Dr. Amal Sufuih Ajrash

Multimedia Syllabus

- Introduction to Multimedia.
- Graphics and Image Data Representation.
- Fundamentals Concept in video.
- Basic of Digital audio.
- Multimedia Animation.
- Multimedia and Network.

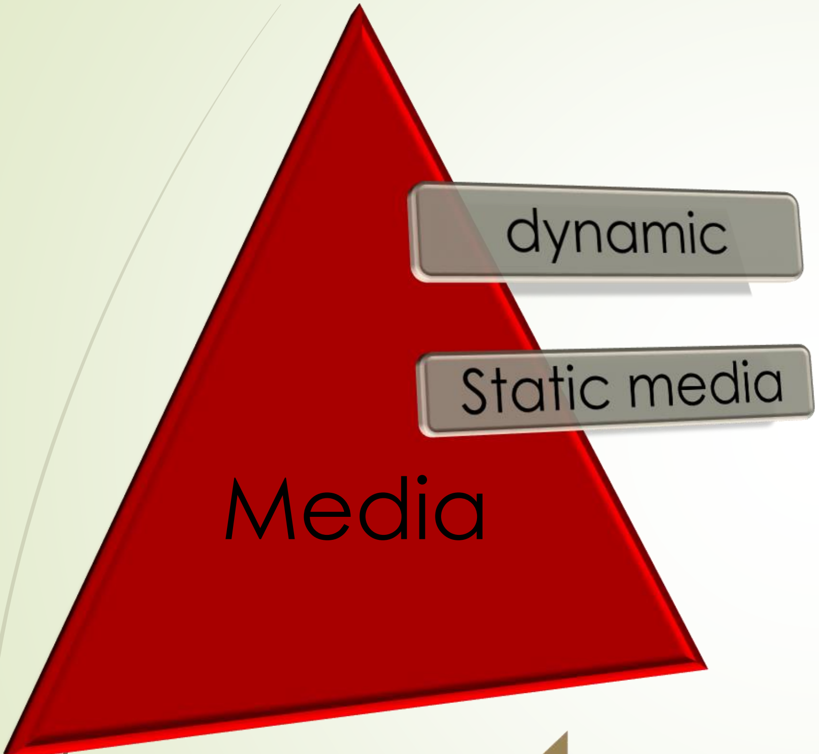
Reference

- Digital Multimedia, Nigel Chapman, Jenny Chapman john wiley & sons, Ltd., 2007.
- Fundamentals of Multimedia, Second Edition, Li,Drew & Liu, 2004.

Multimedia

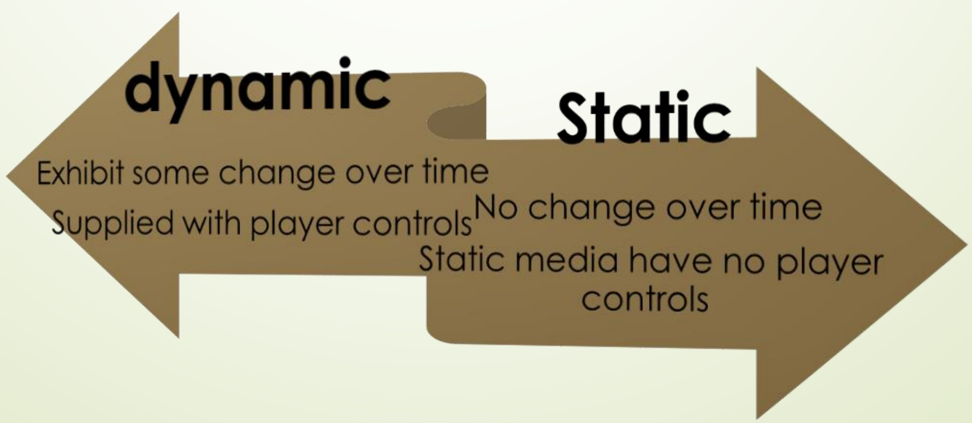
- It can be useful to distinguish between multimedia and multiple media. The distinction is best understood from the point of view of the user. It is common place that we perceive different media in different ways: we just referred to reading text, looking at pictures, listening to sound, and so on. Cognitive scientists call these different sorts of perception modalities. True multimedia requires us to combine modalities.
- Digital multimedia is any combination of two or more media, represented in a digital form, sufficiently well integrated to be presented via a single interface, or manipulated by a single computer program.

Media



Dynamic Media
Video
Animation
sound

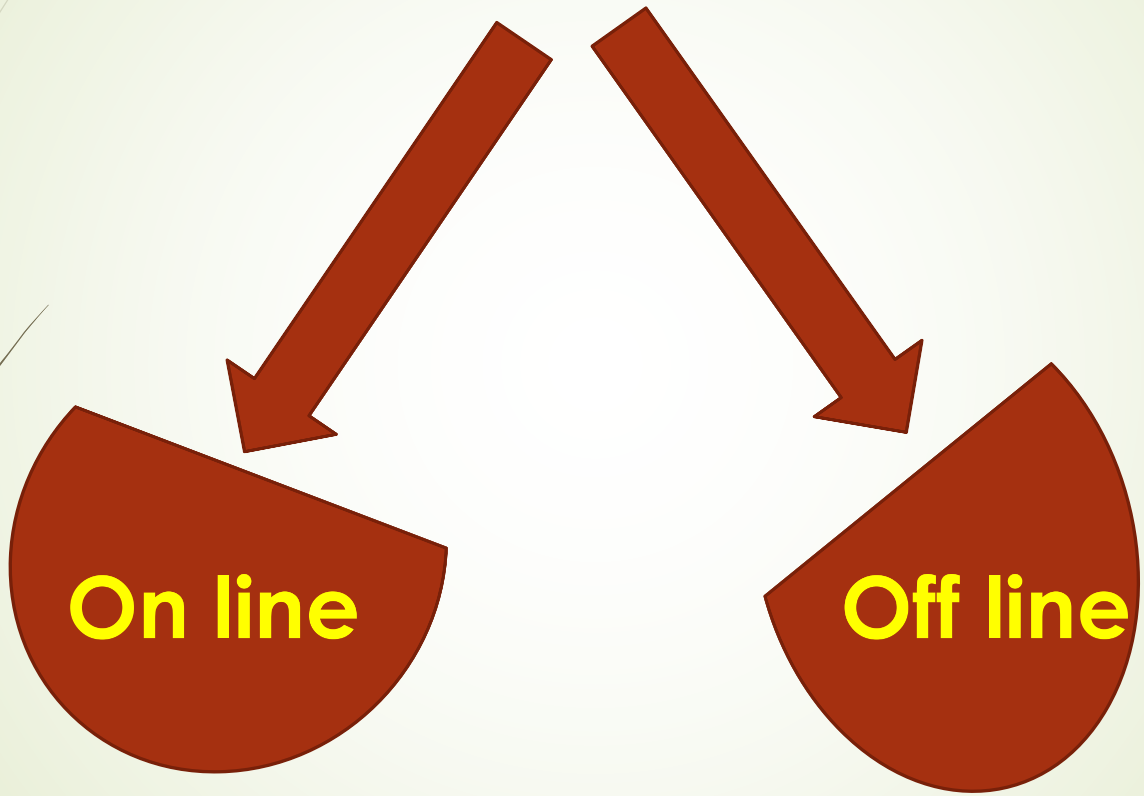
Static Media
Image
Text



Multimedia

- Multimedia combines text, photographic images, video, animations and sound to provide the means to visualize and interact with information in a meaningful way.
- The most familiar example of digital multimedia is the world wide web which combines the different media in a network of pages connected by links. Also magazines, newspapers and books combine text and images, TV combines Sound , image, video and text.
- There are two models currently in use for combining elements of different media types:
- Page-based: In the page_based model, text, images and video are laid out in books and magazines.
- Time-based: The time_ based elements such as video clips and sound are embedded in the page as if they were, images, occupying a fixed area; controls may be provided within the page to start and stop their playback.

Delivery of multimedia



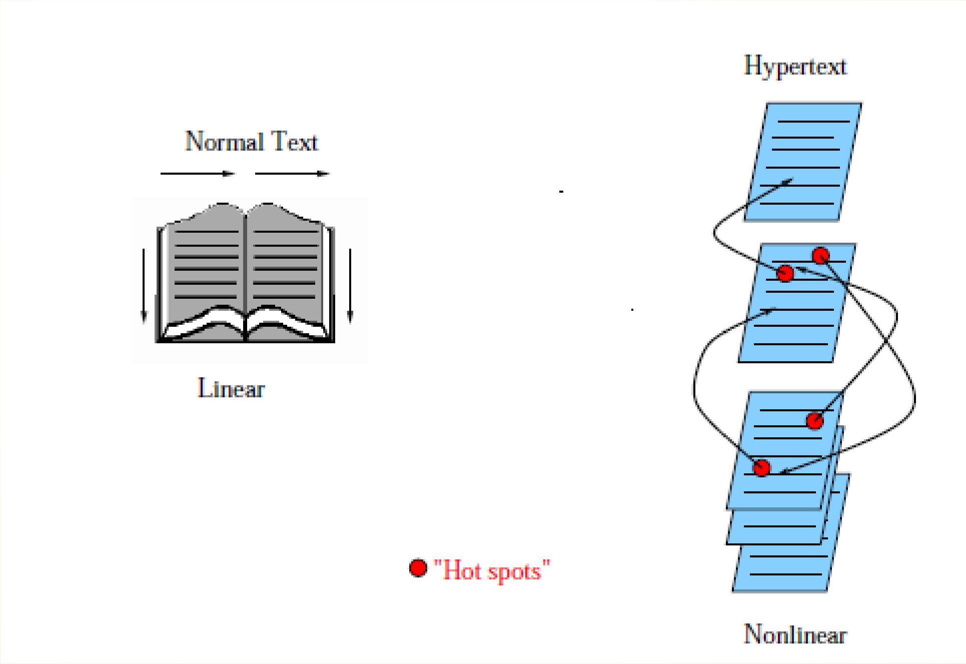
Delivery of multimedia

- Online delivery uses a network to send information from one computer to another. The network may be LAN or WAN (internet).
- Offline delivery used removable storage, CDs, etc.

Authoring *Systems*

- The making of multimedia requires software not only for the preparation of individual media elements, but for their integration into a finished production. Programs that allow a designer to assemble different media elements in space and time, and add interactive behavior to them, are usually called *authoring systems*

Linear and Non-linear Multimedia



Hypertext and Hypermedia

- Individual pages can be combined using links, allowing the user to jump from one page to another by clicking on a representation of the link.
- Linked pages of text are called Hypertext.
- When other media can be embedded in the linked pages the combination is called Hypermedia

Disadvantages of Multimedia

- High monetary cost
  - May require reformatting of information
  - Requires high amount of disk space
  - Requires powerful computer
  - Large amount of varied hardware is necessary
    - CDROM drive
    - Sound card
    - Speakers
    - Scanner
- Technical expertise needed to *set up* multimedia system
- Encourages reliance (يشجع الأتكال)on technology
  - Users may begin to rely on the presence of technology
  - Can be seen as *only* source of information

Digitization of multimedia data

- In multimedia, we encounter values of several kinds that change continuously, either because they originated in physical phenomena or because they exist in some analogue representation .
- For example,
- The amplitude of (volume) of a sound wave varies continuously over time, as does the amplitude of an electrical signal produced by a microphone in response to time wave.
- The color of the image formatted inside a camera by its lens varies continuously across the image plane.

Digitization Definition

- When we have a continuously varying signal, both the value we measure and the intervals at which we can measure it, can vary infinitesimally.
- In contrast, if we were to convert it to a digital signal, we would have to restrict both of these to a set of discrete values that could be represented in some fixed number of bits. Therefore, Digitization is the process of converting a signal from analogue to digital form. it consists of two steps:
- Sampling, when we measure the signal's value at discrete intervals.
- Quantization, when we restrict the value to a fixed set of levels.

Sampling and quantization can be carried out in either order. Sampling rate is the number of samples in a fixed amount of time or space. Quantization level is the equally spaced levels to which a signal is quantized.

Digital versus analogue Representation

1. **Noise Resistance:**

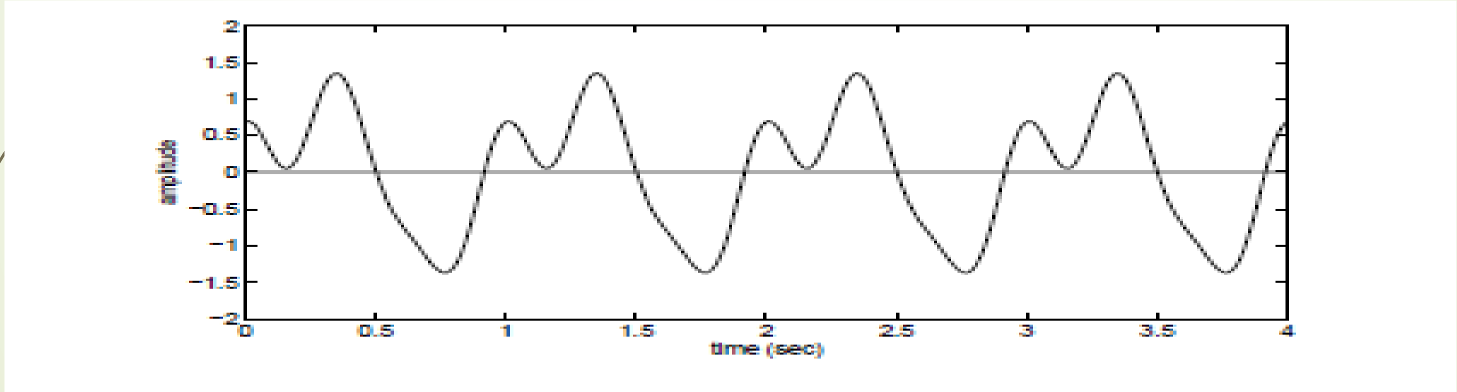
- One of the great advantages that the digital representations have over analogue ones stems from the fact that only certain signal values –those at the quantization levels- are valid. If a signal is transmitted over a wire or stored on some physical medium, inevitably some random noise is introduced, either because of interference from stray magnetic fields, or simply because of the unavoidable fluctuations in thermal energy of the transmission medium. This noise will cause the signal value to be changed.
- If the signal is an analogue one, these changes will not be detectable.
- If the signal is a digital one, any minor functions caused by noise will usually transform a legal value into an illegal one that lies between quantization levels.
- Digital signals are therefore much more strong than analogue ones, and do not suffer degradation when they are copied, or transmitted over noisy media.

2. Accuracy

- Some information must have been lost during the digitization process. The only meaningful measure of accuracy must be how closely the original can be reconstructed. In order to reconstruct an analogue.
- If the original samples were too far apart, any reconstruction is going to be inadequate, because there may be details in the analogue signal that, as it were, slipped between samples.
- It is easily to see that if the sampling rate is too low some detail will be lost in the sampling.

Frequency

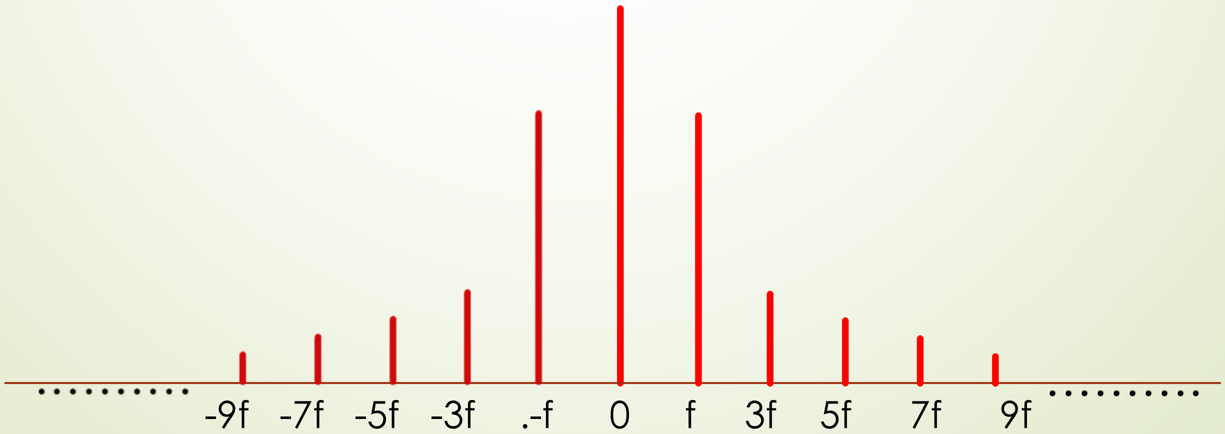
- Any periodic waveform can be decomposed into a collection of different frequency components, each a pure sine wave. In general frequencies, like signals, may be either temporal or spatial. A spatial varying signal may vary in two dimensions.



