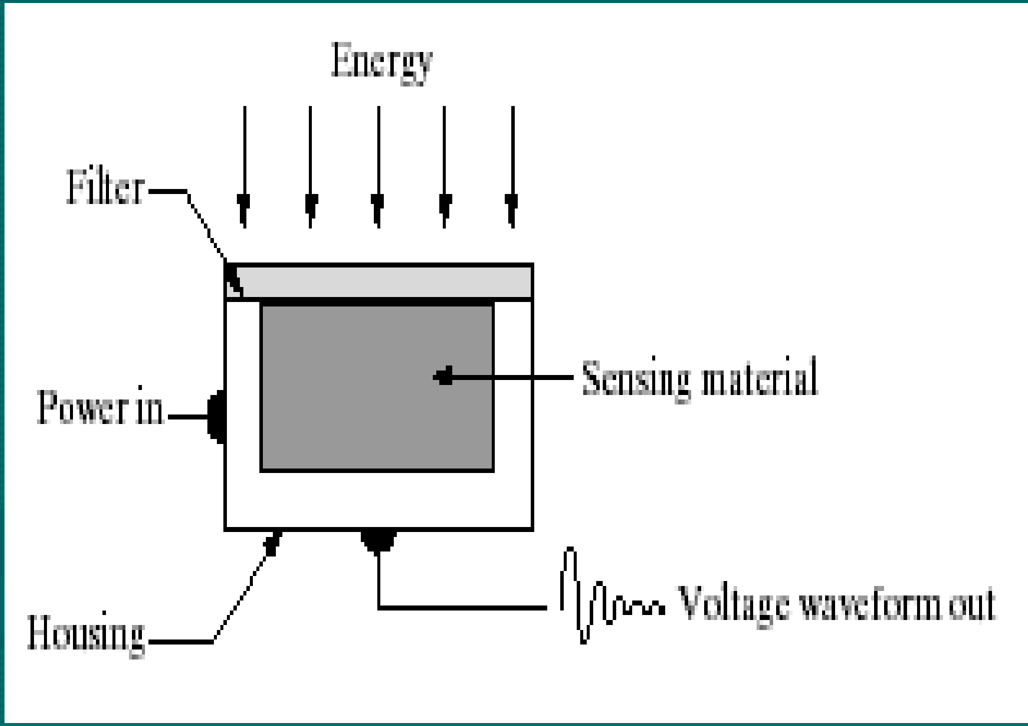
If $M=N$, $b = N^2 m$

Nonuniform Sampling and Quantization

- Adaptive scheme of sampling can be used to improve the appearance of an image for a fixed spatial resolution.
- Here sampling process depends on the characteristics of the image.
  - Fine sampling in the neighbourhood of sharp gray-level transitions
  - Coarse sampling in relatively smooth regions
- E.g. A face superimposed in a uniform background.
- Nonuniform quantization: Considering the relative frequency of gray levels
  - Frequent gray levels (within a range) – finely spaced
  - Gray levels outside this range – coarsely spaced.

Sensors

Sensors are used to transform illumination energy into digital images. Sensors are three types:

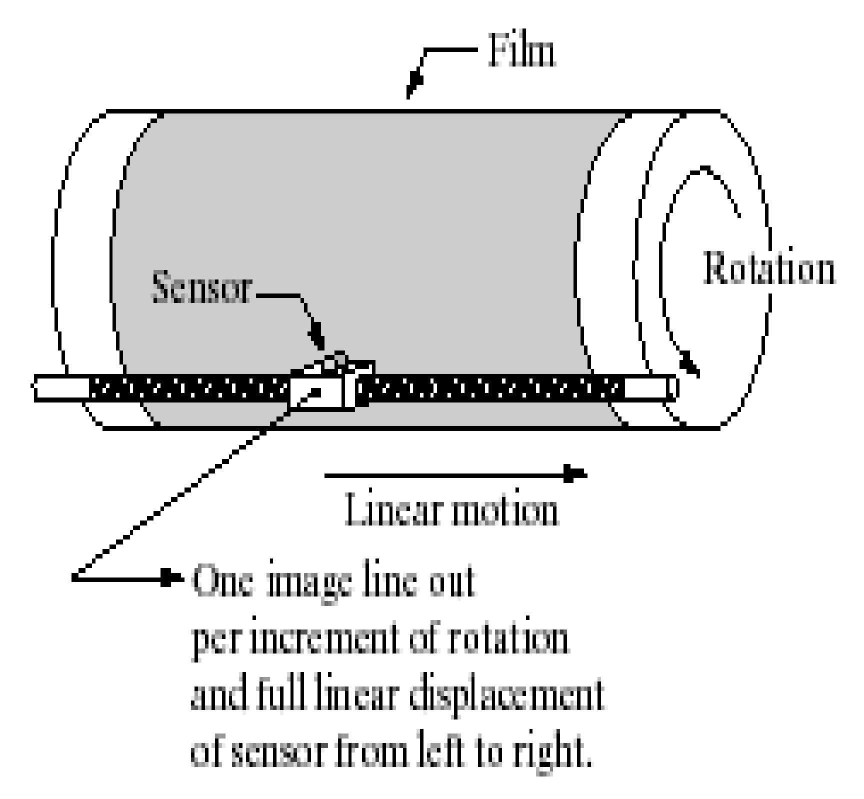


- Single Imaging Sensor
- Line sensor
- Array Sensor

Single Sensors

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.

