Unit 1: Introduction to Computers.
(Multimedia)

Multimedia (Introduction)

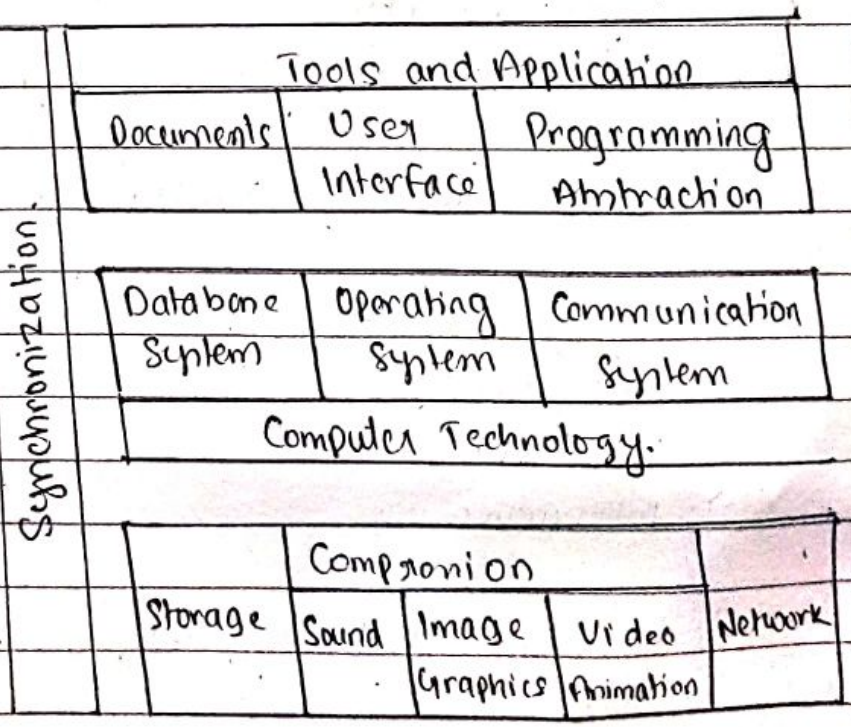
→ Multimedia is the field concerned with the computer controlled integeration of text, graphics, drawings, still and moving images (video), animation, audio and so on.

Multimedia: Global structure of Multimedia.

1] Application domain:
→ provides function to user to develop & present multimedia projects. This includes software tools & multimedia project development methodology.

2] System Domain:
→ It supports for using the functions of the device domain.
Ex: Operating system, communica and database system.

3] Device Domain:
→ Basic concept and skill for processing various multimedia elements and for handling physical device.



- Multimedia Applications

→ Digital video Editing.
→ Electronic newspaper / magzines.
→ WWW
→ Home shopping
→ Interactive TV.
→ Video Conferencing
→ Gaming
→ Virtual Reality
→ Video on demand. etc

- Characteristics of Multimedia.

→ Computer Controlled
→ Integrated
→ Digitally
→ Interactive

- Components of Multimedia.

→ Capture devices,
→ Storage devices,
→ Communication Network
→ Computer Systems
→ Display Devices.

- Challenges of Multimedia System.

→ Distributed Network.
→ Automatically analyzing, indexing and organizing information in audio, image and video is much harder than from text.
→ Synchronization
→ Sequencing
→ Inter-media scheduling.
→ Multimedia involves many different research areas that needs more efficient algorithms and hardware platforms.

Unit 2: Sound / Audio System:

Sound is a physical phenomenon produced by vibration of matter and is transmitted as waves.

Sound waves are characterized via;

1] Period
2] Frequency [pitch] inversly
3] Amplitude [loudness]
4] Bandwidth.

1] Period:

It is the interval at which periodic signal repeats regularly

2] Frequency:

→ It is the ~~ratio of~~ reciprocal value of period.
→ It is the reciprocal value of period.

i.e. $f = 1/p.$

→ Unit of frequency is Heartz or kiloheartz (Hz / KHz)

According to frequency sound can be distributed as.

Sound	frequency.
Infra-Sound	0-20 Hz
Audible Sound (human)	20 Hz - 20 kHz
Ultra-sound	20 kHz - 1 GHz
Hyper-Sound.	1 GHz - 10 THz.

Human ear is most sensitive at 600-6000 Hz

iii] Amplitude

→ It is the loudness of a sound.

iv] Bandwidth

→ It is range of frequency a device can produce ~~on~~ that a human can hear.

* Computer Representation of sound:

→ To store sound wave in a computer, samples of wave are taken and each sample is represented by a number or the code.

→ For computer representation of sound it requires the digitisation of sound.

Digitization of sound.

→ It is the process of converting analog signals to a digital signal.

→ Digitisation of sound consist of following 3 steps

i] Sampling rate,

ii] Quantization and.

iii] ~~Encoding~~ Sound Hardware.

1] Sampling Rate :

→ The rate at which continuous waveform is sampled is called Sampling (Rate).

→ No. of samples per second is called sampling rate.

2] Quantization.

→ Sampling of amplitude is called quantization

→ Procen of representing amplitude of each sample as integer or number is called quantization.

3] Saund Hardware

→ Devices that are connected to ADC and DAC (Analog to Digital Converter and Digital to Analog converter) for input and output of audio to the computers.

* MUSIC and SPEECH.

MIDI [Musical Instrument Digital Interface]

→ It is communication standard developed in early 1980s for electro-inst & computers.

→ MIDI is an equipment that plugs directly into the computer's serial port allowing transmission of music signals.

→ MIDI is the Interface between electronic musical instrument and computers.

→ MIDI is universally accepted standard for communicating information about musical performance by digital means.

→ Common MIDI devices includes

i) electronic music synthesizers
ii) Modules
iii) (MIDI devices in common) sound cards.
iv) drum machines

* MIDI DEVICES.

→ As MIDI is interface between musical instrument & computer. It consist of lots of musical devices some of them are

SyCoSeNe

I] Synthesizer
II] Controllers
III] Sequencers
IV] Networks.

I] Synthesizer [incorporates only with MIDI software]

→ Synthesizer is a component that generate sound based on the MIDI menage.

TYPES of MIDI synthesizer [Examples]

I] Integrated keyboard and Synthesizer
II] Rack-Mounted synthesizer
III] Drum Machine.

→ Synthesizer are not necessarily integrated with a keyboard. such as Rack-mounted MIDI synthesizer.

→ MIDI Synthesizer are used in creating pitched and percussion sound.

II] Controllers. [Incorporates with only MIDI software]

→ MIDI controllers are the devices for manipulating the generated MIDI software manages.

→ MIDI controller can take the form of almost all acoustic or electronic instrument such as keyboard, guitar, drums, sets, drum pads.

→ MIDI controller never synthesize or generate audible music

→ MIDI controller are commonly integrated with synthesizer to maximize the application.

Example: Ztar, Alesis Q,

III] Sequencer [Incorporates with both MIDI software & hardware]

→ MIDI sequencer is electronic device incorporating with both MIDI software and MIDI hardware.

→ It is used for storing and replaying MIDI software manage sequence.

Functions of MIDI Sequencer.

1] Record / store MIDI software menage sequences.
2] Replays the MIDI software menage sequence in approp. timing.
3] provides some sort of editing capabilities.

IV] Networks.

2 → MIDI Network is a combination of hard ware and software to interconnect group of MIDI devices such an synthesizer, controller and sequencer.

3 → There are one or multiple ports or MIDI ports to connect dozens of MIDI devices into a network.

1 → In general the MIDI network is just a generalization of a device called Patch Bay.

4 → All the MIDI devices consist of one or multiple ports (i.e. MIDI-In, MIDI-Out) that are used to interconnect different devices so that a MIDI network could be formed.

* MIDI Menages;

→ MIDI menage transmit Infoomation between different MIDI devices. which determine the kind of munical event to be pamed from device to device.

→ MIDI menage commit of
(i) Status byte
(ii) data byte

1) Status byte:
→ fiorst byte of MIDI menage that describes the menage kind.

11) Data byte
→ data byte follows the status byte

TYPES OF MIDI Menage.

1] Channel Menage
2] Sytem Menage.

1] Channel Menage.

→ Channel Menage go only to the specific device carrying certain infomation.

TYPES OF Channel Menage.
i) Channel Voice Menage
ii) Channd Mode Menage.

i) Channel Voice Menage.

→ Channel voice menage send actual performance data between MIDI devices describing keyboard action, control action and control panel changes.