Consider for example, a signal is composed by adding together sin waves at frequencies of 1 Hz, 2 Hz, and 3 Hz

Time or spatial Domain Versus Frequency Domain

- Normally, the signal is represented in time domain by showing how the amplitude varies with time and space.
- In frequency domain, we can use the frequencies and amplitudes of its components to represent the signals.
- Frequency domain can be computed using a mathematical operation known as the Fourier Transform. The resultant signal is represented in the form of a graph. The horizontal axis represents frequency and vertical axis represents amplitude.
- Frequency Spectrum

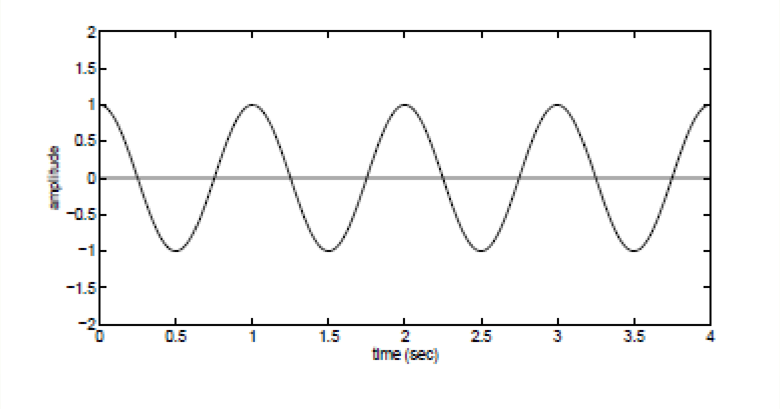


What is the sampling rate that guarantees accurate reconstruction of a signal ?

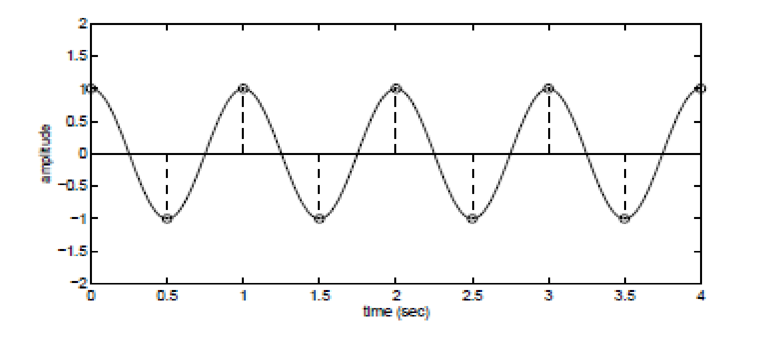
The sampling theorem states that, if the highest frequency component of a signal is at f_h , the signal can be properly reconstructed if it has been sampled at a frequency greater than $2f_h$. This is known as the Nyquist rate.

Example

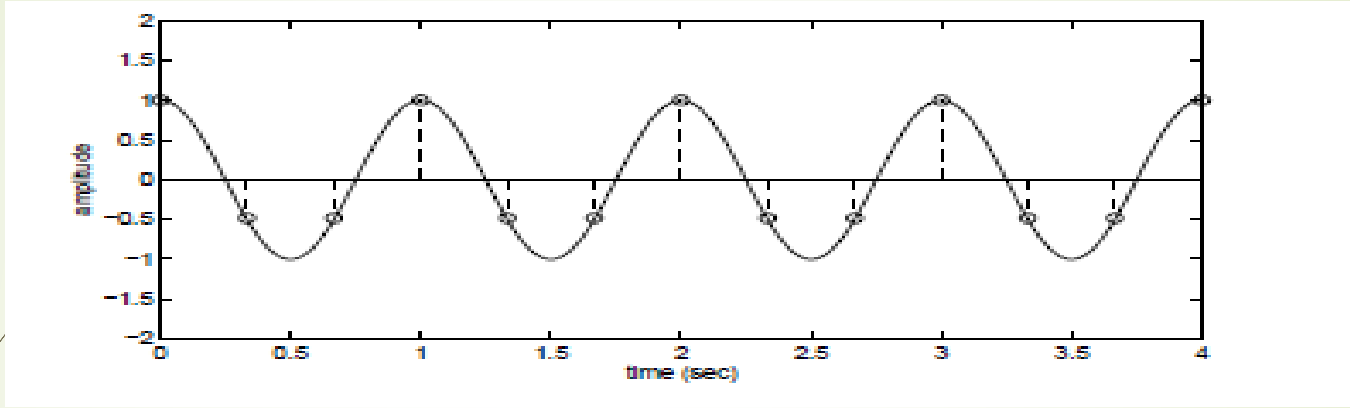
Consider for example a signal composed of a single sine wave at a frequency of 1 Hz :



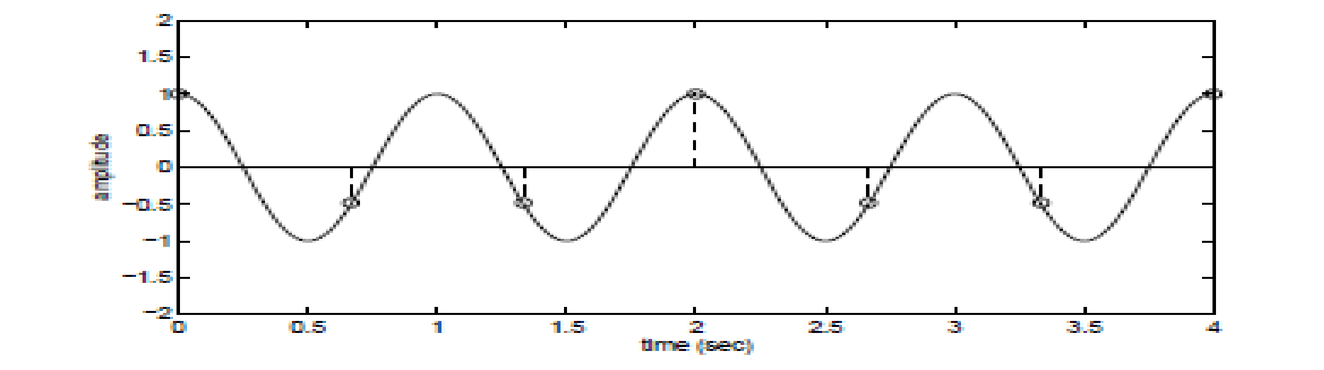
If we sample this waveform at 2 Hz (as dictated by Nyquest theorem) , that is sufficient to capture each peak and through of the signal



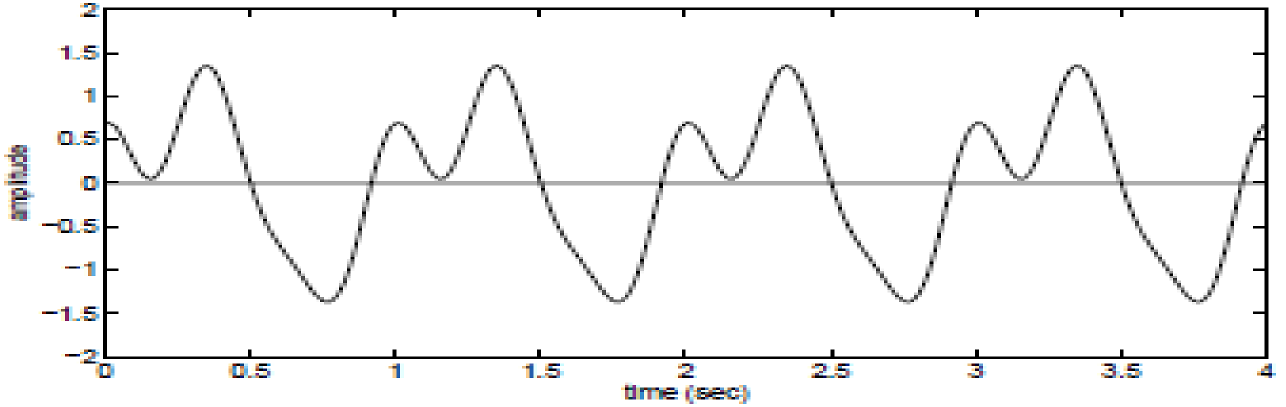
If we sample this waveform at 3 Hz there are more than enough samples to capture the signal



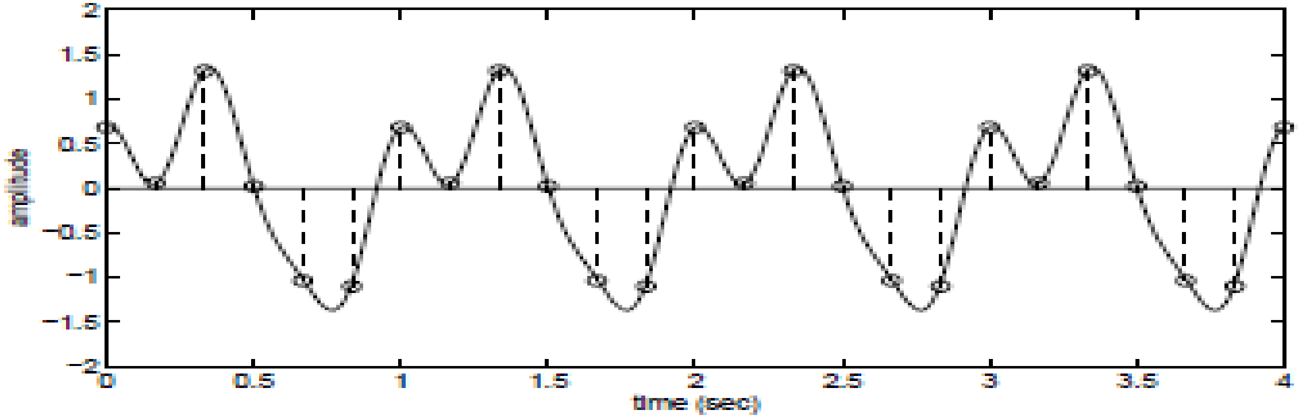
If we sample this waveform at 1.5 Hz then there are no enough samples to capture each peak and through of the signal



Example



Consider again the example of a signal which is composed by adding together sin waves at frequencies of 1 Hz, 2 Hz, and 3 Hz then the sampling rate have to be



Under Quantization Levels

- The effect of an insufficient number of quantization levels are generally easier to grasp than those of inadequate sampling rate. If we can only represent a limited number of different values, we will be unable to make fine distinction among those that fall between.

Example:

- In image, insufficient number of quantization levels as if we were forced to make do with only a few different colors, and so hard to use. This effect is called posterization or brightness contouring, where color areas coalesce, somewhat like a cheaply printed poster.
- When sound is quantized to too few amplitude levels, the result is percieved as a form of distortion, sometimes referred to as quantization noise, it also leads to the loss of quiet passages, and a general fuzziness in sound(rather like a mobile phone in an area of low signal strength).

Original 8-bit image,
256 gray levels

Quantized to 6 bits,
64 gray levels

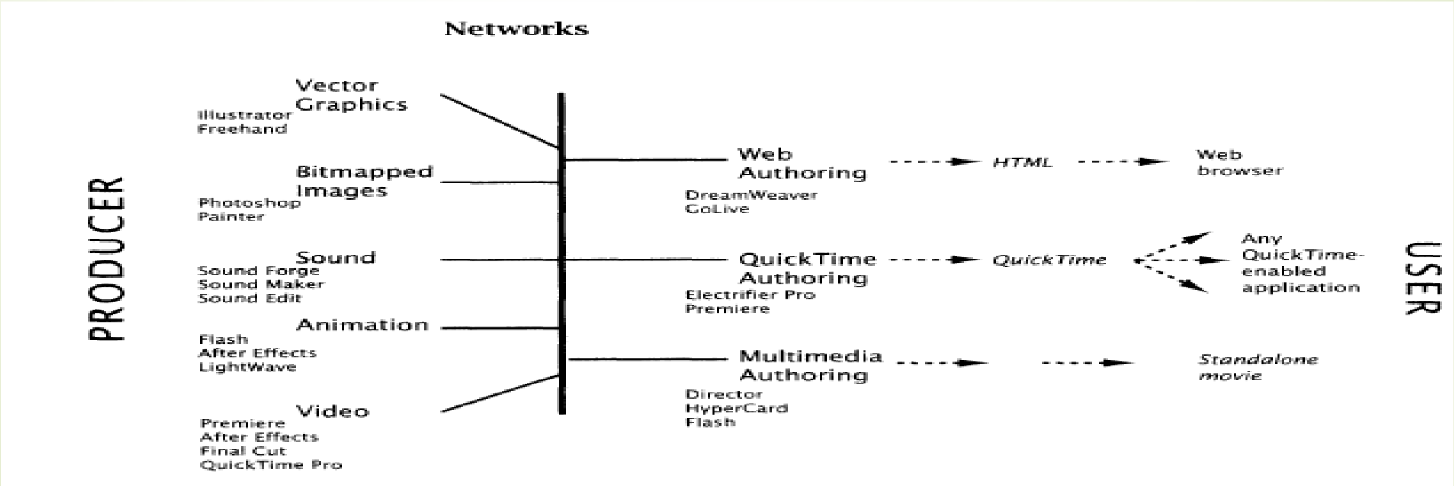
Quantized to 3 bits,

Quantized to 1 bits,

Software

- An essential part of our definition of multimedia is the condition that the combined media be sufficiently well integrated to be presented via a unified interface, and manipulated by a single computer program. This is in complete contrast to the situation we have just described with respect to multimedia production, with its multitude of different software tools and team input. The key to integration is a framework that can accommodate a multiplicity of media and present them to the user. Three distinct approaches can be distinguished.
- The World Wide Web that can accommodate different media, and view it using a dedicated browser.

- Architecture, comprising a format that can contain different media types, together with an API (Application Programming Interface) that provides a rich set of functions to manipulate data in that format. The prime example here is QuickTime. A QuickTime movie can contain video, sound, animation, images, virtual reality panoramas, and interactive elements, all synchronized and coordinated.
- The third approach is quite the opposite to the previous one: deliver the multimedia production in a 'stand alone' form, that needs no additional software to be used.



Networks

- Online distribution of multimedia over LANs or the Internet is almost always based on the client/server model of distributed computation. In this model, programs called *servers 'listen'* on a communication channel for *requests from other programs, called clients, which are generally running on a different machine* elsewhere on the network. Whenever a server receives a request, it sends a *response, which provides some service or data to the client*.
- The requests and responses conform to a *protocol, a set of rules* governing their format and the actions to be taken by a server or client when it receives a request or response. The most popular form of online multimedia delivery is the World Wide Web, whose implementation is an example of the client/server
- model. Web servers and clients communicate with each other using the *HyperText Transfer Protocol, usually abbreviated to HTTP*.

BITMAP IMAGES

BY

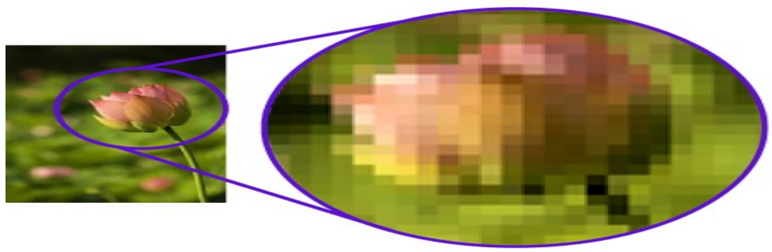
Dr. Amal S. Ajrash

Lecture2

Bitmap images

1. **What is bitmap image?**

- A bitmap is a type of memory organization or image file format used to store digital images. The term *bitmap* comes from the computer programming terminology, meaning just a *map of bits*
- A bitmap, then, is a simple matrix of the tiny dots that form an image and are displayed on a computer screen or printed.



2. Pixels & Colors

- Color is represented on a computer by using varying amounts of red, green and blue light. These are the primary colors of what's called "additive" color - by adding percentages of red, green and blue, any color can be created.
- In this sort of bitmap, each pixel requires one byte for each color index for a total of three bytes per pixel.
- As a byte represents eight bits, each pixel requires 24 bits to store all its color red, green, and blue information. This defines the maximum number of discrete colors this sort of bitmap can represent as 2^{24}, or 16,777,216.
- Such graphics are referred to as “True Color" images, or just as "24-bit" graphics

3. Bit-Depth = Color-Depth

Number of Colors = 2^(Bit-depth) Bit-depth is the number of bits.

It is also called “Color resolution”.

