Multimedia – An Introduction

Lesson 1

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

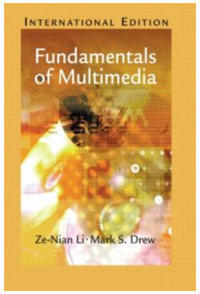
To discuss…

- ✓ Course Issues
- ✓ Multimedia – Definitions
- ✓ Multimedia System
- ✓ Data Stream & continuous media
- ✓ Streaming Media
- ✓ Multimedia - Applications

Course Issues

- **Text Books**

1) Ze-Nian Li & Mark S. Drew, "Fundamentals of Multimedia", Pearson Education, 2004



2) Susanne Weixel & Jennifer Fulton, "Multimedia BASICS", 2nd Edition, 2010

In Addition to reference books, Additional readings & the class notes.

Course Issues

- Coverage

Topics	Number of Lectures
Introduction to Multimedia & Media Basics	2
Digital Images representation and processing, basic relationships between pixels	3
Colors , color science, Human visual system HVS, color models in image	3
Spatial Filtering	3
Video, Types of video signals, Analog video, Digital video	2
Audio, Sound, Types of Audio	3
Data compression, some basic methods	4
TOTAL	20

Course Issues

- Evaluation
  - Test-1 - 13%
  - Test-2 - 13%
  - Test-3 - Optional.
  - Quizzes - 9%
  - Lab / Assignments /Project – - 15%
  - In Class representation - 5% Extra (Optional)
  - Final Exam - 50%

CLASSROOM ETIQUETTE

- At all times be considerate to your classmates and to your instructor.
- Come to class on time, ready to ask questions about previous lessons/assignments.
- Ask pertinent questions; contribute to discussions; avoid "private" conversations that distract the instructor and other students.
- Any student that disrupts the class will lose the lecture and/or be asked to leave the room.
- Remember that the instructor is the one to end the class– do not prepare to leave early.

KEYS TO SUCCESS

- Have a positive attitude about learning and the class.
- Attend all class sessions and be punctual.
- Complete reading assignments and handouts before beginning lab.
- Do your own work. Work with your assigned partner. Ask for help when needed.
- Don't expect to understand every topic the first time it is presented; review often; spend as much time as necessary to master the material.
- Be flexible.
- Enjoy the class!

Multimedia- Definitions

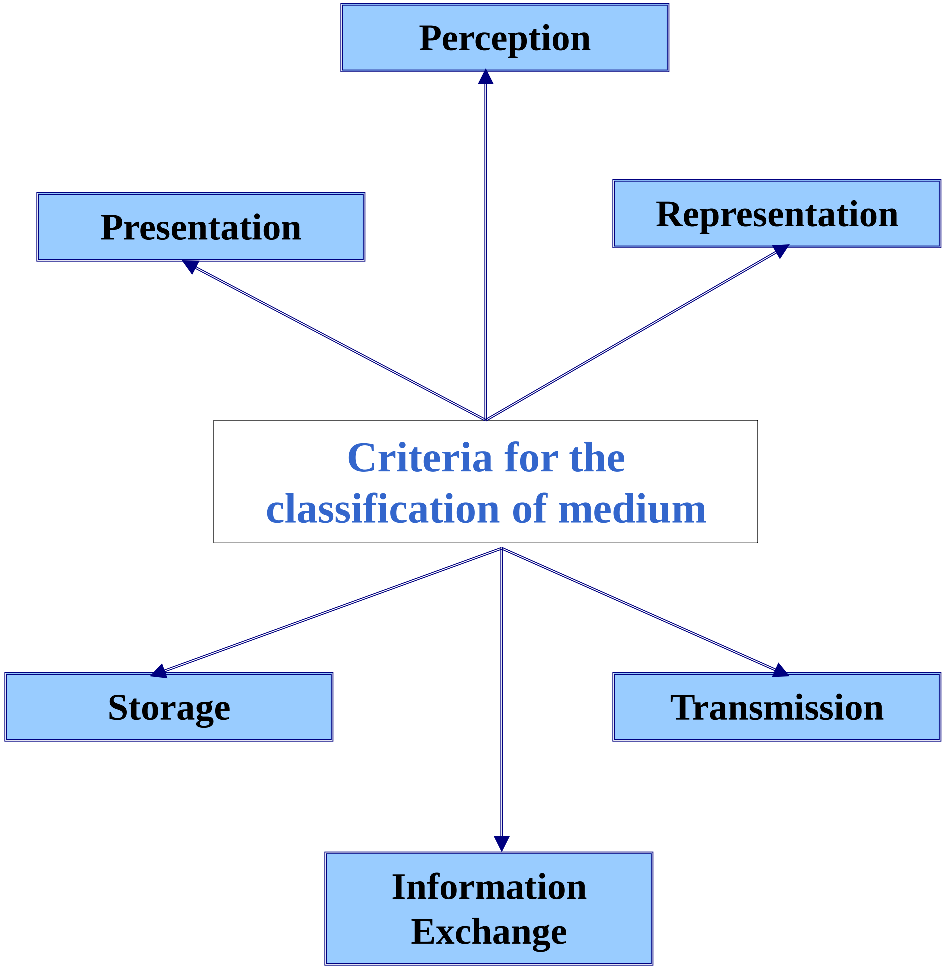
– *Multi* - many; much; multiple
– *Medium* - a substance regarded as the means of transmission of a force or effect; a channel or system of communication, information, or entertainment

(Webster Dictionary)

- So, Multimedia???
  – The terms Multi & Medium don’t seem fit well
  – Term Medium needs more clarification !!!

Multimedia- Definitions

- Medium
  - Means for distribution and presentation of information
- Classification based on perception (text, audio, video) is appropriate for defining multimedia



Multimedia- Definitions

- Time always takes separate dimension in the media representation
- Based on time-dimension in the representation space, media can be
  - Time-independent (*Discrete*)
    - Text, Graphics
  - Time dependent (*Continuous*)
    - Audio, Video
    - Video, sequence of frames (images) presented to the user *periodically*.
    - Time dependent aperiodic media is not continuous!!
  - Discrete & Continuous here has no connection with internal representation !! (relates to the viewers impression…)

Multimedia- Definitions

- Multimedia is *any* combination of *digitally manipulated* text, art, sound, animation and video.
- A more strict version of the definition of multimedia do not allow just any combination of media.
- It requires
  - Both continuous & discrete media to be utilized
  - Significant level of independence between media being used
- The less stricter version of the definition is used in practice.

Multimedia- Definitions

- Multimedia elements are composed into a project using *authoring tools*.

- *Multimedia Authoring tools* are those programs that provide the capability for creating a complete multimedia presentations by linking together objects such as a paragraph of text (song), an illustration, an audio, with appropriate interactive user control.

Multimedia- Definitions

- By defining the objects' relationships to each other, and by sequencing them in an appropriate order, *authors* (those who use authoring tools) can produce attractive and useful graphical applications.
- To name a few authoring tools
  - Photoshop
  - Macromedia Flash
  - Macromedia Director
  - Authorware
- The hardware and the software that govern the limits of what can happen are *multimedia platform* or *environment*

Multimedia- Definitions

- Multimedia is interactive when the end-user is allowed to control what and when the elements are delivered.

- Interactive Multimedia is Hypermedia, when the end-user is provided with the structure of linked elements through which he/she can navigate.

Multimedia- Definitions

- Multimedia is *linear*, when it is not interactive and the users just sit and watch as if it is a movie.
- Multimedia is nonlinear, when the users are given the navigational control and can browse the contents at will.

Multimedia System

- Following the dictionary definitions, *Multimedia system is any system that supports more than a single kind of media*
  - Implies any system processing text and image will be a multimedia system!!!
  - Note, the definition is quantitative. A qualitative definition would be more appropriate.
  - The kind of media supported should be considered, rather the number of media

Multimedia System

- A multimedia system is characterized by computer-controlled, integrated production, manipulation, storage and communication of independent information, which is encoded at least through a continuous (time-dependent) and a discrete (time-independent) medium.

Data streams

- *Data Stream* is any sequence of individual packets transmitted in a time-dependent fashion
  - Packets can carry information of either continuous or discrete media
- Transmission modes
  - Asynchronous
    - Packets can reach receiver as fast as possible.
    - Suited for discrete media
    - Additional time constraints must be imposed for continuous media

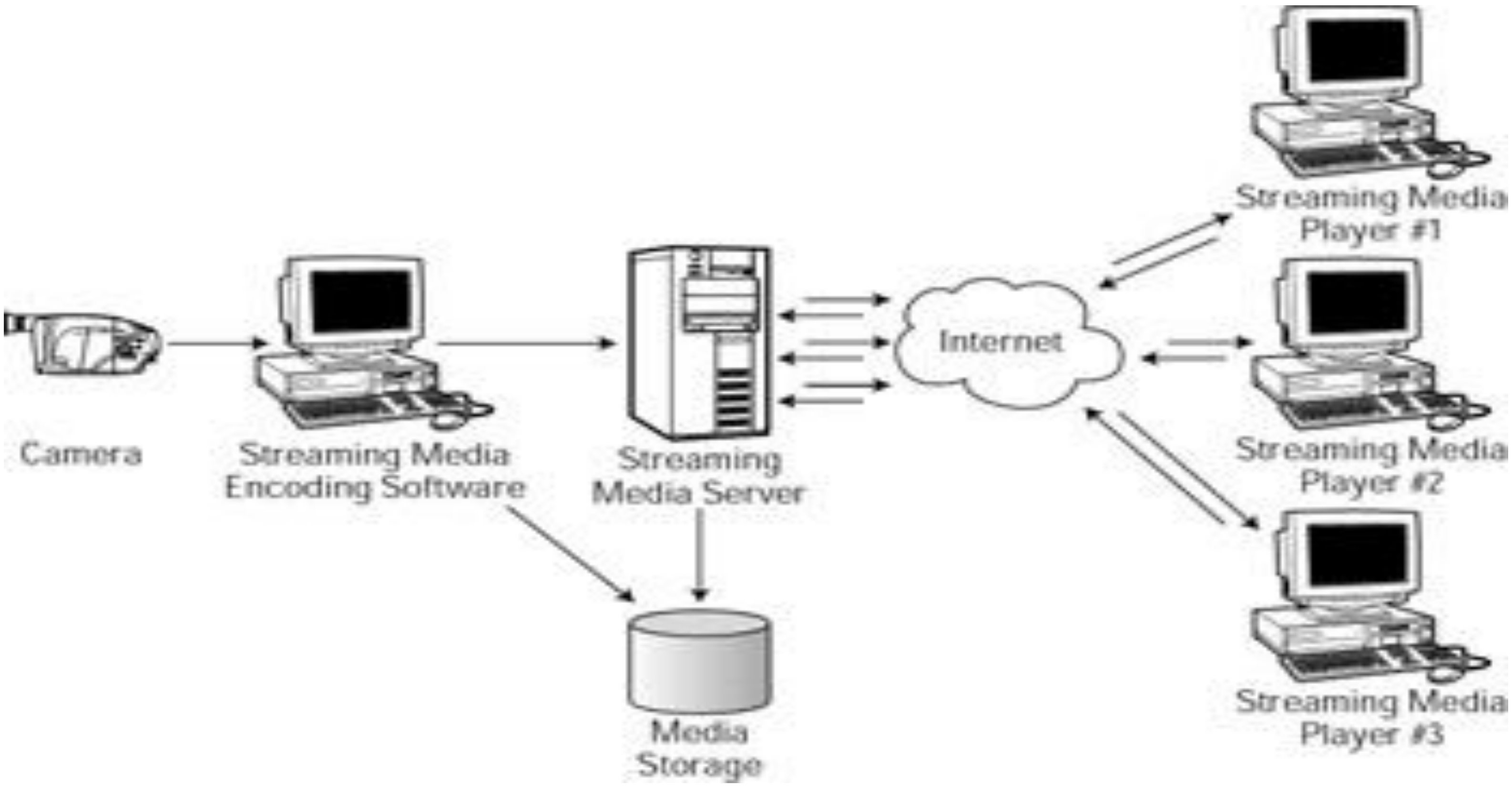
Data streams

- Synchronous
  - Defines maximum end-to-end delay
  - Packets can be received at an arbitrarily earlier time
  - For retrieving uncompressed video at data rate 140Mbits/s & maximal end-to-end delay 1 second the receiver should have temporary storage 17.5 Mbytes
- Isochronous
  - Defines maximum & minimum end-to-end delay
  - Storage requirements at the receiver reduces

Streaming Media

- Popular approach to continuous media over the internet
- Playback at users computer is done while the media is being transferred (no waiting till complete download!!!)
- You can find streaming in
  - Internet radio stations
  - Distance learning
  - Movie Trailers

Streaming Media



Multimedia- Applications

Multimedia plays major role in following areas

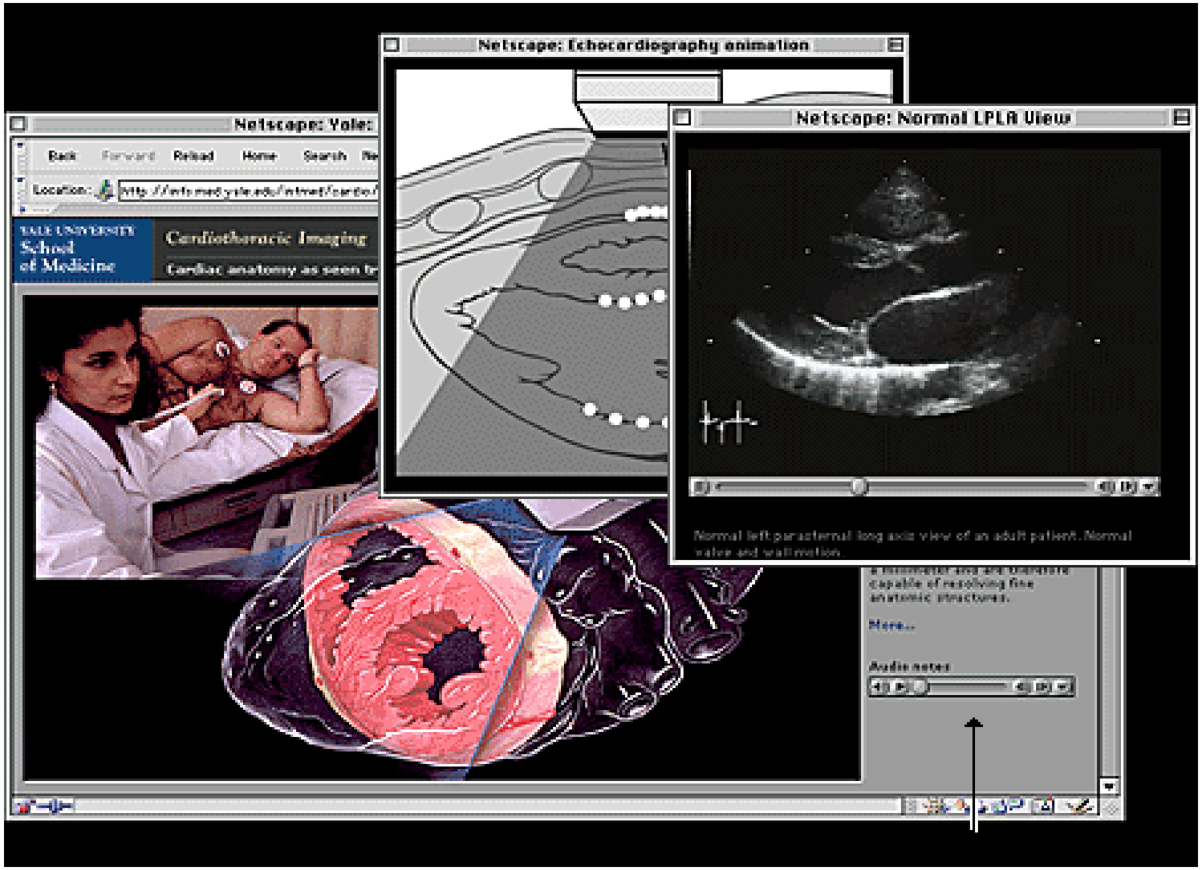
- Instruction
- Business
  - Advertisements
  - Training materials
  - Presentations
  - Customer support services
- Entertainment
  - Interactive Games

Multimedia- Applications

- Enabling Technology
  - Accessibility to web based materials
  - Teaching-learning disabled children & adults

- Fine Arts & Humanities
  - Museum tours
  - Art exhibitions
  - Presentations of literature

Multimedia- Applications



In Medicine

Source:
Cardiac Imaging, YALE centre for advanced cardiac imaging

Multimedia- Applications

In training

Multimedia- Applications

Public awareness campaign

Source

Interactive Multimedia Project
Department of food science& nutrition, Colorado State Univ

Digital Image Representation & Processing

Lesson-2

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

To discuss…

- ✓ Image fundamentals
- ✓ Image Formation
- ✓ 1-bit & 8 bit image
- ✓ Color image
- ✓ Color Lookup Table

Digital Image

- Image can be defined as a 2-D function f(x,y), where x and y are spatial coordinates and the amplitude of f at any pair of coordinates (x,y) is called the *intensity/gray level* of the image at that point
  - When the image is gray scale, intensity values represent the range of shades from black to white.
  - For a color image the intensity values are represented as a combination of R, G, B
- Can be considered as comprising x **x** y number of elements (*picture elements*, *image elements*, *pels*, *pixels*), each of which has a location and value.