→ Channel voice menage describes music by defining pitch, amplitude, timbre, duration and other sound qualities.

→ Each menage has at leant one and usually two data bytes that accompany/helps statin byte to describe sound quality.

Example: Note On, Note Off, channel Pressure, Control change etc

ii) Channel Mode Menage.

→ Channel Mode Menage determine the way that a recieving MIDI device respond to channel voice menages.

→ MIDI channel receiving mode is set for different MIDI devices.

Example: Local Control, All Notes OFF, Omni Mode OFF etc.

2] System Menage.

→ System Menages go to all the MIDI devices because no channel numbers are specified.

TYPES OF SYSTEM Menage

i) System Real-Time Menage,
ii) System common menage and
iii) System exclusive menage.

i) System Real-Time Menage:

→ These are very short and simple consisting of single byte.
→ They carry data with them to synchronise the timing of MIDI devices in performance.
→ Therefore, It is important that these menage are sent precisely
→ To avoid delays, these menage are sent in middle of the menage

Example: System Reset, Timing clock.

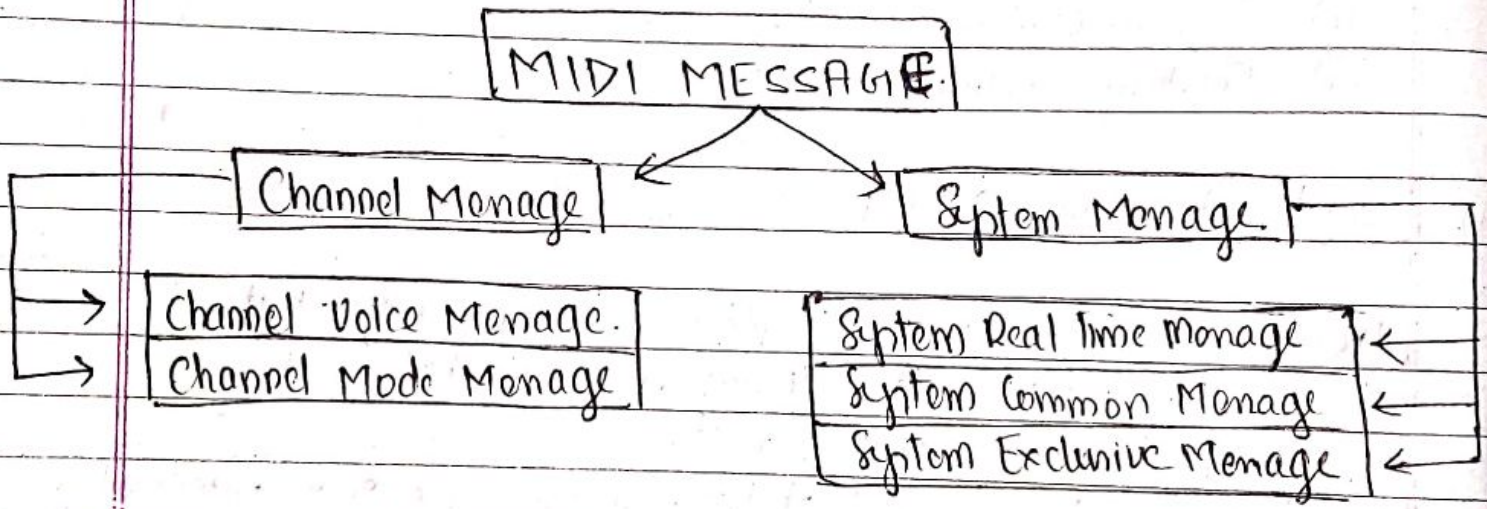
ii) System Common Menage:

→ These are commands that prepare sequencers and synthesizers to play a song.
→ Various menage enables to select the song, find common starting place in song and tune synthesizer if needed.

Example: Song Select, Tune Select.

iii) System Exclusive Menage:

→ It allows MIDI manufacturers to create customized MIDI menage to send between MIDI devices.
→ This menage starts with system exclusive menage where manufacturer is specified and ends with end-of-exclusive menage.



* MIDI and SMPTE timing Standard:

→ MIDI reproduces traditional note length using MIDI clocks. which are represented through timing clock menages.
→ Using MIDI clock, reciever can synchronize with clock cycles of the sender.
→ for example, a MIDI clock helps keep separate sequencer in the same MIDI system playing at the ~~tempo~~. same tempo.
→ In order to keep standard timing reference, the MIDI specifications state that 24 MIDI clock equal one quarter note.

→ As an alternative, SMPTE timing standard (Society of Motion Pictures and Television Engineers) can be used

→ SMPTE, originally developed by NASA as a way to mark in coming data from different tracking stations so that receiving computers could tell exactly what time each piece of data was created.

→ The film and video promoted by SMPTE, the SMPTE timing standard acts as very precise clock that stamps a time reading on each frame and fraction of time, counting from begining of a film or video.

→ SMPTE, for more precise reading of time, the SMPTE consists of hours: minutes: seconds: frames; bits (example; 30 frames per second), uses a 24 hour clock that counts from 0 to 23.

→ for even more precise time SMPTE breaks each frame into 80 bits.

→ SMPTE divides time into segments as small as one twenty-five hundredth of a second.

* MIDI Software.

→ Once a computer is connected to MIDI system, a variety of MIDI software applications can run on it.

→ The software application generally falls into four major categories.

- Music recording and performance application.
- Musical notation and painting application.
- Synthesizer patch editors and librarians
- Music education applications.

* Speech Generation.

Basic notions.

1] Speech consist of fundamental frequency as it's component as it is used to present a voiced sound.
(allophone)

2] It consist of morph, phone as smallest speech unit.

3] A voiced sound is generated via vocal cords.

4] It consist of vowels, comtants as it's component.

* Reproduce Speech Output/Generation.

→ The eary way to reproduce the speech output is to use the previously recorded speech and playing it back.
→ Speech can be recorded as PCM (Pulse code Modulation) or using the procon of digitization of sound.

Speech Output / Generation
→ To reproduce the speech output we can use following techniques.

I] Time dependent sound Concatenation / Series.
II] Frequency dependent sound concatenation / series.

* Speech Analysis:

→ The primary goal of speech analysis is to correctly determine individual word.
→ Speech Analysis fails to achieve it's goal ~~when~~ due to following factors.
    1] Ambient Noise.
    2] Word ambiguity (example; there/their, bear/bare.
    3] Dialects and stress.

* Speech Recognition.
→ Capability of electronic device to understand words spoken.
→ A microphone records a person's voice and hardware ~~can~~ digitize it. finally a software ~~inpret~~ interprets with digitized sound as individual word.

* Speech Analysis;

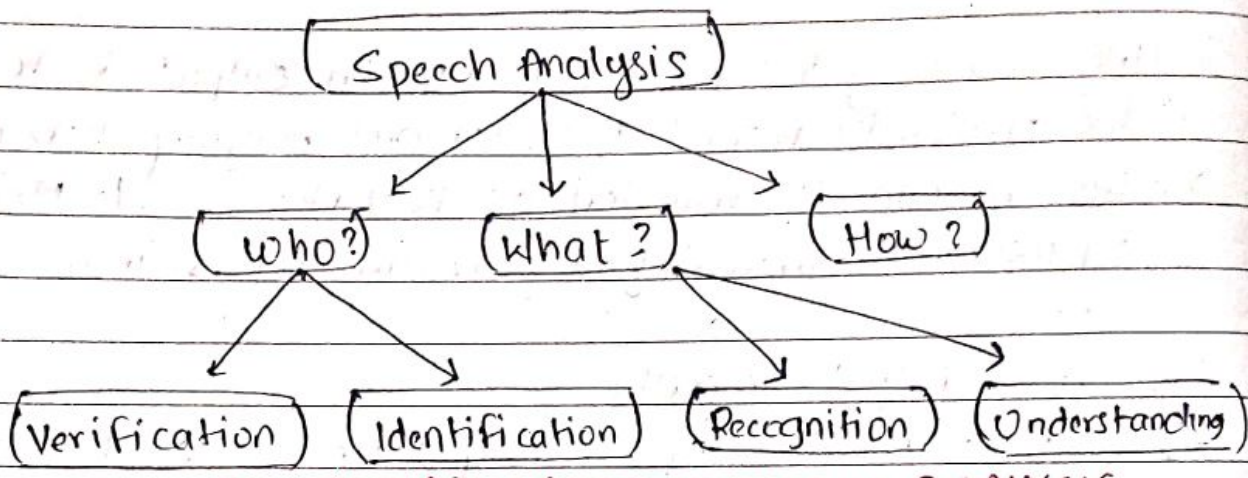


FIG: RESEARCH AREA OF SPEECH ANALYSIS

→ The diagram gives the idea regarding research area concerned with speech analysis.