Bit depth	Color resolution	Calcuation
1-bit	2 colors	2^1 = 2
2-bit	4 colors	2^2 = 4
3-bit	8 colors	2^3 = 8
4-bit	16 colors	2^4 = 16
8-bits	256 colors	2^8 = 256
16-bits	65,536 colors	2^16 = 65536
24-bits	16,777,215 colors	2^24 = 16.7 million

4. Resolution

There are three types of resolution measuring different aspects of the quality, detail and size of an image:

- **Color resolution :-** describe the **number of bits** used to **represent the colour** of a single pixel.
- **Image resolution:-**Image resolution measures the pixel dimension of an overall image or how many pixel the image has. For example, 100 * 100-pixel image has a total of 10,000 pixels.
- **Display resolution:-** It simply means how many pixels can be displayed on the computer screen

Display resolution normally uses a setting of 640x480(VGA), 800x600 (SVGA), 1024x768.

5. Pixel point processing

- Image processing is performed by computing a new value for each pixel in an image. The simplest methods compute a pixel's new value solely on the basis of its old value, without regard to any other pixel.
- So increasing the brightness makes every pixel lighter, decreasing it makes every pixel darker.

6. Pixel group processing

- A second class of processing transformations works by computing each pixel's new value as a function not just of its old value, but also of the values of neighboring pixels. Functions of this

Sort perform *pixel group processing,* which produces qualitatively different effects from the pixel point processing operations. Such *filtering* operations can be implemented as operations that combine the value of a pixel with those of its neighbors, it turns out that, instead of transforming our image to the frequency domain (for example, using a DCT) and performing a filtering.

So what is convolution?

Convolutions are operations in which new pixel values are determined from linear combinations (for example addition and multiplication) of a pixel value with its neighbors.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} * \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = (1*a)+(2*b)+(3*c)+(4*d)+(5*e)+(6*f)+(7*g)+(8*h)+(9*i$$

7. Smoothing spatial filters

- Smoothing filters are used for blurring and for noise reduction.
- The output of the smoothing filter is the average of the pixels contained in the neighborhood of the filter mask.(so sometimes called average filter or low pass filters)

$$\frac{1}{9}\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \qquad \frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Original image blurred image

8. Sharpening spatial filter

- The objective of sharpening is to highlight transitions in intensity.
- Sharpening is based on first and second order derivative
- First derivative sharpening make

❑ zero areas of constant intensity

❑ Non zero at the areas of intensity step or ramp

- Second order sharpening make

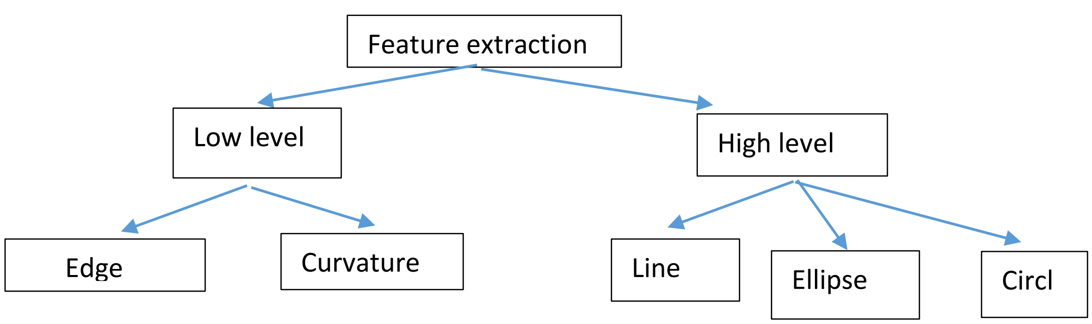
❑ Zero in constant areas

❑ Non zero at onset and end of intensity step or ramp

❑ Zero along ramps of constant slop

Before sharpening **After sharpening**

9. Feature extraction and analysis



9.1 What is feature extraction?

Is a part of data reduction process and is followed by feature analysis. Feature is a “point of interest” for image description

9.2 Properties of a good feature

- Consistent over several images of the same scene

- Invariant towards certain transformations

- Insensitive to noise - “salient”

10. Types of image features extraction

Two types of image features extraction:-

10. A) Low –level feature extraction

Low-level *features* to be those basic features that can be extracted automatically from an image without any shape information (information about *spatial* relationships).

10. A.1) Edge detection

It is well known that we can recognize people from caricaturists’ portraits. It is called *edge detection* and it aims to produce a *line drawing which can be done using* ***First*** order edge detection operators. Edge detection highlights Image *contrast*. Like (prewitt , sobel and canny edge operators) and second order detection operators like (laplacian , and Marr-Hildreth)

10. A.2) detecting image curvature (corner extraction)

What is curvature? Curvature is the rate of change in edge direction. This rate of change characterizes the points in a curve; points where the edge direction changes rapidly are *corners*, such extreme points are very useful for shape description and matching, since they represent significant information with reduced data. Curvature detection is done using:-

Computing differences in edge detection

It computes the *difference* in edge *direction* between connected pixels forming a discrete curve. As such, curvature is simply given by:-

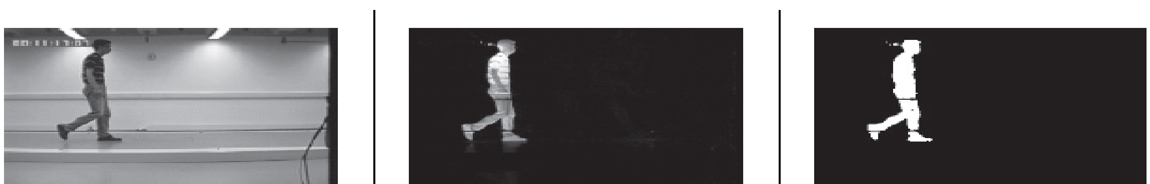
$$k(t) = \varphi_{t+1} - \varphi_{t-1}$$

Where the sequence $\ldots \varphi_{t-1}, \varphi_t, \varphi_{t+1}, \varphi_{t+2} \ldots$ represents the gradient direction of a sequence of pixels defining a curve segment.

10. B) High level feature extraction

High-level *feature extraction* concerns finding shapes in computer images. To be able to recognize Faces automatically, for example, one approach is to extract the component features. This requires extraction of, say, the eyes, the ears and the nose, which are the major facial features. To find them, we can use their shape: the white part of the eyes is ellipsoidal.

- **Thresholding and subtraction**
  Thresholding is a simple shape extraction technique, where the images could be viewed as the result of trying to separate the eye from the background. If it can be assumed that the shape to be extracted is defined by its brightness, then thresholding an image at that brightness level should find the shape

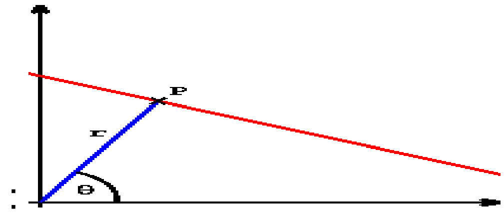


(a) Image of walking subject | (b) After background subtraction | (c) After thresholding

- **Hough transform**

The *Hough transform* (HT) (Hough, 1962) is a technique that locates shapes in images. In Particular, it has been used to extract LINES The general formula of a line is the well-known equation, where *m* denotes the slope, while *b* is the point where the line crosses the y-axis:
$y = m \cdot x + b$

Although this is a simple and easy-to-use formula, vertical lines cannot be described with it. Now have a look at the following image:



Here a polar coordinate describes the line: *r* is the distance of the origin and the line, while *θ* is the angle of the vector pointing from the pole to the closest *P* point of the line. Using these notations, the equation of the line becomes the following:

$y=-\cos(\theta)\sin(\theta)x+r\sin(\theta)$

Which can be rearranged to:

$r=x\cdot\cos(\theta)+y\cdot\sin(\theta)$

Suppose, that we have an arbitrary *P* point on an image. To get all possible lines going through it, simply rotate a line crossing *P* from 0 to 180 degrees: as *θ* changes *r* will change the two points are on the same line. The *θ* and *r* coordinates of the crossing point exactly describe the line on which both points are found. This method can be extended to find circles and elliptical shapes.

11. Binary object features

- In order to extract object features, we need an image that has undergone image segmentation and any necessary morphological filtering.
- After all the binary objects in the image are labeled, we can treat each object as a binary image.
- The labeled object has a value of '1' and everything else is '0'.
- The labeling process goes as follows:
  − Define the desired connectivity.
  − Scan the image and label connected objects with the same symbol.

The binary object features include

1. Area.
2. Center of area.
3. Axis of least second moment.
4. Perimeter.

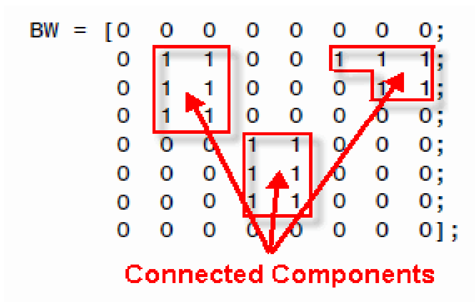
Tells us about object location

5. Euler number.
6. Projections.
7. Thinness ration.
8. Aspect ratio.

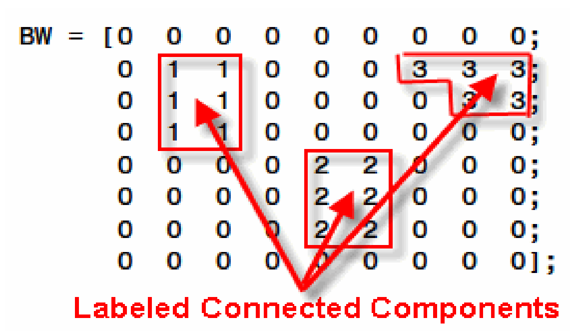
Tells us about the shape of the object

11.A) Understanding Connected-Component Labeling

- A connected component in a binary image is a set of pixels that form a connected group. For example, the binary image below has three connected components.



Connected component labeling is the process of identifying the connected components in an image and assigning each one a unique label, like this:



1. Area

The area of the *i*th object is defined as follows:

$$A_i = \sum_{r=0}^{height-1} \sum_{c=0}^{width-1} I_i(r,c)$$

The area A*i* is measured in pixels and indicates the relative size of the object.

2. Center of Area

The center of area is defined as follows:

$$\bar{r}_i = \frac{1}{A_i} \sum_{r=0}^{height-1} \sum_{c=0}^{width-1} r I_i(r,c)$$

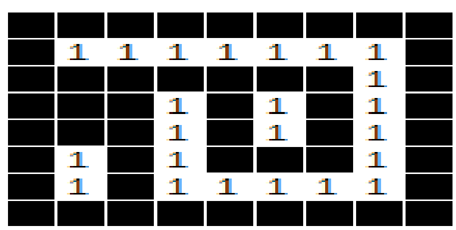
$$\bar{c}_i = \frac{1}{A_i} \sum_{r=0}^{height-1} \sum_{c=0}^{width-1} c I_i(r,c)$$

These correspond to the row and column coordinate of the center of the *i*th object

3. Perimeter

- The perimeter is defined as the total pixels that constitute the edge of the object.
- Perimeters can be found by counting the number of ‘1’ pixels that have ‘0’ pixels as neighbors.
- Perimeter can also be found by applying an edge detector to the Object, followed by counting the ‘1’ pixels

Thresholded image:



4. Thinness Ratio

The thinness ratio can be calculated from perimeter and area.

$$T_i = 4\pi\left(\frac{A_i}{P_i^2}\right)$$

The thinness ratio is used as a measure of roundness.

– It has a maximum value of 1, which corresponds to a circle.

5. Aspect Ratio

This can be found by scanning the image and finding the minimum and maximum values on the row and column where the object lies.

$$\frac{c_{max} - c_{min} + 1}{r_{max} - r_{min} + 1}$$

High aspect ratio indicates the object spread more towards **Horizontal direction.**

6. Euler Number

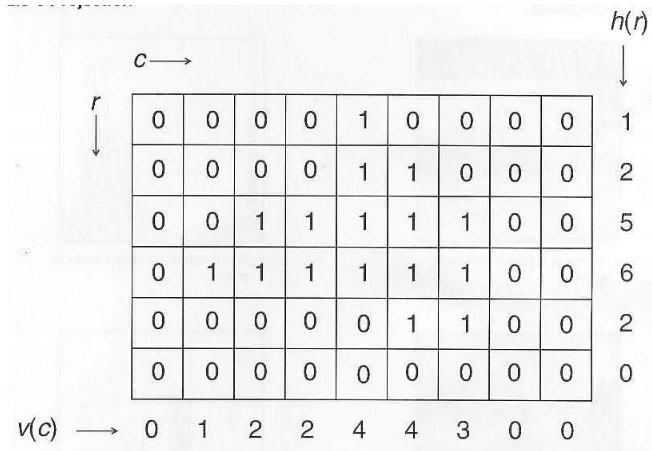
Euler number is defined as the difference between the number of objects and the number of holes.

Euler number = num of object – number of holes

In the case of a single object, the Euler number indicates how many closed curves (holes) the object contains. The Euler number of an image is an important feature that can be used to describe the topological structure of that image.

11.B) Projection

- The projection of a binary object, may provide useful information related to object's shape
- It can be obtained by summing all 1 of an object in horizontal and vertical directions.



									$h(r)$
0	0	0	0	1	0	0	0	0	1
0	0	0	0	1	1	0	0	0	2
0	0	1	1	1	1	1	0	0	5
0	1	1	1	1	1	1	0	0	6
0	0	0	0	0	1	1	0	0	2
0	0	0	0	0	0	0	0	0	0
$v(c)$: 0	1	2	2	4	4	3	0	0	

To find the projections, we sum the number of 1s in the rows and columns.

12. Features analysis using distance and similarity measures

In feature extraction process, we might need to compare two feature vectors. The primary methods to do this are either to measure the difference between the two or to measure the similarity, in feature extraction process, we might need to compare two feature vectors. The primary methods to do this are either to measure the difference between the two or to measure the similarity.

A. Distance Measures

Several metric measurement
1. Euclidean distance.
2. Range-normalized Euclidean distance.
3. City block or absolute value metric.
4. Maximum value.

Euclidean distance is the most common metric for measuring the distance between two vectors.
Given two vectors A and B, where:

$$A = [a_1 \quad a_2 \quad \ldots \quad a_n]$$
$$B = [b_1 \quad b_2 \quad \ldots \quad b_n]$$

The Euclidean distance is given by:

$$\sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2}$$

This measure may be biased as a result of the varying range on different components of the vector.
Another Distance measure, called the **City Block** or absolute value metric, is defined as follows:

$$\sum_{i=1}^{n} |a_i - b_i|$$

The final distance metric considered here is the **maximum value metric** defined by:

$$\max\{|a_1 - b_1|, |a_2 - b_2|, \ldots, |a_n - b_n|\}$$

B. Similarity Measures

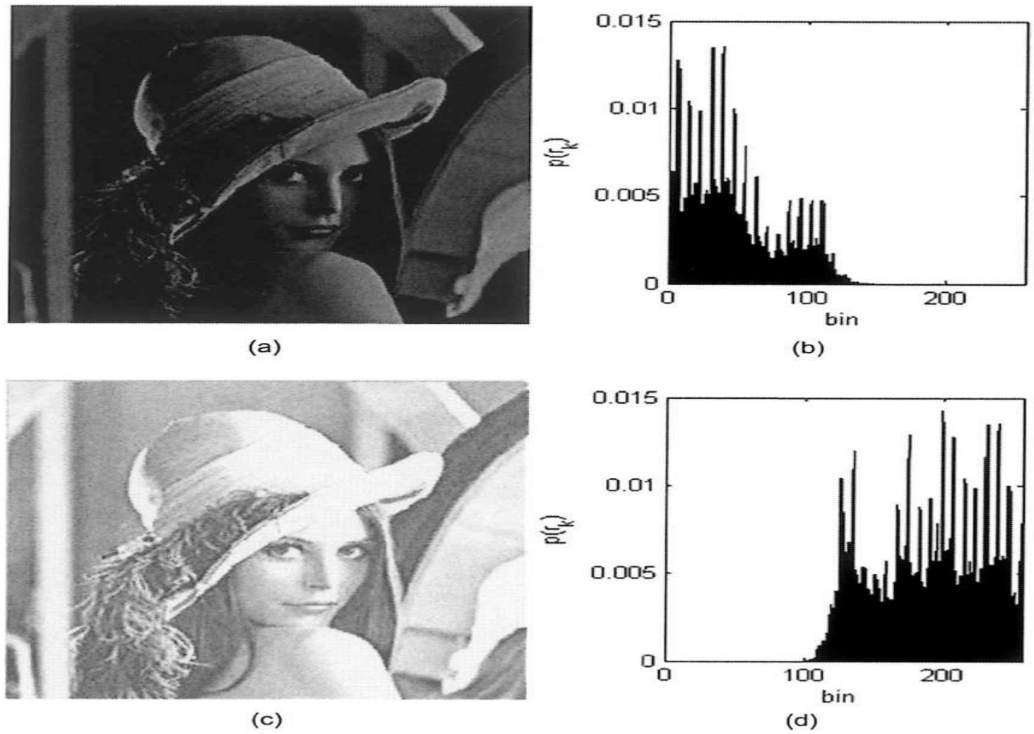
The second type of metric used for comparing two feature vectors is the similarity measure.
The most common form of the similarity measure is the vector inner product.
Using our definition of vector A and B, the vector inner product can be defined by the following equation:

$$\sum_{i=1}^{n} a_i b_i = (a_1 b_1 + a_2 b_2 + \ldots + a_n b_n)$$

When selecting a feature for use in a computer imaging application, an important factor is the robustness of the feature. A feature is robust if it will provide consistent results across the entire application domain.