Image Formation

- Pixel values are proportional to the energy/ electromagnetic waves radiated from the source
  - It implies this value cannot be negative, ranges from 0 to +ve infinity
- Function f(x,y) characterized by components
  - Illumination   i(x,y), value ranging from 0 to infinity
  - Reflectance   r(x,y), value ranging from 0 to 1
  - f(x,y)= i(x,y) **x** r(x,y)
- f(x,y) lies between L_{min} to L_{max} scaled to [0,L-1], where 0 representing black and L-1 representing white, the intermediate values are the shades of gray from black to white

Image Formation [2]

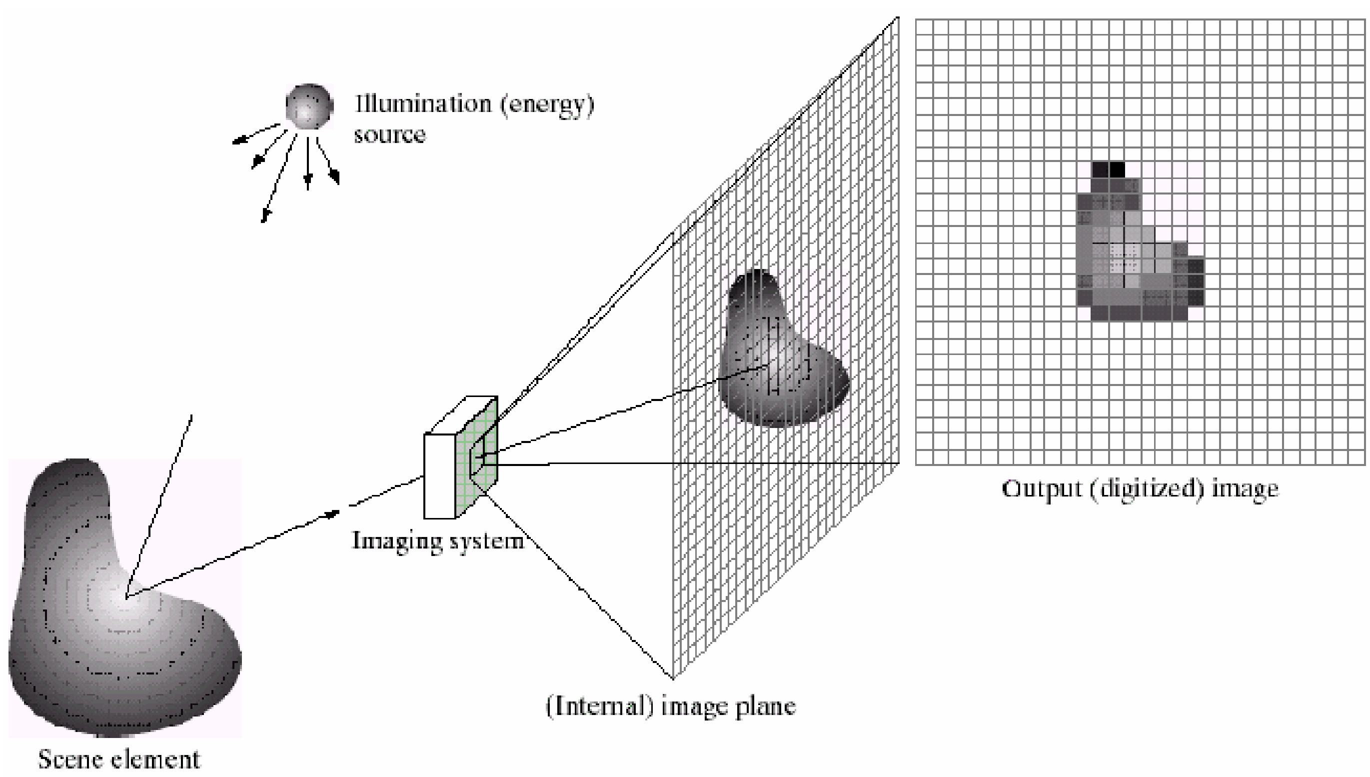


Image Formation [3]

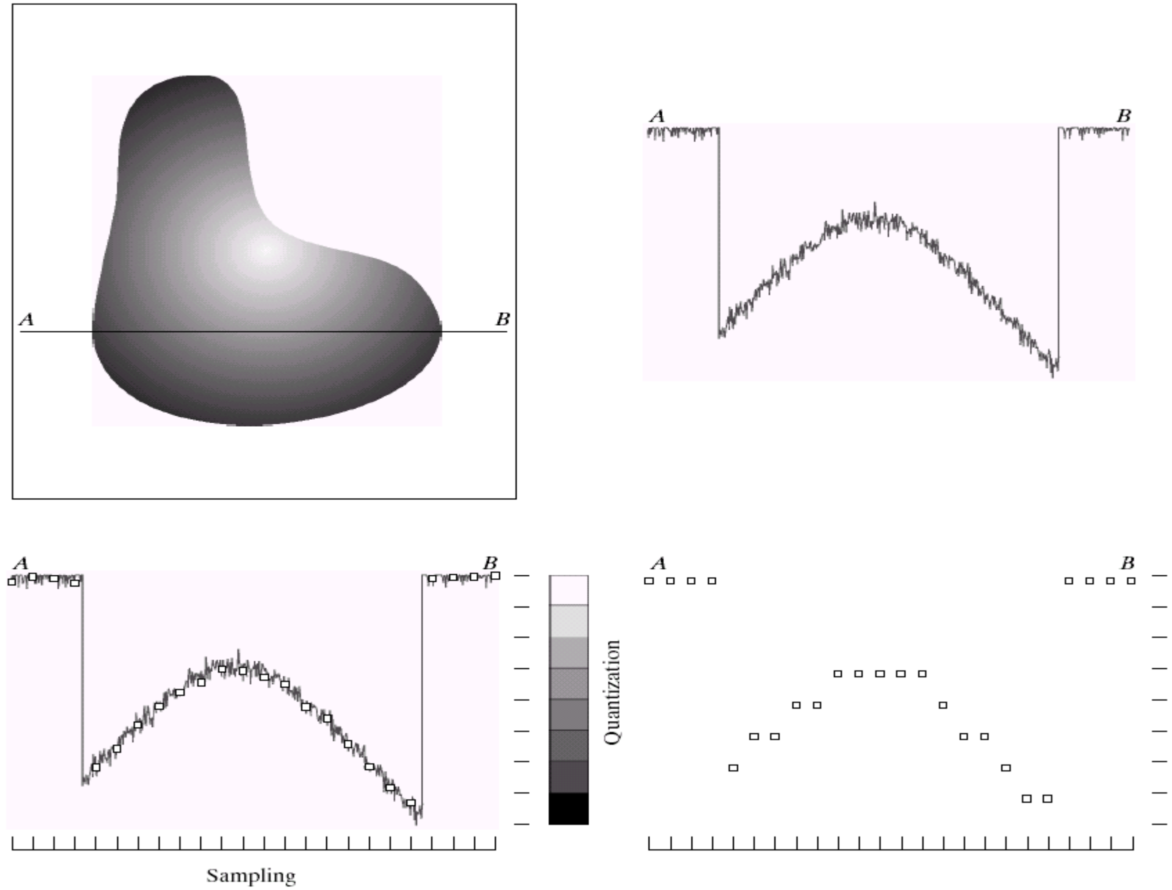


Image Formation [4]

256 gray levels (8bits/pixel)

32 gray levels (5 bits/pixel)

16 gray levels (4 bits/pixel)

8 gray levels (3 bits/pixel)

4 gray levels (2 bits/pixel)

2 gray levels (1 bit/pixel)

Image Formation[5]

- Output of image sensors are continuous voltage waveform, digitization is necessary for further processing
- Digitizing the coordinate positions are called *sampling*
- Digitizing the amplitude values are called *quantization*
  - Number of gray levels will be in an integer power of 2
  - $L=2^k$, [0…L-1]
  - Number of bits needed to store an image b=M x N
- Image is k bit image if it has 2^k gray levels
  - 8 bit image has 256 gray levels

1-bit image

- Simplest type of image
- Each pixel consist of only ON / OFF information
- Called 1-bit monochrome (since no color) image
- Suitable for simple graphics & text
  - JBIG (**J**oint **B**i-level **I**mage experts **G**roup), A compression standard for binary image

8-bit image

- Gray levels between 0 to 255 (black to white)
- *Image resolution* refers the number of pixels in an image. The higher the resolution, the more pixels in the image. Higher resolution allows for more detail and subtle color transitions in an image
- Shown is 512x512 byte image

Color image

- 24- bit color image
  - Each pixel is represented by 3 bytes, RGB
  - Each R, G, B are in the range 0-255
  - 256 x 256 x 256 possible colors
  - If space is a concern, reasonably accurate color image can be obtained by quantizing the color information
- 8- bit color image
  - Carefully chosen 256 colors represent the image
  - We get information can be received from the color histogram

Color image [2]

- For 640 x 480 image represented with
  - 24 bits requires 921.6 kbytes
  - 8 bit requires 300 kbytes
- The 8-bit color image stores only the index of the color, the file header will contain the mapping information.
- The table where the color information for all the 256 indices is called *color lookup table* (LUT)

Basic Relationships Between Pixels

Lesson-3

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

To discuss…

- ✓ Neighborhood
- ✓ Adjacency
- ✓ Connectivity
- ✓ Paths
- ✓ Regions and boundaries

Neighbors of a Pixel

- Any pixel p*(x, y) has two vertical and two* horizontal neighbors, given by

  $(x+1, y), (x-1, y), (x, y+1), (x, y-1)$
- This set of pixels are called the 4-neighbors of P, and is denoted by N4(P).
- Each of them are at a unit distance from P.

Neighbors of a Pixel [2]

- The four diagonal neighbors of p*(x,y) are* given by,

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

  This set is denoted by $N_D(P)$.
- Each of them are at Euclidean distance of 1.414 from P.

Neighbors of a Pixel [3]

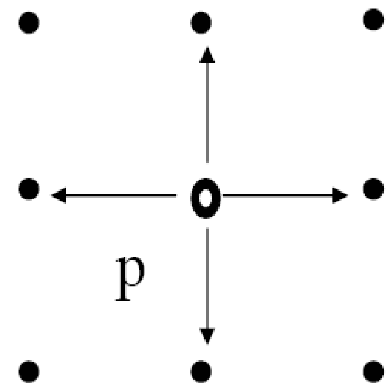
- The points $N_D(P)$ and $N_4(P)$ are together known as 8-neighbors of the point P, denoted by $N_8(P)$.

- Some of the points in the N_4, N_D and N_8 may fall outside image when P lies on the border of image.

Neighbors of a Pixel [4]

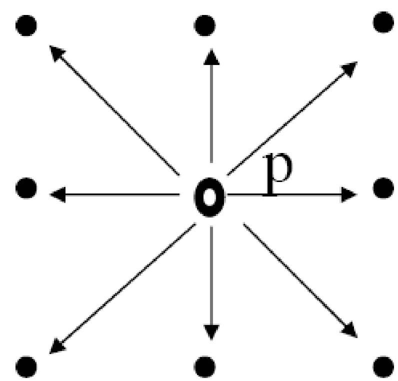
- Neighbors of a pixel

a. 4-neighbors **of a pixel p are its vertical and horizontal neighbors denoted by N_4(p)**



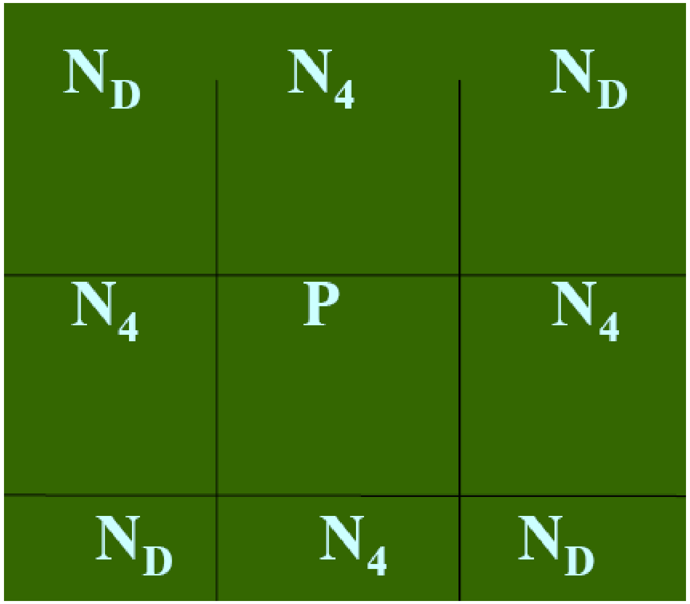
$N_4(p)$

b. 8-neighbors **of a pixel p are its vertical horizontal and 4 diagonal neighbors denoted by N_8(p)**



$N_8(p)$

Neighbors of a Pixel [5]



- •N_4 - 4-neighbors
- •N_D - diagonal neighbors
- •N_8 - 8-neighbors (N_4 U N_D)

Adjacency

- **Two pixels are connected if they are neighbors and their gray levels satisfy some specified criterion of similarity.**

- **For example, in a binary image two pixels are connected if they are 4-neighbors and have same value (0/1).**

Adjacency

- Let V be set of gray levels values used to define adjacency.
- <u>4-adjacency</u>: Two pixels *p and q with values from V are 4-* adjacent if q is in the set $N_4(p)$.
- <u>8-adjacency</u>: Two pixels *p and q with values from V are 8-* adjacent if q is in the set $N_8(p)$.
- <u>m-adjacency</u>: Two pixels *p and q with values from V are m-adjacent* if,

  – q is in $N_4(p)$.

  – q is in $N_D(p)$ and the set $[\, N_4(p) \cap N_4(q) \,]$ is empty (has no pixels whose values are from V).

Connectivity :

- To determine whether the pixels are adjacent in some sense.
- Let V be the set of gray-level values used to define connectivity; then Two pixels p, q that have values from the set V are:

  a. 4-connected, if q is in the set $N_4(p)$

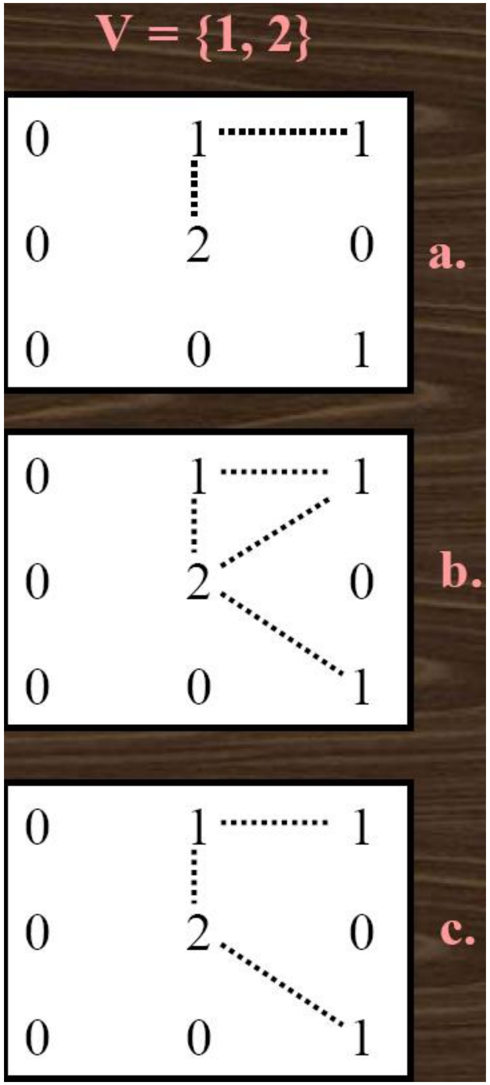
  b. 8-connected, if q is in the set $N_8(p)$

  c. m-connected, iff

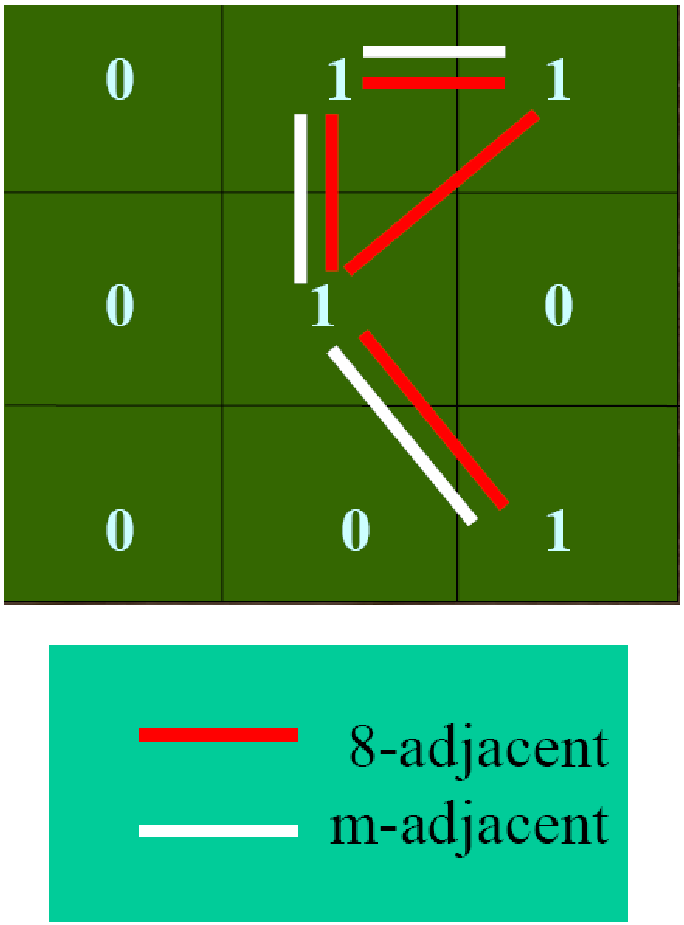
  i. q is in $N_4(p)$ or

  ii. q is in $N_D(p)$ and the set

  $[\ N_4(p) \cap N_4(q)\ ]$ is empty

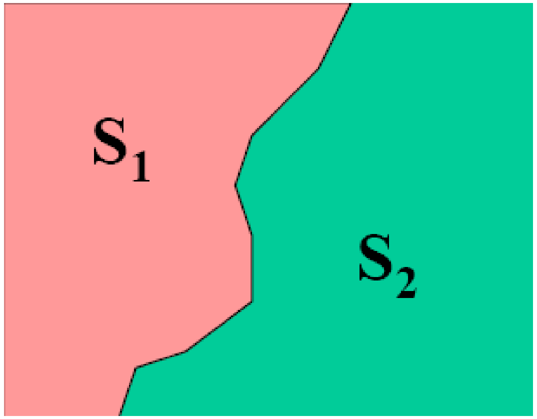


Adjacency/Connectivity



Adjacency/Connectivity

- Pixel p is *adjacent* to pixel q if they are connected.
- Two image subsets S1 and S2 are *adjacent* if some pixel in S1 is adjacent to some pixel in S2



Paths & Path lengths

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

  $(x_0, y_0), (x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$,

  where $(x_0, y_0)=(x, y)$ *and* $(x_n, y_n)=(s, t)$;

  (x_i, y_i) *is adjacent to* (x_{i-1}, y_{i-1})

  Here *n is the length of the path.*

- We can define 4-, 8-, and m-paths based on type of adjacency used.