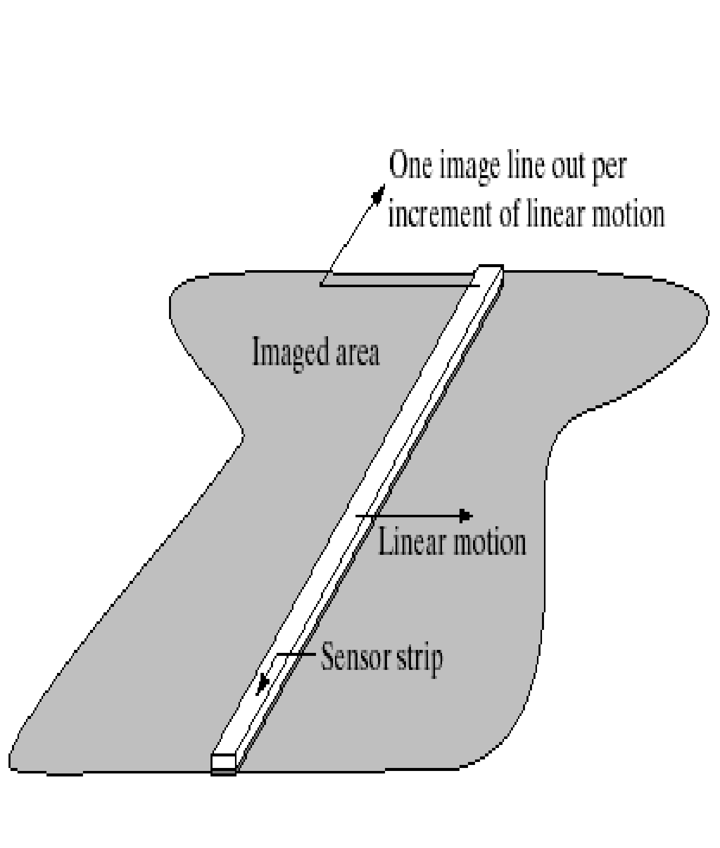
A film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Since mechanical motion can be controlled with high precision, this method is an inexpensive (but slow) way to obtain high-resolution images.



Linear Sensor

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. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction.

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. Lenses or other focusing schemes are used to project the area to be scanned onto the sensors.