C. Histogram Features

- The histogram of an image is a plot of the gray-level values versus the number of pixels at that value.
- The shape of the histogram provides us with information about the nature of the image.
- The characteristics of the histogram have close relationship with characteristic of image such as Brightness and Contrast.



$$P(g) = \frac{N(g)}{M}$$

P(g) : probability of gray level *g* in image
N(g) : number of pixel with gray level *g* in image
M : total number of pixel in image

The features based on the first-order histogram probability are
– Mean
– Standard deviation
– Skew
– Energy
– Entropy.

Mean:
- The mean is the average value, tells us something about general brightness of the image.

- A bright image has a high mean.

- A dark image has a low mean.

-The mean can be defined as follows:

$$\overline{g} = \sum_{g=0}^{L-1} gP(g) = \sum_{r}\sum_{c}\frac{I(r,c)}{M}$$

Standard Deviation:
- The standard deviation, which is also known as the square root of the variance, tells us something about the contrast. It describes the spread in the data. Image with high contrast should have a high standard deviation.

- The standard deviation is defined as follows:

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1}(g-\overline{g})^2 P(g)}$$

Skew
The skew measures the asymmetry about the mean in the gray-level distribution. It is defined as

$$\text{SKEW} = \frac{1}{\sigma_g^{\ 3}}\sum_{g=0}^{L-1}(g-\overline{g})^3 P(g)$$

Energy
- The energy measure tells us something about how gray levels are distributed.

- The equation for energy is as follows:

$$ENERGY = \sum_{g=0}^{L-1}[P(g)]^2$$

- The energy measure has a value of 1 for an image with a constant value.

- This value gets smaller as the pixel values are distributed across more gray level values.

- A high energy means the number of gray levels in the image is few, therefore it is easier to compress the image data.

Image Measurements Examples

Lecture 4

By

Dr. Amal S. Ajrash

Example①:- suppose we have an image with 8bits/Pixel.

4	2	2
2	4	3
2	3	4

(3X3)

find 1) Mean 2) standard Deviation 3) skew
4) Energy.

Solu.//

* we must find the probability of each Pixel using the following formula:-

$$P(g) = \frac{N(g)}{M}$$

g:- gray pixel {2, 3, 4}

P(g):- probability of gray level in image.

N(g):- number of Pixel with gray level g in image.

$g=2 \longrightarrow N(2)=4$

$g=3 \longrightarrow N(3)=2$

$g=4 \rightarrow N(4)=3$

M:- total number of pixel in image.

$M=9$

$\therefore P(2)=4/9=0.4$

$P(3)=2/9=0.2$

$P(4)=3/9=0.3$

1) Mean

$$\bar{g} = \sum_{g=0}^{L-1} g * P(g) \quad (L = 256 \text{ because } 8\text{bits/Pixel})$$

$$\bar{g} = (4*0.3) + (2*0.4) + (2*0.4) + (2*0.4) + (4*0.3) + (3*0.2) + (2*0.4) + (3*0.2) + (4*0.3)$$

$$\bar{g} = 1.2 + 0.8 + 0.8 + 0.8 + 1.2 + 0.6 + 0.8 + 0.6 + 1.2$$

$$\bar{g} = 3.6 + 3.2 + 1.2$$

$$\boxed{\bar{g} = 8}$$

2) Standard deviation

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1} (g-\bar{g})^2 \, p(g)}$$

$$\sigma = \sqrt{((4-8)^2*0.3) + ((2-8)^2*0.4) + ((2-8)^2*0.4) + ((2-8)^2*0.4) + ((4-8)^2*0.3) + ((3-8)^2*0.2) + ((2-8)^2*0.4) + ((3-8)^2*0.2) + ((4-8)^2*0.3)}$$

$$\sigma = \sqrt{16*0.3 + 36*0.4 + 36*0.4 + 36*0.4 + 16*0.3 + 25*0.2 + 36*0.4 + 25*0.3 + 16*0.3}$$

$$\boxed{\sigma = 9.3}$$

3) Skew

$$Skew = \frac{1}{\sigma^3} \sum_{g=0}^{L-1} (g - \bar{g})^3 * P(g)$$

تمثل عدد مرات تكرار pixel

$$= \frac{1}{(9.3)^3} \left[((4-8)^3 * 0.3) * 3 + ((2-8)^3 * 0.4) * 4 + ((3-8)^3 * 0.2) * 2 \right]$$

$$= \frac{1}{(9.3)^3} \left[-57.6 + (-345.6) + (-50) \right]$$

$$= \frac{-453.2}{804.4}$$

$$\boxed{Skew = -0.6}$$

4) Energy

$$Energy = \sum_{g=0}^{L-1} [P(g)]^2$$

$$= \underset{P(4)}{(0.3)^2} * 3 + \underset{P(2)}{(0.4)^2} * 4 + \underset{P(3)}{(0.2)^2} * 2$$

$$\boxed{Energy = 0.99}$$

Example ②: Suppose we have an image with 8 bits/pixel

vector(1)	10	2	11	3	0
vector(2)	4	1	15	3	8
(3)	2	12	2	10	9
(4)	5	15	7	14	1
(5)	9	8	10	2	5

(5x5)

Find/ 1) Euclidean Distance 2) City block
3) maximum value metric. 4) inner Product

1) $E.D. = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$

calculate the E.D. for vector(1) and vector(5)

$E.D. = \sqrt{(10-9)^2 + (2-8)^2 + (11-10)^2 + (3-2)^2 + (0-5)^2}$

$E.D. = \sqrt{64} = \boxed{8}$

2) City Block or absolute value metric

$C.B. = \sum_{i=1}^{n} |a_i - b_i|$

$= |10-9| + |2-8| + |11-10| + |3-2| + |0-5|$

$= 14$

3) Maximum value metric.

$\max\{|a_1-b_1|, |a_2-b_2|, \ldots, |a_n-b_n|\}$

$= \max\{|10-9|, |2-8|, |11-10|, |3-2|, |0-5|\}$

$= \max\{1, 6, 1, 1, 5\}$

$\therefore \max = 6$

4) inner Product.

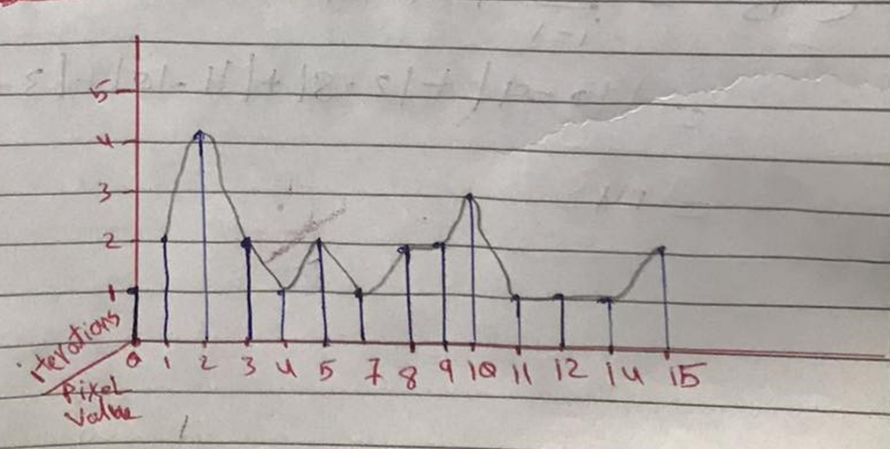
$I.P. = \sum_{i=1}^{n} a_i b_i$

$= (10*9) + (2*8) + (11*10) + (3*2) + (0*5)$

$= 90 + 16 + 110 + 6 + 0$

$I.P. = 222$

Histogram هو تكرارات كل pixel في الصورة



Example ③: If you have the following matrix:

						h(r)
1	1	0	0	0	0	*
0	1	1	1	0	0	*
0	1	1	1	0	0	*
0	0	1	1	1	0	*
0	0	0	0	1	0	*
0	0	0	0	0	1	*

v(c)

Find the horizontal & vertical projection

projection:-

① summing the row values give the horizontal projection

2 3 3 3 1 1

② summing the column values give the vertical projection

1 3 3 3 2 1

H.W

For the above matrix find the Area and center of Area

$$\text{Area} = \sum_{r=0}^{H} \sum_{c=0}^{W} I(r,c)$$

$$\hat{r} = \frac{1}{A} \sum_{r=0}^{H} \sum_{c=0}^{W} r\, I(r,c)$$

$$\hat{c} = \frac{1}{A} \sum_{r=0}^{H} \sum_{c=0}^{W} c\, I(r,c)$$

Fundamental Concept of Video

By
Dr. Amal S. Ajrash

1. Introduction

Video is the technology that captures moving images electronically. Those moving images are really just a series of still images that change so fast that it looks like the image is moving.

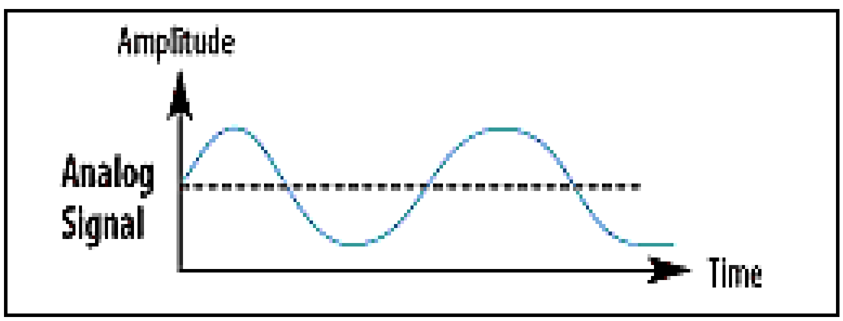
Video is an excellent tool for delivering multimedia because it contains all the components of multimedia (image, sound, and text). It is recorded from a live source.

2. Video standard

2.1. Analog Video

- Analog is the process of taking audio or video signal and translating it into electrical pulses.

In an analog system, the output of the CCD **(Charge Coupled Devices**) is processed by the camera into three channels of color information and synchronization pulses (sync) and the signals are recorded onto magnetic tape.

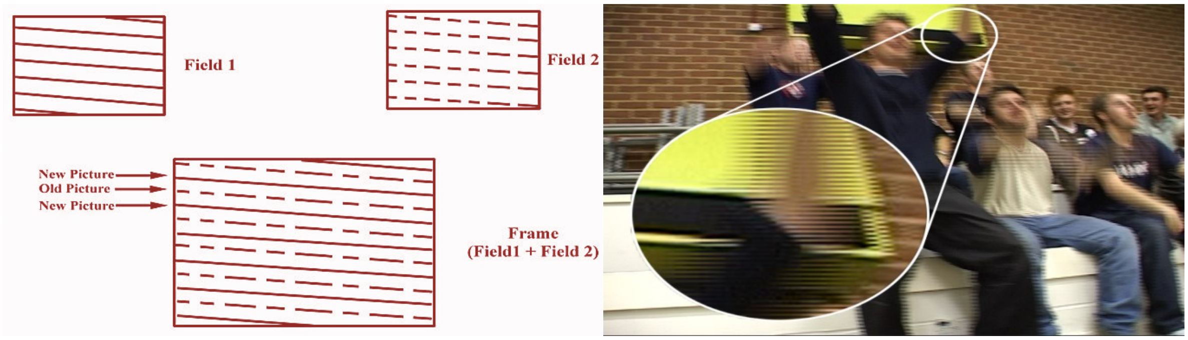


Three analog broadcast video standards are commonly in use around the world:

➢ **NTSC**

- National Television Standards Committee (NTSC).
- Used in United States, Japan, Canada and many other countries.
- a single frame of video was made up of 525 horizontal scan lines , first it laid down all the odd-numbered lines, and then all the even-numbered lines.
- Each of these passes painted a field, and the two fields were then combined to create a single frame.

- In NTSC system for television and video, frames are displayed at a rate of 30 per second.



➤ PAL

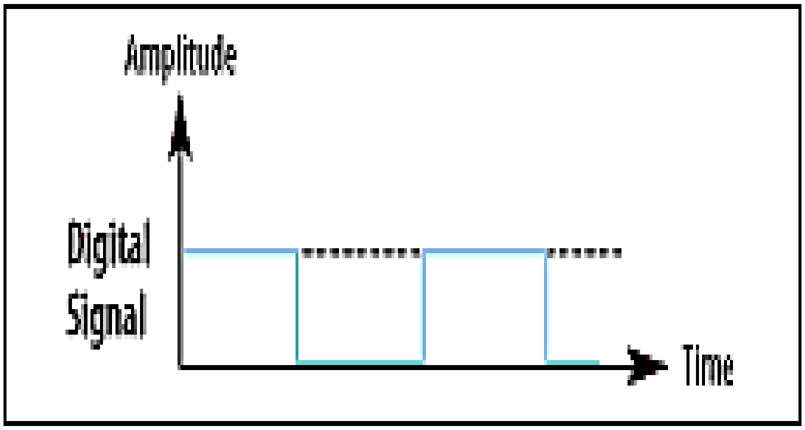
- Phase Alternating Line (PAL)
- Used in the United Kingdom, South Africa, China, and Australia.
- PAL increased the screen resolution to 625 horizontal lines.
- 25 frames per second.

➤ SECAM

- Sequential Color and Memory (SECAM)
- Used in France, Eastern Europe, and a few other countries.
- SECAM is a 625-line.
- 25 frames per second.

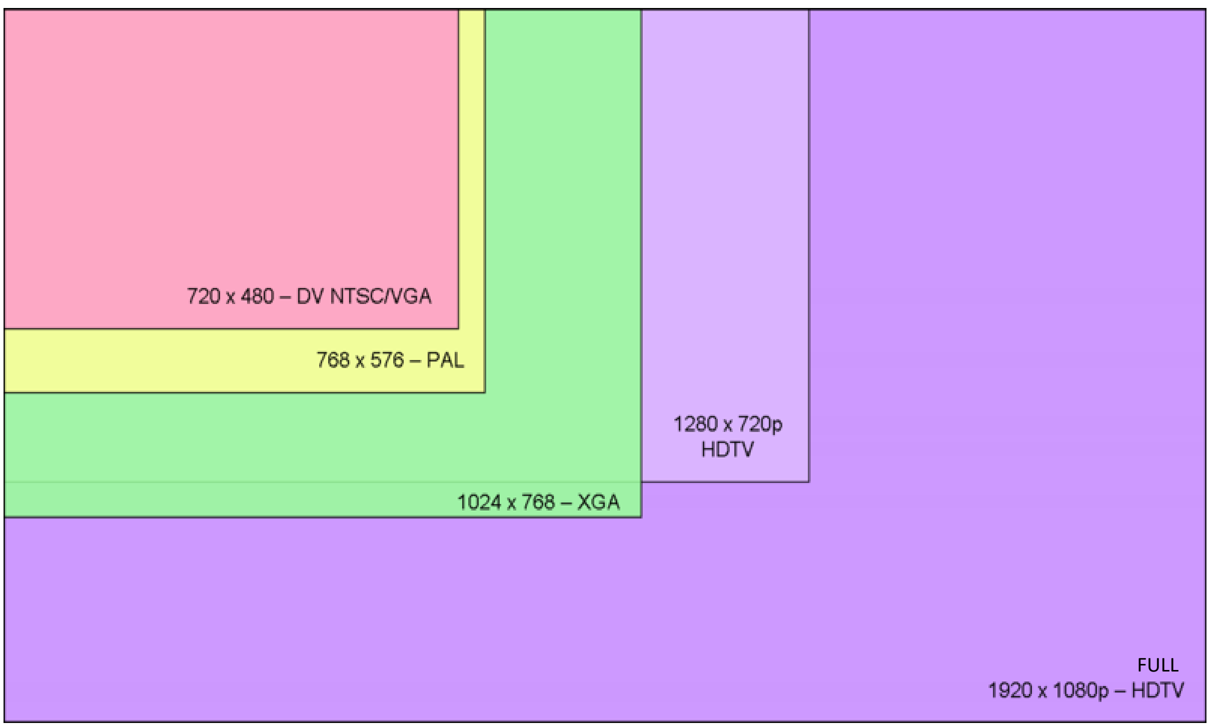
2.2. Digital Video

Digital video is a **sequence of picture** signals (**frames**) that are represented by binary data (bits) that describe a finite **set of color and luminance levels**. The digital information contains characteristics of the video signal and the position of the image (bit location) that will be displayed. The basic process of creating digital video is the image **digitization process, compression analysis that produces key frames and difference frames, and formatting the data into files or streams (video formats).**



HD

- High Definition formats.
- Here Modified the Digital format Standard and moved from an analog to a digital standard.
- HD frames may have width 720 x 1080 pixels.