Connected Components

- If *p and q are pixels of an image subset S* then *p is connected to q in S if there is a* path from *p to q consisting entirely of* pixels in S.
- For every pixel *p in S, the set of pixels in* S that are connected to *p is called a connected component of S.*
- If S has only one connected component then S is called *Connected Set.*

Regions and Boundaries

- **A subset R of pixels in an image is called a *Region of the image if R is a* connected set.**
- **The *boundary of the region R is the* set of pixels in the region that have one or more neighbors that are not in R.**

Distance Measures

- **Given pixels** ***p, q and z with coordinates (x, y), (s, t), (u, v) respectively, the distance*** **function D has following properties:**

a. $D(p, q) \geq 0$, $[D(p, q) = 0$, iff $p = q]$

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

c. $D(p, z) \leq D(p, q) + D(q, z)$

- The following are the different Distance measures:

a. Euclidean Distance :

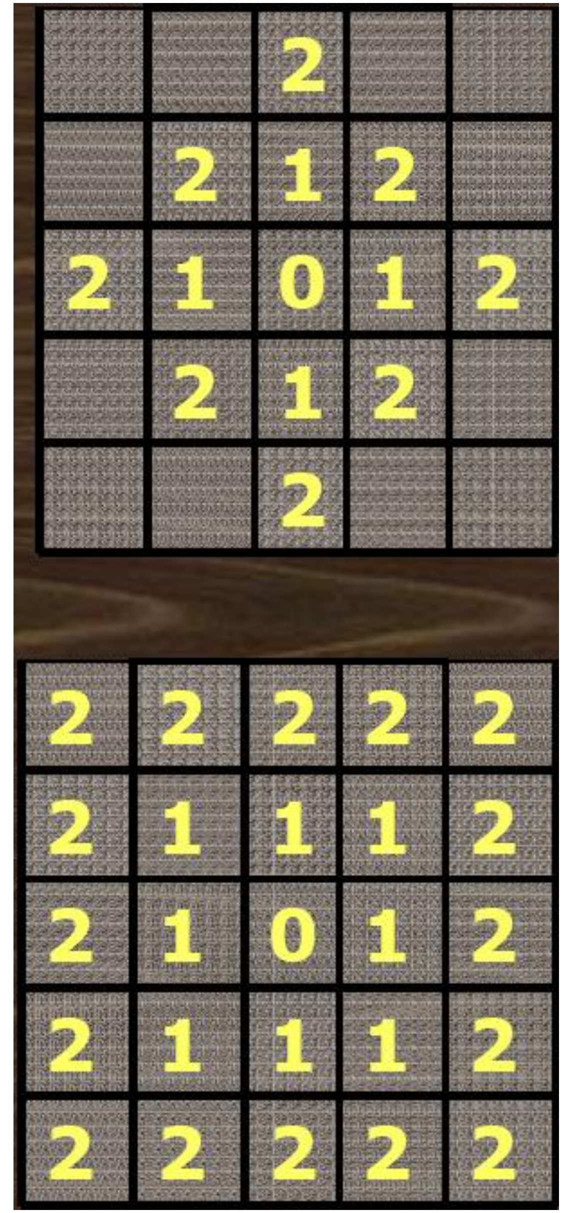
$D_e(p, q) = SQRT\,[(x-s)^2 + (y-t)^2]$

b. City Block Distance:

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

c. Chess Board Distance:

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



Relationship between pixels

- Arithmetic/Logic Operations:

  - Addition : $p + q$

  – Subtraction: $p - q$

  – Multiplication: $p * q$

  – Division: p/q

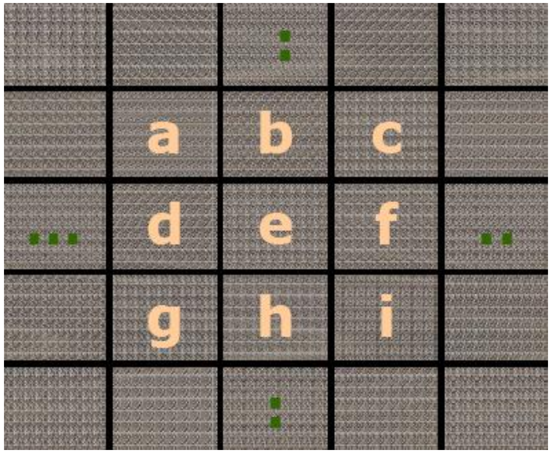
  – AND: p *AND* q

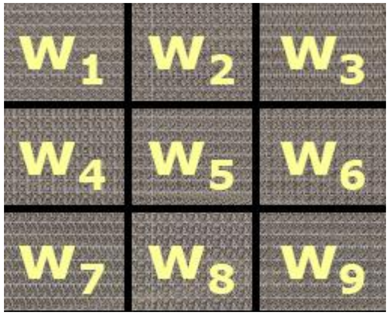
  – OR : p *OR* q

  – Complement: **NOT**(q)

Neighborhood based arithmetic/Logic

Value assigned to a pixel at position '*e' is a* function of its neighbors and a set of window functions.





$$p = (w_1 a + w_2 b + w_3 c + w_4 d + w_5 e + w_6 f + w_7 g + w_8 h + w_9 i)$$
$$= \sum w_i f_i$$

Arithmetic/Logic Operations

- Tasks done using neighborhood processing:
  – Smoothing / averaging
  – Noise removal / filtering
  – Edge detection
  – Contrast enhancement

Color

Lesson-4

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

To discuss…

- ✓ Color Science
- ✓ Human Visual Perception
- ✓ Color Models in image

Color Science – Light & Spectra

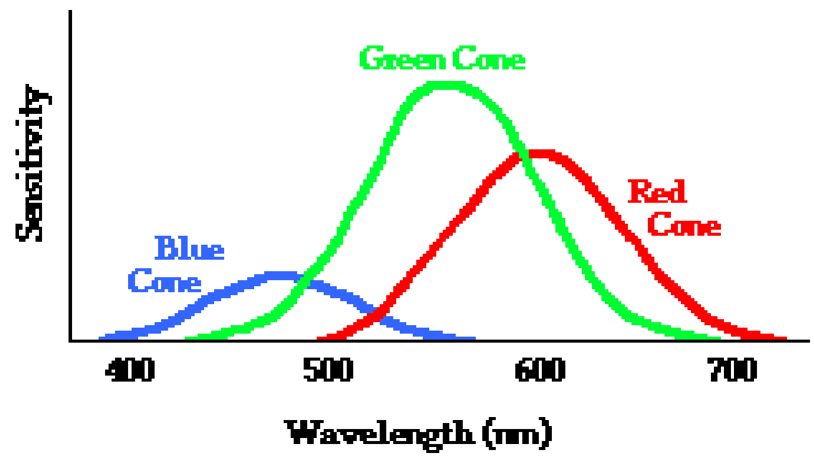
- Light is an electromagnetic wave
  - It's color is characterized by it's wavelength
- Most light sources produce contributions over many wavelengths, contributions fall in the visible wavelength can be seen
- λ vs Spectral power curve is called *spectral power distribution* E(λ)
- Light from 400 to 700 nanometer (10^{-9} meter)

Color Science – Light & Spectra

- Red light has longer wavelength in the visible light & blue the shorter
- The shorter the wavelength, higher the vibration & energy
- Red photons carry around 1.8eV & blue 3.1eV (1 electron volt = $1.60217646 \times 10^{-19}$ joules) unit of energy
- The RGB in the image files are converted to analog & drive the electron guns of CRT (Cathode Ray Tube)

Color Science – Vision & Sensitivity [2]

- Eye is most sensitive to the middle of the visible spectrum
- Let us denote the spectral sensitivity of R, G, B cones as a vector $q(\lambda)$



- $q(\lambda)=[q^R(\lambda), q^G(\lambda), q^B(\lambda)]^T$

Color Science – Vision & Sensitivity [3]

- The sensitivity of each cones can be specified as
  - $R = \int E(\lambda)\, q^R(\lambda)\, d\,\lambda$ ---------- (1)
  - $G = \int E(\lambda)\, q^G(\lambda)\, d\,\lambda$ ---------- (2)
  - $B = \int E(\lambda)\, q^B(\lambda)\, d\,\lambda$ ---------- (3)
    - » $\int$ - integral
- Equations 1, 2, 3 quantify the signals transmitted to the brain

Human Visual Perception

- Human perception encompasses both the physiological and psychological aspects.
- We will focus more on physiological aspects, which are more easily quantifiable and hence, analyzed.

Human Visual Perception

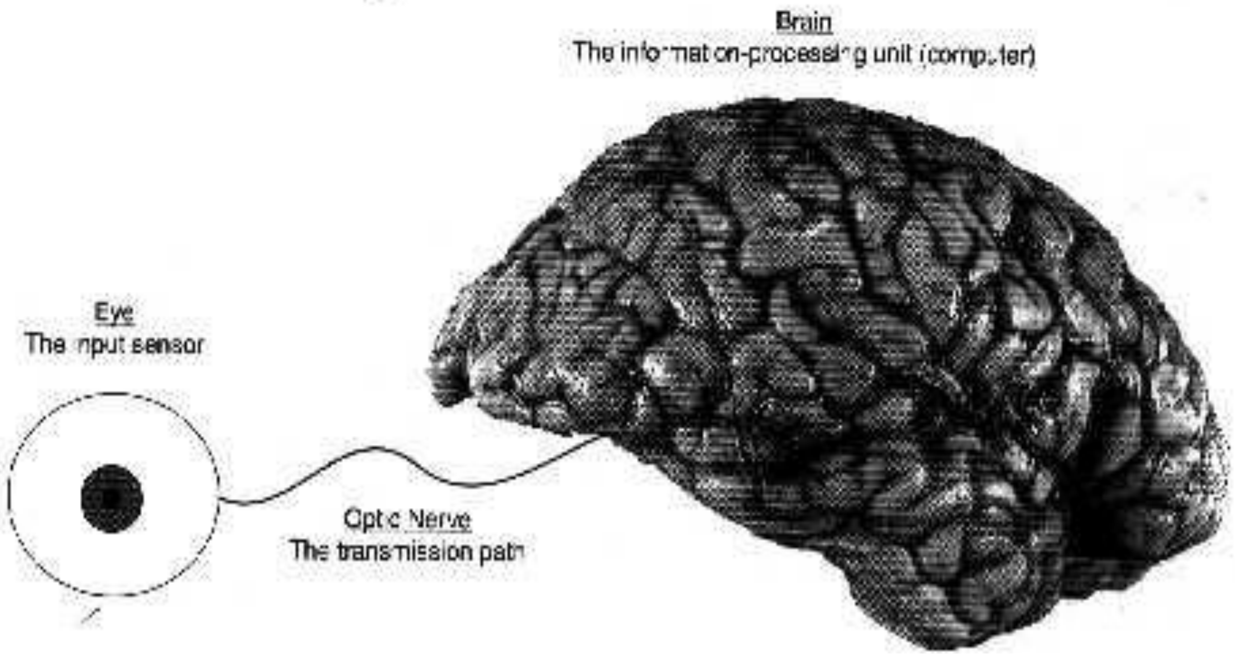
- Why study visual perception?
  – Image processing algorithms are designed based on how our visual system works.
  – In image compression, we need to know what information is not perceptually important and can be ignored.
  – In image enhancement, we need to know what types of operations that are likely to improve an image visually.

The Human Visual System

- The human visual system consists of two primary components – the eye and the brain, which are connected by the optic nerve.
  - Eye – receiving sensor (camera, scanner).
  - Brain – information processing unit (computer system).
  - Optic nerve – connection cable (physical wire).

The Human Visual System

Figure 1.6-1 The Human Visual System



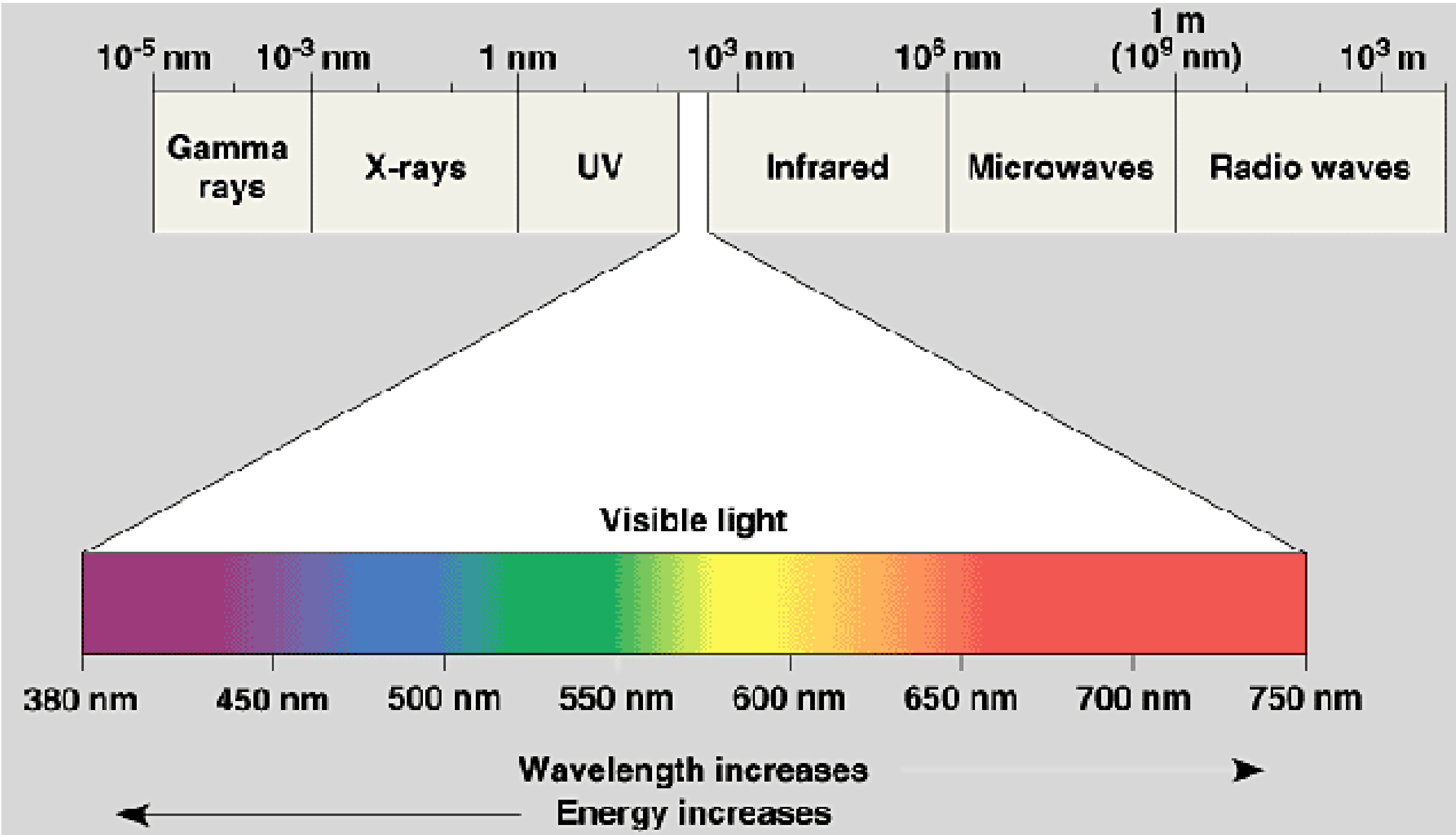
The Human Visual System

- This is how human visual system works:
  - Light energy is focused by the lens of the eye into sensors and retina.
  - The sensors respond to the light by an electrochemical reaction that sends an electrical signal to the brain (through the optic nerve).
  - The brain uses the signals to create neurological patterns that we perceive as images.

The Human Visual System

- The visible light is an electromagnetic wave with wavelength range of about 380 to 825 nanometers.
  - However, response above 700 nanometers is minimal.
- We cannot “see” many parts of the electromagnetic spectrum.

The Human Visual System

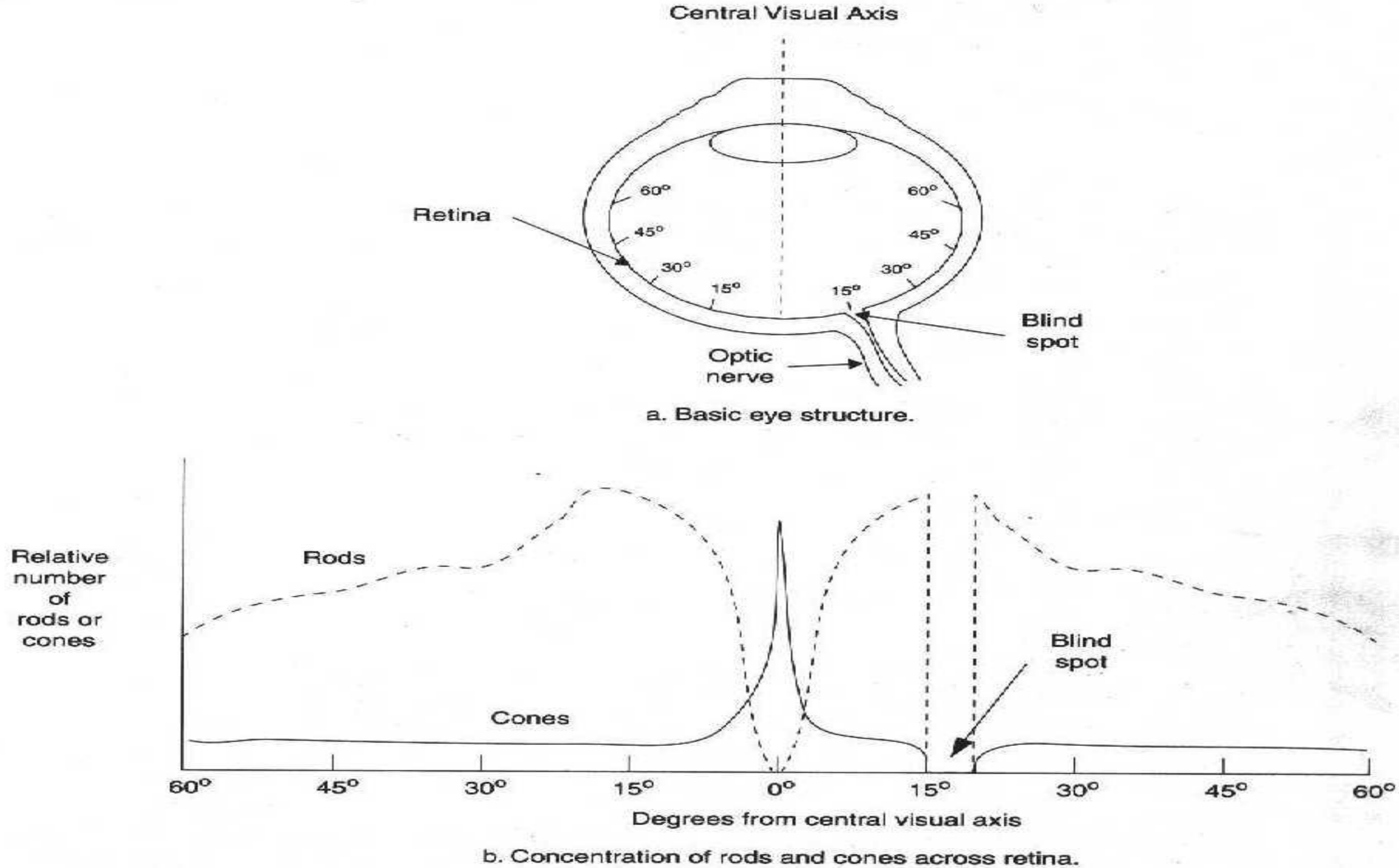


The Human Visual System

- The visible spectrum can be divided into three bands:
  - Blue (400 to 500 nm).
  - Green (500 to 600 nm).
  - Red (600 to 700 nm).
- The sensors are distributed across retina.
- 3 kinds of cones are more sensitive to R, G & B present in the ratios 40:20:1

The Human Visual System

Figure 1.6-3 The Human Eye



The Human Visual System

- There are two types of sensors: rods and cones.
- Rods:
  - For night vision.
  - See only brightness (gray level) and not color.
  - Distributed across retina.
  - Medium and low level resolution.