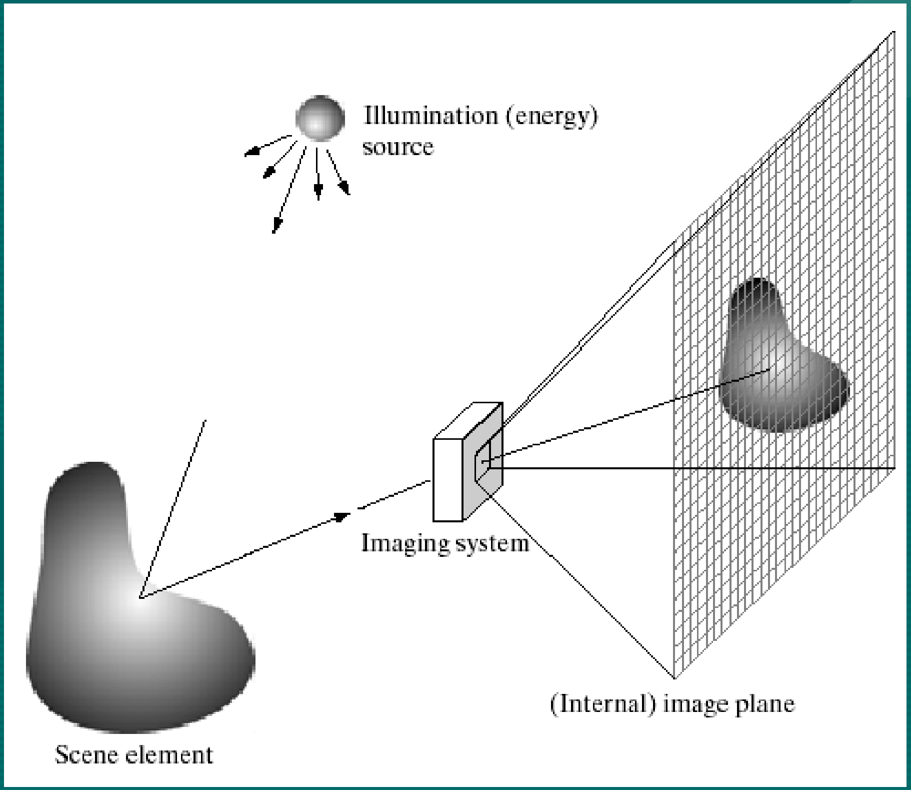


Array Sensor

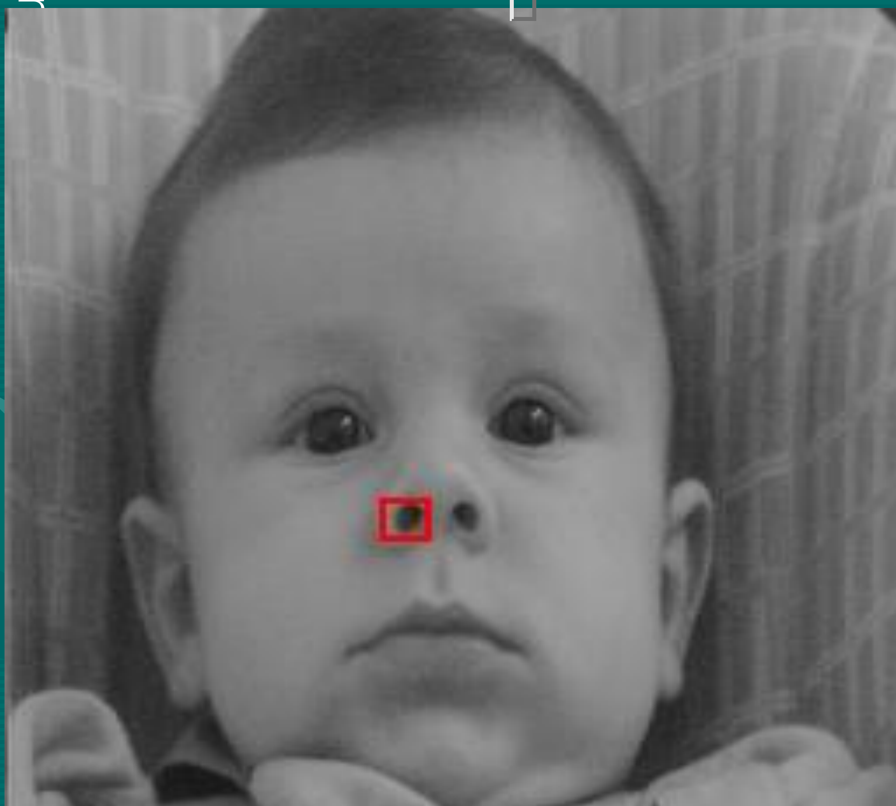
Numerous electromagnetic and some ultrasonic sensing devices frequently are arranged in an array format. This arrangement is also found in digital cameras.

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 illumination source reflects the energy from a scene element and 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.



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

Number of storage bits for various values of N and k

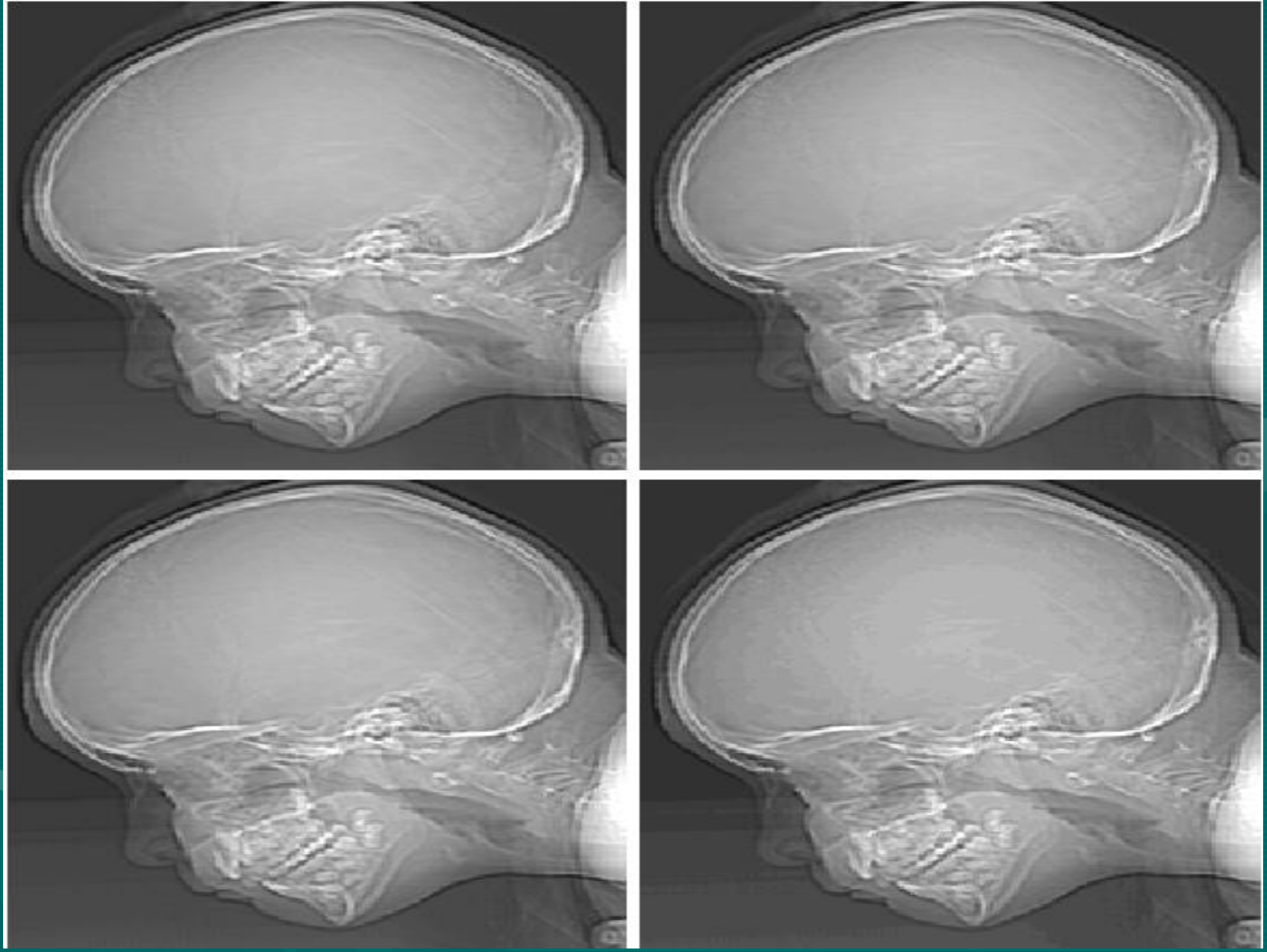
N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Resolution