3. The Main Difference Between the Digital and Analog Medium

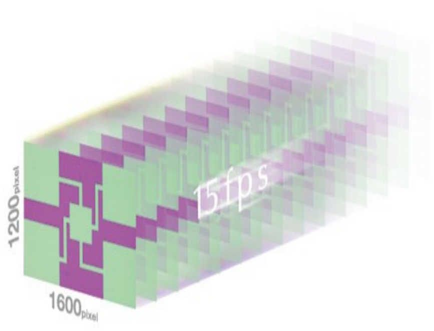
- In digital format Random Access enables us to quickly jump to any point in a movie. In analog format, we have to wind the tape backward and forward to reach a point in the movie.
- It is easy to duplicate digital video without loss of quality, were video producers can convert real-life video that they have shot into digital format and edit it without losing the original quality of the film. In the analog tapes, video producers lose some quality after the first time they edit the video.
- Digital format also allows us to quickly cut, paste, or edit video in other ways. It also allows easy addition of special effects.

4. Basic Concept of Video.

1- Bit Rate is amount of data that can be carried from one point to another in a given time period (usually a second). Bit rate is sometimes called data rate or transfer rate or bandwidth**.**

2- Aspect Ratio This is the ratio of width to height video will be encoded. This information is present in the output video stream and used by the decoder to display the video at the correct aspect ratio. The computer display is designed for an aspect ratio of 1.33:1, which means that the width of the display area is only 1.33 times the height, almost *square*.

3- Frame means one still picture. By changing still pictures (frames) quickly, human eye "thinks" that the video is smooth and can't separate pictures from each other’s and instead sees smooth video.



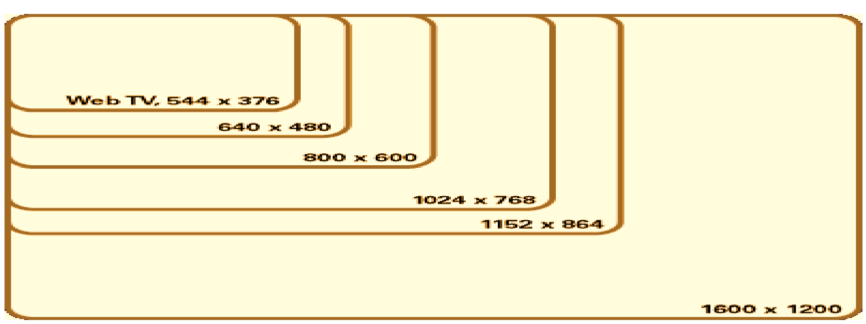
4- Frame Rate is the number of video frames (complete pictures) that will be presented to the viewer each second. Human eye can sees smooth video with the frame rate more than ~24 fps (frames per second). The normal human eye sees 250 frames per second.

Number of Frame	Clarity and smoothness
5 or 10	Intermittent and non-smooth
24	smoothly without roughness (Cinematography)
30	smoothly without roughness (TV work)
≥ 60	Slow footage

5- Frame Buffer is a special memory to hold the complete digital representation of the frame to be displayed on a computer screen.

6- Color Depth or bit depth is the number of bits used to represent the color of a single pixel in a bitmapped image or video frame buffer. It is known as bits per pixel (bpp). The quality of video is dependent on the color quality.

7- Image Size a standard full screen resolution is 640x480 pixels . New high-definition televisions (HDTV) are capable of resolutions up to *1920×1080p60*, 1920 pixels per scan line by 1080 scan lines, progressive, at 60 frames per second.



8- File Size Several elements determine file size:

- Frame rate
- Image size
- Color depth

To determine file size use the following formula:

File size = Frames Per Second (FPS) * image size * color depth / 8

5. Types of Video Signals

Video signals can be organized in three different ways:

5.1. Component video

- Higher-end video systems make use of three separate video signals for the red, green, and blue image planes. Each color channel is sent as a separate video signal.
- Best color reproduction.
- Component video requires more bandwidth and good synchronization of the three components.

5.2. Composite Video (1 Signal)

Color (“chrominance”) and intensity (“luminance”) signals are mixed into a single carrier wave.

a) **Chrominance** is a composition of two color components (I and Q, or U and V).

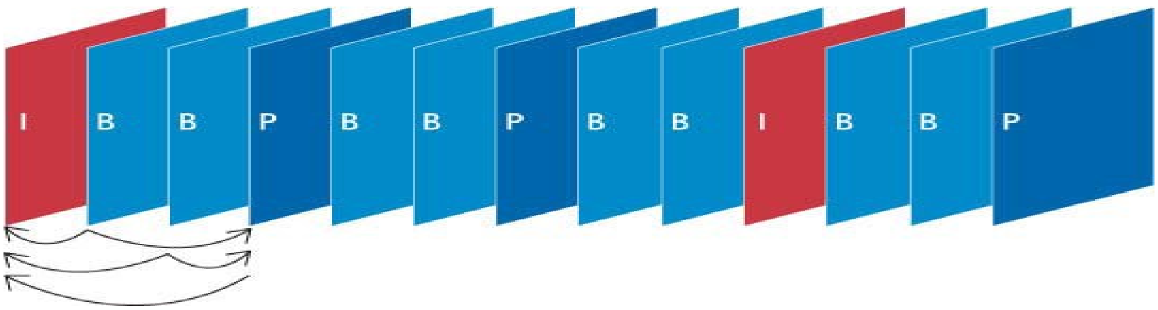
b) The chrominance and luminance components can be separated at the receiver end and then the two color components can be further recovered.

c) Since color and intensity are wrapped into the same signal, some interference between the luminance and chrominance signals is inevitable.

5.3. S-Video (2 Signals)

As a compromise, (**Separated video**, or **Super video**, e.g., in S-VHS) uses two wires, one for luminance and another for a composite chrominance signal. There is less crosstalk between the color information and the crucial gray-scale information.

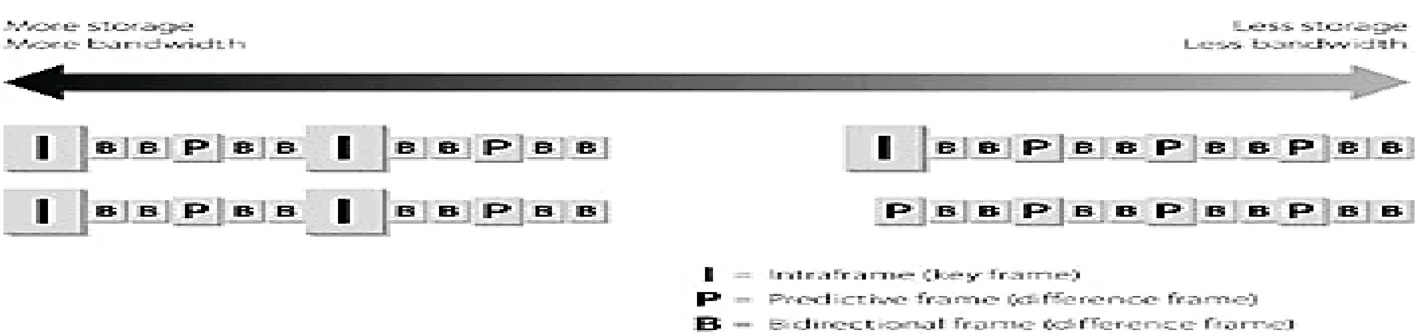
6. Frames Types



- I Frame (Intra Frame) This is a type of key frame. Key frames derive directly from the video and are not calculated from other frames. I frames are the largest frames and must store the most data.

- P Frame (Predictive Frame) This frame is derived from the frame before it and specifies how it differs from the previous frame. P frames are smaller than I frames, requiring much less data storage. P frames are a type of difference frame. All difference frames are calculated from other frames, so they store much less data per frame.

- B Frame (Bidirectional Frame) This frame is computed from both the frames before and after it. B frames are the smallest of the three frame types. Like P frames, B frames are difference frames.



7. Compression

Compression Is technique of reducing the amount of storage required to hold a digital file to reduce the disk space the file requires and allows it to processed and transmitted more quickly.

The techniques are employed to achieve desirable reductions in image data:

- Reduce color nuances within the image.
- Reduce the color resolution with respect to the prevailing light intensity.
- Remove small, invisible parts, of the picture.
- Compare adjacent images and remove details that are unchanged.

7.1. Video Compression

Refers to reducing the quantity of data used to represent digital video image two key compression factors into consideration:

Lossiness refers to a quality decrease produced by the compression process. It means that the picture you see when you decompress won't be exactly the same picture you originally compressed.

Time A compressed video needs to be restored (decompressed) to be viewed. It takes time—often a lot of it—to decompress your video.

- Used compression to reduce the redundancy there are two types:

1- Temporal Redundancy can be reduced by registering differences between frames (B and P frame).

2- Spatial Redundancy is reduced by registering differences between parts of a single frame (I- frames).

7.2. Types of Data Compression

There are two categories of data compression

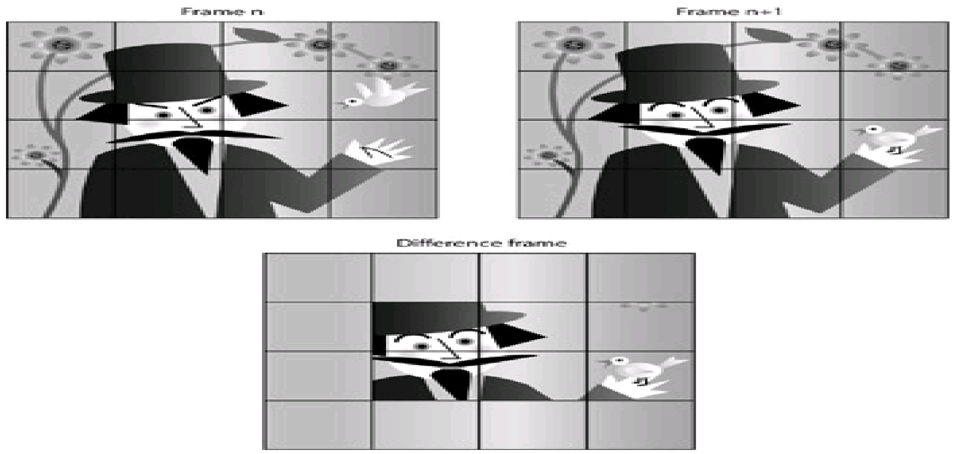
Lossy Compression: in which the compression processes remove some of the data.

Lossless Compression: the compression process allows for subsequent decompression of the data with no loss of the original data. Lossless compression is used for program.

7.3. Compression / Coding Standards

The various Motion Picture Experts Group (MPEG) formats are good examples of Intra frame and the table below summarize some common video compression format.

MPEG-1	MPEG-2	MPEG-4
Released in 1992	Released in 1994	Released in 1998
A standard for coded representation of Moving pictures, Audio, and Combination of above.	Used in NTSC/PAL with bit rates target between 2-10 Mbit/s, It also supports HDTV applications.	Algorithms and tools for coding and flexible representation of audio/video.
Typical application video CD (VCD).	Applications: Digital video tape recorder (VTR) , DVD.	Typical application in mobile phone and HDTV.
	key frame rates increase, so do image quality, storage, and bandwidth.	
	Dividing pictures into blocks and comperes. When successive blocks prove sufficiently dissimilar, the change is recorded onto a difference frame, either a **B** or **P** frame. The difference frame stores only those parts of the picture that have changed.	

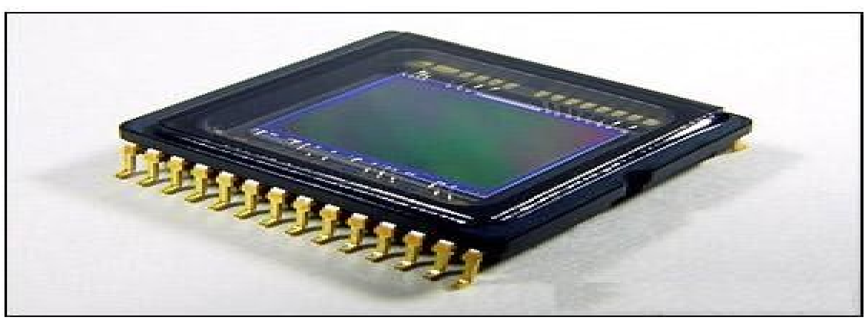


Compression in MPEG-2

8. Hardware for video capture

8.1. Digital Video camera

A digital video camera (camcorder) functions in the same way as a still digital camera. The user points the camcorder at the scene they wish to capture and a lens focuses the image onto an **array of CCDs** (**Charge Coupled Devices**) each CCD corresponds to one pixel in the image being recorded. The CCD is a sensor which changes the light striking it into an electrical signal.

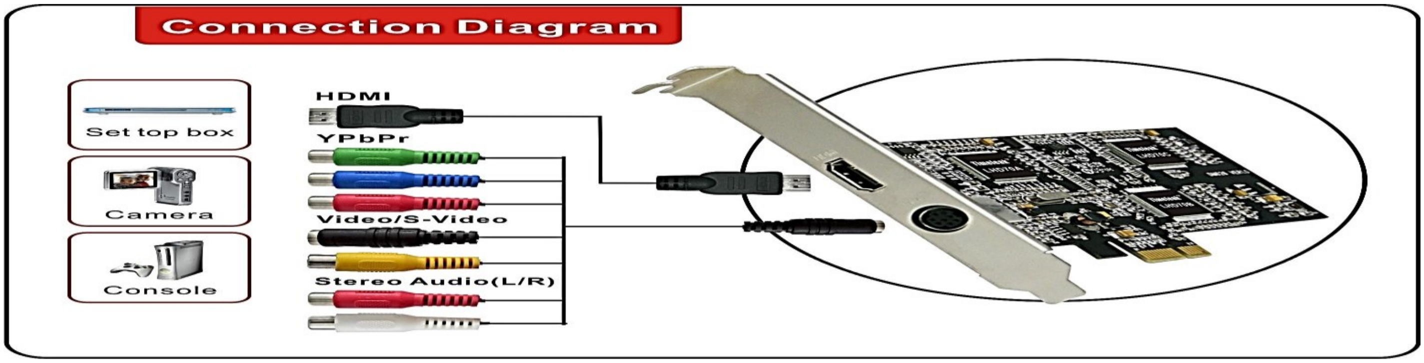


8.2. Webcams

- Are used for real-time chat, including instant messaging (IM), telephone calls over the internet, live broadcasts and video conferencing.
- Webcams are a lot less expensive than camcorders.
- There are fewer CCDs.
- They have no backing storage facility.

8.3 Video Capture Card

- If a recording was made with an analogue video camcorder then the conversion from analogue and compression will not have taken place.
- Receives the video data, converts it to digital and compresses it.
- Can be used to receive data from an analogue video recorder or analogue TV.



9. Advantages of Using Video

- Captures interest.
- Increase retention.
- Clarifies complex physical actions and relationships.
- Video is the most powerful communicative tool in history, it allow you to communicate your message quickly and effectively.
- With video, you can expect nearly 100% viewership.
- video presents standardized information for every viewer, every time can incorporate other media.

10. Disadvantages of Using Video

- Is expensive to produce.
- Requires extensive memory and storage.
- Requires special equipment.
- Does not effectively illustrate abstract concepts and static situations.

Sound

By

Dr. Amal S. Ajrash

1. Sound

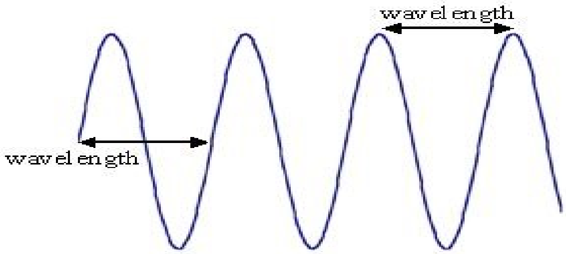
Sound is a wave phenomenon like light, but is macroscopic and involves molecules of air being compressed and expanded under the action of some physical device.

For example, a speaker in an audio system vibrates back and forth and produces a longitudinal pressure wave that we perceive as sound.