The Human Visual System

- Cones:
  - For daylight vision.
  - Sensitive to color.
  - Concentrated in the central region of eye.
  - High resolution capability (differentiate small changes).

The Human Visual System

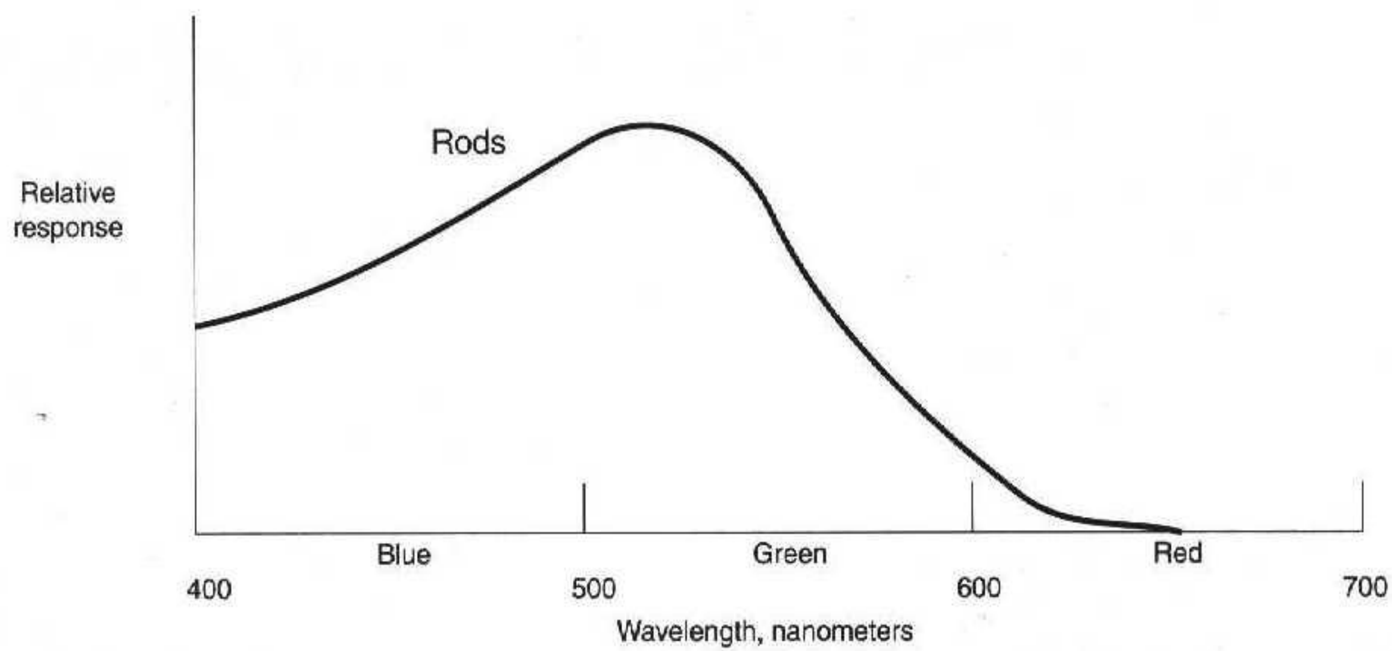
- Blind spot:
  - – No sensors.
  - – Place for optic nerve.
  - – We do not perceive it as a blind spot because the brain fills in the missing visual information.
- Why does an object should be in center field of vision in order to perceive it in fine detail?
  - – This is where the cones are concentrated.

The Human Visual System

- Cones have higher resolution than rods because they have individual nerves tied to each sensor.
- Rods have multiple sensors tied to each nerve.
- Rods react even in low light but see only a single spectral band. They cannot distinguish color.

The Human Visual System

Figure 1.6-4 Relative Responses of Rods and Cones

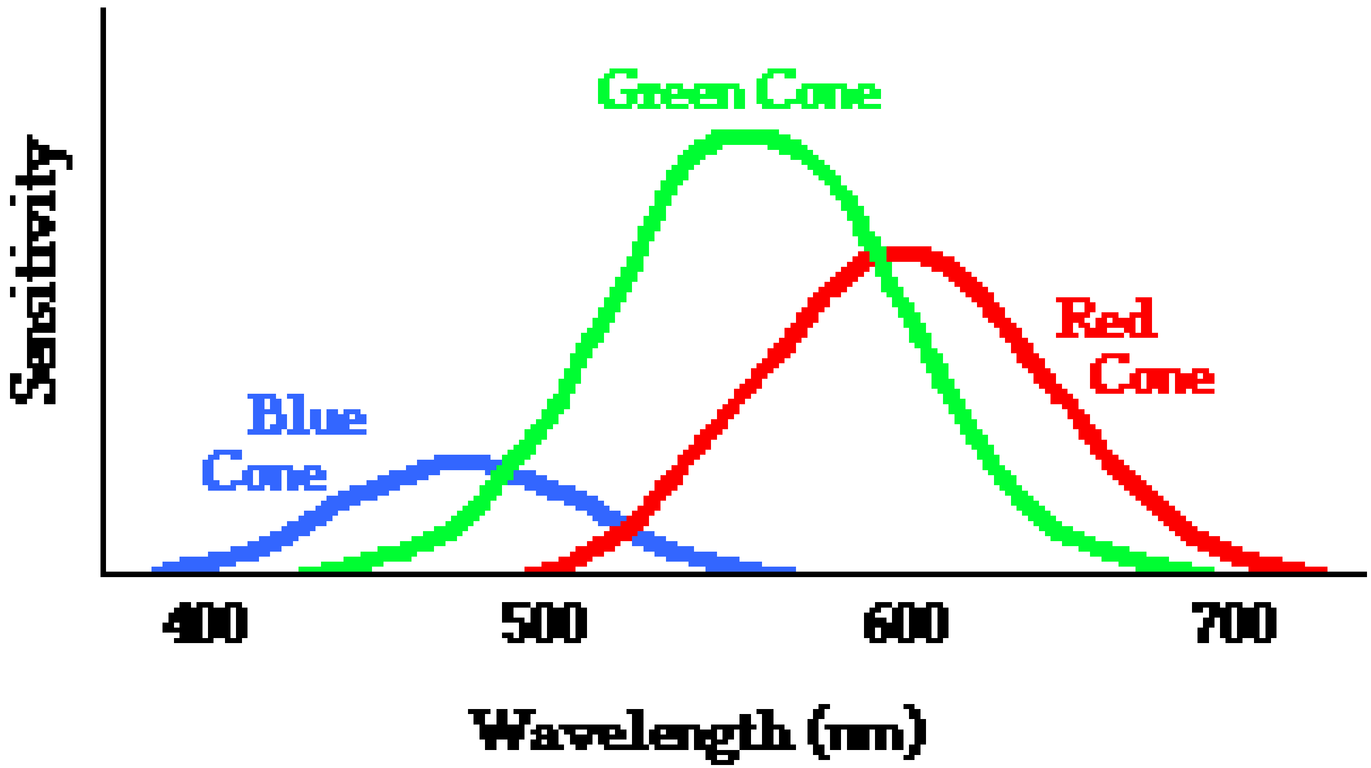


a. Rods react even in low light levels but see only a single spectral band; they cannot distinguish colors.

The Human Visual System

- There are three types of cones. Each responding to different wavelengths of light energy.
- The colors that we perceive are the combined result of the response of the three cones.

The Human Visual System



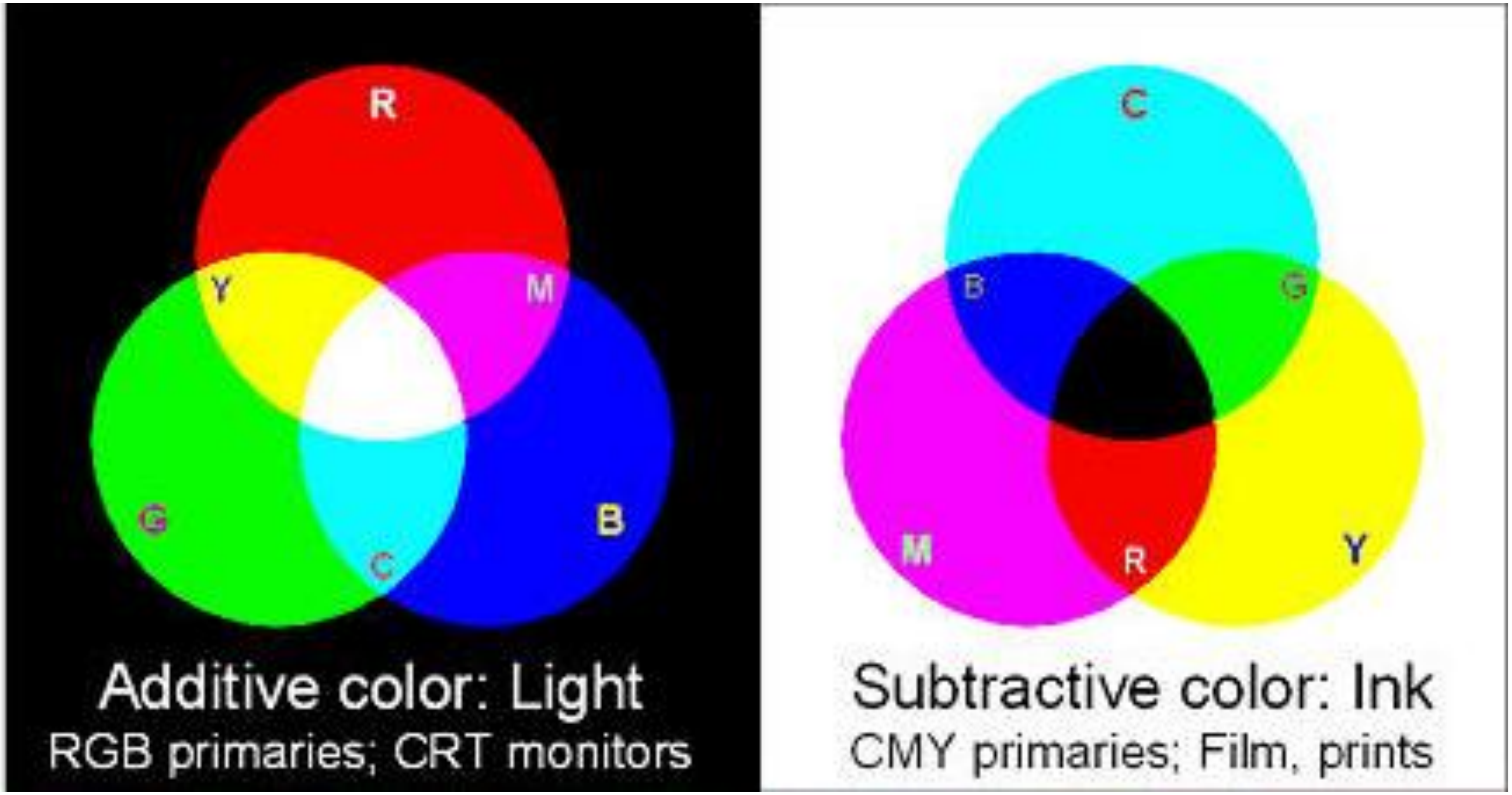
Color Science – Other Color Coordinate Systems

- CMY (Cyan-Magenta-Yellow)
- HSL(Hue-Saturation-Lightness)
- HSV(Hue-Saturation-Value)
- HSI(Hue-Saturation-Intensity)
- HCI(Hue-Chroma-Intensity)
- HVC(Hue-Value-Intensity)
- HSD(Hue-Saturation-Darkness)

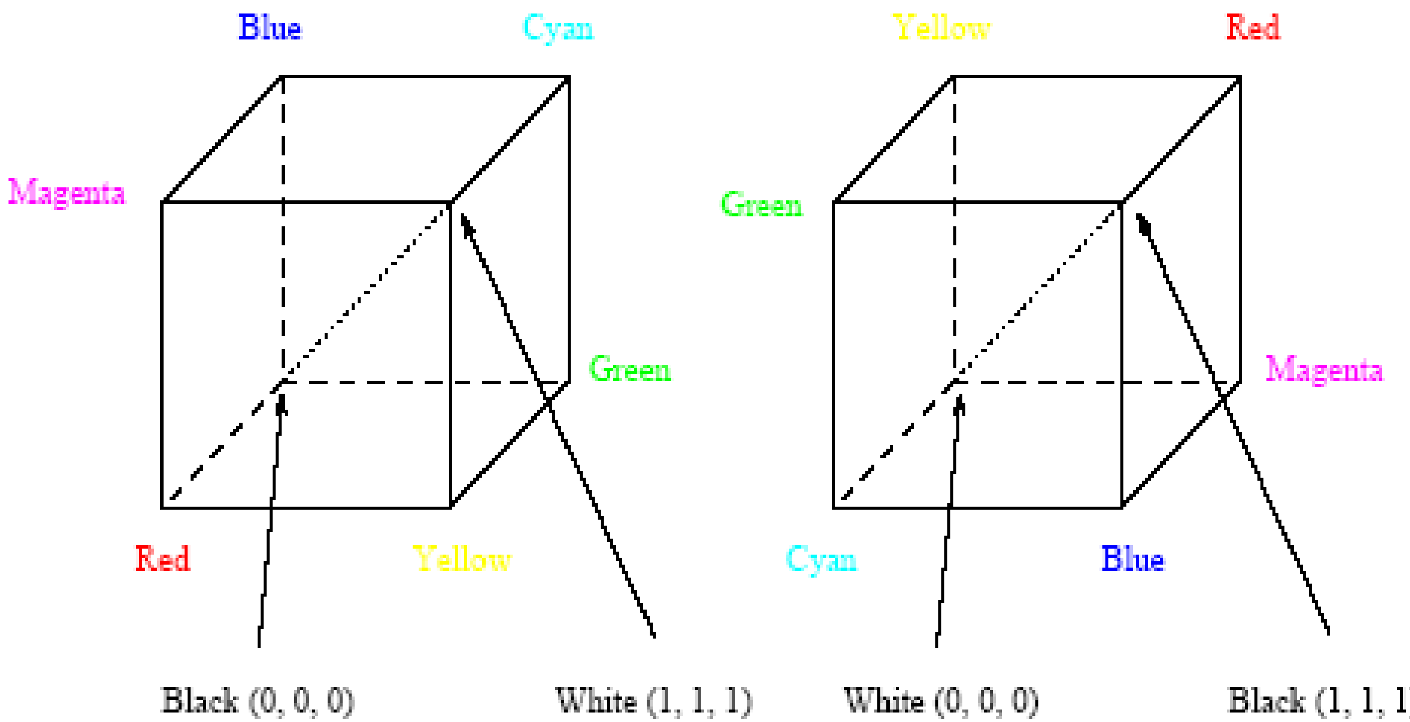
Color Models for Image – RGB Vs CMY

- Additive Vs Subtractive Models
- Additive model
  - Used in computer displays, Uses light to display color, Colors result from transmitted light
  - Red+Green+Blue=White
- Subtractive Models
  - Used in printed materials, Uses ink to display color, Colors result from reflected light
  - Cyan+Magenta+Yellow=Black

Color Models for Image – RGB Vs CMY [2]



Color Models for Image – RGB Vs CMY [3]



RGB & CMY Cubes

Color Models for Image – RGB Vs CMY [4]

- Conversion From RGB to CMY

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- **Conversion From CMY to RGB**

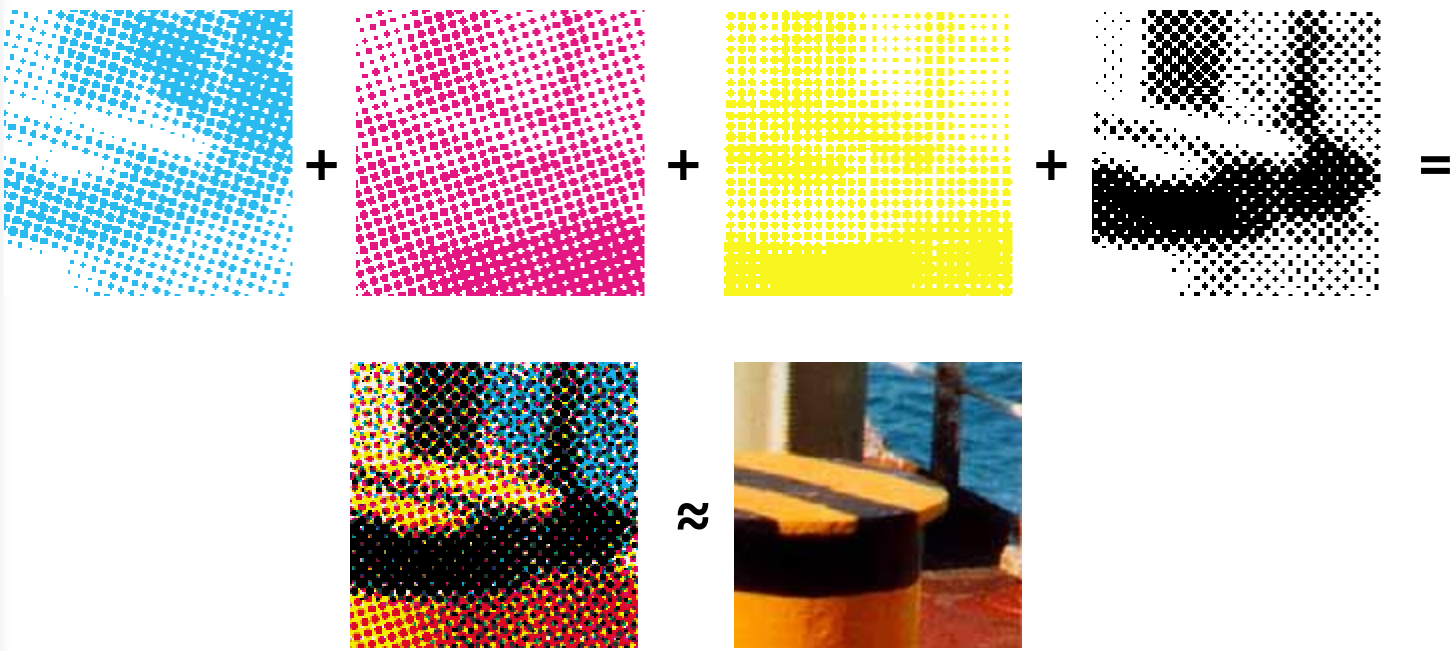
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$

Color Models for Image – CMYK

- Eliminating amounts of yellow, magenta, and cyan that would have added to a dark neutral (black) and replacing them with black ink
- Four-color printing uses black ink(K) in addition to the subtractive primaries yellow, magenta, and cyan.
- Reasons for Black addition includes
  - CMY Mixture rarely produces pure black
  - Text is typically printed in black and includes fine detail
  - Cost saving : Unit amount of black ink rather than three unit amounts of CMY

Color Models for Image – CMYK[2]

- **Used especially in the printing of images**



Spatial Filtering

Lesson 5

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

Background

- Filter term in “Digital image processing” is referred to the subimage
- There are others term to call subimage such as mask, kernel, template, or window
- The value in a filter subimage are referred as coefficients, rather than pixels.