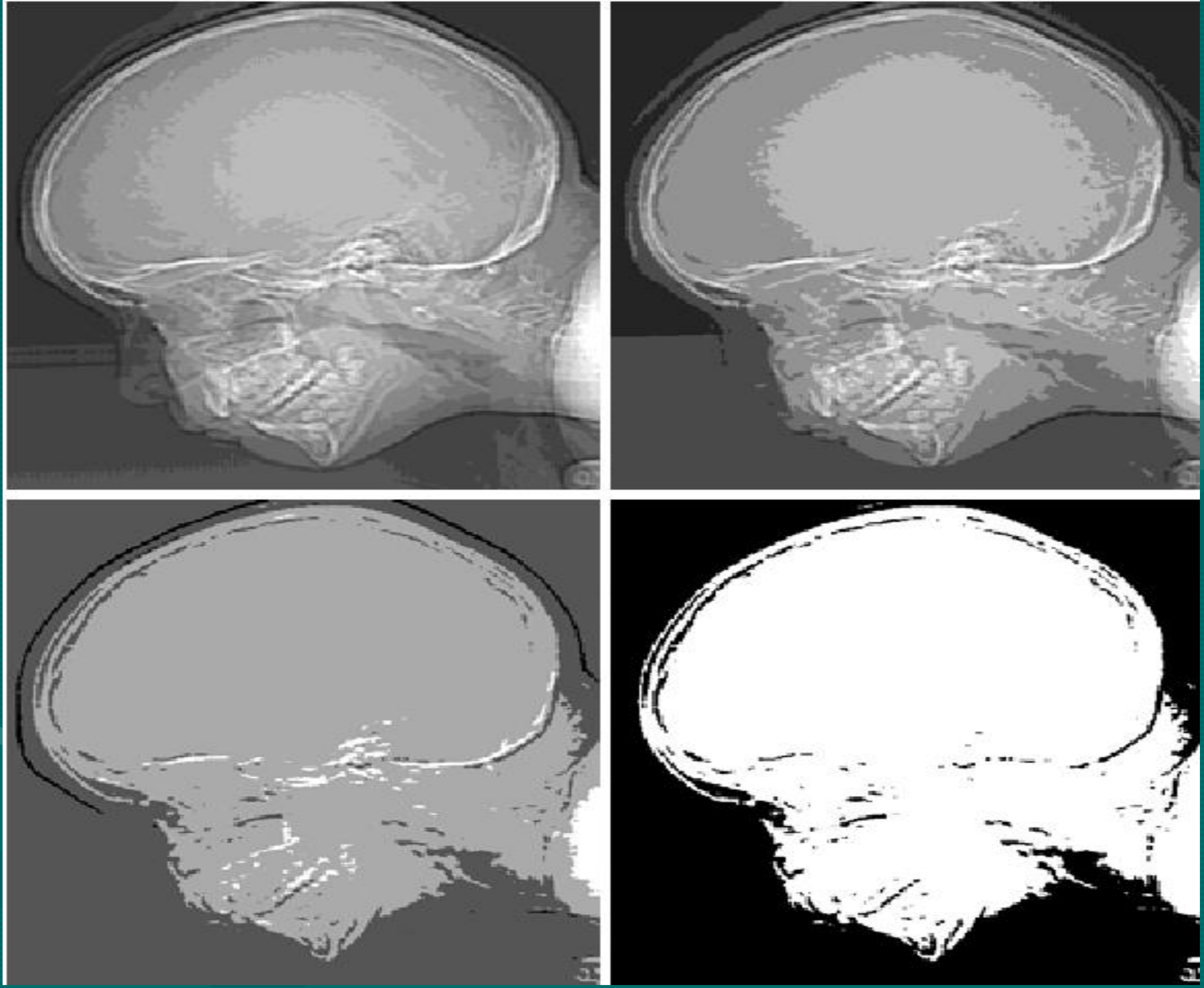
- **Spatial Resolution**: Spatial resolution is the smallest detectable detail in an image.
- **Grey level Resolution:** *Gray-level resolution* similarly refers to the smallest detectable change in gray level.