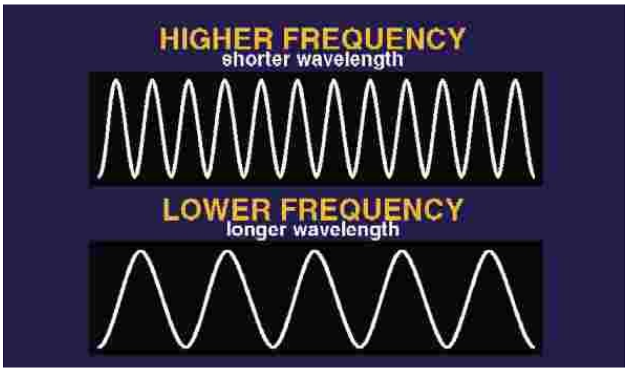
- Sound is a continuous wave that travels through the air .
- the wave is made up of pressure level at a location.
- Sound waves have normal wave properties such as reflection, refraction(change the angle when entering a medium with different density) and diffraction(bending around an obstacle).

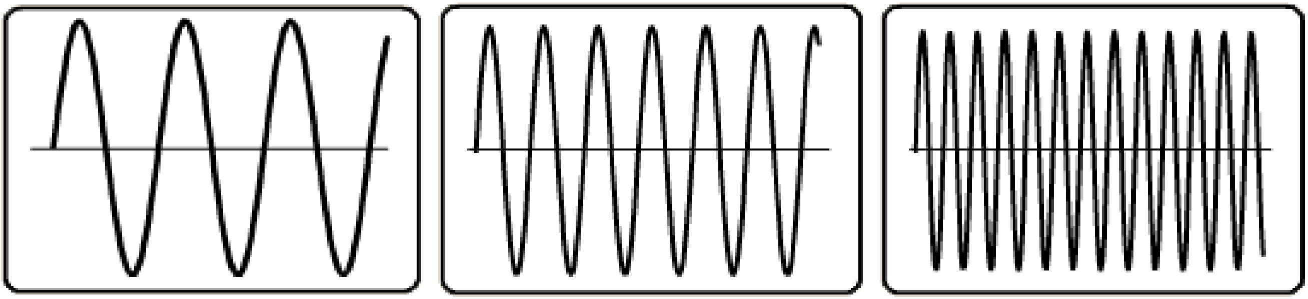
Wave length & frequency

- Wave length : is the distance between one part of a wave and the same part of the next wave



- Frequency is the number of occurrences of a repeating event per unit of time. The unit of frequency is the Hertz(Hz).





- Low pitch
- Low frequency
- Longer wavelength

- High pitch
- High frequency
- Shorter wavelength

1.1 Digitizing Sound

Contemporary formats for digital audio are influenced by the CD format, which dominated audio for over two decades. For example, the sampling rate and number of quantization levels used for high quality audio is almost always the same as that used for CD.

1.2 Digital audio

Digital audio is created when you represent the characteristics of a sound wave using numbers. One can digitized sound from a microphone, synthesizer, existing recordings, live radio and television broadcasts and popular CD and DVDs. Digitized sound is sampled sound. Every *n*th fraction of a second, a

Sample of sound is taken and stored as digital information in bits and bytes. The quality of this digital recording depends upon how often the samples are taken (**sampling rate** or frequency, measured in kilohertz, or thousands of samples per second) and how many numbers are used to represent the value of each sample (**bit depth**, **sample size**, resolution, or dynamic range).

1.3 Sampling

If high fidelity audio reproduction is desired, a sampling rate must be chosen that will preserve at least the full range of audible frequencies. If the limit of hearing is taken to be 20 kHz, a minimum rate of 40 kHz. Is required by the sampling theorem. The sampling rate used for audio CDs is 44.1 kHz (the precise figure being chosen by manufacturers to produce a desired playing time given the size of the medium). Because of the ubiquity of the audio CD, the same rate is commonly used by the sound cards fitted to computers, to provide compatibility.

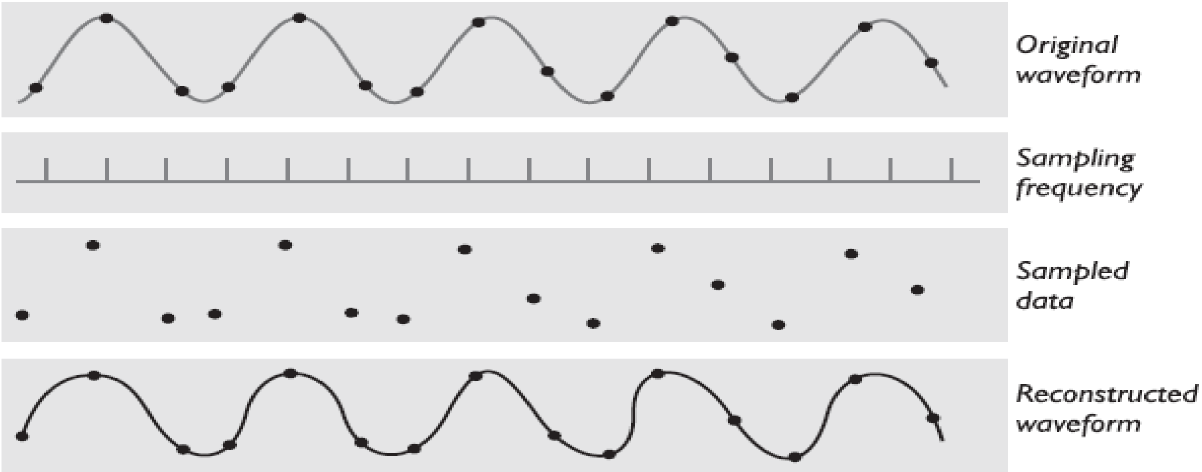
Where a lower sound quality is acceptable, or is demanded by limited bandwidth, sub multiples of 44.1 kHz are used: 22.05 kHz is commonly used for audio distained for delivery over the internet, while 11.025kHz is used for speech. Some professional recording devices use sample rate that are multiples of 48 kHz.

Sampling relies on highly accurate clock pulses to determine the intervals between samples. If the clock drifts, so will the intervals. Such timing variations are called jitter. The effect of jitter is to introduce noise into the reconstructed signal. There is a little tolerance for jitter: it has been estimated that for CD-quality sound, the jitter in the ADC must be less than 200 picoseconds ($200 * 10^{-12}$ seconds) .

1.4 Quantization

The number of quantization levels for analogue to digital conversion in any medium is usually chosen to fit into convenient number of bits. For sounds, the most common choice of sample size is 16 bits as used for CD audio giving 65.536 quantization levels.

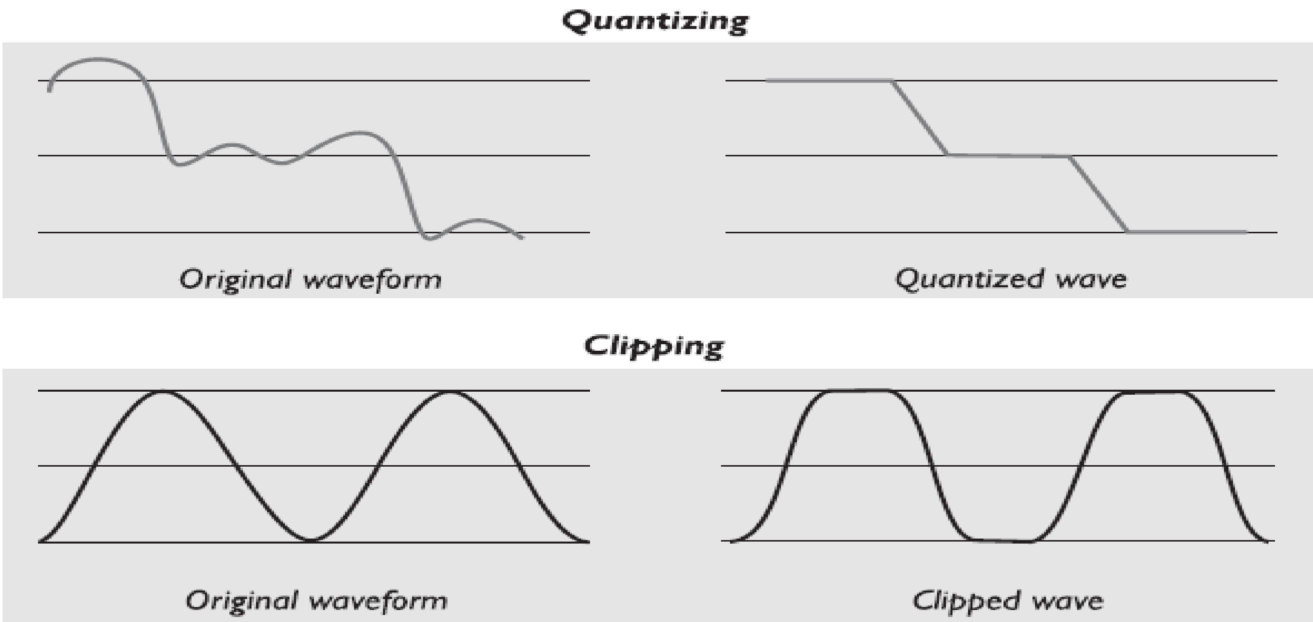
- A smaller samples size are sometimes needed to maintain small file sizes and bit rates. The minimum acceptable is 8-bit sound. In the search for higher fidelity reproduction, as many as 24 bits are sometimes used to record audio samples.
- Quantization noise will be worse for signals of small amplitude.
- Before sampling, a small amount of random noise is added to the analogue signal. This deliberately added noise is known as dither.
- The presence of noise has caused the samples to alternate rapidly between quantization levels, instead of jumping cleanly and abruptly from one to the next level. The sharp transition have been soften, and the quantization error has been randomized.



Slices of analog waveforms are sampled at various frequencies, and each discrete sample is then stored either as 8 bits or 16 bits (or more) of data.

The value of each sample is rounded off to the nearest integer (**quantization**), and if the amplitude is greater than the intervals available, clipping of the top and bottom of the wave occurs (see figure below). Quantization can produce an unwanted background hissing noise, and clipping may severely distort the sound.

Examples of quantizing &clipping



1.5 Recording Sound

There are two types of recording such as :-

1) Stereo recordings are made by recording on two channels, and are lifelike and realistic.

2) Mono sounds are less realistic, flat, and not as dramatic, but they have a smaller file size.

Stereo sounds require twice the space as compared to mono recordings. To calculate the storage space required, the following formula are used:

Mono Recording:

File size = Sampling rate x duration of recording in seconds x (bits per sample/8) x 1

Stereo Recording:

File size = Sampling rate x duration of recording in seconds x (bits per sample/8) x 2

1.5 Sound File Attribute

1) Hz

Hz, pronounced "Hertz" is a kind of measurement. Hz measures cycles of something happening. If a sound wave goes up and down, that's one Hz. Directly translated into sound, a Hz is one vibration. A standard music CD holds sound at a level of 44,100 Hz. This means that the sound can change, or vibrate, 44,100 times every second.

2) Bit Rate

"Bit" means binary digit, and is numerically represented by a zero or a one. Computers use zeros and ones to make up their whole language. When referring to sound, bits are used to tell the computer how much electricity to feed into your speakers to create a vibration. The level of electrical signal dictates the height of each wave. If you used two bits, you would be able to come up with four different levels to describe the signal strength of each sound wave (Hz):00, 01, 11, 10 Standard music CD's use 16 bit sound, which means that 16 zeros and ones are used to describe the level (height) of each sound wave.

1.6 Sound File Type

Sounds, like other computer files, can be saved in different formats. You'll learn how to create sounds in each of the following formats.

1) WAV: WAV is a sound file developed by Microsoft for use on windows based machines. WAV is also the file format for standard music CDs. The WAV file uses interesting algorithms to compress raw sound without a loss in quality.

2) AIFF: Originally developed by Apple, the "Audio Interchange File Format" is mostly used by Silicon Graphics and Macintosh machines. AIFF files are easily converted to other file formats, but can be quite large. One minute of 16-bit stereo audio sampled at 44.1kHz usually takes up about 10 megabytes. AIFF is often used in high end applications where storage space is not a consideration.

3) MP3: MP3 is a type of compression that can dramatically reduce file size without drastically reducing sound quality. MP3 works by, among other things, chopping off all sounds that are outside of the normal human range of hearing. You will need a sound player to listen to MP3's.

4) RealAudio: Developed by Progressive Networks, RealAudio was the first format to allow for real time streaming of music and sound over the web. Listeners are required to download the Real player to enjoy sound in RealAudio Format. The Real player can also stream video and is currently in use by millions of Internet users worldwide.

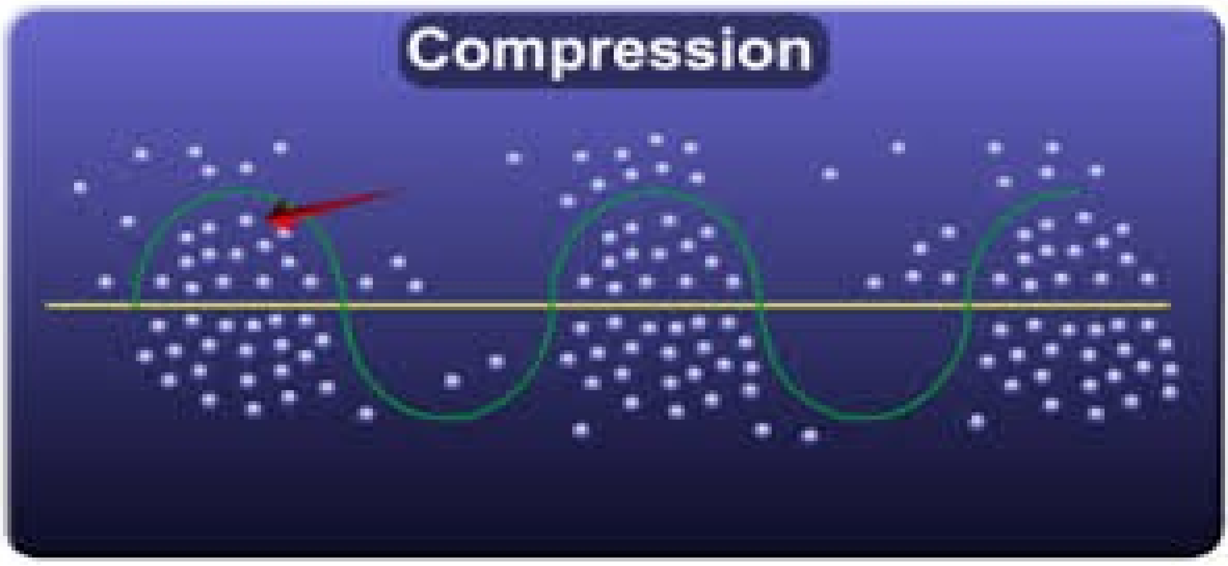
1.7 Streaming Audio

Streaming audio means sound is deliver over a network and played as it arrives without having to be stored on the computer. Live transmission and playing of files that are too big to be held on an average sized hard disk.

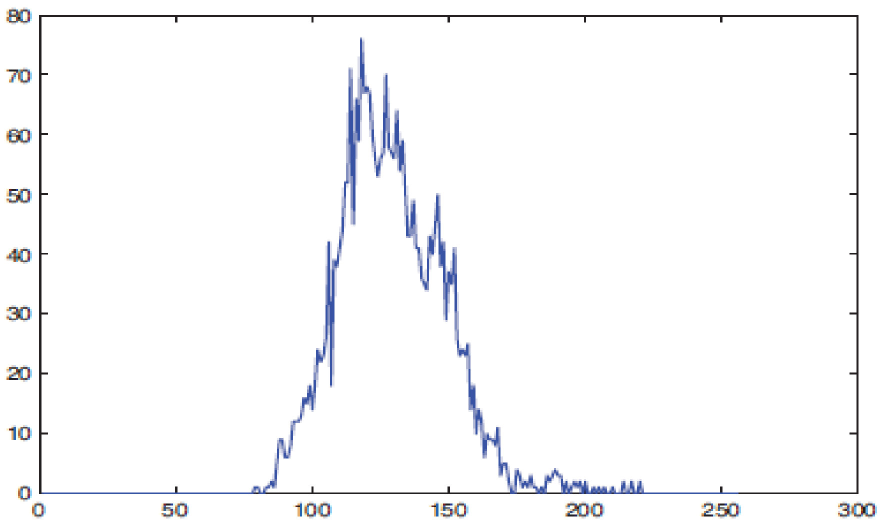
Because of lower bandwidth required by audio, streaming is more successful for sound than it is for video. Example for streaming audio format: Real Network, RealAudio, streaming QuickTime. Software required for playing streaming audio : RealPlayer , Winamp, windows media Player.