Basics of Spatial Filtering

- The concept of filtering has its roots in the use of the Fourier Transform for signal processing in the so-called frequency domain.
- Spatial filtering term is the filtering operations that are performed directly on the pixels of an image

Mechanics of spatial filtering

- The process consists simply of moving the filter mask from point to point in an image.
- At each point (x,y) the response of the filter at that point is calculated using a predefined relationship

Linear spatial filtering

Pixels of image

	w(-1,-1) f(x-1,y-1)	w(-1,0) f(x-1,y)	w(-1,1) f(x-1,y+1)	
	w(0,-1) f(x,y-1)	w(0,0) f(x,y)	w(0,1) f(x,y+1)	
	w(1,-1) f(x+1,y-1)	w(1,0) f(x+1,y)	w(1,1) f(x+1,y+1)	

The result is the sum of products of the mask coefficients with the corresponding pixels directly under the mask

Mask coefficients

w(-1,-1)	w(-1,0)	w(-1,1)
w(0,-1)	w(0,0)	w(0,1)
w(1,-1)	w(1,0)	w(1,1)

$$
\begin{aligned}
f(x,y) = {} & w(-1,-1)\,f(x-1,y-1) + w(-1,0)\,f(x-1,y) + w(-1,1)\,f(x-1,y+1) + \\
& w(0,-1)\,f(x,y-1) + w(0,0)\,f(x,y) + w(0,1)\,f(x,y+1) + \\
& w(1,-1)\,f(x+1,y-1) + w(1,0)\,f(x+1,y) + w(1,1)\,f(x+1,y+1)
\end{aligned}
$$

Note: Linear filtering

- The coefficient w(0,0) coincides with image value f(x,y), indicating that the mask is centered at (x,y) when the computation of sum of products takes place.
- For a mask of size mxn, we assume that m-2a+1 and n=2b+1, where a and b are nonnegative integer. Then m and n are odd.

Linear filtering

- In general, linear filtering of an image f of size MxN with a filter mask of size mxn is given by the expression:

$$g(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b}w(s,t)f(x+s,y+t)$$

Discussion

- The process of linear filtering similar to a frequency domain concept called "*convolution*"

Simplify expression

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_{mn} z_{mn} = \sum_{i=1}^{mn} w_i z_i$$

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Where the w's are mask coefficients, the z's are the value of the image gray levels corresponding to those coefficients

Nonlinear spatial filtering

- Nonlinear spatial filters also operate on neighborhoods, and the mechanics of sliding a mask past an image are the same as was just outlined.
- The filtering operation is based conditionally on the values of the pixels in the neighborhood under consideration

Smoothing Spatial Filters

- Smoothing filters are used for blurring and for noise reduction.
  - Blurring is used in preprocessing steps, such as removal of small details from an image prior to object extraction, and bridging of small gaps in lines or curves
  - Noise reduction can be accomplished by blurring

Type of smoothing filtering

- There are 2 ways of smoothing spatial filters
  - Smoothing Linear Filters
  - Order-Statistics Filters

Smoothing Linear Filters

- Linear spatial filter is simply the average of the pixels contained in the neighborhood of the filter mask.
- Sometimes called “averaging filters”.
- The idea is replacing the value of every pixel in an image by the average of the gray levels in the neighborhood defined by the filter mask.

Two 3x3 Smoothing Linear Filters

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Standard average

$$\frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Weighted average

5x5 Smoothing Linear Filters

$$\frac{1}{25} \times$$

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

Smoothing Linear Filters

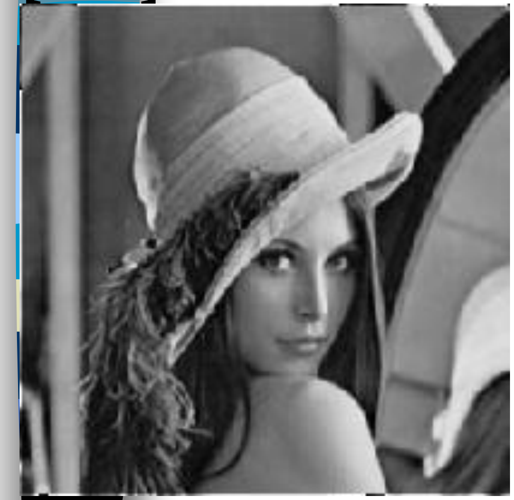
- The general implementation for filtering an MxN image with a weighted averaging filter of size mxn is given by the expression

$$g(x,y)=\frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)\,f(x+s,y+t)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)}$$

Result of Smoothing Linear Filters

Original Image

[3x3]



[5x5]

[7x7]

Order-Statistics Filters

- Order-statistics filters are nonlinear spatial filters whose response is based on ordering (ranking) the pixels contained in the image area encompassed by the filter, and then replacing the value of the center pixel with the value determined by the ranking result.
- Best-known “median filter”

Process of Median filter

10	15	20
20	100	20
20	20	25

10, 15, 20, 20, 20, 20, 20, 25, 100

5^{th}

- Corp region of neighborhood
- Sort the values of the pixel in our region
- In the MxN mask the median is MxN div 2 +1

Result of median filter

Noise from Glass effect

Remove noise by median filter

Sharpening Spatial Filters

- The principal objective of sharpening is to highlight fine detail in an image or to enhance detail that has been blurred, either in error or as an natural effect of a particular method of image acquisition.

Introduction

- The image blurring is accomplished in the spatial domain by pixel averaging in a neighborhood.
- Since averaging is analogous to integration.
- Sharpening could be accomplished by spatial differentiation.

Foundation

- We are interested in the behavior of these derivatives in areas of constant gray level(flat segments), at the onset and end of discontinuities(step and ramp discontinuities), and along gray-level ramps.
- These types of discontinuities can be noise points, lines, and edges.

Definition for a first derivative

- Must be zero in flat segments
- Must be nonzero at the onset of a gray-level step or ramp; and
- Must be nonzero along ramps.

Definition for a second derivative

- Must be zero in flat areas;
- Must be nonzero at the onset and end of a gray-level step or ramp;
- Must be zero along ramps of constant slope

Definition of the 1^{st}-order derivative

- A basic definition of the first-order derivative of a one-dimensional function f(x) is

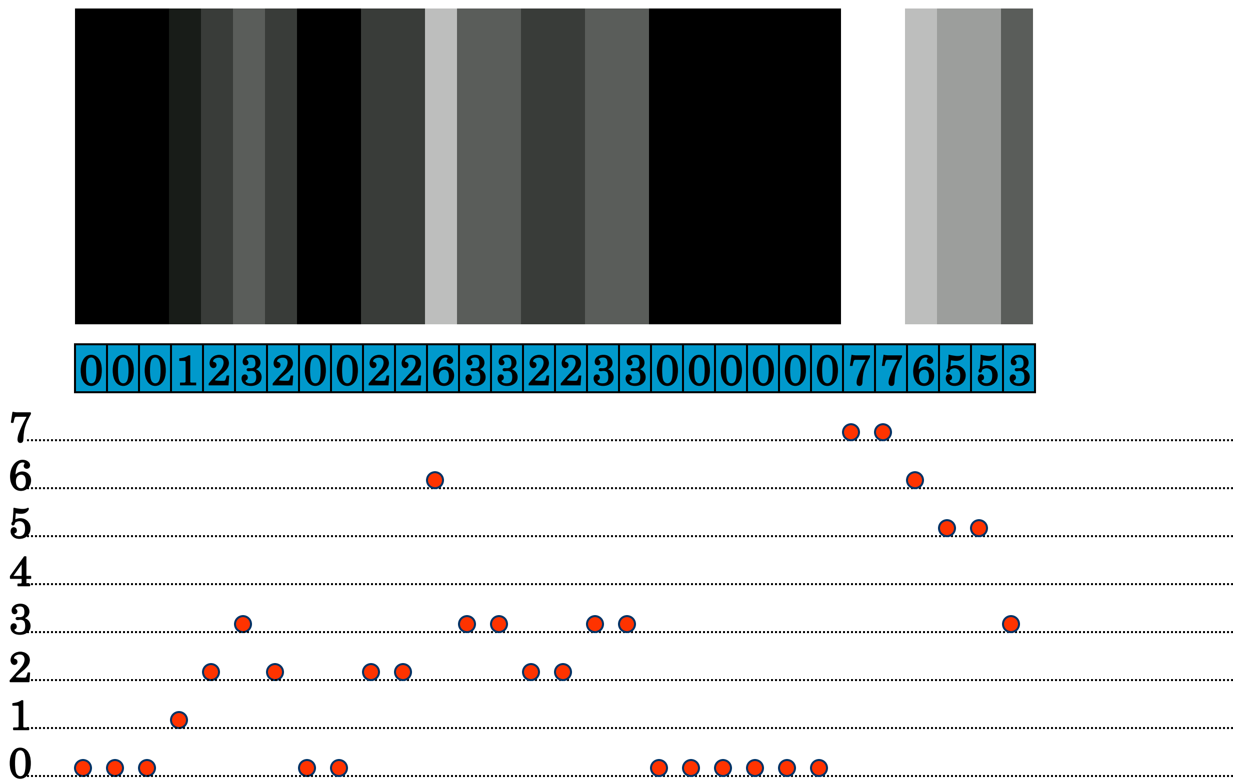
$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

Definition of the 2nd-order derivative

- We define a second-order derivative as the difference

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x).$$

Gray-level profile



Derivative of image profile

```
      0  0  0  1  2  3  2  0  0  2  2  6  3  3  2  2  3  3  0  0  0  0  0  0  7  7  6  5  5  3
fi       0  0  1  1  1 -1 -2  0  2  0  4 -3  0 -1  0  1  0 -3  0  0  0  0  0 -7  0 -1 -1  0 -2
second      0 -1  0  0 -2 -1  2  2 -2  4 -7  3 -1  1  1 -1 -3  3  0  0  0  0 -7  7 -1  0  1 -2
```

Analyze

- The 1^{st}-order derivative is nonzero along the entire ramp, while the 2^{nd}-order derivative is nonzero only at the onset and end of the ramp.

1^{st} make thick edge and 2^{nd} make thin edge

- The response at and around the point is much stronger for the 2^{nd}- than for the 1^{st}-order derivative

The Laplacian (2^{nd} order derivative)

- Shown by Rosenfeld and Kak[1982] that the simplest isotropic derivative operator is the Laplacian is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Discrete form of derivative

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

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

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

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

2-Dimentional Laplacian

The digital implementation of the 2-Dimensional Laplacian is obtained by summing 2 components

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x^2}$$

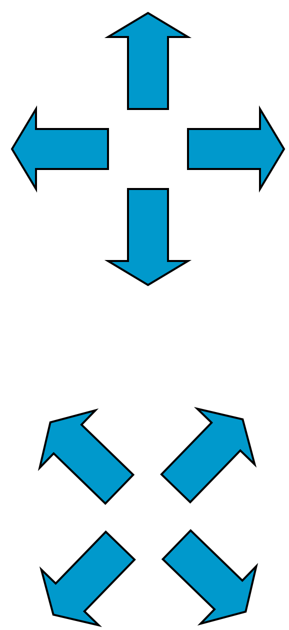
$$\nabla^2 f = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

	1	
1	**-4**	**1**
	1	

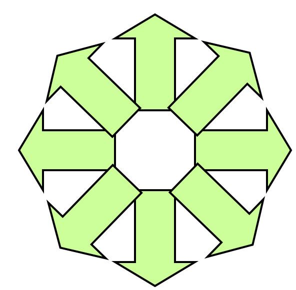
Laplacian

0	1	0
1	-4	1
0	1	0

1	0	1
0	-4	0
1	0	1

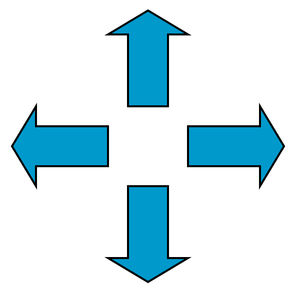


1	1	1
1	-8	1
1	1	1

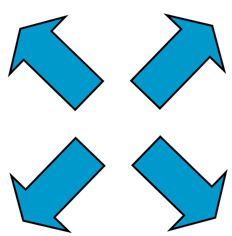


Laplacian

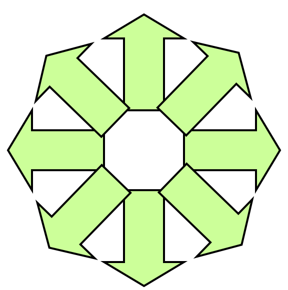
0	-1	0
-1	4	-1
0	-1	0



-1	0	-1
0	4	0
-1	0	-1



-1	-1	-1
-1	8	-1
-1	-1	-1



Implementation

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f(x,y) & \textbf{\textit{If the center coefficient is negative}} \\ f(x,y) + \nabla^2 f(x,y) & \textbf{\textit{If the center coefficient is positive}} \end{cases}$$

Where f(x,y) is the original image

$\nabla^2 f(x,y)$ **is Laplacian filtered image**

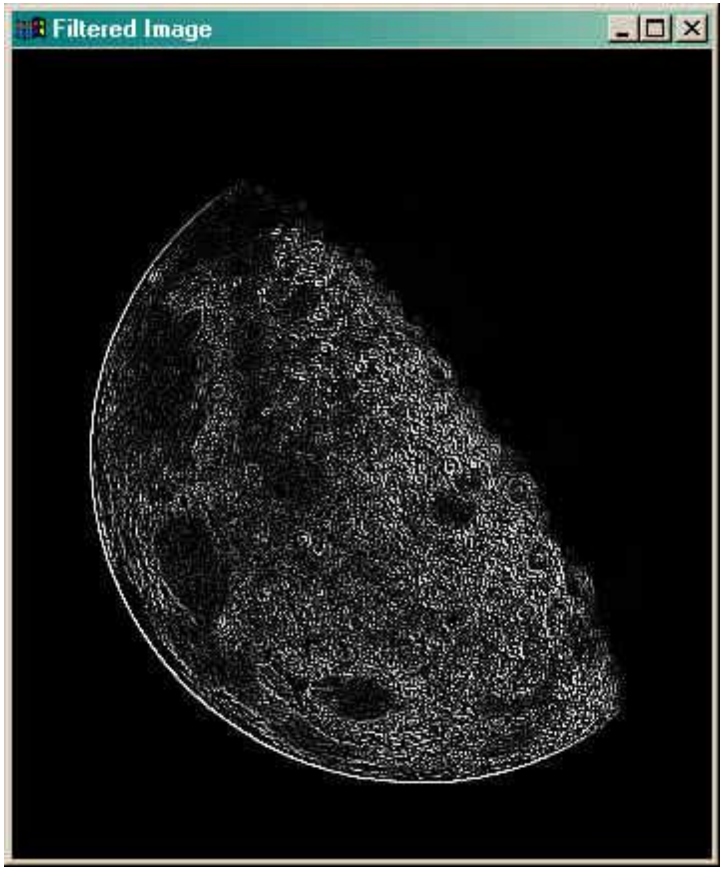
g(x,y) is the sharpen image

Implementation

-1	-1	-1
-1	8	-1
-1	-1	-1

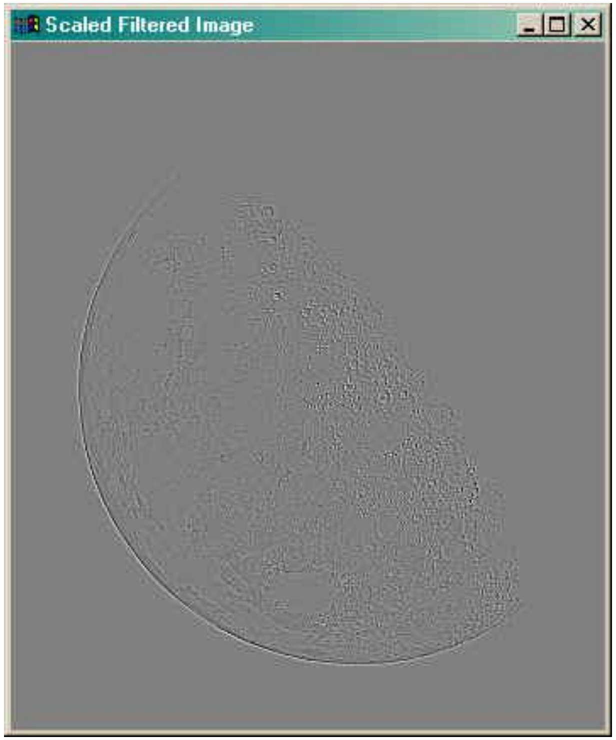
Implementation

Filtered = Conv(image,mask)