A 1024*1024, 8-bit image subsampled down to size 32*32 pixels.The number of allowable gray levels was kept at 256

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



(a) 452*374, 256-level image (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.



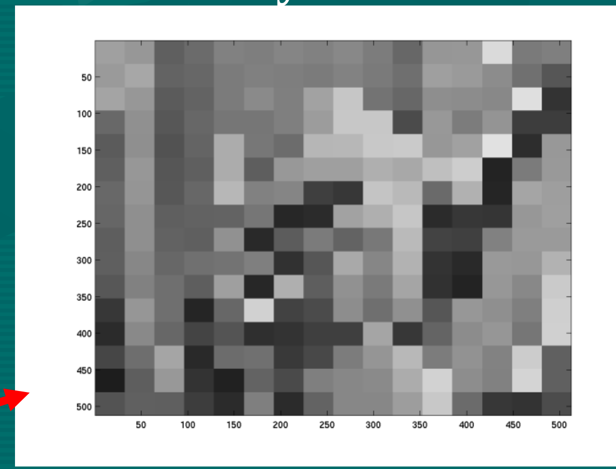
Effects of Reducing Spatial Resolution

Original

Reduced by 2 in each direction

Reduced by 8 in each direction

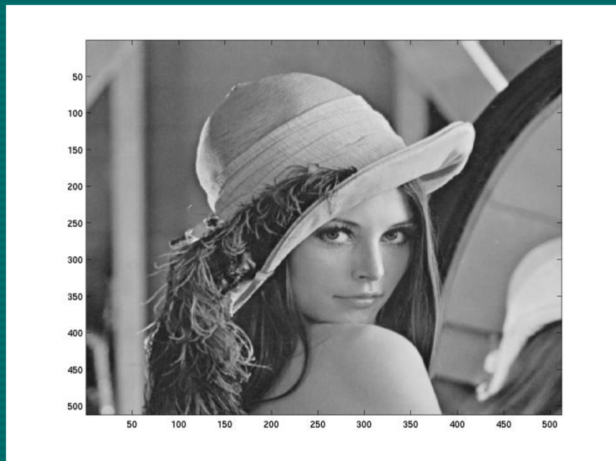
Reduced by 32 in each direction



Checkerboard Effect

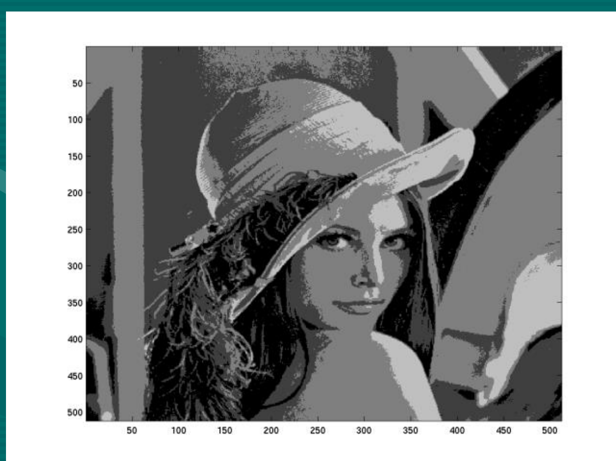
Effects of Reducing Gray Levels

Original (256 levels)



64 levels

4 levels



2 levels

Effects of Reducing Gray Levels

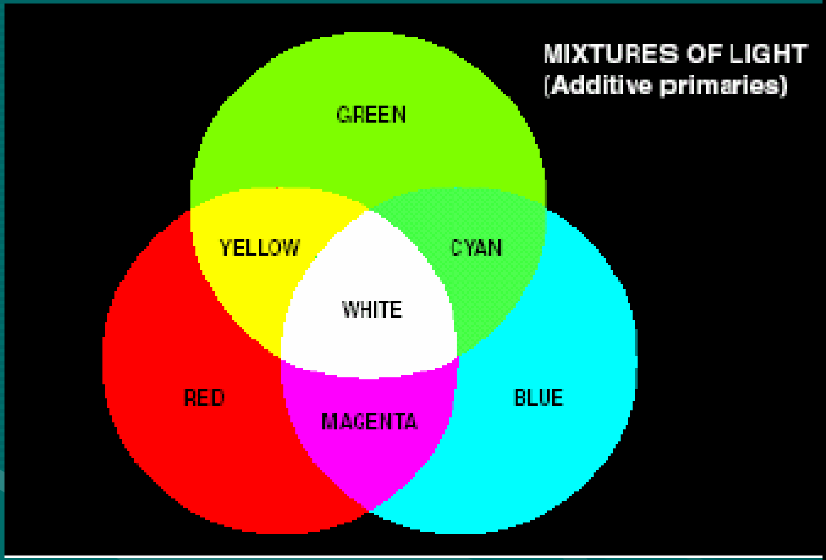
When the number of gray level values are reduced, very fine ridge like structures develop in the areas of gray levels.