1]. A human speaker has certain characteristics, hence speech analysis needs to recognize who is the speaker, ie. to know ~~identify~~ the speaker's verification and Identification.
- Computer identifies and verifies the speaker using an acoustic fingerprint which is digitally stored along with certain statement of a person.

Example: Company for speech Analysis verifies and identifies the employee with certain statement spoken by the him/her. The computer system verifies and Identify the statement spoken to match in it's database.

2] Another main task of Speech analysis is to deal with is what has been said. ie. to recognise and understand the speech itself.

3. final area of speech analysis tries to deal with "how the statement was said.

Example: Statement spoken sounds differently if a person is calm or angry.

This research field/area of speech analysis as "how" is also used in a lie detector systems.

* Speech Recognition.

→ Capability of electronic device to understand the spoken words.

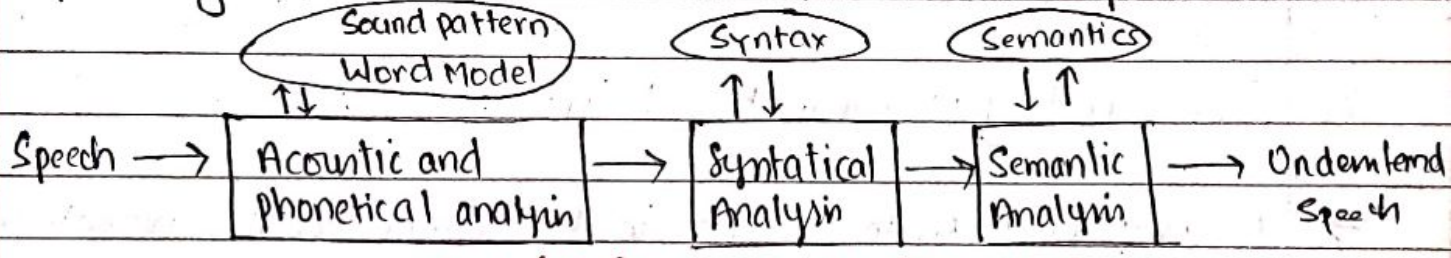


FIG: COMPONENTS OF SPEECH RECOGNITION.

1. Initially the speech generated wents for acountic and phonetical analysis.
2. Then goes for syntatical analysis so that the errors of previous step be recognised and syntatical analysis is made providing additional decision aid to produce recognised speech.
3. Third step deals with semantic analysis to the recognised speech that further filters the errors of previous and finally understable speech is generated.

TYPES OF SPEECH RECOGNITION SYSTEM.

[i] Speaker Dependent Recognition system.
[ii] Speaker Independent " ".

* Speech Transmision.

→ Procen of sending the speech/audio from sender to recicver with fundamental goal to provide the same speech/audio (sound quality) on was genorated at sender side.

1] Signal form Coding.

→ It is technique to achieve most efficient coding of the audio signal without considering speech property & parameter.

2] Source coding in parametrized system.

→ Parameterized system works with source coding algorithm.

→ It uses the speech/audio characteristics for data rate reductio

Example: Channel vo-coder.

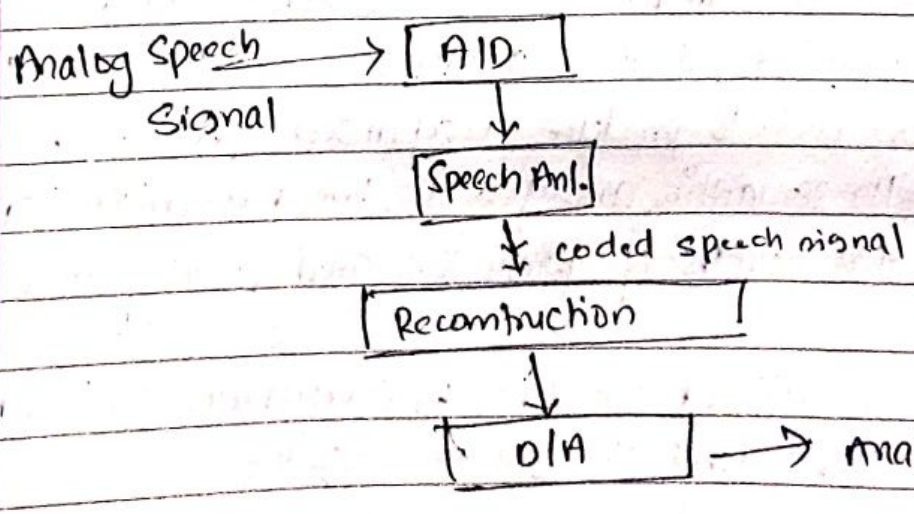


fig: Components of speech transmision via source coding in parametrized System.

3] Recognition / synthesis System.

→ Only the characteristics of the speech elements are transmitted.
→ Used in data rate reduction.

Speech Coding:
→ Application for data compression of digital audio signal

Unit III: Image and Graphics.

Digital Image.

→ An Image composed of pixels in an electronic / digital system. is said to be digital Image.
→ Each pixel is represented by numerical value.
→ Digital Image made with more pixels is more clear at basic level. Digital Image is represented by a matrix of numeric values.
→ Digital Image with higher dimension consist of higher no. of pixels and vice versa which is directly propotional to the quality of a Image.

(Image file format)

	Format	Description.
1]	JPEG [Joint Photographic Experts Group]	→ Common file format with highly supported format → Uses lossy comprenion → Quality degrade is not much noticeable after comprenion.
2]	GIF [Graphics Interchangeable format]	→ supports only 256 colors. → can be animated → can add transparency
3]	PNG [Portable Network Graphics]	→ combination of good features of JPEG and GIF → file size is bigger → not commonly used / used for special purpose.
5]	TIFF Tagged Image file format	→ Uncomprened Image / Loss less comprenion → Used by experts, high quality.

Image format:

→ Image format can be of two types.
1. Captured Image format.
2. Storage Image format

1] Captured Image format

→ Captured Image format is specified by two main parameters as

(I) Spatial resolution.
(II) Color encoding.

2] Stored Image format

→ format, while storing the Image

i) JPEG (Joint Photographic Expert Group)
ii) GIF (Graphics Interchangeable format)
iii) PNG (Portable Network Graphics)
iv) TIFF (Tagged Image file format)
v) Bitmap (.bmp developed by Microsoft for windows)
vi) EPS (Encapsulated PostScript)
vii) RAW (.raw, .cr2)

Image file format	Description.
1] JPEG (.jpg, .jpeg) Joint Photographic Experts Groups	• Common file format commonly used for digital cameras. • Compression: ~~None~~ Lossy • Best for: Web Images, Email, powerpoint. • Special Attribute: amount of comprenion can be allocated.
2] GIF (.gif) Graphics Interchangable format.	• Widely used for web graphics. and supports only 256 colors. • Compression: Lossless. • Best for: Web Images. • Special Attributes: can be animated can add transparency
3] PNG (.png) Portable Network graphics.	• combination of good features of JPEG and gif with bigger file size. and supports 16 M colors, used for special purpose. • Compression: Lossless. • Best for: Web Images. Special Attribute: can add transparency.
4] TIFF (.tif, .tiff) Tagged Image file format.	• widely used for high quality prints, professional publication • Comprenion: Lossless • Best for: high quality prnts. • Special Attribute: can add transparency.

Graphics Format:

→ Basically graphic image are specified through graphic primitives and their attributes, that may consist of lines, rectangles, circles, ellipses etc.

In general following are considered as graphic format.

1) Raster graphics (Bitmap graphics)
2) Vector graphics
3) Metafiles.

	Raster Graphics	Vector Graphics.
1)	Are composed of pixels.	• Stores images as mathematical representation. such as lines.
2)	Pixel dependent for the quality of an Image i.e. higher the no. of pixel ~~more~~ produces an image of high quality and vice-versa.	• Pixel Independent
3)	An image loses the quality while it's rescaled	• Rescaled without loss of quality of an Image
4)	Supports compression.	Usually don't support compression.