Implementation

filtered = filtered - Min(filtered)
filtered = filtered * (255.0/Max(filtered))



Implementation

sharpened = image + filtered
sharpened = sharpened - Min(sharpened)
sharpened = sharpened * (255.0/Max(sharpened))

Algorithm

- Using Laplacian filter to original image
- And then add the image result from step 1 and the original image

Simplification

- We will apply two steps to be one mask

$$g(x,y) = f(x,y) - f(x+1,y) - f(x-1,y) - f(x,y+1) - f(x,y-1) + 4f(x,y)$$

$$g(x,y) = 5f(x,y) - f(x+1,y) - f(x-1,y) - f(x,y+1) - f(x,y-1)$$

0	**-1**	**0**
-1	**5**	**-1**
0	**-1**	**0**

-1	**-1**	**-1**
-1	**9**	**-1**
-1	**-1**	**-1**

Unsharp masking

- A process to sharpen images consists of subtracting a blurred version of an image from the image itself. This process, called *unsharp masking*, is expressed as

$$f_s(x, y) = f(x, y) - \bar{f}(x, y)$$

Where $f_s(x, y)$ **denotes the sharpened image obtained by unsharp masking, and** $\bar{f}(x, y)$ **is a blurred version of** $f(x, y)$

High-boost filtering

- A high-boost filtered image, f_{hb} is defined at any point (x,y) as

$$f_{hb}(x,y) = Af(x,y) - \bar{f}(x,y) \quad where\ A \geq 1$$

$$f_{hb}(x,y) = (A-1)f(x,y) + f(x,y) - \bar{f}(x,y)$$

$$f_{hb}(x,y) = (A-1)f(x,y) + f_s(x,y)$$

This equation is applicable general and does not state explicity how the sharp image is obtained

High-boost filtering and Laplacian

- If we choose to use the Laplacian, then we know $f_s(x,y)$

$$f_{hb} = \begin{cases} Af(x,y) - \nabla^2 f(x,y) & \textit{If the center coefficient is negative} \\ Af(x,y) + \nabla^2 f(x,y) & \textit{If the center coefficient is positive} \end{cases}$$

0	**-1**	**0**
-1	**A+4**	**-1**
0	**-1**	**0**

-1	**-1**	**-1**
-1	**A+8**	**-1**
-1	**-1**	**-1**

The Gradient (1st order derivative)

- First Derivatives in image processing are implemented using the magnitude of the gradient.
- The gradient of function f(x,y) is

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$$

Gradient

- The magnitude of this vector is given by

$$mag(\nabla f) = \sqrt{G_x^2 + G_y^2} \approx |G_x| + |G_y|$$

$\mathbf{G_x}$

-1	**1**

$\mathbf{G_y}$

1
-1

This mask is simple, and no isotropic. Its result only horizontal and vertical.

Robert's Method

- The simplest approximations to a first-order derivative that satisfy the conditions stated in that section are

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

$G_x = (z_9 - z_5)$ and $G_y = (z_8 - z_6)$

$$\nabla f = \sqrt{(z_9 - z_5)^2 + (z_8 - z_6)^2}$$

$$\nabla f \approx |z_9 - z_5| + |z_8 - z_6|$$

Robert's Method

- These mask are referred to as the Roberts cross-gradient operators.

-1	**0**
0	**1**

0	**-1**
1	**0**

Sobel's Method

- Mask of even size are awkward to apply.
- The smallest filter mask should be 3x3.
- The difference between the third and first rows of the 3x3 Image region approximate derivative in x-direction, and the difference between the third and first column approximate derivative in y-direction.

Sobel's Method

- Using this equation

$$\nabla f \approx \left|(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)\right| + \left|(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)\right|$$

-1	**-2**	**-1**
0	**0**	**0**
1	**2**	**1**

-1	**0**	**1**
-2	**0**	**2**
-1	**0**	**1**

Video

Lesson-6

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

To discuss…

- ✓ Types of video signals
- ✓ Analog Video
- ✓ Digital Video

Types of Video Signals

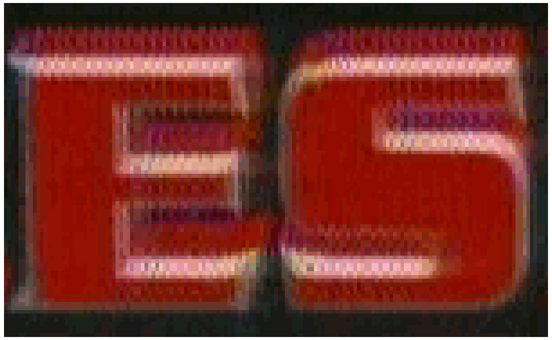
Video Signals can be classified as

1. Composite Video
2. S-Video
3. Component Video

Types - Composite Video

- Used in broadcast TV's
- Compatible with B/W TV
- Chrominance (I & Q or U & V) & Luminance signals are mixed into a single carrier wave, which can be separated at the receiver end
- Mixing of signals leads interference & create ***dot crawl***

Male F-Connector, Connecting co-axial cable with the device



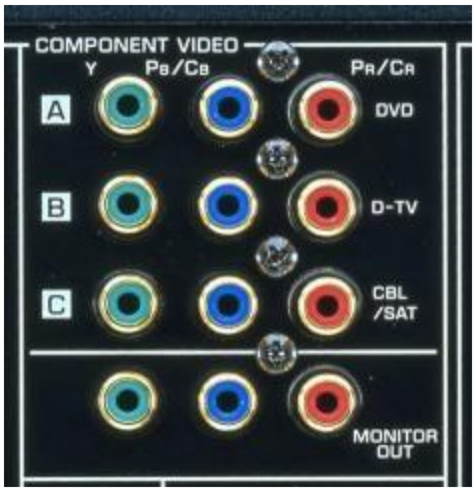
Dot Crawl, due to interference in composite video

Types - S-Video

- S stands *Super / Separated*
- **Uses 2 wires**, one for *luminance* & the other for *chrominance* signals
- Humans are able to differentiate spatial resolution in gray-scale images with a much higher acuity than for the color part of color images.
- As a result, we can send less accurate color information than must be sent for intensity information

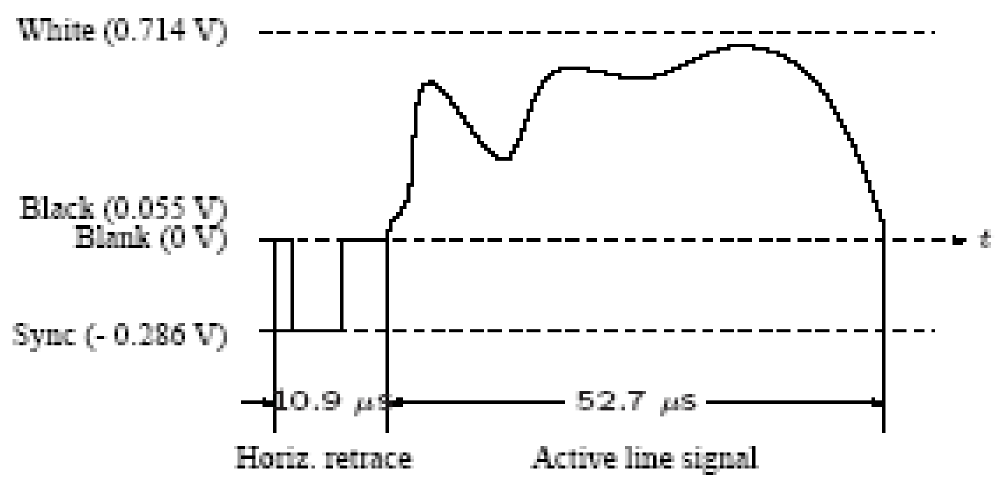
Types - Component Video

- Each primary is sent as a separate video signal.
  - The primaries can either be RGB or a luminance-chrominance transformation of them (e.g., YIQ, YUV).
  - Best color reproduction
  - Requires more bandwidth and good synchronization of the three components



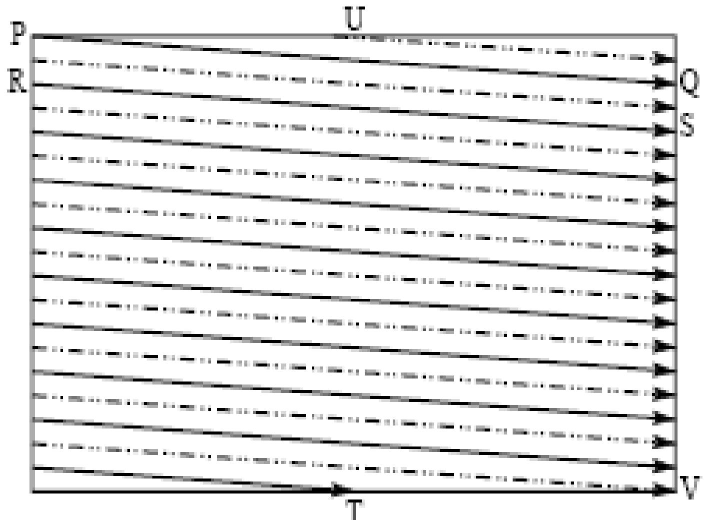
Analog Video

- Represented as a continuous (time varying) signal



Analog Video [2]

With interlaced scan, the odd and even lines are displaced in time from each other.



Interlaced Scan

Analog Video [3]

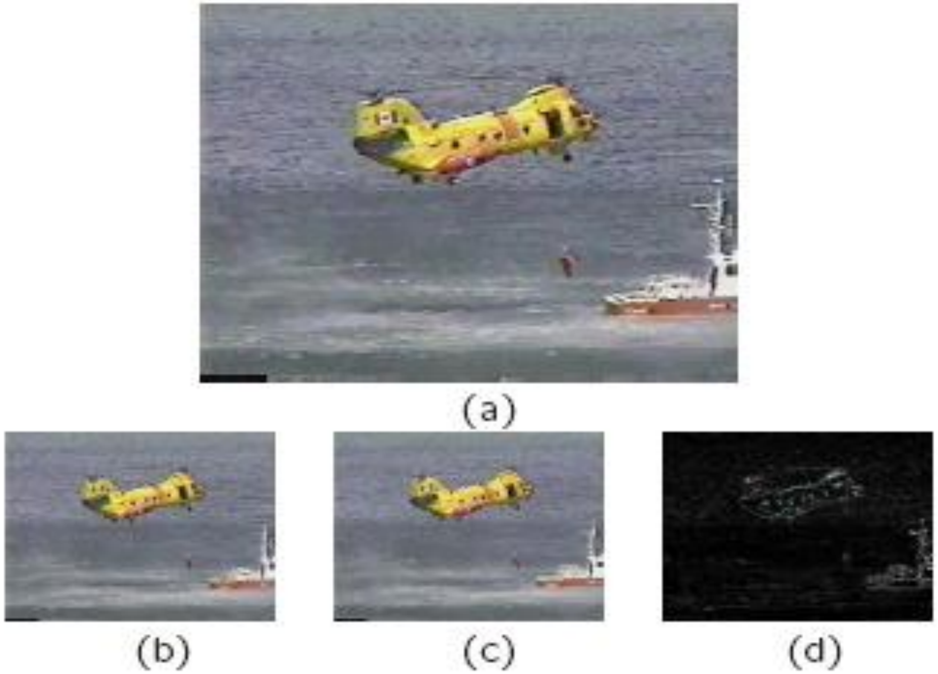


Fig. 5.2: Interlaced scan produces two fields for each frame. (a) The video frame, (b) Field 1, (c) Field 2, (d) Difference of Fields

NTSC (National Television System Committee)

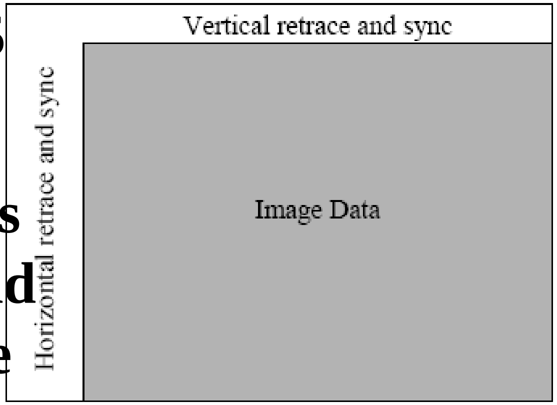
- It uses the familiar **4:3 aspect ratio** (i.e., the ratio of picture width to its height)
- Uses **525 scan lines per frame** at **30 frames per second** (fps).
- NTSC follows the **interlaced scanning system**, and each frame is divided into two fields, with 262.5 lines/field.
- Thus the horizontal sweep frequency is 525x 29.97 =15,734 lines/sec, so that **each line is swept out in 63.6 μ sec**
  $(1/15.734 \times 10^3$ sec $)$
- **63.6 μ sec = 10.9 μ sec for Horizontal retrace + 52.7 μ sec active line signal**
- For the active line signal during which image data is displayed

NTSC (National Television System Committee) [2]

- 20 lines at the beginning of every field is for Vertical retrace control information leaving 485 lines per frame

- 1/6 of the raster at the left side is blanked for horizontal retrace and sync. The non-blanking pixels are called active pixels.

- Pixels often fall in-between the scan lines. NTSC TV is only capable of showing about 340 (visually distinct) lines



Video raster, including retrace and sync data

NTSC (National Television System Committee) [3]

- NTSC video is an analog signal with **no fixed horizontal resolution**
- Pixel clock is used to divide each horizontal line of video into samples. Different video formats provide different numbers of samples per line
- **Uses YIQ Color Model**
- Quadrature Modulation is used to combine I & Q to produce a single chroma signal

$$C = I\cos(F_{sc}t) + Q\sin(F_{sc}t)$$

NTSC (National Television System Committee) [4]

- Fsc is 3.58MHz
- Composite signal is formed by

$$\text{composite} = Y + C = Y + I\cos(F_{sc}t) + Q\sin(F_{sc}t)$$

- The available bandwidth is 6MHz, in which the audio is signal centered at 5.75MHz and the lower spectrum carries picture information

PAL (Phase Alternating Line)

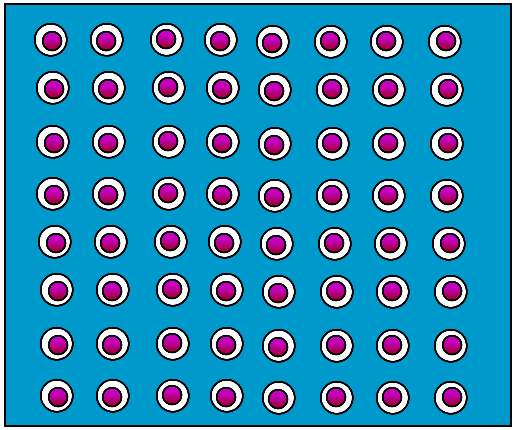
- Widely used in Western Europe, China, India, and many other parts of the world.
- Uses **625 scan lines per frame**, at 25 frames/second, with a **4:3 aspect ratio** and **interlaced fields**
- Uses the **YUV color model**
- Uses an 8 MHz channel and allocates a bandwidth of 5.5 MHz to Y, and 1.8 MHz each to U and V.

Digital Video

- *Advantages over analog:*
  – Direct random access --> good for nonlinear video editing
  – No problem for repeated recording
  – No need for blanking and sync pulse
- Almost all digital video uses *component video*

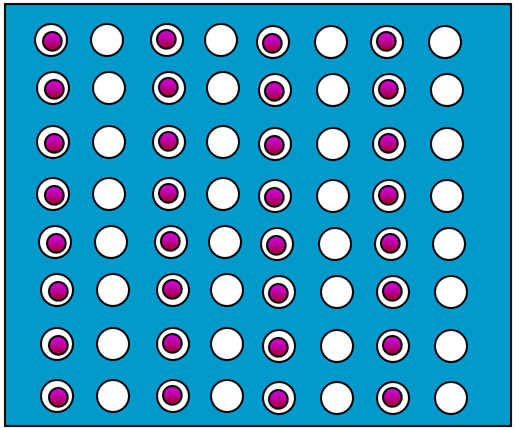
Chroma Subsampling

- The human eye responds more precisely to brightness information than it does to color, *chroma subsampling* (*decimating*) takes advantage of this.
  - In a 4:4:4 scheme, each 8×8 matrix of RGB pixels converts to three YCrCb 8×8 matrices: one for luminance (Y) and one for each of the two chrominance bands (Cr and Cb)



Chroma Subsampling [2]

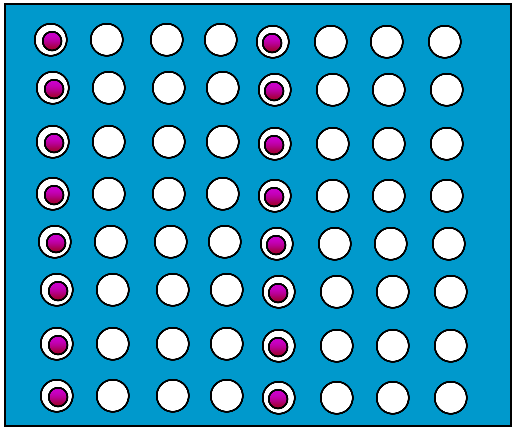
– A 4:2:2 scheme also creates one 8×8 luminance matrix but decimates every two horizontal pixels to create each chrominance-matrix entry. Thus reducing the amount of data to 2/3rds of a 4:4:4 scheme.