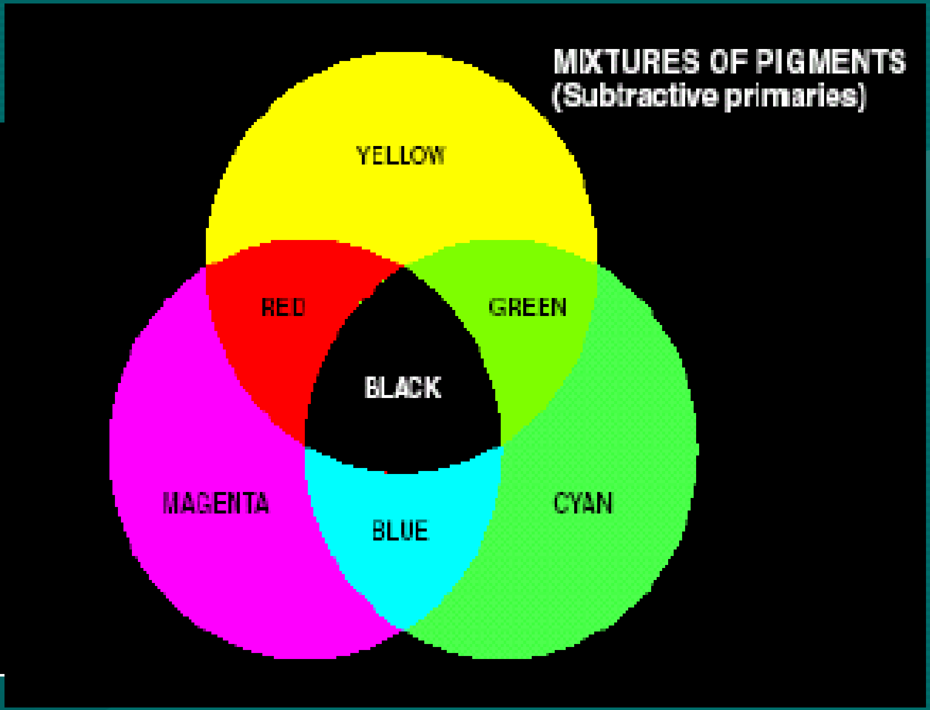
This effect is known as ***false contouring*** and is caused by the insufficient number of gray levels in smooth areas of the image.

Color Models

- **RGB (red, green, blue)**
- **CMY (cyan, magenta, yellow)**
- **YIQ (luminance, inphase, quadrature)**
- **HSI (hue, saturation, intensity)**
- **HSV (hue, saturation, value)**