5)	large in size	• far smaller in size compared to Raster.
6)	Cost less	• Comparatively higher cost
7)	Commonly used file extensions; .JPG, .TIFF, .GIF, .PNG, .BMP	• Commonly used file extensions; .SVG, .EPS, .DWG, .DXF, .DGN.

3) Metafiles.

→ It is the developed or advanced version that allows both raster and vector graphics to reside within a same file.

→ It was developed for transfering of image data, & it is used within specific sector with specific application.

→ Some common extensions: CGM, WMF, EMF, WPG.

CGM → Computer Graphics Metafile.
WMF → Windows Metafile
EMF → Windows enhanced Metafile
WPG → WordPerfect Graphics Metafile.

Unit IV; Video and Animation.

* Video Signal Representation:

→ Video Signal Representation is the procen of presenting the visual signals. In conventional TV set or monitors the video signal is displayed using CRT (cathode Ray Tube)

→ An electron beam sweeps the screen from top to bottom beam carrying the corrosponding pattern information.

The major aspects of video signal representation are.

1) Visual Representation.
2) Transmission
3) Digitalization.

1) Visual Representation.

→ Visual Representation is one of major aspect of video signal representation which further Includes following Important measures.

a) Vertical Detail and viewing distance.

→ The geometry of the TV image depends upon aspect ratio. The conventional aspect ratio is 4/3.

$$\text{Aspect ratio} = \frac{\text{width}}{\text{height}} = \frac{w}{h}$$

b) Horizontal detail and Picture Width.

→ The picture width for conventional TV (T) is 4:3 of picture height.

$$\text{width} = \frac{4}{3} \text{ of picture height.}$$

c) Total Detail Content of Image.

→ i) Vertical resolution = no. of pixels in picture ht.
ii) NO. of pixel in width of picture = vertical resol X aspect ratio.

d) Perception of Depth.

→ PoD depends upon angular separation of ~~two~~ the image recieved by two eyes of the viewer.

→ In case of flat screen of TV, a considerable degree of depth is inferred from the perspective appearance of the object matter.

e) Luminance and Chrominance.

→ Color vision is achieved via Intensity of RGB in each portion/part of screen.

→ However, during the transmission of signal from camera to display, a different color encoding that uses luminance and two chrominance signals are used.

f) Temporal Aspects of Illuminance. | Illumination.

→ In contrast to continuous pressure, wave of an acoustic signal, a discrete sequence of individual pictures can be percieved/recognised as a continuous sequence.

→ To represent visual reality, following two conditions needs to be satisfied.

(1) Rate of repetition of images must be high enough to guarantee smooth motion from frame to frame.

(2) The rate must be high enough so that persistence of vision extends over intervals between flashes.

g) Continuity of Motion.

→ Continuity of motion is recognised at any frame rate faster than 15 frames per second,
→ Smooth video motion is achieved at 30 fps.
→ Movies, basically uses 24 frame per second.
→ NTSC (National Television Systems Committee) specified 30 fps for motion videos whereas European PAL system specified @ 25 fps.

h) Flickering
→ Periodic fluctuation of brightnen
i) Temporal Aspect of Video Bandwidth.

2) Transmission.

→ NTSC (National Television Systems Committee) is the oldest standard for transmission and reception of video signals.

- for transmission of video signal, it consist of one luminance and two chrominance signals.
- To encode color, a video signal is composed of three different signals.

→ Approaches for color encoding.

a) RGB Signal:

→ Consist of Red, Green and blue color. from the combo of these three colors different other colors are coded.

b) YUV Signal.

→ Human perception for brightnen is more sensitive compared to other chrominance information.

→ YUV signal is technique that separates brightness information (luminance - Y) from color information (chrominance channels U and V).

Component division for YUV signal;

$$Y = 0.30R + 0.59G + 0.11B$$

$$U = (B-Y) \times 0.493 \qquad V = (R-Y) \times 0.877.$$

c) YIQ Signal.

→ YIQ Signal is similar to YUV signal with following NTSC format.

$$Y = 0.30R + 0.59G + 0.11B$$
$$I = 0.60R - 0.28G - 0.32B$$
$$Q = 0.21R - 0.52G + 0.31B$$

3) Digitalization.

→ It consist of sampling the gray (color) level in the picture at M x N array of points. and converts it to bit streams.

→ The next step in creation of digital motion video is to digitize pictures in time and get a sequence of digital images per second that approximates analog motion video.

* Computer Video formats.

	Formats	Description
1.	CGA [Color Graphics Adaptor]	It has resolution of 320 x 200 pixels with simultaneous presentation of 4 colors. And storage capacity per image is: $320 \times 200\text{p} \times \frac{2\text{bit/pixel}}{8\text{bit/pixel}} = 16000\text{ bytes}$
2.	EGA [Enhanced Graphics Adaptor]	Supports resolution of 640 x 350 pixel with 16 color presentation with storage capacity of 112,000 byte. $640 \times 350\text{p} \times \frac{4\text{ bits/pixel}}{8\text{ bits/pixel}}$
3.	VGA [Video Graphics Adaptor]	Supports 640 x 480 pixels and can display 256 colors simultaneously. Monitor is controlled via RGB output. Storage capacity per image is 307,200 bytes. $640 \times 480\text{p} \times \frac{8\text{ bits/pixel}}{8\text{ bits/pixel}}$
4.	SVGA [Syper VGA]	Supports resolution upto 1024 x 768 pixels and color format upto 24 bits per second, with storage capacity per image 2,359,296 bytes.

* Computer based Animation.

→ Computer based Animation is an animation performed by a computer using graphical tools to provide visual effects. To animate something is like injecting a life to it.

Basic steps for computer based Animation.

1] Input Process

→ At the initial stage of generating animation, the drawings must be digitized in order to generate / create key frames. For digitizing the drawings the method of optical scanning, tracing along with data tablet is used.

→ Thus digitized image should be kept in key frames at extreme or characteristics position that has to be animated.

2] Composition stage.

→ In this stage the foreground and background figures are combined to generate the individual frames for the final animation.

3] In between Process

→ The animation of movement from one position to another needs a composition of frames with intermediate frames in between on key frames. This is called in between process, which is created

out by the interpolation method. In

→ In Interpolation method; in this process system gets only starting and ending position.
The easiest Interpolation method is linear Interpolation called lerping with some limitations.

for example; Inorder to calculate intermediate position of a thrown ball (vertically) in air using the sequence of three key frames as starting, mid and final position, then the lerping or linear interpolation is unable to calculate the intermediate position. In this case splines are invoked.

4] Changing colors:

→ for changing colors, computer aided animation uses lut or CLUT (color lookup Table) in a frame buffer and the process of double buffering. The lut animation is generated by manipulating the color lookup table. The simplest method is to cycle the colors in the lut.

* Animation language:

→ Different languages for describing animation, are;

1] Linear-list Notations.

→ Linear list notation language is animation supporting language where each event in the animation is described by starting and ending frame number and an action that has to take place

→ SCEFO (scene format) is an example of linear-list notation

→ Example;

45, 53, B, rotate, "palm", 1, 30.

here,

starting frame no : 45
ending frame no : 53
B => table
Rotate => Action
Palm => Object
1 => straight angle
30 => end angle

OR.

from 42 frame no to 53, rotate the object called palm about axis 1 by 30 degrees.

2] Graphical language

→ It is also a high level ~~programming~~ computer language to develop animation as well graphics.
Example: Autocad / AutoCAD.

3] General purpose language:

→ It is also a high level ~~proo~~ computer language developed for normal application software development that support the features of animation along with graphic drawing.
Example: QBASIC, C, C++, java etc.

* Methods for controlling Animation.

Different methods for controlling animation are

1] full explicit control

→ It is the simplest type of control. In this animator provides the entire information of entire events that could occur in the animation.
The animator may specify changes such as translation, scaling, rotation.

2] Procedural Control:

→ It is based on communication between various objects to determine their properties.
In particular,
- physically-based system, the position of one object may influence the motion of another object
  example: balls cannot pan through walls.
- actor-based syntams, the individual actors may pan their position to other actors to affect theim behaviours.

3] Constraint-based systems.

→ It is used to specify an animation sequence.

4] Tracking live action:

→ live action can be tracked to generate trajectories of objects in the course of an animation.

5] kinematics and Dynamics.

→ kinematics refers to the position and velocity of points. kinematics describes the scene. And dynamics takes the account of physical laws of movement that governs the kinematics.

* Display of Animation.

→ To display animations with raster systems, animated objects must be scan-converted in their pixmap in frame buffer.