1.8 Audio Compression

Is a technique for converting or encoding audio (sound) information so that a smaller amount of information elements or reduced bandwidth is required to represent, store or transfer audio signals. Audio compression coders and decoders (codecs) analyze digital audio signals to remove signal redundancies and sounds that cannot be heard by humans. Some of the basic coding processes include waveform coding, perceptual coding and voice coding.



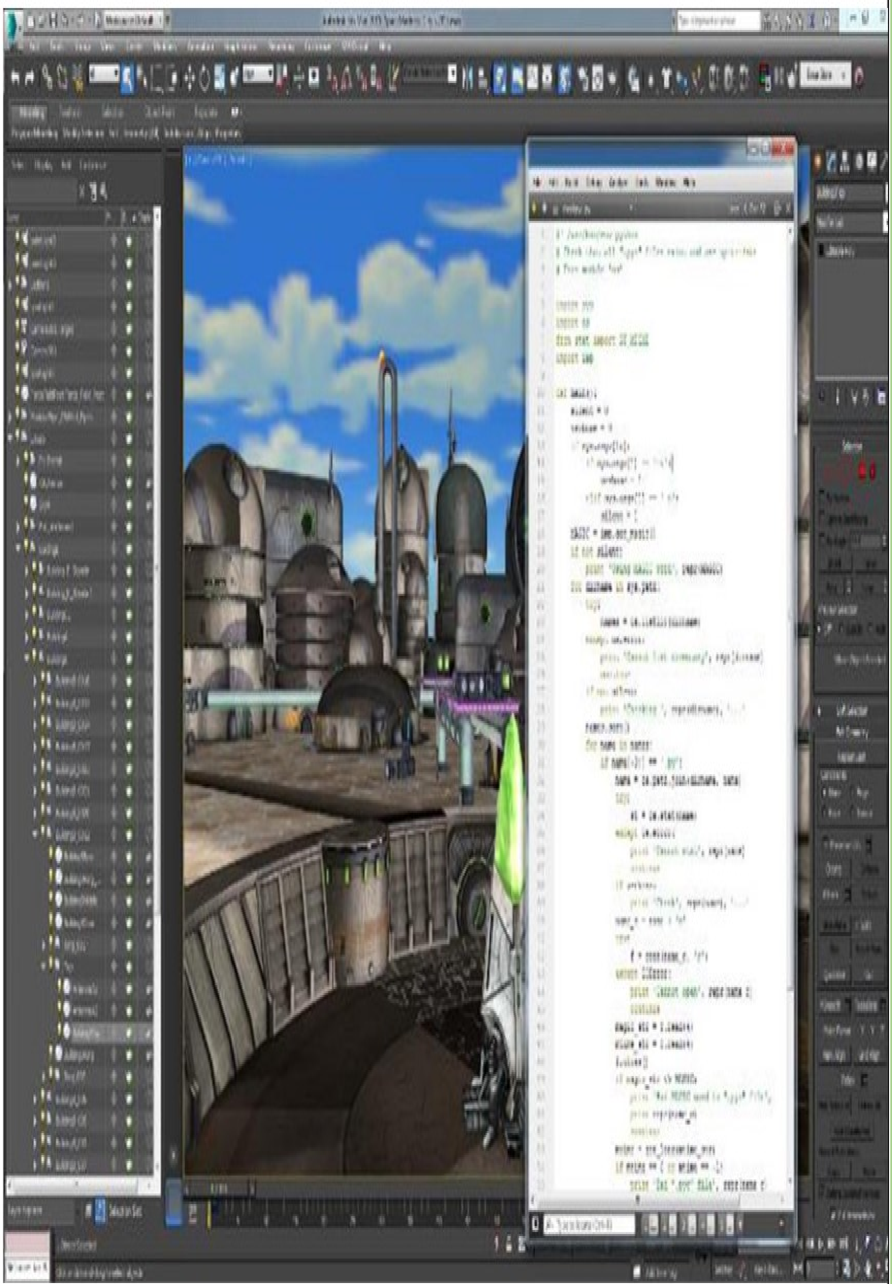
The frequency range of the human ear is from about 20 Hz to about 20,000 Hz, but the ear's sensitivity to sound is not uniform. It depends on the frequency, and experiments indicate that in a quiet environment the ear's sensitivity is maximal for frequencies in the range 2 kHz to 4 kHz. Figure a shows the hearing threshold for quit environment.



(a)

It should also be noted that the range of the human voice is much more limited. It is only from about 500 Hz to about 2 kHz. The existence of the hearing threshold suggests an approach to lossy audio compression. Just delete any audio samples that are below the threshold. Since the threshold depends on the frequency, the encoder needs to know the frequency spectrum of the sound being compressed at any time.

Animation & Text



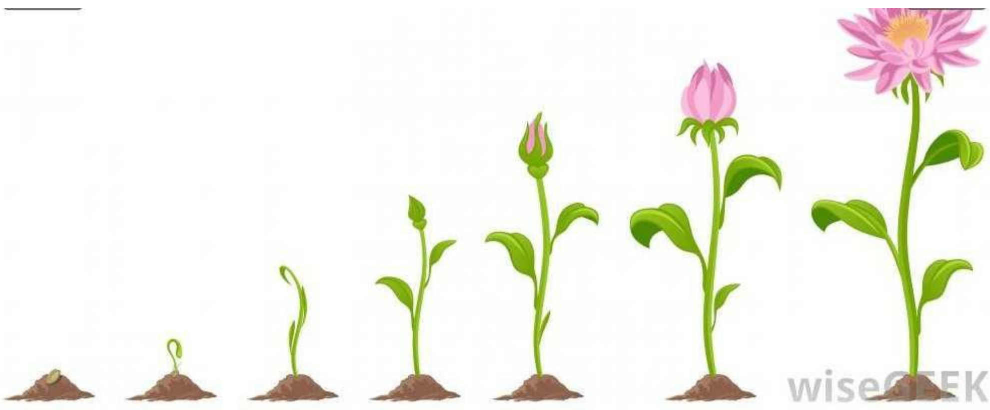
By

Dr. Amal Sufuih Ajrash

Animation

1. Introduction

- Animation is the art of creating an illusion of movement from a series of still drawings, Animation may be defined as the creation of moving pictures one frame at a time (**frame is a single complete image out of the sequence of images comprising an animation.**); the word is also used to mean the sequences produced in this way. Throughout the twentieth century, animation has been used for entertainment, advertising, instruction, art and propaganda on film, and latterly on video; it is now also widely employed on the World Wide Web and in multimedia presentations.
- Animation makes static presentations come alive. It is visual change over time and can add great power to your multimedia projects and web pages. Many multimedia applications for both Macintosh and Windows provide animation tools.



Animation involves displaying still images, one after another, to create the illusion of movement. The following images could be shown sequentially to give the illusion of a fast-growing flower.
Image 1 of 4

- Depending on the size of the project, you can animate the whole thing or you can just animate parts of it.
- Visual effects such as wipes, fades, zooms, and dissolves, available in most authoring packages, are a simple form of animation.
- Animation is more than wipes, fades, and zooms. Animation is an object actually moving across or into or out of the screen; a spinning globe of our earth; a car driving along a line-art highway Animations were the primary source of dynamic action in multimedia presentations.
- Animator is an artist who creates multiple images, known as frames, that give an illusion of movement called animation when displayed in rapid sequence.
- With animation, a series of images are changed very slightly and very rapidly, one after the other, seemingly blending together into a visual illusion of movement.

- Digital display video builds 24, 30, or 60 entire frames or pictures every second. Movies on film are typically shot at a shutter rate of 24 frames per second.

2. Traditional Methods of Animation

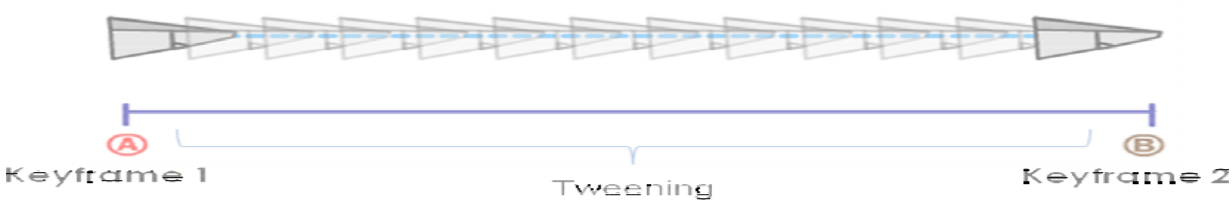
Before the advent of computer animation, all the frames in animation had to be drawn by hand. Considering that each second of animation contains 24 frames (film), one can imagine the tremendous amount of work that had to go into creating even the shortest of animated films! Different techniques were developed for creating animation by hand. These are:

Key frames: these are the first and last frames of an action, these are the key points of the animation.

Last frames First frames

Cel animation: an animation technique made famous by Disney, uses a series of progressively different graphics on each frame of movie film, which plays at 24 frames per second). A minute of animation may thus require as many as 1,440 separate frames, and each frame may be composed of many layers of cels. The term cel derives from the clear celluloid sheets that were used for drawing each frame, which have been replaced today by layers of digital Cel animation art work begins with keyframes.

The series of frames in between the keyframes are drawn in a process called tweening. **Tweening** is an action that requires calculating the number of frames between keyframes and the path the action takes.



Rotascoping: is the process of copying images from moving video to create animation, one disadvantage of rotascoping is that you have to get the exact video you want to animate. All these animation techniques are great, but when they are most useful is when they are all used together. Cel animation by itself would not help much if it was not for the key frames and the ability to distribute the workload across many people.

3. Process of Animation

In general, an animation sequence is designed with the following steps:

1. Story board layout—Give the outline of the action. Define the motion sequence as a set of basic events (it can also contain rough sketches).
2. Object definitions—Define the movement for each object.
3. Key frame specifications—Give the detailed drawing of the scene of animation at a particular time.
4. Generation of in-between frames—Define the number of intermediate frames between key frames.

4. Computer Animation

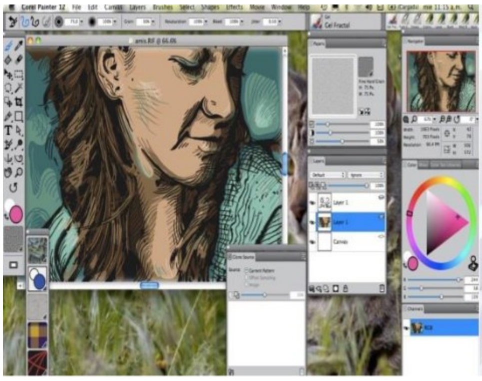
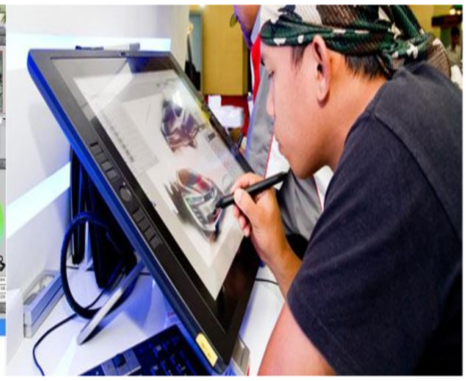
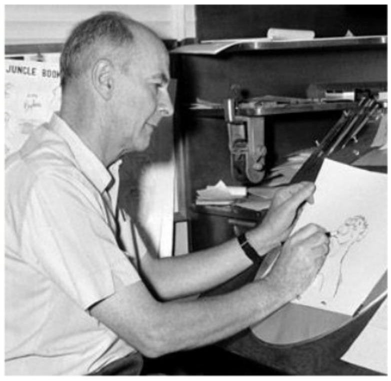
Computer animation programs typically employ the same logic and procedural concepts as cel animation and use the vocabulary of classic cel animation—terms such as layer, key frame, and tweening. The primary difference among animation software programs is in how much must be drawn by the animator and how much is automatically generated by the software.

Using appropriate software and techniques, you can animate visual images in many ways. The simplest animations occur in two-dimensional (2-D) space; more complicated animations occur in an intermediate "2½-D" space (where shadowing, highlights, and forced perspective provide an illusion of depth, the third dimension); and the most realistic animations occur in three-dimensional (3-D) space.

2D animation

In 2-D space, the visual changes that bring an image alive occur on the flat Cartesian x and y axes of the screen. A blinking word, a color-cycling logo (where the colors of an image are rapidly altered according to a formula), or a button or tab that changes state on mouse rollover to let a user know it is active are all examples of 2-D animations. These are simple and static, not changing their position on the screen.

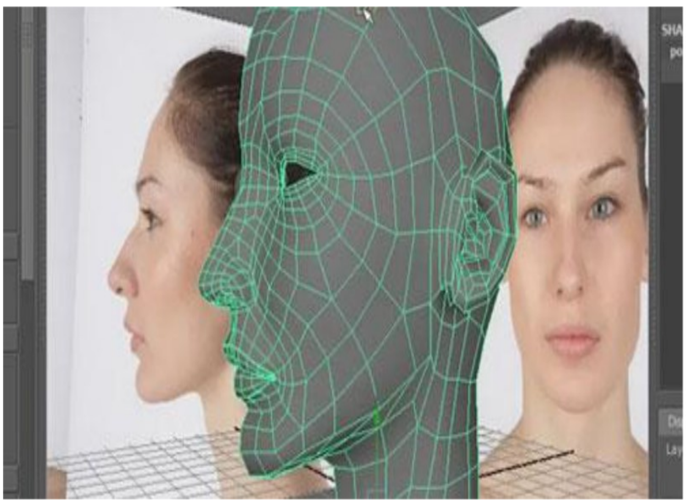
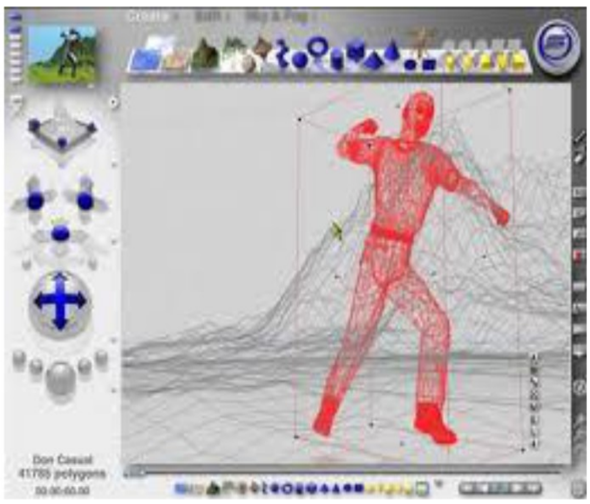
Path animation in 2-D space increases the complexity of an animation and provides motion, changing the location of an image along a predetermined path (position) during a specified amount of time (speed). Authoring and presentation software such as Flash or PowerPoint provide user-friendly tools to compute position changes and redraw an image in a new location, allowing you to generate a bouncing ball. Combining changes in an image with changes in its position allows you to "walk" your corporate mascot onto the stage.

In 2½-D animation, an illusion of depth (the z axis) is added to an image through shadowing and highlighting, but the image itself still rests on the flat x and y axes in two dimensions. Embossing, shadowing, beveling, and highlighting provide a sense of depth by raising an image or cutting it into a background. Zaxwerks' 3D Invigorator, for example, provides 3-D effects for text and images and, while calling itself "3D," works within the 2-D space of image editors and drawing programs such as Adobe Illustrator, Photoshop, Fireworks, and After Effects.

3D animation

In 3-D animation, software creates a virtual realm in three dimensions, and changes (motion) are calculated along all three axes (x, y, and z), allowing an image or object that itself is created with a front, back, sides, top, and bottom to move toward or away from the viewer, or, in this virtual space of light sources and points of view, allowing the viewer to wander around and get a look at all the object's parts from all angles. Such animations are typically rendered frame by frame by high-end 3-D animation programs such as NewTek's Lightwave or AutoDesk's Maya. Today, computers have taken the handwork out of the animation and rendering process, and commercial films such as Shrek, Coraline, Toy Story, and Avatar have utilized the power of computers. The main difference is the tools that are used to create animations, the effort and the price.

5. Adding Audio to Animation

Packages like Aawin, Animator player, and Adobe premier can be used to attach audio to animation. The steps for adding audio to animation are:

1. Creating the animation.
2. Selecting audiofile (midi, wav, or any other).
3. Editing audiofile (to fit the animation).
4. Linking the audio file to the animation by using either a package (for example Adobe Premier) or a programming language.

6. Animation File Formats

- Some file formats are designed specifically to contain animations, so they can be ported among applications and platforms with the proper translators. Those formats include Director (.dir and .dcr), AnimatorPro (.fli and .flc), 3D Studio Max (.max), GIF89a (.gif), and Flash (.fla and .swf).
- Because file size is a critical factor when downloading animations to play on web pages, file compression is an essential part of preparing animation files for the Web. A Director's native movie file (.dir), for example, must be preprocessed and compressed into a proprietary Shockwave animation file (.dcr) for the Web. Compression for Director movies is as much as 75 percent or more with this tool, turning 100K files into 25K files and significantly speeding up download/display times on the Internet.