Additive and Subtractive colors





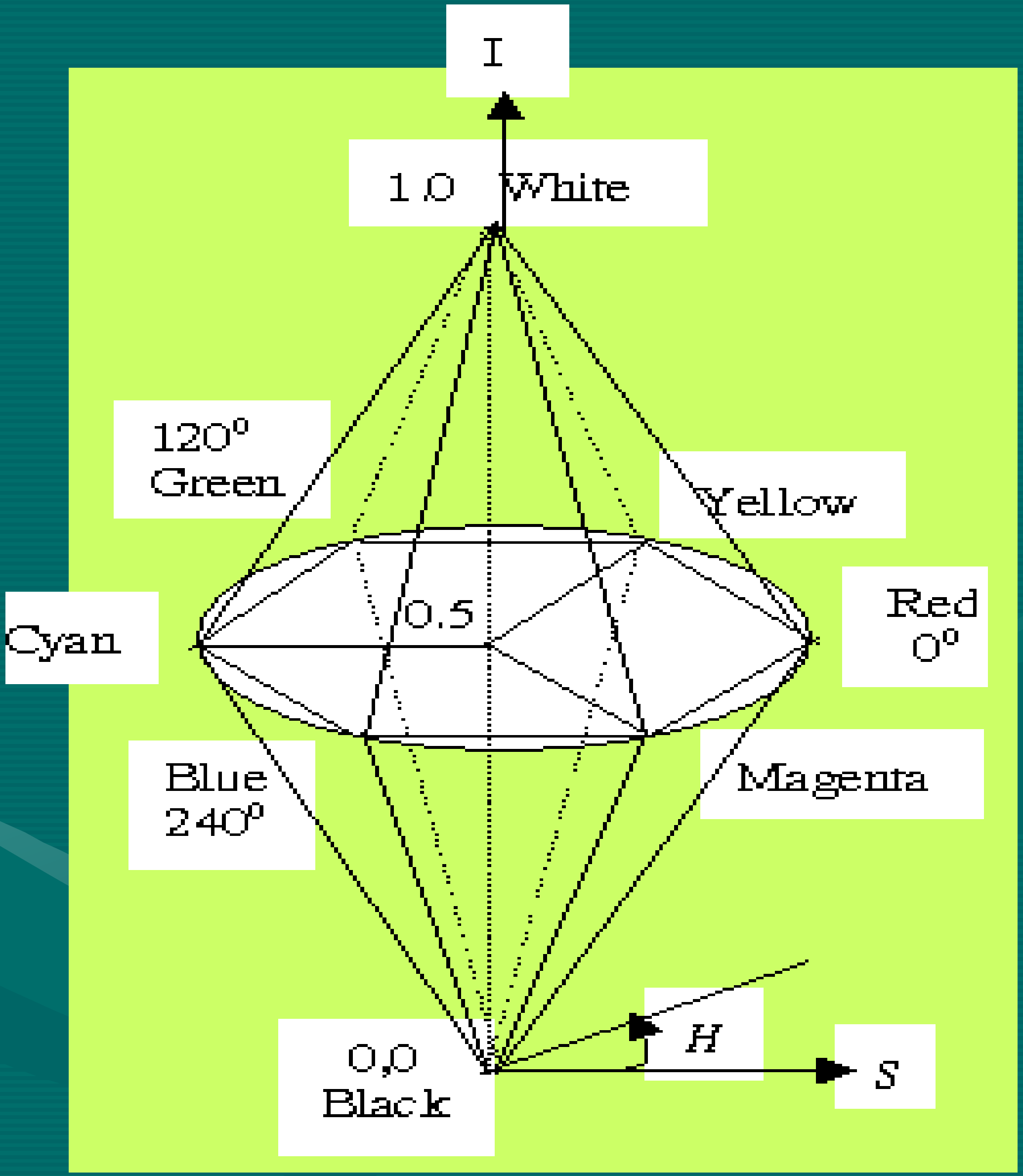
YIQ Model

- Used by US commercial color television broadcasting (used by NTSC standard)
- Y: encodes luminance

  I, Q: encode color (chromaticity)
- For black and white TV, only the Y channel is used.
- People are more sensitive to the luminance difference, so We can use more bits (bandwidth) to encode Y and less bits to encode I and Q

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.3 & 0.59 & 0.11 \\ 0.6 & -0.28 & -0.32 \\ 0.21 & -0.52 & 0.31 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

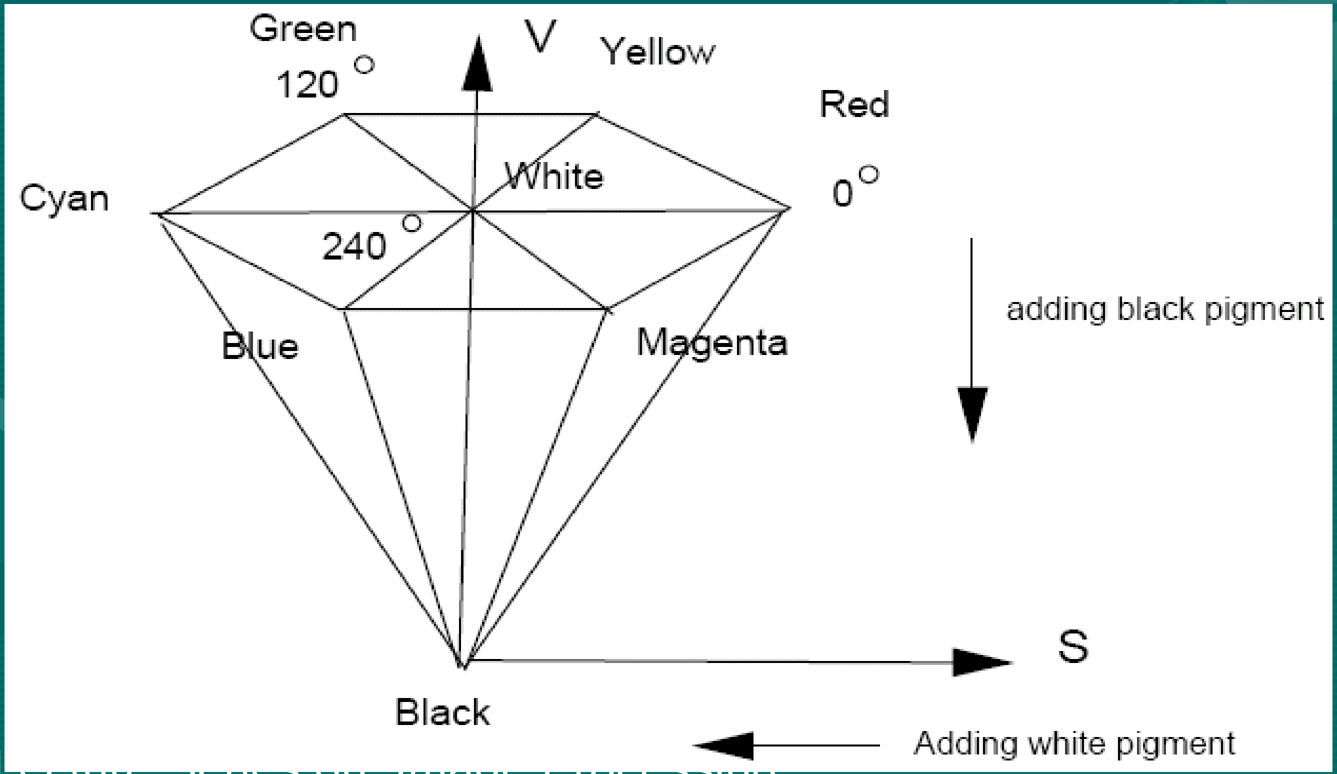
HSI Model



I
1.0 White
120°
Green
Yellow
Cyan
0.5
Red
0°
Blue
240°
Magenta
0,0
Black
H
S

HSV Model

- User oriented color model.
- *H (Hue)*: 0−360 , representing different colors
- *S (Saturation)*: 0−1 (adding white pigment = reduce S)
- *V (Value)*: 0−1 (adding black pigment = reduce V)
- S = 1, V = 1 corresponds to an artist's pure pigment