→ To show a rotating object, we can scan-convert into the successive pixmap views.

→ Scan-conversion provides a smooth effect

→ The frame buffer is divided into two ~~halves~~ images, each with half of bits per pixel of overall frame butter.

Let us assume $image_0$ and $image_1$ as two halves of overall frame buffer (pixmap).

- Load look-up table to display values as background colon
- Scan-convert object into $image_0$
- Load look-up table to display only $image_0$

REPEAT

- Scan convert object into $image_1$
- Load look-up table to display only $image_1$
- Rotate object data structure description
- Scan convert object into $image_0$
- Load look-up table to display only $image_0$
- Rotate object data structure description.

UNTIL (termination condition)

* Transmission of Animation.

→ for the transmission of animation following two different techniques are used.

1] Symbolic Representation:

→ It is the graphical description (circle) of an animated object (ball) along with operation command (roll).
→ Animation is displayed at the reciever by scan-conversion of objects into pixmap.
→ Transmission rate depends upon
- size of symbolic representation structure, size of operation structure
- number of animated objects and of commands.

→ Transmission time is short as symbolic representation is smaller than pixmap representation but display time is longer since scan conversion has to be done at receiver's end.

2] Pixmap Representation.

→ Longer transmission time than symbolic representation because of large data size in pixmap.
→ Shorter display time because no scan-conversion has to be done at receiver's end.
→ Transmission rate = size of pixmap x frame rate (fixed transmission rate)

Unit V: Data Compression.

* Storage Space.

→ The storage space requirement for the uncompressed data is far huge as compared to the comprened data. The uncomprened data taking more storage space while subjected to transmision over digital network is highly complicated in the sense that the uncomprened data requires higher bandwidth for a single point-to-point communication. And to eradicate this complication along with more cost efficiency, multimedia system uses compressed formats of digital audio and video stream.

* Coding Requirement:

→ Images requires higher storage compared to text. Audios and videos have even more demanding properties for data storage along with data rates for communication.

→ Comprenion / coding in multimedia is subjected to certain limitations / constraints. While decoding / decomprening the comprened data, it should be as good as ponible.

1] Dialogue Mode Application
2] Retrieval Mode Application.

1] Dialogue Mode Application.

→ Interaction among humans / users via multimedia information.

→ Requirement for comprenion and decomprenion.

- End-to-end delay lower than 150 ms.
- End-to-end delay of 50 ms for face-to face dialogue applications.

2] Retrieval Mode Applications.

→ A human user retrieves information from a multimedia database.

→ Requirements.

- fast forward and backward data retrieval with simultaneous display.
- fast search for information in multimedia databases.
- Random access to single images and audio frames with an access time len than 0.5 second.
- Decomprenion without a link to other data units for random accon and editing.

* Source, Entropy and hybrid coding.

1] Entropy Coding. (statistical)

→ Lossless coding, decompression process regenerates the data completely..
→ Used regardless of the media's specific characteristics.
→ Data stream is considered to be a simple digital sequence.
Examples: Run-length encoding, Huffman encoding, Arithmetic encoding.

2] Source Encoding.

→ Lossy coding.
→ Semantics of the data are considered.
→ Degree of compression depends on data contents.
→ Examples: Pulse Code Modulation (PCM)
Differential Pulse Code Modulation (DPCM)
Discrete Cosine Transformation (DCT).

Entropy Coding	Source Coding.
• Lossless coding, decompression process completely generates data completely.	• Lossy coding.
• Independent of data characteristics.	• May consider semantics of data
• Simple digital sequence of data stream	• Degree of compression depends on data content.
Ex: Huffman coding, Arithmetic encoding.	• Ex; DPCM, DCT.

Entropy Coding
- Huffman coding.

- Huffman Coding.

→ Huffman coding is a lossless coding, where during the decoding of data completely regenerates the original data.

→ Purpose of Huffman Coding.
i) Data comprenion
ii) Data encoding
iii) lossless data comprenion.

Prefix Code:
→ (It means that codes) Prefix codes means that the codes (bit sequence) are anigned in such a way that the code assigned to one character is not the prefix code assigned to any other characters.

OR prefix code are the unique sequence of bit or code that is assigned to a certain character so that no other character have the same code.

→ Prefix code is made up of 0's and 1's.

Example:

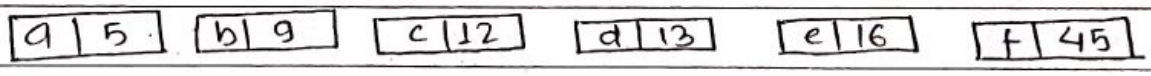
character	frequency
a	05
b	09
c	12
d	13
e	16
f	45

here, no. of character = 6.

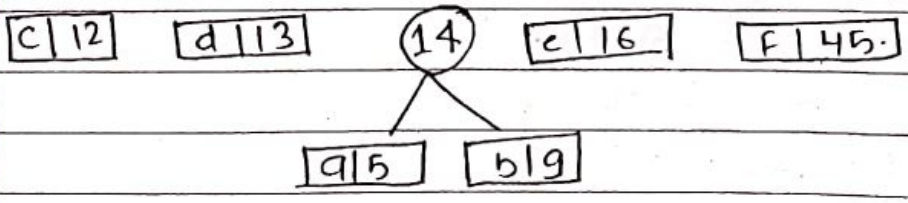
i.e. code must be of length 3

[$2^2 < 6$ (not suitable to use)
$2^3 < 6$ (suitable to use)]

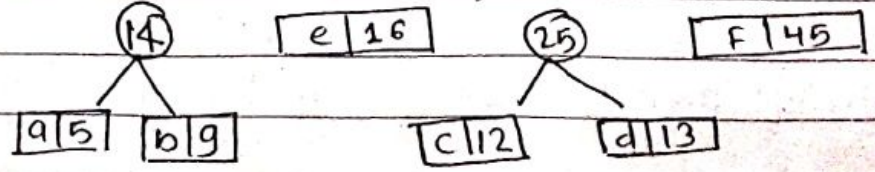
STEP 1:



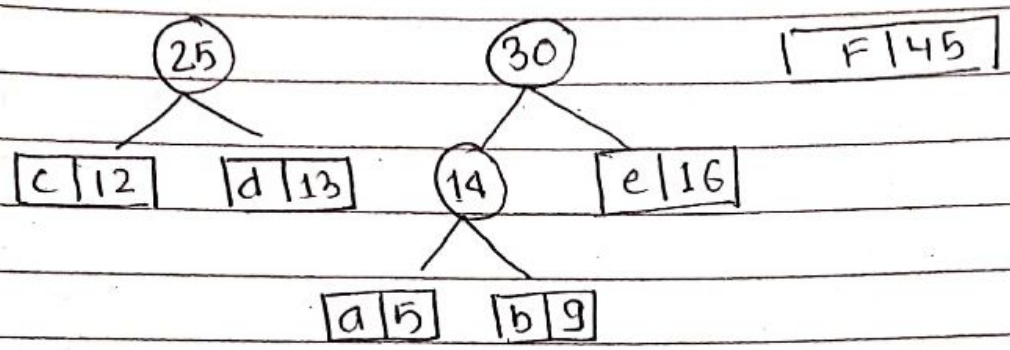
STEP 2: Select two leaf containing the least values.
(Always select character (two) that gives minimum output)
selecting [a 5] [b 9] (place smaller at right side)



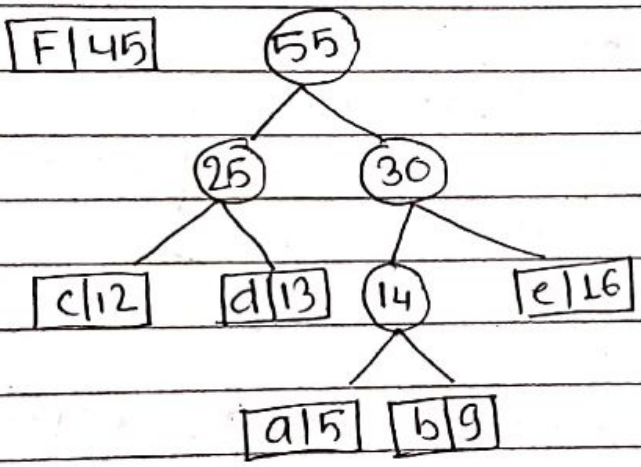
STEP 3: Select two leaf with minimum output
(place smaller at left side)