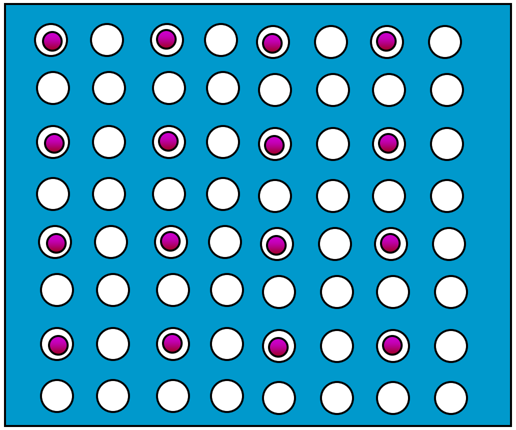
4 : 2 : 2

Chroma Subsampling [3]

– Ratios of 4:2:0 decimate chrominance both horizontally and vertically, resulting in four Y, one Cr, and one Cb 8×8 matrix for every four 8×8 pixel-matrix sources. This conversion creates half the data required in a 4:4:4 chroma ratio

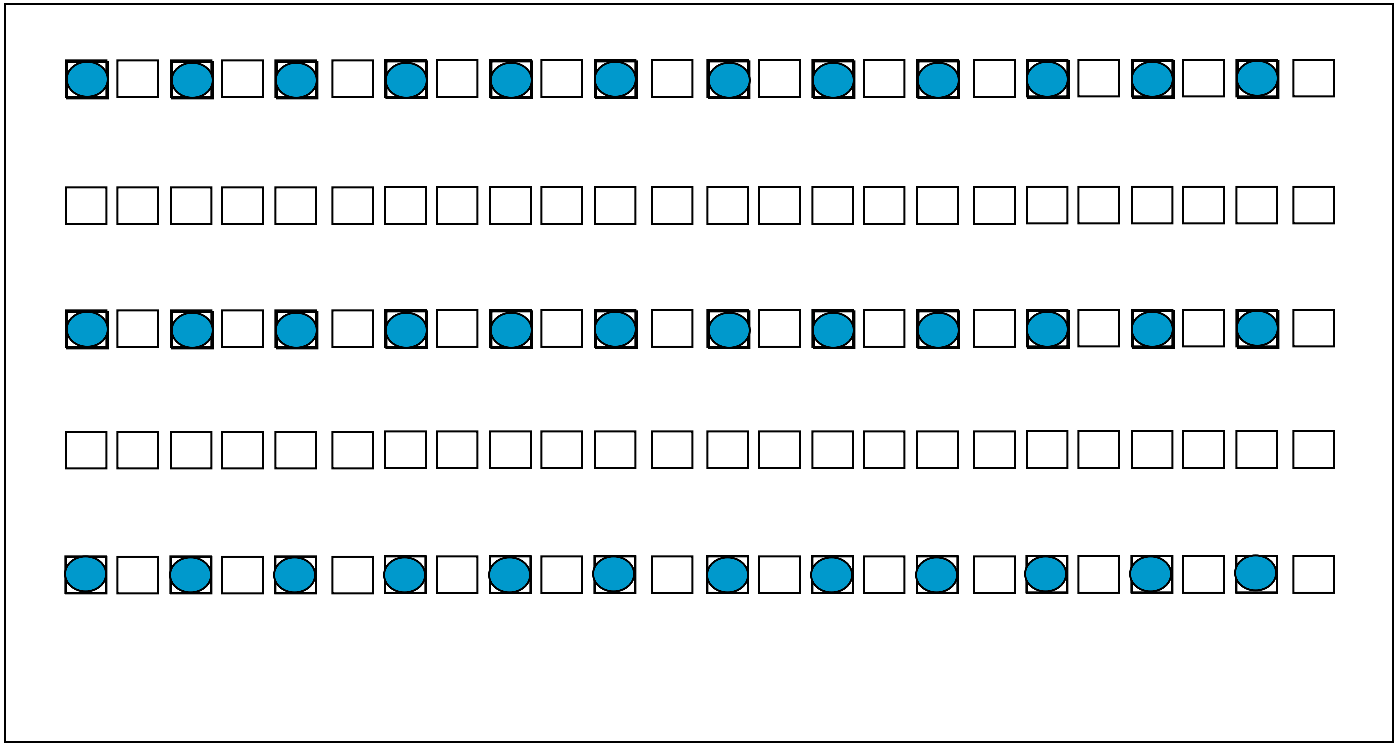


4 : 1 : 1



4 : 2 : 0

Chroma Subsampling [4]



Luma sample

Chroma sample

4:2:0

Chroma Subsampling [5]

- The 4:1:1 and 4:2:0 are used in JPEG and MPEG
- 256-level gray-scale JPEG images aren't usually much smaller than their 24-bit color counterparts, because most JPEG implementations aggressively subsample the color information. Color data therefore represents a small percentage of the total file size

High Definition TV (HDTV)

- The main thrust of **HDTV** (High Definition TV) is not to increase the definition in each unit area, but rather to increase the visual field especially in its width.
  - The first generation of HDTV was based on an analog technology developed by Sony and NHK in Japan in the late 1970s.
  - Uncompressed HDTV will demand more than 20 MHz bandwidth, which will not fit in the current 6 MHz or 8 MHz channels
  - More than one channels even after compression.

High Definition TV (HDTV) [3]

- The salient difference between conventional TV and HDTV:
  – HDTV has a much wider aspect ratio of 16:9 instead of 4:3.
  – HDTV moves toward progressive (non-interlaced) scan. The rationale is that interlacing introduces serrated edges to moving objects and flickers along horizontal edges.

AUDIO

Lesson-7

Assistant Prof. Dr. Ayad A. Salam
College of Education for women
University of Baghdad

To discuss…

- What is sound?
  – Waveforms and attributes of sound
- Capturing digital audio
  – Sampling
- MIDI (Musical Instrument Digital Interface)

What is Sound?

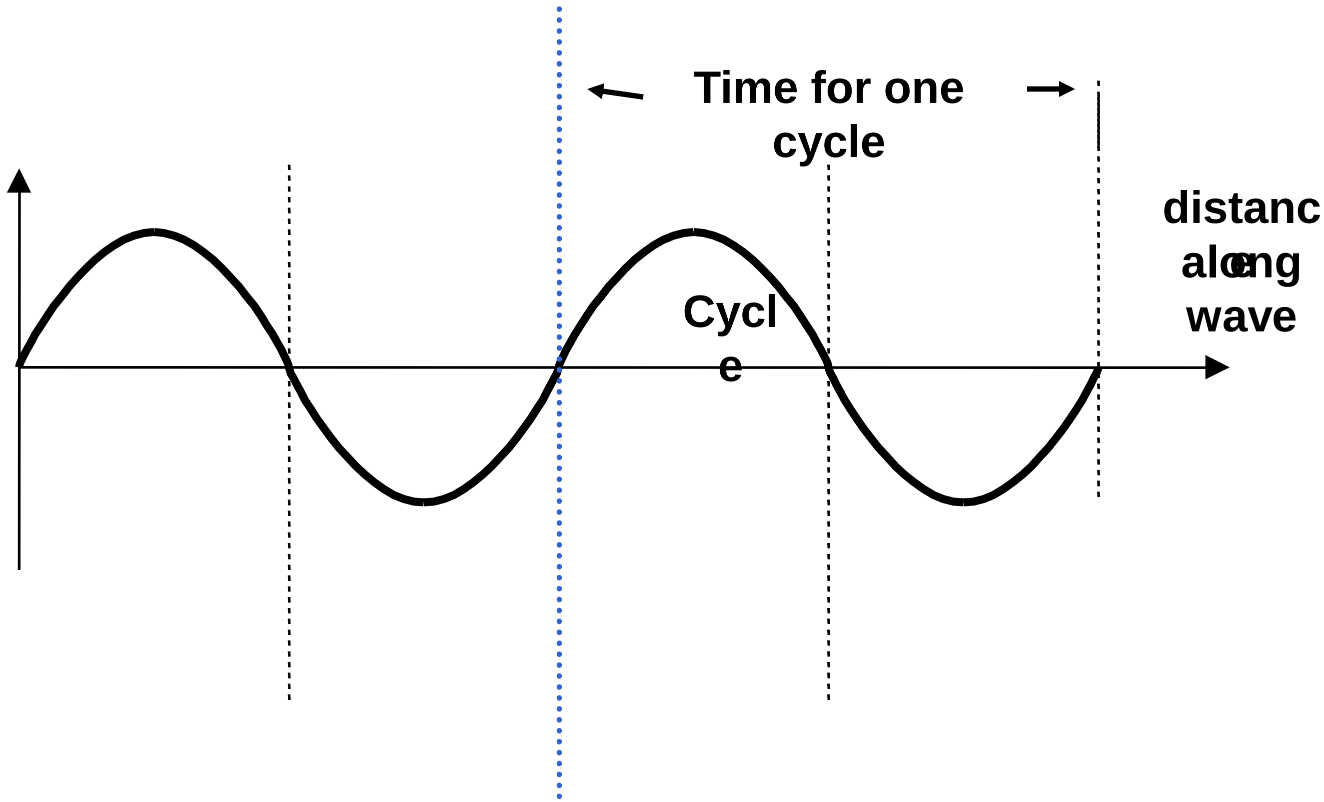
- Sound comprises the spoken word, voices, music and even noise.
- **Sound is a pressure wave, taking continuous values**
- It is a complex relationship involving a **vibrating object** (sound source), a **transmission medium** (usually air), a **receiver** (ear) and a **perceptor** (brain). Example banging drum.

- Increase / decrease in pressure can be measured in amplitude, which can be digitized
- Measure the amplitude at equally spaced time intervals (*sampling*) and represent it with one of finite digital values (*quantization*)
- Sampling frequency refers the rate at which the sampling is performed

Waveforms

- Sound waves are manifest as **waveforms**
  - A waveform that repeats itself at regular intervals is called a **periodic waveform**
  - Waveforms that do not exhibit regularity are called **noise**
- The unit of regularity is called a **cycle**
  - This is known as **Hertz** (or **Hz**) after Heinrich Hertz
    - One cycle = 1 Hz
    - Sometimes written as **kHz** or **kiloHertz** (1 kHz = 1000 Hz)

Waveforms



The characteristics of sound waves

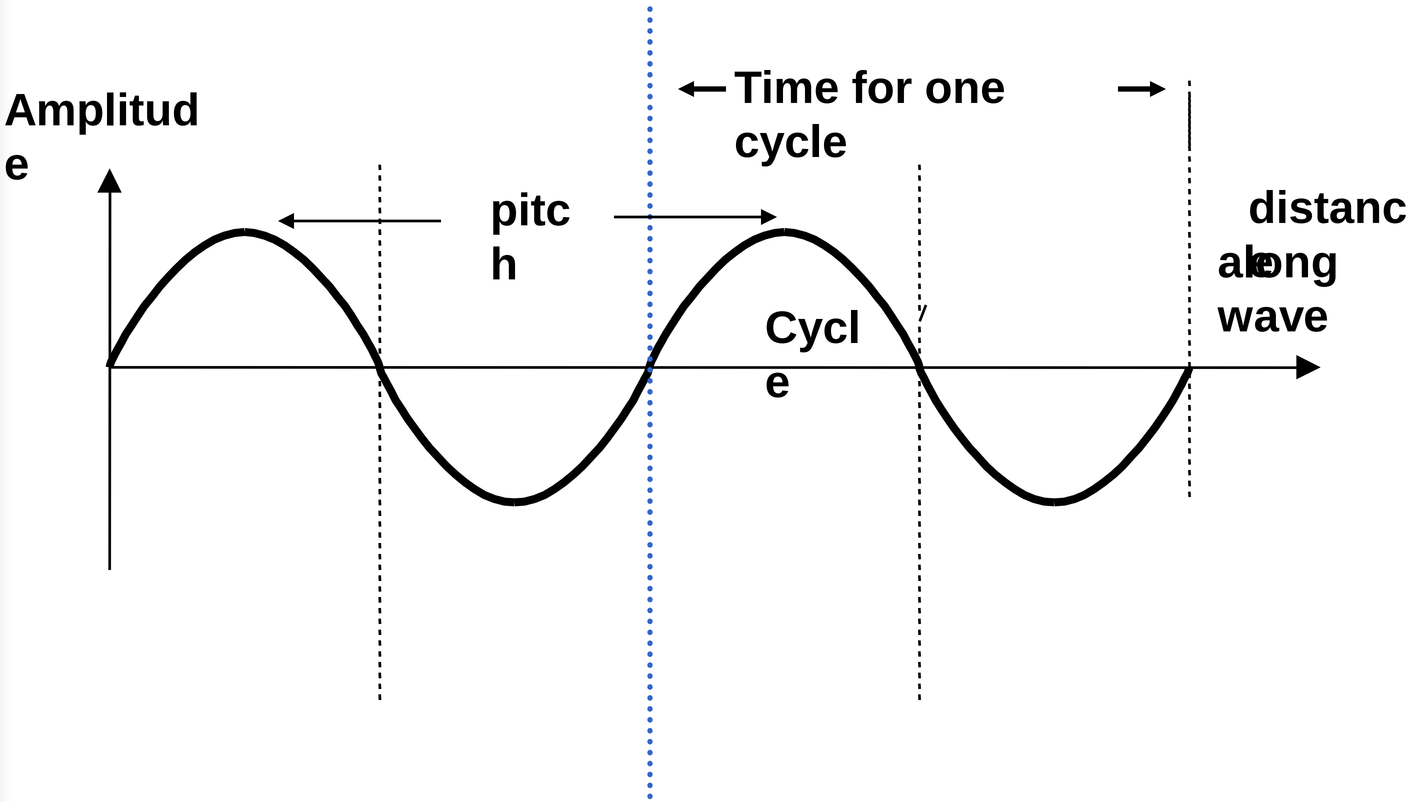
- Sound is described in terms of two characteristics:
  - **Frequency**
  - **Amplitude** (or loudness)

- **Frequency**
  - the rate at which sound is measured
  - Number of cycles per second or Hertz (Hz)
  - Determines the **pitch** of the sound as heard by our ears
  - The higher frequency, the clearer and sharper the sound → the higher pitch of sound

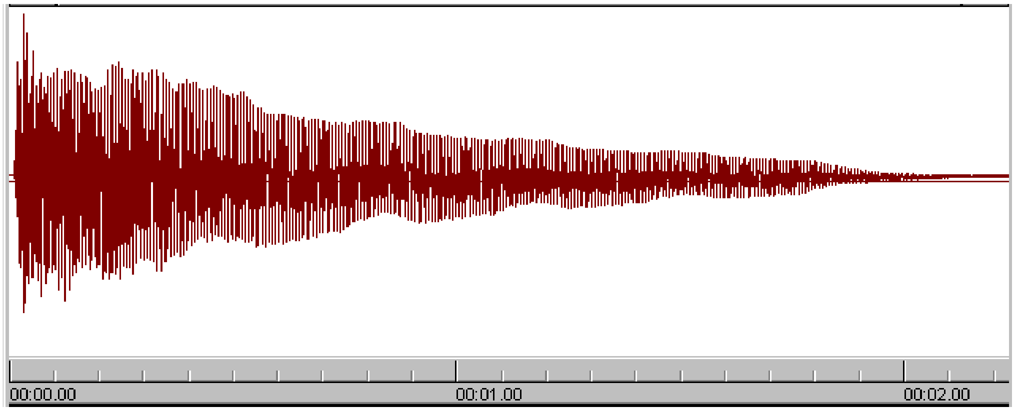
The characteristics of sound waves

- **Amplitude**
  - Sound’s intensity or loudness
  - The louder the sound, the larger amplitude.

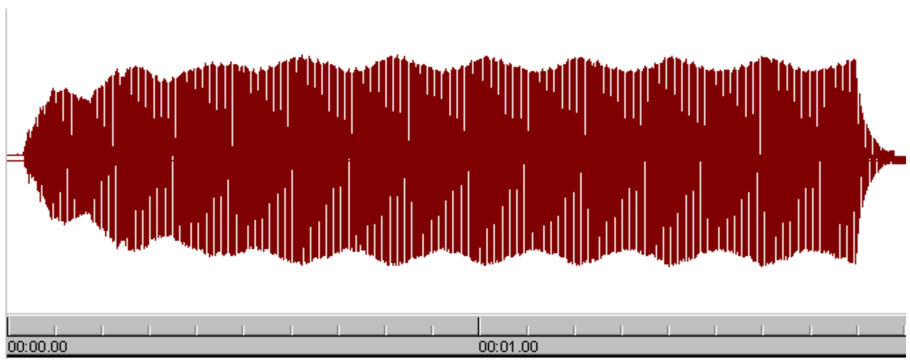
The characteristics of sound waves



Example waveforms

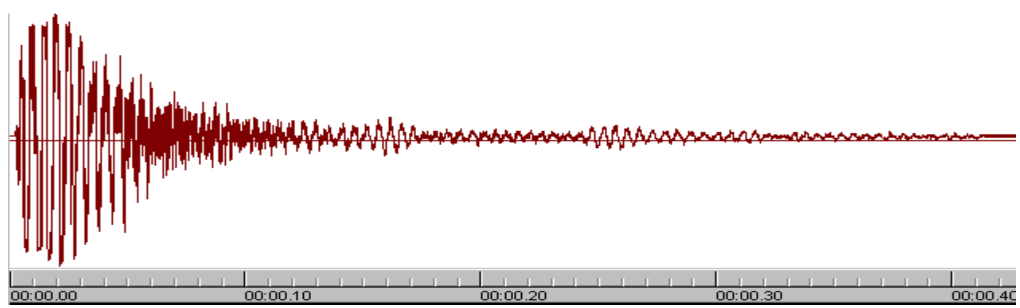


Piano

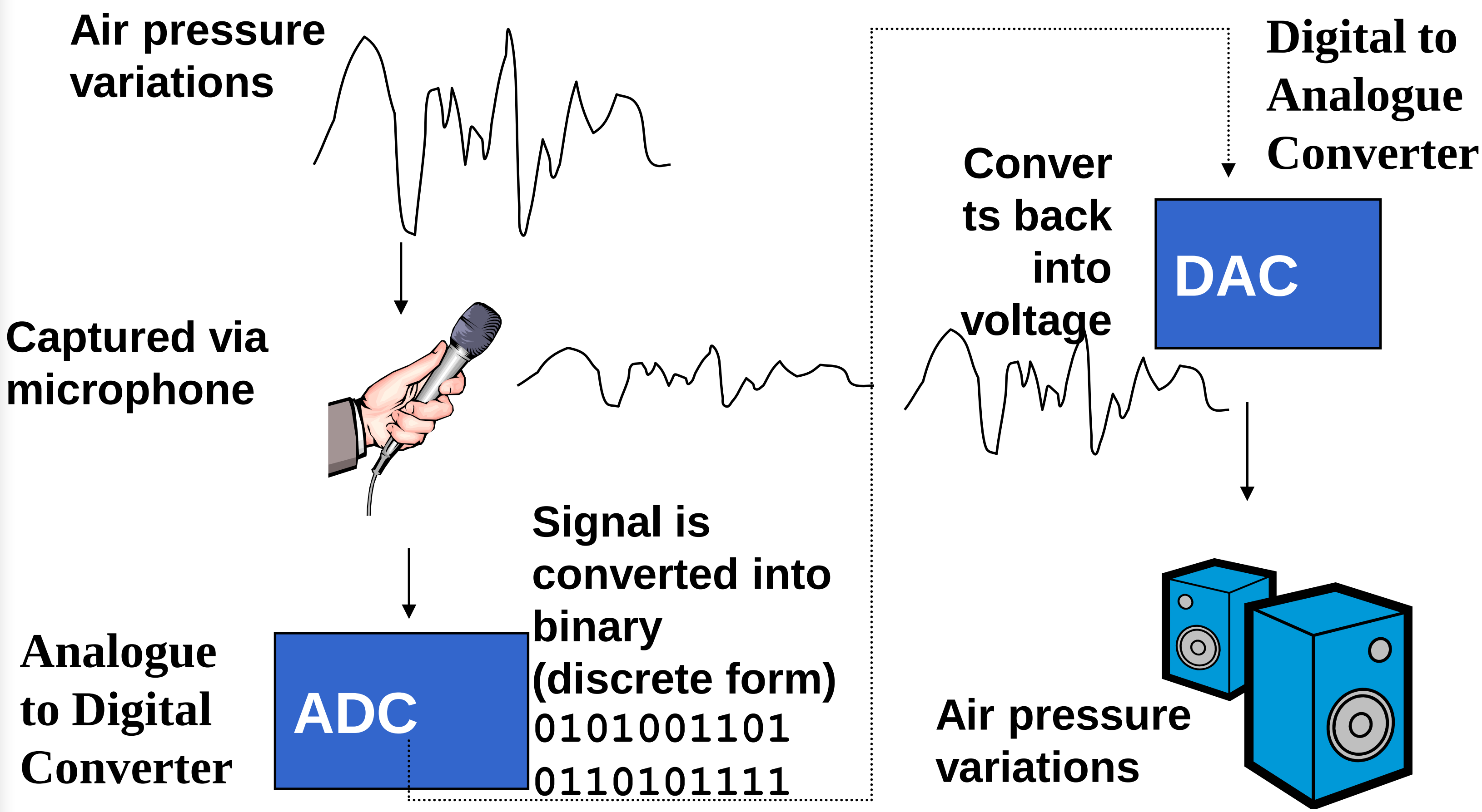


Pan flute

Snare drum



Capture and playback of digital audio



The Analogue to Digital Converter (ADC)