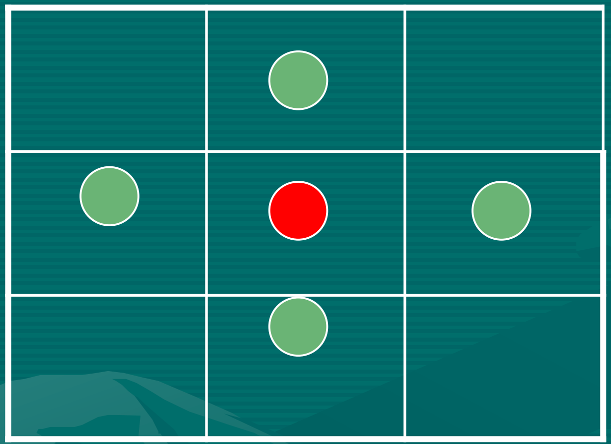


Relationship between pixels

Neighbours of a Pixel:

A pixel ***p*** at coordinates ***(x,y)*** has ***four(4) horizontal and vertical*** neighbours. The coordinates of these neighbours are given by

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



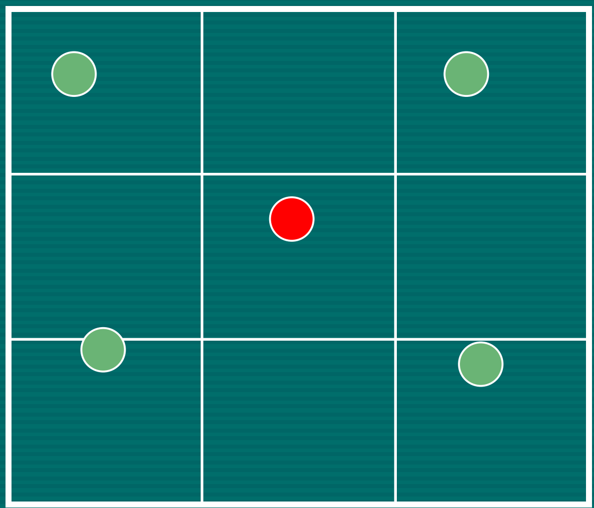
The above set of pixels is called the ***4-neighbours of p –*** $N_4(p)$

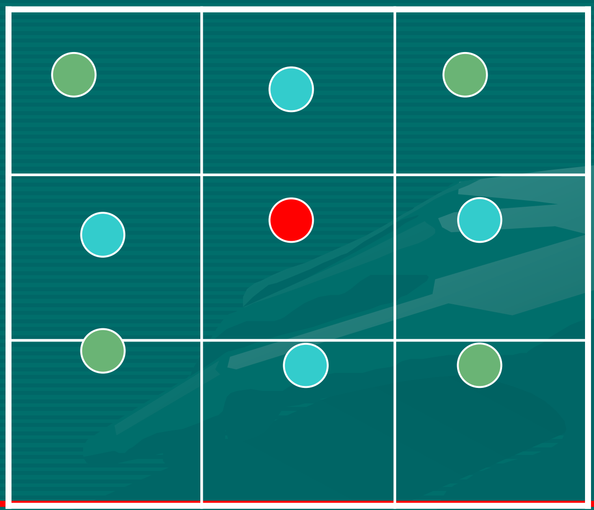
Each pixel is a unit distance from ***(x,y)***

Neighbours of a Pixel

The ***four (4) diagonal neighbours*** of ***p*** have coordinates

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





$N_8(p)$

The above set of pixels is denoted by – $N_D(p)$.

These pixels, together with 4-neighbours are called ***8-neighbours of p*** and is denoted by $N_8(p)$.

$N_8(p) = N_4(p) \cup N_D(p)$.

Connectivity

Connectivity forms the basis for establishing the ***boundaries of an objects*** and also components of regions in an image.

To establish whether two pixels are connected:

1. Whether the ***pixels are adjacent*** (E.g. are they 4-neighbours)
2. Whether their gray levels satisfy ***a specified criterion of similarity*** (E.g. equal or belongs to a set – falls within a given range)

- ***4-connectivity*** – Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$
- ***8 connectivity*** – Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$
- ***m-connectivity*** - Two pixels p and q with values from V are m-adjacent if: 1. q is in $N_4(p)$ or

  2. q is in $N_D(p)$ ***and*** the intersection of ($N_4(p)$ and $N_4(q)$) is empty

Distance Measures

- For pixels p,q, and z with coordinates (x,y),(s,t)and (v,w), respectively, D is a distance functions if:

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

Euclidean Distance between p and q is defined as :

$$D_e(p,q) = [x-s]^2 + (y-t)^2]^{1/2}$$

City-Block Distance (D_4) between p and q is defined as

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

Distance Measures …

Chessboard Distance (D_8) between p and q is defined as

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

D_e – 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)***

D_4 – Pixels having a ***D_4*** distance from ***(x,y)*** less than or equal to some value ***r*** form a diamond centered at ***(x,y)*.**

D_8 – Pixels having a ***D_8*** distance from ***(x,y)*** less than or equal to some value ***r*** form a square centered at ***(x,y)*.**