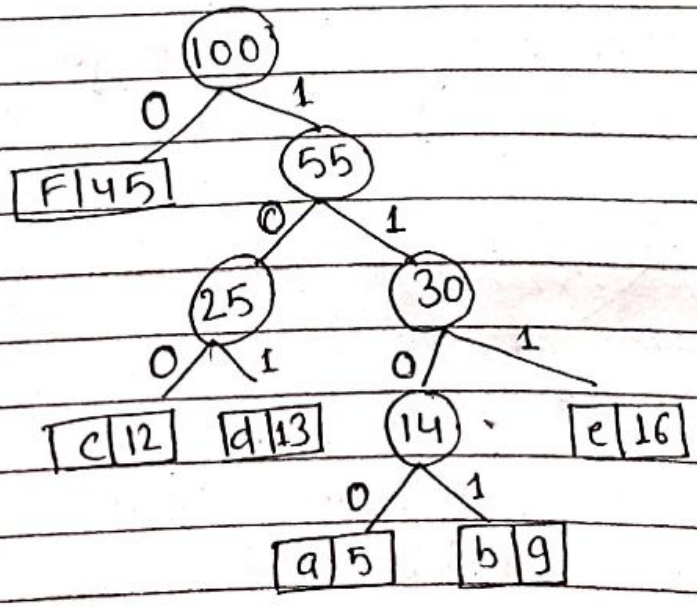
STEP 4:



STEP 5:



STEP 6:



Amign 0 to left and 1 to right leaf for entire tree.

Character	frequency	P. code.
a	5	1100
b	9	1101
c	12	100
d	13	101
e	16	111
f	45	0

Bits used before coding : $5\times3 + 9\times3 + 12\times3 + 13\times3 + 16\times3 + 45\times3 = 300$.

Bits used after coding : $5\times4 + 9\times4 + 12\times3 + 13\times3 + 16\times3 + 45\times0 = 179$

$\therefore \frac{300-179}{300} \times 100 = 40.3\%$ of space is saved or compressed after using huffman coding.

* Arithmetic Encoding

→ A.E. is a form of entropy encoding used in lossless data comp.

→ A string of characters is represented using fixed number of bits per character.

→ When a string is converted to arithmetic encoding, frequently used characters will be stored with fewer bits and characters not occuring frequently will be stored with more bits.

→ Arithmetic encoding encodes the entire message into single number, a fraction n where $(0.0 \leq n < 1.0)$ (or from 0 to 1)

* Run length Encoding.

→ It is the simplest form of entropy coding used in lossless compression of data. allowing 4-259 byte compression into 3 bytes.

→ In this encoding, frequently (at least 4 times) repeating byte are replaced depending on their occurance into three bytes, where first byte depict the repeating byte, second is the exclamation and third byte depicts the no. of occurance.

Syntax: x!x

- x → repeating character
- ! → Exclamation (flag)
- x → No. of repitition

Example.

a) ABCCCCCDDDDF. - 12 → compressed

RLE: ABC!5D!4F. - 9

→ Procedure of RLE.

- A byte should occur at least 4 consecutive times.
- Comprened data contains following format.

x!x.

- x → repeating byte.
- ! → exclamation (flag)
- x → No. of repitition

Examples.

a) ABBBBBBBBBBBCCDDDDDDDDE

11 8

↓ RLE.

A01B11C02D08E01

b) ABBBBBCCCCCCDEEEEFF

↓ RLE

AB!5C!6DE!4FF OR A01B05C06D01E04F02

c) In terms of binary format.

0000000000000000	0000	00000000
14	4	12

↓ R.L.E.

1110	0100	1100

* Drawback of RLE.

→ When no. repeating value then RLE generates more of output value (bigger size) then Input value.

i,e. Comprenion of a data will Increase the size ~~th~~ compared to original data.

Example:

i) AABCCD

RLE: A02B01C02D01

ii) XYZ

RLE: X01Y01Z01.

* Hybrid Coding.

→ Combination of techniques or compression algorithms which works more efficiently inorder to generate compressed data.

→ Hybrid Coding uses the same format of data compression for compressing data.

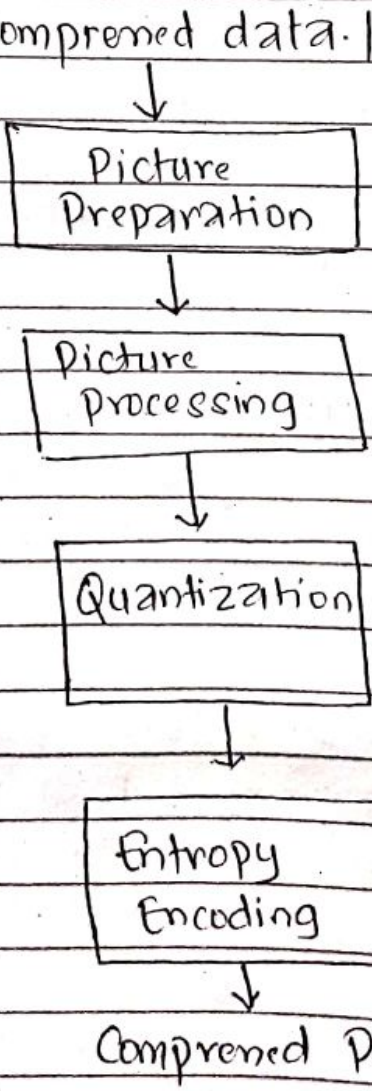


fig: MAJOR STEPS OF DATA COMPRESSION.

DCT- Discrete cosine Transform.

* Lossy sequential DCT based Mode.

→

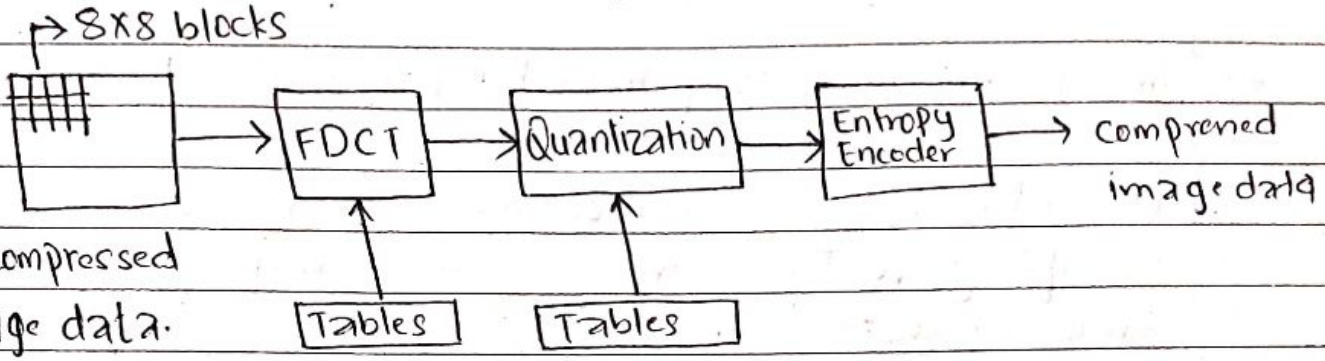


fig: lomy Sequential DCT based mode.

$$T(u,v) = \frac{1}{4} C_u C_v \sum_{x=0}^{7} \sum_{y=0}^{7} \cos\frac{(2x+1)u\pi}{16} \cdot \cos\frac{(2y+1)v\pi}{v\pi}$$

where,

$C_u, C_v = 1/\sqrt{2}$ for $u,v = 0$, otherwise $C_u, C_v = 1$

1] Image Processing.

→ After Image preparation, the uncomprened image data are grouped into ~~8x8~~ units of 8x8 pixels.

→ Each sample is enconded using $p = 8$ bit.

→ Each pixel is an integer between 0 to 255

Image procening is carried out as follows.

i) DCT based transformation coding is carried out.

ii) A forward DCT (FDCT) is applied

iii) for laterreconstruction, the decoder uses IDCT.

2] Quantization.

→ Image procesing is followed by quantization.

- The entire lossy process occurs in quantization.
- Specific frequencies are given more importance than others.
- Tables are used for quantization and dequantization
- Image quality may decrease due to quantization.

3] Entropy Encoding.

→ Quantization is followed by Entropy Encoding. (using Huffman coding) (only)

- DC coefficients are encoded by subtracting the DC coefficients of previous unit.
- Huffman coding is prefered as it is free (not patented)
- Coding tables for each DC and AC coefficients must be provided.
- AC coefficients are procesed using zig-zag sequence

* Expanded Lossy DCT Mode.