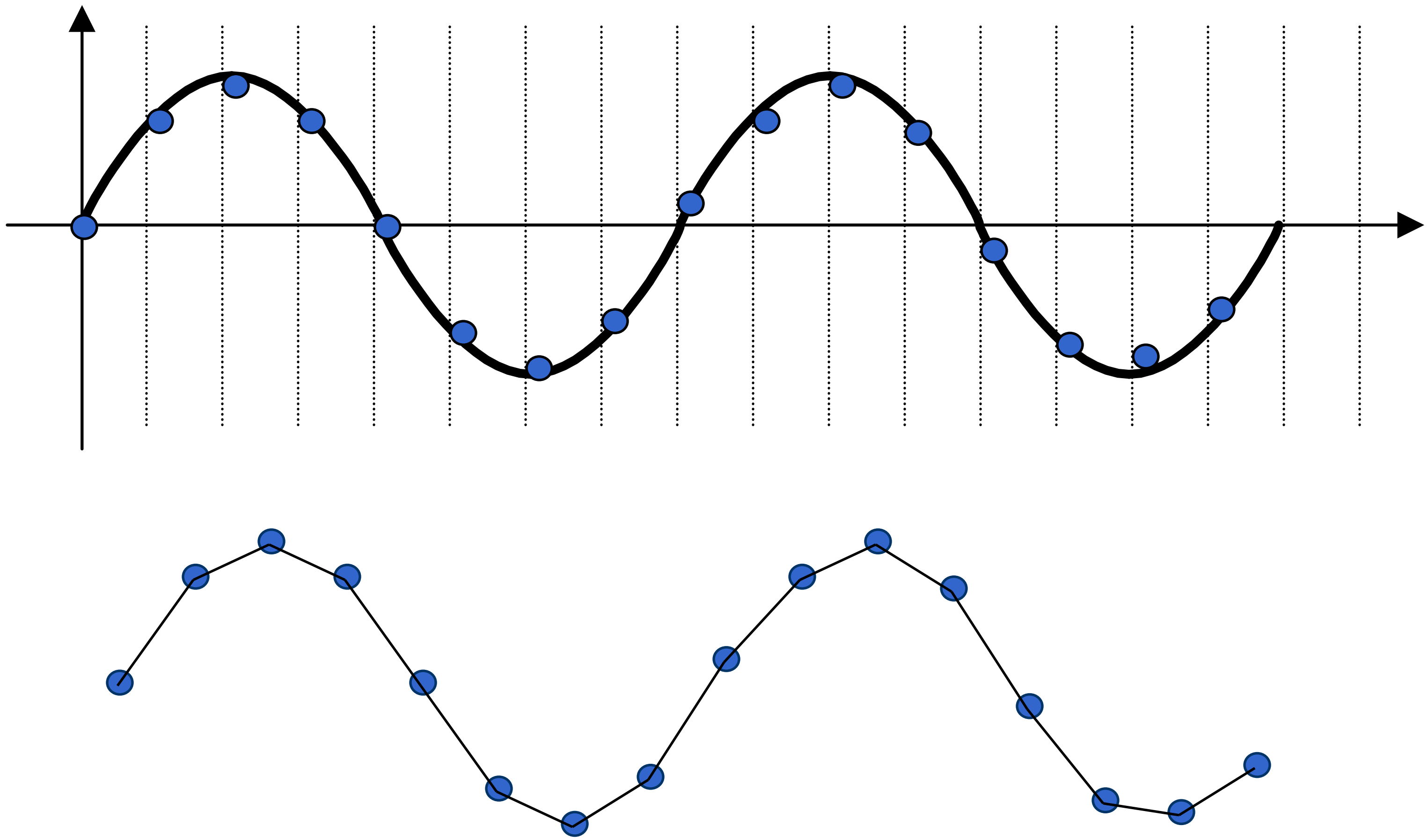
- An **ADC** is a device that converts analogue signals into digital signals
- An analogue signal is a **continuous value**
  - It can have any single value on an infinite scale
- A digital signal is a **discrete value**
  - It has a finite value (usually an integer)
- An ADC is synchronised to some **clock**

The Analogue to Digital Converter (ADC)

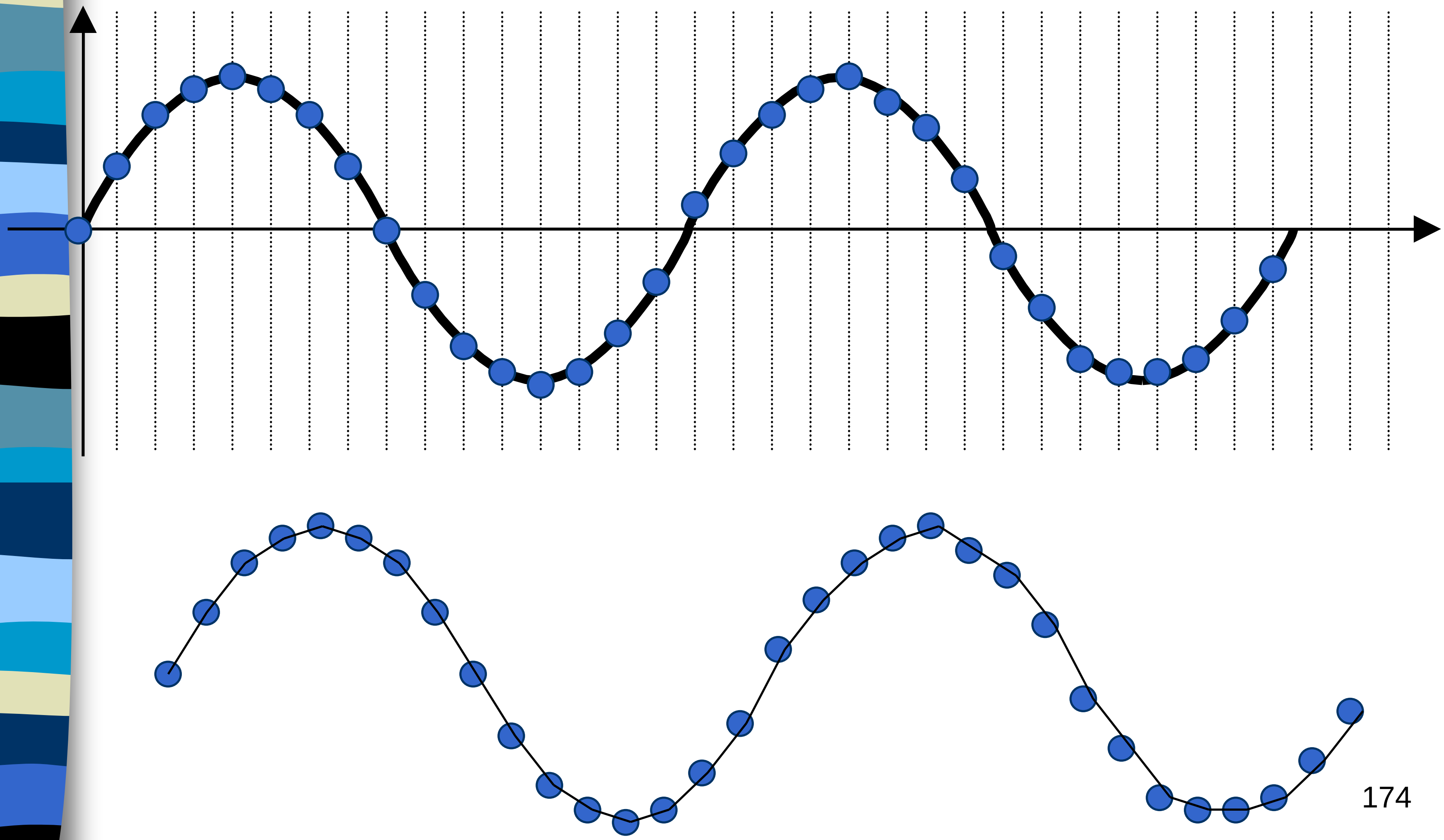
- It will monitor the continuous analogue signal at a set **rate** and convert what it sees into a discrete value at that specific moment in time

- The process to convert the analogue to digital sound is called **Sampling**. Use PCM (Pulse Code Modulation)

Digital sampling
Sampling frequency



Digital sampling
Sampling frequency



Sampling

- Two parameters:

  Sampling Rate

  - Frequency of sampling
  - Measure in Hertz
  - The higher sampling rate, higher quality sound but size storage is big.
  - Standard Sampling rate:

    - 44.1 KHz for CD Audio

    - 22.05 KHz

    - 11.025 KHz for spoken

    - 5.1025 KHz for audio effect

Sampling

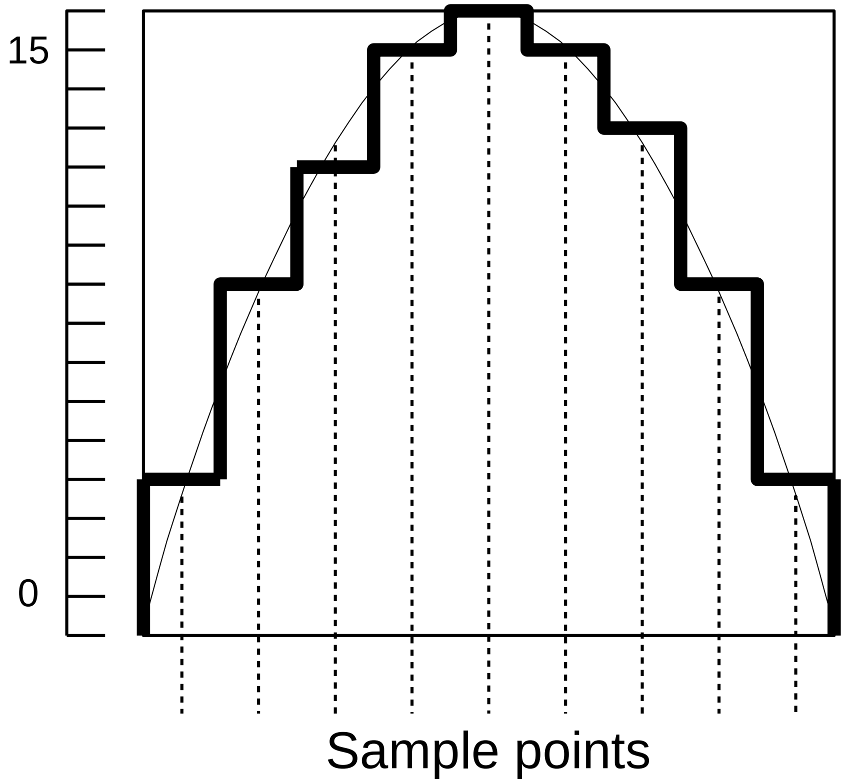
Sample Size

The **resolution** of a sample is the number of bits it uses to store a given **amplitude value**, e.g.

- 8 bits (256 different values)
- 16 bits (65536 different values)
- A higher resolution will give higher quality but will require more memory (or disk storage)

Quantisation

- Samples are usually represented the audio sample as a integers(discrete number) or digital



Calculating the size of digital audio

- The formula is as follows:

$$\frac{rate \times duration \times resolution \times number\ of\ channels}{8}$$

- The answer will be in bytes
- Where:
  - sampling rate is in *Hz*
  - Duration/time is in *seconds*
  - resolution is in *bits* (1 for 8 bits, 2 for 16 bits)
  - number of channels = 1 for mono, 2 for stereo, etc.

Calculating the size of digital audio

- **Example:**

Calculate the file size for 1 minute, 44.1 KHz, 16 bits, stereo sound

$$\frac{rate \times duration \times resolution \times number\ of\ channels}{8}$$

- **Where:**
  - **sampling rate is 44,100 Hz**
  - **Duration/time is 60 seconds**
  - **resolution is 16 bits**
  - **number of channels for stereo is 2**

44100 * 60 * 16 *2 / 8

Mono Recording

- Mono simply indicates the use of a single channel. Mono includes the use of a single microphone used to record a sound, which is then played through a single channel through a speaker.
- The easiest way to check if a sound is a mono recording is through a set of headphones, incidentally, you can easily distinguish whether or not the sound plays through one headphone and not the other.
- Mono recording was typically used before the development of stereo recording.

Mono Recording

Advantages

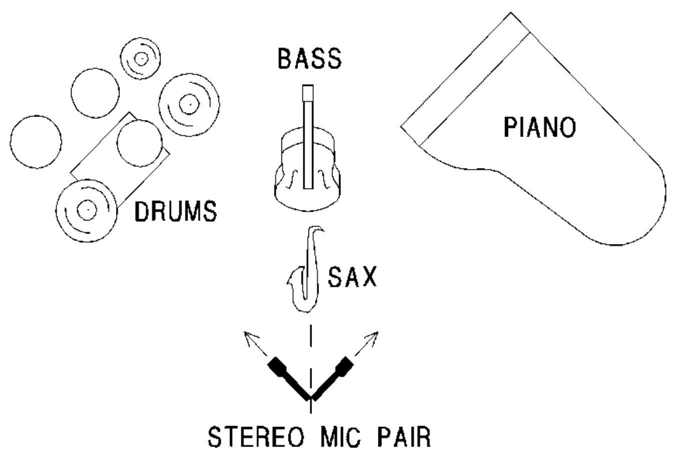
- Mono file sizes are around half the size of their stereo counterparts
- High resolution files can be recorded with relatively low file sizes
- High resolution files can be recorded around the same size file as a low resolution stereo file
- It is much easier to mix mono sounds than it is with stereo
- Mono sounds are much easier to manipulate in editing programmes
- Everyone hears the exact same signal
- Mono systems are suitable for reinforcing a sound

Disadvantages

- There is no sound perspective
- It is impossible to tell whether or not a sound has been recorded from a distance
- Films that use mono sounds do not provide as much as an impact than if the film was recorded using stereo sounds

Stereo Recording

- Whereas mono recording has one independent audio channel, stereo has two channels.
- Signals that are reproduced through stereo recording have an exact correlation with each other, so when the sound is played back through either speakers or headphones, the sound is a mirrored representation of the original recording.
- Stereo recording would be useful in situations that require the use of sound perspective, for instance the clear location of instruments on a stage.
- The stereo systems must have an equal cover over the two audio channels.



Stereo Recording

Advantages

- Provides sound perspective
- Gives an idea of the direction the sound is coming from, or how it has been recorded
- Provides better experience when listening to songs or films
- It is possible to tell whether or not the sound has been recorded from a distance
- Offers the possibility of multi-track recordings

Disadvantages

- Since stereo files use two audio channels instead of one, the files sizes are going to be a lot bigger
- High resolution stereo files are relatively big files
- Mono sound files can be recorded at high resolutions for half the file sizes of stereo files
- Stereo files are harder to edit than mono files as there are two channels to work with
- The sound is played equally over two channels, therefore if one channel is broken, the sound quality is not only played through one speaker, it is also halved or some of the audio is missed out altogether
- Stereo is a lot more expensive to set up

Audio formats

- Depend on O/S. For examples:
  - AIFF (Audio Interchange File Format)
  - SOU
    - For Macintosh
  - .WAV
    - Waveform file format. For Windows/Microsoft
  - .VOC
    - Sound Blaster Card

What is WAV?

- WAV is an audio file format that was developed by Microsoft.
- It is so wide spread today that it is called a standard PC audio file format.
- A Wave file is identified by a file name extension of WAV (.wav).
- Used primarily in PCs, the Wave file format has been accepted as a viable interchange medium for other computer platforms, such as Macintosh.

What is WAV?

- This allows content developers to freely move audio files between platforms for processing,
- For example. The Wave file format stores information about
  - the file's number of tracks (mono or stereo),
  - sample rate
  - bit depth

MIDI (Musical Instrument Digital Interface)

- **MIDI** is a standard for specifying a musical performance
- Rather than send raw digital audio, it sends **instructions** to musical instruments telling them what note to play, at what volume, using what sound, etc.
- The **synthesiser** that receives the MIDI events is responsible for generating the actual sounds. Example: Keyboard Piano

MIDI Versus Wav

- Quality recording, MIDI depend to the tools
- Audio .wav easier to create compare than MIDI

- **MIDI Advantages**
  - Small File Size
  - Size Storage also small

Advantages and Disadvantages of using audio

Sound adds life to any multimedia application and plays important role in effective marketing presentations.

- **Advantages**
  - Ensure important information is noticed
  - Add interest
  - Can communicate more directly than other media
- **Disadvantages**
  - Easily overused
  - Requires special equipment for quality production
  - Not as memorable as visual media