- svg (scalable vector graphics) file, where graphic elements can be programmed to change over time

Text

1. Introduction

Words and symbols in any form, spoken or written, are the most common system of communication. They deliver the most widely understood meaning to the greatest number of people accurately and in detail. Because of this, they are vital elements of multimedia menus, navigation systems, document titles, keyword lists, and content.

- **Dual nature of Text**
  - Visual representation of language (content).

    Need to relate bit patterns stored in a computer's memory or transmitted over a network to symbols of a written language.
  - Graphic element (appearance).

    Precise shapes of characters, spacing and layout (typography).
- Each abstract character may have many different graphic representations.
- Abstract characters are grouped into alphabets.

The difference between the content and appearances of texts are:

It is convenient to distinguish between the lexical content of a piece of text and its appearance. By content we mean the characters that make up the words and other units, such as punctuation or mathematical symbols. The appearance of the text comprises its visual attributes, such as the precise shape of the characters, their size, and the way the content is arranged on the page or screen. For example, the content of the following two sentences is identical, their appearance is not:

The authors of digital multimedia's book are Nigel Chapman & Jenny Chapman.

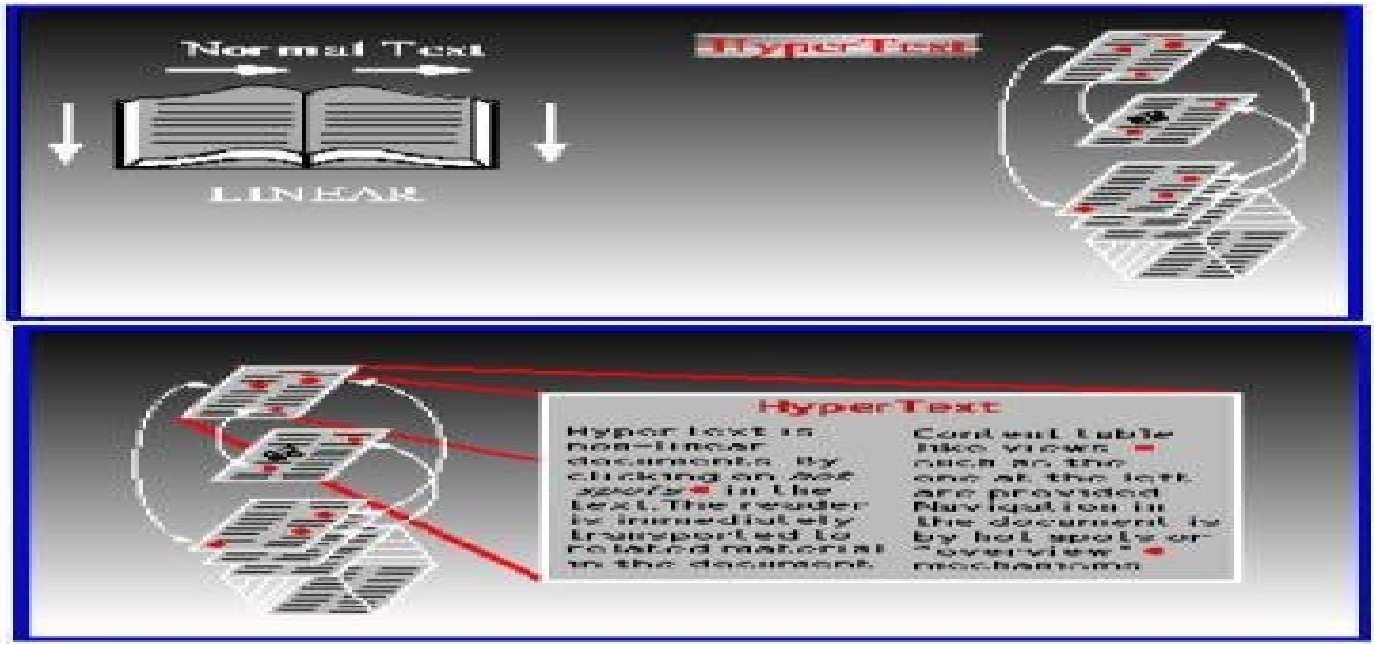
THE AUTHORS OF DIGITAL MULTIMEDIA'S BOOK ARE NIGEL CHAPMAN & JENNY CHAPMAN.

We could say that the content is the part of a text that carries its meaning or semantics, while the appearance is a surface attribute that may affect how easy the text is to read, or how pleasant it is to look at, but does not substantially alter its meaning.

Multimedia products depend on text for many things:

1. To explain how the application work.
2. To guide the user in navigating through the application.
3. Deliver the information for which the application was designed.

Texts consist of two structures as shown in figure below Linear and Non-Linear nonlinear text structure.



2. Hypertext Systems

a. A system in which words are keyed or indexed to other words in a non-linear way.

b. Using hypertext systems – Information management and hypertext programs present electronic text, images, and other elements in a database fashion.

c. Software robots visit web pages and index entire web sites.

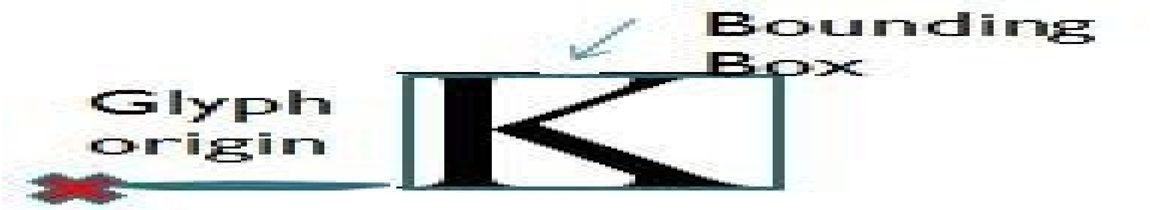
d. Hypertext databases make use of proprietary indexing systems.

e. Server-based hypertext and database engines are widely available.

f. Some of Hypertext systems uses: Electronic publishing and reference works and Electronic catalogs.

3. Font Typography

Typography is the process of selecting and arranging typefaces, sizes, and spacing requirements for the layout of multimedia or the Web. Typography gives a page a certain personality (formal or informal, modern or classic). A **glyph** is a graphic representation of a character's shape. A character may be represented by many different glyphs. Glyphs are arranged into collections called fonts. Fonts are stored in specified locations on a computer system, may be embedded in documents. If font is not embedded, document may not display properly on systems where that font is not installed.



Rasterization

Is a process that converts the letter **A** from a mathematical representation to a recognizable symbol displayed on the screen or in printed output. the computer must know how to represent the letter using tiny square pixels (picture elements), or dots. It does this according to the hardware available and your specification, from a choice of available typefaces and fonts.

4. HTML Documents

- Since the explosion of the Internet and the World Wide Web, text has become more important than ever. The native language of the Web is HTML (Hypertext Markup Language) the standard document format used for displaying text pages on the Web. Originally designed to display simple text documents on computer screens, with occasional graphic images thrown in. Academic papers, magazine articles, complex instruction manuals, and even the contents of entire books are now available for reading with a web browser. Also a built-in function can be added that links, with a click of the mouse, selected words and phrases to other related and perhaps more detailed material (the ―hypertext‖ part of HTML).
- HTML documents are marked using *tags* –annotations that control a text document's layout and formatting or indicate its structure.
- An advanced form of HTML is DHTML stands for Dynamic Hypertext Markup Language.
- DHTML uses Cascading Style Sheets (CSS).
- In visual markup, tags specify aspects of the text's appearance. In structural markup, they identify logical elements, such as paragraphs, lists or headings.

Multimedia and Networking

Dr. Amal S. Ajrash

1. Introduction of Delivering Multimedia

The network and transport protocols do no more than deliver packets of data more or less reliably to their designated destinations. Higher-level protocols must run on top of them to provide services suitable for distributed multimedia applications. We will describe many protocol such protocols: HTTP, which is the basis of the World Wide Web, and **RTSP (Real Time Streaming Protocol)**, a protocol designed to control streamed media. Our description is intended to show what is involved in mapping the requirements of some kinds of distributed multimedia applications on to the transport facilities provided by the protocols.

2. Communication Service Types

1. **Broadcast:** single source many users like TV**,** one way distribution which means the transmission is from one device to all device in the network.
2. **Multicast**: like broadcast but users sign to a session, one way connection means one device to some device.
3. **Retrieval, unicast**: a symmetric connection, users receive much more than send the transmission is from one device to another one.

3.Streaming

Streaming means audio and video data are send synchronously in time and will reach user preserving their time organization. No data loss or delays are allowed. Example telephone network in multimedia streaming is absolutely necessary.

4. Key features of the applications of streaming audio and video data have the following

1. Stored media, the contents have been prerecorded and is stored at the server. So, a user may pause, rewind, or fast-forward the multimedia contents. The response time to the above actions should be in the order of 1-10 seconds.

2. Streaming, a user starts payout a few seconds after it begins receiving the file from the server. So, a user payout the audio/video from one location in the file while it is receiving later parts of the file from the server. This technique is called streaming and avoids having downloaded the entire file before starting payout.
3. Continuous payout, once payout begins; it should proceed based on the original timing of the recording. This requires high quality on the end-to-end delay.

5. Streaming stored audio and video file (client store its media file to server)

- The clients request audio/video data stored at servers.
- Servers send the data into a socket connection for transmission Both TCP and UDP socket connections have been used in practice.
- The data are segmented and the segments are encapsulated with special headers appropriate for audio/video traffic.
- The real time protocol (RTP) is a public-domain standard for encapsulating such segments.
- Audio/video streaming applications usually provide user interactivity which requires a protocol for client/server interaction. The real time streaming protocol (RTSP) is a public-domain protocol for this purpose.
- Clients often request data through a Web browser. A separate helper application (called media player) is required for paying out the audio/video. Well used helpers include RealPlayer and Media Player

6. Access audio/video through Web server (client download media file from server)

- The stored audio/video files can be delivered by a Web server or by an audio/video streaming server.
- When an audio file is delivered by a Web server, the file is treated as an ordinary object in the server's file system, like HTML and JEPG files.
- To get the file, a client establishes a TCP connection with the server and sends an HTTP request for the object.

- On receiving the request, the Web server encapsulates the audio file in an HTTP response message and sends the message back to the TCP connection. It is more complicated for the video case because usually the sounds (audio) and images are stored in two different files.
- Then, a client sends two HTTP requests over two separate TCP connections and the server sends two responses, one for sounds and the other for images, to the client in parallel. It is up to the client to synchronize the two streams.

7. Network Performance Terms and Definitions

Best effort service is a network service in which the network does not provide any guarantees that data is delivered (meaning that they obtain unspecified variable bit rate and delivery time depending on the current traffic load), the following features used to measuring he Quality of Service **QoS** .

a. **Packet loss:** Measure of network reliability with respect to loss or de-sequencing of (unreliable) packet transport, taken in %. IP provides the best-effort service but does not guarantee the delivery of packets. Packets may be discarded due to congestion.

b. **End-to-end delay**: delay is the transit time for a packet within the network. IP does not guarantee the end-to-end delay either. The time for transmitting a packet may vary due to the conditions of the network. Also, in order to guarantee delivery, positive acknowledgment and retransmission are used in TCP. The cost of realizing the reliable transmission in TCP is a longer end-to-end delay.

c. **Packet jitter:** Delay variation in packet arrival. since the end-to-end delay for each packet may depend on the conditions of the network, the delays of packets in the same packet stream may vary. Especially, the packets may arrive at the receiver in the wrong order.

d. **Bandwidth** : Average throughput capacity of the network.

e**. End-to-End Latency**: Time needed for a packet to travel between application end-points. stream Latency -Relative latencies between synchronized streams (e.g. audio and video)

9. Removing jitter at the receiver for audio

In applications like Internet phones or audio-on-demand, it is up to the receiver to remove the jitters. Common techniques used include a sequence number, timestamp, and delaying payout. The sender can put a sequence number on every packet sent and the receiver can use the sequence number to recover the correct order of the received packets. The timestamp is similar to the sequence number, the sender stamps each packet with the time at which the packet is generated. In order to get the correct order from the sequence number and timestamp for a sequence of packets, the receiver needs to receive all of the packets in the sequence. Payout delay is used for this purpose.

The payout delay should be long enough to receive all packets in a subsequence of packets that can be paid. On the other hand, the delay should be short enough so that the user will not notice the delay. The payout delay can be either fixed or adaptive.

10. Packet Switching IP internet

Packet switching data are organized in packets, each packet carries addresses In packet switching, packets are flowing in little controlled way making streaming problematic, and interactivity is fast The difference is that only circuit switching fully guarantees streaming but it is expensive In standard packet switching there is no network resource reservation before the packets. the basic method of transferring packets on the internet is the:

UDP - User Datagram Protocol only tries its best to deliver datagrams. It does not offer reliable delivery, so it has less overhead than TCP, which makes it more suitable for streaming video and audio.

IP - Internet Protocol only provides a mechanism for getting datagrams from their source to their destination through a network of networks. Each host is identified by a unique IP address.

11. Protocols for Real-Time Interactive Applications

a. Real-Time Transport Protocol (RTP)

RTP runs on top of UDP, adding features for synchronization, sequencing, and identifying different payloads. In this protocol, each packet gets a timestamp and number when it is sent. On the receiving side, it is possible to detect if packets are lost and what are the packets delays the complete packet looks like this.

RTP protocol Provides end-to-end network transport functions suitable for applications transmitting real-time data, such as audio or video. It is Independent of the underlying transport and network layers.

b. **Real-Time Streaming Protocol (RTSP)**

RTSP is a protocol that allows a media player to control the transmission of a media stream. The control actions include pause/resume, repositioning of playback, fast-forward, and rewind. RTSP messages use a different port number from that used in the media stream and can be transmitted on UDP or TCP.

c. **Session Initiation Protocol (SIP)**

- application-layer control protocol for creating, modifying, and terminating sessions with one or more participants
- Internet multimedia conference, Internet telephone calls, and multimedia distribution.
- supports name mapping and redirection services
- Internet telephony gateways that connect Public Switched Telephone Network (PSTN) parties can also use SIP to set up calls.
- invite both persons and" robots", such as a media storage service.
- can be used to initiate sessions as well as invite members to sessions that have been advertised and established by other means

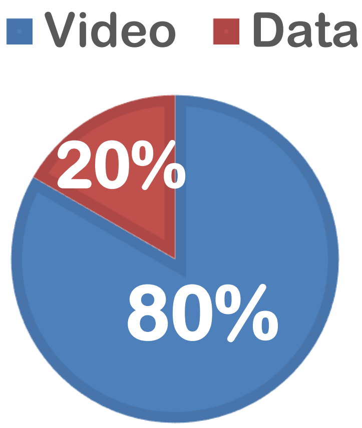
d. **Hypertext Transfer Protocol (HTTP)**

- a Simple protocol designed for the fast transmission of hypermedia between Web servers and clients (e.g. browsers).
- It is a very simple way to stream media files.
- Interaction between a Web client and server over HTTP takes the form of a disciplined conversation, with the client sending requests which are met by responses from the server.

e. **Transmission Control Protocol TCP/IP**

TCP is layered on top of IP to provide reliable delivery of sequenced packets, using a system of acknowledgments with a sliding window of unacknowledged packets.TCP uses transport addresses, consisting of an IP address and a port number, to provide connections between programs running on different hosts.

12. Quality of Service and Quality of Experience



Digital video is driving the future of Communication and Internet. Internet traffic is dominated by video content, there is a need for technology which reduces bandwidth. Video service providers aim at cutting down cost but ensuring good video quality. Delivering video streams at high quality and minimal delay is a major technical challenge.

12.1 Methodologies for assessing video QoE

In principle, assessment of video Quality of Experience (QoE) must be performed using subjective tests, with metrics such as the mean opinion score (MOS). However, it is also possible to estimate QoE based on objective measurement of key parameters, and associated quality estimation models

1.**Subjective Testing**

Subjective testing needs more resources and effort because it requires human subjects, and is not so convenient in a live-service setting.

2.Objective measurement and QoE calculation

Objective measurement and QoE calculation is generally much faster and more convenient, especially in a live-service environment, but the accuracy of the QoE estimation depends on the complexity of the models, and an understanding of the important human factors.

12.2 Quality of Experience (QoE)

QoE assessment methods are very useful for in-service quality monitoring and management, as well as codec optimization, codec selection and quality design of networks or terminals.

- QoE is measure both network and content impairments (application-driven).
- QoE is measure how end-user perceives service issues (user-oriented).
- QoE is end-to-end quality measurement.
  - Cover different impairment sources.
  - Identify problem causes.