- Image preparation must be of $p = 12$ bit instead of $p = 8$ bit per pixel.
- Arithmetic encoding can be used in addition to Huffman coding in entropy encoding section.

NOTE: Only above two points differentiate lossy Sequential DCT based mode with expanded lossy DCT mode. [Rest is same]

* JPEG and MPEG Comprenion.

Unit 6: User Interface.

* Basic Design Issues.

→ Basic Design Issues consist of ~~atti~~ architectural issues.

(a) Architectural Issues

→ An effective presentation design should support an interactive action along with sequential flow of actions. Each presentation requires continuous and Instantaneous feedback while making decisions regarding media and modalities.

- Information Characteristics for Presentation.
- Effective human-computer Interaction
- Audio and video at user Interface.
- User friendliness (primary goal)

* Information Characteristics for Presentation.

→ A complete set of Information is necenary for appropnate mapping in between Information and presentation technique. The Information characteristics specify.

1] Types
2] Relational structures
3] Multi-domain relation
4] Large data sets.

1] Types.

→ There are two types of ordered data.

1] Coordinates versus amount.
2] Intervals versus ~~amou~~ ratio.

2] Relational structures.

→ Relational structures refers to the way in which a relation maps among its domain sets (dependency).
→ It consist of two different dependencies.
a) functional Dependency.
Example: bar chart.
b) Non-functional Dependency.
Example: student entry in relational database.

3] Mulli-domain Relations.

→ multiple attribute of single object such as position, color.
→ mulliple object sets such as graphical symbols on a map.
→ multiple displays. such as multiple windows.

4] Large data sets.

→ Refers to the numerous attribute of collection of heterogeneous objects such as presentation of semantic networks.

* Presentation function.

→ Presentation function is something that displays an objects as per the user requirement. such as printf for display of an object or character.

→ It is important to specify presentation function Independent to presentation form, style or information it conveys.

* Presentation Design knowledge

→ for presenting a design following things/topics needs to be considered.

1] Content Selection

→ content selection is the key to presentation design. The more clear and simple content in presentation leads to higher efficacy of presentation and easy to convey the desired/appropriate Information to the user.

2] Media selection.

→ Along with content selection. media selection should be made appropriate to convey the desired Information to user. with higher efficacy. for example, a presentor has to demonstrate the size and structure of a playground (football) then he/she should use graphics, videos rather than text describing the length and dimension.

3] Coordination:

→ Coordination simply refers to the collection of different aspects which helps the presentor to convey the desired information to the user with clear vision.

Coordination may combine following aspects for making a better persentation.

I) Encoding techniques.
II) Presentation object
III) Multiple Displays.

* Effective human-computer Interaction.

→ The most important issue regarding multimedia Interface is effective human computer Interface i.e. user friendliness.

The main goal of multimedia is considered as effective human-computer Interaction / user friendliness.

for being user friendly following criteria should match.

✓ 1) Easy to learn Instruction	VI) Entry elements
II) Easy to remember Instruction	VII) Meaningful location of function
III) Effective Instruction	VIII) Presentation ✓
IV) Aesthetics.	IX) Dialogue boxes ✓
V) Effective Implementation support	X) Additional design criteria. ✓
	XI) Design specific criteria. ✓

* Classification of software

1] Easy to learn Instruction.

→ MUI for being a user-friendly it should posses a property that the user can easily remember the application instruction rules. Any of MUI having easy and simple instruction helps user to interact with computers or devices. without any complications which will further lead to friendliness to MUI.

→ Inorder to maintain the friendliness, the Interface should have those instruction that are simpler, easier and also the prevailing instruction some other means which are in general use of public humans.

2] Presentation.

→ It is another criteria that leads to user friendliness. In general presentation ~~lead~~ can have following variants.

i) Abbreviated text
ii) Full text
iii) Icons i,e. graphics
iv) Micons i,e. motion videos

Each variants of presentation may or maynot be easily understood to the user. Example: CW (Call waiting) may be easy to understand to some user but not to all. Call waiting can easily be easily understood but not clear. Some icon like recycle bin may be easily depict it's use to some but not to all.

3] Dialogue Boxes.

→ The dialogue boxes being display with options like ok and abort should always have a strict rule to place them at a specific part of dialogue box. for example: While multiple dialogue box are displayed then the button ok of entire dialogue boxes should be at same position like the underlying dialogue box should have the same positioning of options as of surface / displayed one.

4] Additional Design Criteria.

→ for being user-friendly Interface it is important that certain additional design be included which would be helpful to user while Interacting.

- The cursor acting as a rotating fish Instead of being steady while a task is at progress.
- An entry being highlighted while being selected.
- The details of any image, video, audio being displayed as a cursor stays on it for a while.

5] Design Specific Criteria.

→ Along with additional design criteria, design specific Giteria may increase the Interaction between MUI and humans or user.

• In telephone, mobile networks / or telephone specific network the end device must have functions like dialing a number, arriving a call request over call, as per customer's / user's requirement placing the call on hold or adding a call.

* Audio and video at user Interface

→ Audio can be implemented at the user Interface for application control. Thus speech analysis is necessary.

• Speech Analysis is basically of two type.

a) Speaker-dependent Analysis.

→ It allows the Input of approximately 25,000 different words with a relatively low error rate

→ In Speaker-dependent Analysis a training session is must for speaker-specific characteristics.

b) Speaker-Independent Analysis.

→ It is less advanced than that of Speaker-dependent system in a sense that it can only recognize only few limited words.

→ In this analysis there is no need of training session.

• Dimension of Space

- monophony (all audio sources have same spatial location)
- Sterophony (allows bilateral hearing)
- Quadrophony (concept of two or more channels)

- **Video** at user interface

→ A continuous sequence of atleast 15 images per second gives a rough perception of continuous motion picture.

→ As a user interface video is implemented through a continuous sequence of individual images.
for example; xv user interface software that implements video through continuous sequence of images.

- Remote Camera Control Application.

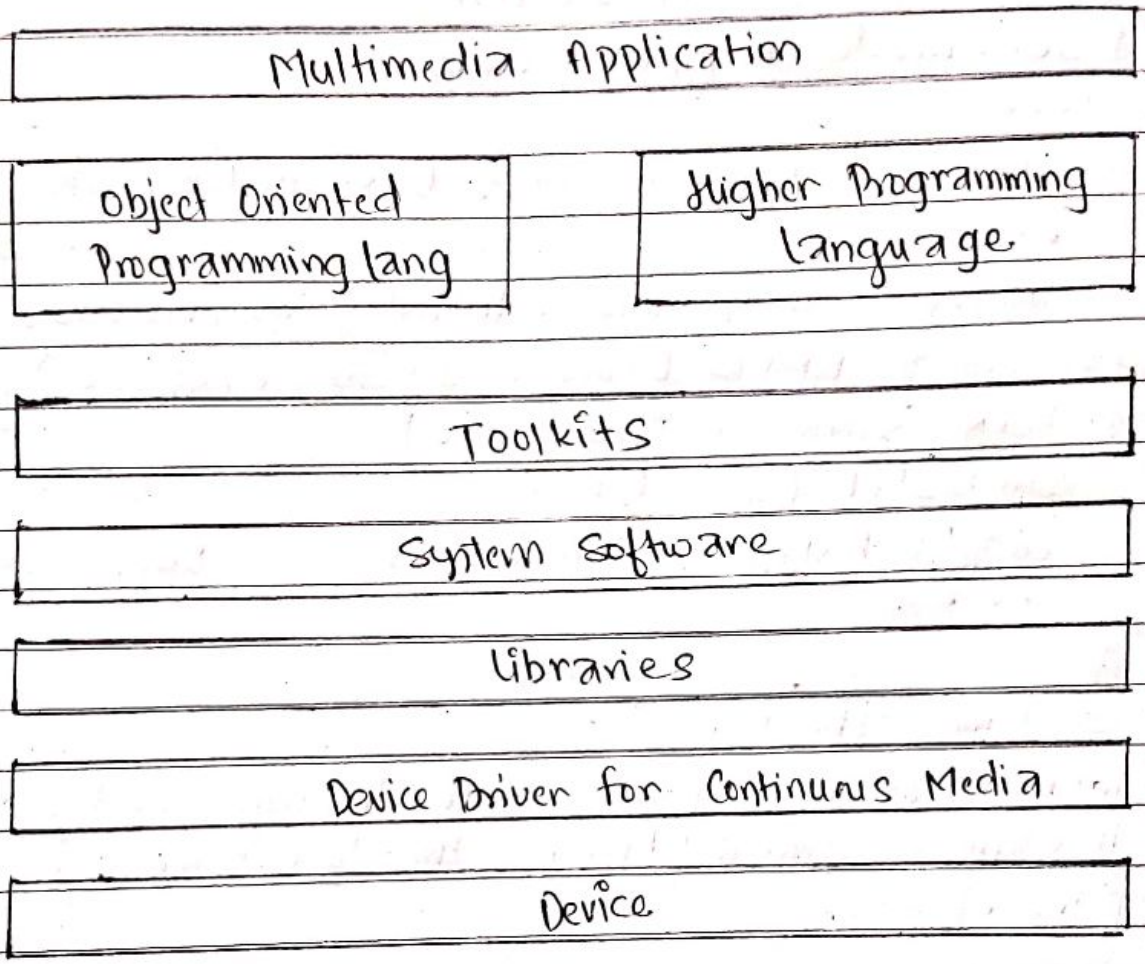
→ It is banically used for surveillance purpose

→ Here camera is remotely controlled by a computer. Each time camera receives control information from computer and work accordingly.

→ Camera can be controlled through
- keyboards
- buttons in window system
- scroll bars
- mouse, joystick etc

System Software	Application software
• Enables entire system for proper functioning	• Software that run on O.S for specific purpose.
• Machine dependent	• Machine Independent
• Direct hardware interaction	• Indirect hardware interaction.
• Written in low-level language	• Written in high-level-language.
• Installed alng with O.S.	• Installed when required.
Ex: Device driver, O.S.	Ex: video editor, spreadsheet.

Unit 7: Abstraction for Programming.



1] Introduction to libraries.

→ The procesing of continuous media is based on function or set of functions embedded into libraries.
→ This is usual solution for programming multimedia data.
→ libraries differ in their degree of abstraction.

Example: IBM's early AVC (Audio Visual connection)

```
acb
abc. channel = AAPI_CHINA
acb. mode = AAPI_PLAY
...
aud_init (& acb)   /* acb is audio control block */
...
audrc = Fab_open (AudioFullFileName, AAFB_OPEN,
        AAFB_EXNO, 0, & Fab, 0, 0, 0, 0);
Fork (START in PARALLEL)
aud_strt (& acb)
displayPosition (Relative Start time, Duration)
...
```

2] System Software.

→ a Instead of Implementing access to multimedia devices through individual libraries, the device access can become part of O.S.

a) Data as Time Capsule. (file extensions)

- Each LDU (Logical Data Unit) carries in its time capsule along with data type, actual value and valid life span.
- Useful concept for video, where each frame has valid span of 40 ms.
- ⓟ changed presentation rate for VCR (Video Cassette Recorder) function like fast forward, slow forward, fast rewind.

b) Data as stream.

→ A stream denotes the continuous flow of audio and video data between a source and sink.

→ Prior to the flow, a stream is established equivalent to the setup of a connection in a networked environment.

3] Toolkits.

→ Toolkit is considered as a simpler way compared to system software. Interface from the user's point of view as abstraction could be maintained many details.

OR

→ Toolkit allows easier abstraction thus it is simpler approach compared to system software Interface.

• Toolkits are used for;

- abstraction from actual physical layer.
- allows uniform Interface for communication with all different devices of continuous media.
- Introduce client-server paradigm.
- can hide process structure.
- can be embedded into programming language or object oriented programming environment.

4] Higher Programming language.

→ Higher Programming language is also called as Higher level language (H.L.L).

a) Media as Types:

- definition of appropriate data type (ex. for audio, video)
- LDU can be smallest unit
- Example of merging text and a motion picture.

b) Media as files:

→ Instead of considering continuous media as data type they can be considered as files.

```
file_h1 = open (MICROPHONE_1, ...)
file_h2 = open (MICROPHONE_2, ...)
file_h3 = open (SPEAKER, ...)
...
read (file_h1)
read (file_h2)
mix (file_h3, file_h2, file_h1)
activate (file_h1, file_h2, file_h3)
...
deactivate (file_h1, file_h2, file_h3)
...
rc1 = close (file_h1)
rc2 = close (file_h2)
rc3 = close (file_h3)
```

c) Media as Procen.

→ It is possible to map continuous media to process and integrate them into H.L.L.

```
PROCESS cont_process_a;
...
on_menage-do
    set_volume...
    set_loudnen...
    ...
...
[main]
pid = create (cont_procen_a).
send (pid, set_volume, 3)
send (pid, set_loudnen, )
...
```

* Programming Language Requirement.

→ HLL should support parallel procening, because the procening of continuum data is.

i) controlled by HLL (via pure asynchronous Intructions)
ii) an integral part of program through identification of media.

* InterProcen Communication Mechanism. [IPC]

→ Different procen must be able to communicate through an Interprocening communication and must be able to

- understand implicitly / prior specified time requirement.
- transmit continuous data according to requirement.
- initiate procening of recieved continuous procen on time.

* Language

→ The Interprocen Communication [IPC] must be designed and implemented in real-time, (the current IPC can be omitted).

→ A simple language should be developed so that it can be designed and be implemented in real time.

Example: OCCAM-2, Parallel C-variant, ADA.

5] OBJECT ORIENTED APPROACHES.

→ OOP is a programming paradigm based on concept of object and is the collabaration of different features such as class and object, Inheritance, polymor- data abstraction, Encapsulation.

a) Class and object

→ Class is a collection of similar objects or class is a user defined data type having data member and member function. It don't allocates memory while it is created.

→ Object is an instance of a class. It allocates the memory when allocated.

Syntax.

```
class class-name
{
// data
// function
};
```

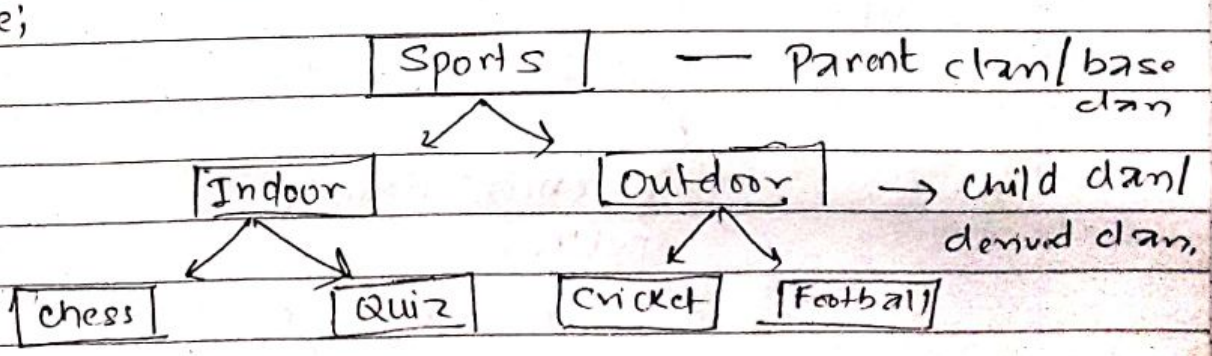
Example:

```
class student
{
public:
char name[20];
int id;
};
```

b) Inheritance

→ It is the process of acquiring the quality from parent/~~sub~~ class to child class.

Example;



c) Polymorphism.

→ Poly means many, morphs means forms. Thus it means having many forms.

Example: A person M could be the brother of X, could be husband of Y, could be son of Z. Here person M is having multiple forms.

Types of Polymorphism.

a) Run-time polymorphism
b) Compile-time Polymorphism.

* Devices as class.

→ Here the devices are anigned as a object to represent behaviour & Interface.

Example:

```
class media_device
{
    char name[10];
public:
    void on(), oFF();
};
```

BMO : Basic Multimedia object
CMO : Compound Multimedia object.

Date
Page

* Processing unit as classes.

→ Three main objects
a) source object
b) destination object
c) combined source - destination object allows data flow path via connection of objects.

* Media as class;

- Here media are assigned as a object that define hierarchical relation for different media.
- different class hierarchies are better suited for different applications.

Example;

Medium
Acoustic_Medium
Music
Opus
Score
Audio_Block
Sample_Block
Speech
...
...
Optical Medium
Video
Video_scene
Video
Video_scene
Image
Image_segment
Pixel
Line
Pixel
Column
Pixel
Animation
...
Text
...