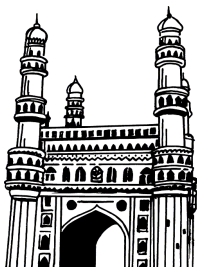
Rahul's ✔
Topper's Voice



B.Com.

III Year VI Semester

Latest 2022 Edition

MULTIMEDIA SYSTEMS

- ☞ **Study Manual**
- ☞ **Important Questions**
- ☞ **Short Question & Answers**
- ☞ **Choose the Correct Answers**
- ☞ **Fill in the blanks**
- ☞ **Solved Model Papers**

- by -

WELL EXPERIENCED LECTURER

₹.199/-

Rahul Publications™
Hyderabad. Ph : 66550071, 9391018098

B.Com.

III Year VI Semester

MULTIMEDIA SYSTEMS

Inspite of many efforts taken to present this book without errors, some errors might have crept in. Therefore we do not take any legal responsibility for such errors and omissions. However, if they are brought to our notice, they will be corrected in the next edition.

Price ` . 199/-

Sole Distributors :

☎ : **66550071, Cell : 9391018098**

VASU BOOK CENTRE

Shop No. 2, Beside Gokul Chat, Koti, Hyderabad.

Maternity Hospital Opp. Lane, Narayan Naik Complex, Koti, Hyderabad.

Near Andhra Bank, Subway, Sultan Bazar, Koti, Hyderabad -195.

MULTIMEDIA SYSTEMS

CONTENTS

STUDY MANUAL

SOLVED MODEL PAPERS

SYLLABUS

UNIT - I

MEDIA AND DATA STREAMS:

Properties of multimedia systems, Data streams characteristics: Digital representation of audio, numeric instruments digital interface Bark concepts, Devices, Messages, Timing Standards Speech generation, analysis and transmission.

UNIT - II

DIGITAL IMAGE&ANIMATIONS:

Digital Image: Analysis, recognition, transmission, **Video**: Representation, Digitalization, transmission.

Animations: Basic concepts, animation languages, animations control transmission.

UNIT - III

DATA COMPRESSION STANDARDS&STORAGE:

Data Compression Standards: JPEG, H-261, MPEG DVI

Optical storage devices and Standards: WORHS, CDDA, CDROM, CDWO, CDMO.

Real Time Multimedia, Multimedia file System.

UNIT - IV

MULTIMEDIA COMMUNICATION SYSTEM, DATABASES&SYNCHRONIZATION:

Multimedia Communication System: Collaborative computing session management, transport subsystem, QOS, resource management.

Multimedia Databases: Characteristics, data structures, operation, integration in a database model.

Synchronization: Issues, presentation requirements, reference to multimedia synchronization, MHEG.

UNIT - V

MULTIMEDIA APPLICATION:

Media preparation, Composition, integration communication, consumption, entertainment.

Contents

UNIT - I

UNIT - II

Topic **Page No.**

UNIT - III

UNIT - IV

Important Questions

UNIT - I

1. What is Multimedia? Explain the properties of multimedia systems.

Ans :

Refer Unit-I, Q.No. 1

---

2. Explain about the global structure of multimedia system.

Ans :

Refer Unit-I, Q.No. 2

---

3. Explain Traditional Data Streams Cha-racteristics.

Ans :

Refer Unit-I, Q.No. 4

---

4. Explain about the relationship between music and computer.

Ans :

Refer Unit-I, Q.No. 7

---

5. Explain various kinds of MIDI messages.

Ans :

Refer Unit-I, Q.No. 9

---

6. What is speech transmission? Explain about the components of speech transmission system.

Ans : **(Imp.)**

Refer Unit-I, Q.No. 23

UNIT - II

1. What is image format? Discuss its types.

Ans :

Refer Unit-II, Q.No. 2

---

2. Explain about various measures taken for visual representation of a video signal.

Ans :

Refer Unit-II, Q.No. 9

3. Explain about computer video format and its standards.

Ans :

Refer Unit-II, Q.No. 13

4. Explain various standards of television conventional system.

Ans :

Refer Unit-II, Q.No. 14

5. Mention the basic concepts in computer based animations.

Ans :

Refer Unit-II, Q.No. 18

6. Explain various types of animation languages.

Ans :

Refer Unit-II, Q.No. 19

7. Explain various methods of controlling animation.

Ans :

Refer Unit-II, Q.No. 20

UNIT - III

1. Explain various types of coding techniques in data compression.

Ans :

Refer Unit-III, Q.No. 1

2. Explain the need of Data Compression.

Ans :

Refer Unit-III, Q.No. 2

3. Explain MPEG video compression standard.

Ans :

Refer Unit-III, Q.No. 8

4. Explain about Write Once, Read Many storage device.

Ans :

Refer Unit-III, Q.No. 11

5. Explain CDWO technology.

Ans :

Refer Unit-III, Q.No. 15

6. Explain about CDMO technology.

Ans :

Refer Unit-III, Q.No. 16

7. Explain Rate Monotonic Algorithm with example.

Ans :

Refer Unit-III, Q.No. 22

7. Why do we need multimedia file systems? Explain.

Ans :

Refer Unit-III, Q.No. 23

UNIT - IV

1. What is Collaborative Computing? Explain the dimensions of Collaborative Computing.

Ans :

Refer Unit-IV, Q.No. 1

2. Explain about session management architecture.

Ans :

Refer Unit-IV, Q.No. 3

3. What are the requirements of the transport sub system.

Ans :

Refer Unit-IV, Q.No. 5

4. Explain how routing is done in network layer and write the supportive protocols to perform routing in network layer.

Ans :

Refer Unit-IV, Q.No. 10

5. Explain several issues which need to be addressed in QoS.

Ans :

Refer Unit-IV, Q.No. 11

6. Explain resource management architecture with a neat diagram.

Ans :

Refer Unit-IV, Q.No. 13

7. Explain about the data structures used for multimedia database with an examples.

Ans :

Refer Unit-IV, Q.No. 18

8. Explain the Integration of MMDBMS?

Ans :

Refer Unit-IV, Q.No. 20

9. Explain about the basic issues in synchronization.

Ans :

Refer Unit-IV, Q.No. 22

10. Explain the reference model of multimedia synchronization.

Ans :

Refer Unit-IV, Q.No. 24

11. What is MHEG? Explain Armedia client architecture.

Ans :

Refer Unit-IV, Q.No. 26

UNIT - V

1. Write about multimedia preparation programs in US and EUROPE

Ans :

Refer Unit-V, Q.No. 1

2. What are the different editors used for media composition?

Ans :

Refer Unit-V, Q.No. 4

3. Define media integration. Write about media editors used for integration.

Ans :

Refer Unit-V, Q.No. 5

4. Write about various sources for media communication.

Ans :

Refer Unit-V, Q.No. 6

5. Write about media entertainment technologies.

Ans :

Refer Unit-V, Q.No. 9

6. Explain the concept of interactive audio and games.

Ans :

Refer Unit-V, Q.No. 11

UNIT I

MEDIA AND DATA STREAMS:

Properties of multimedia systems, Data streams characteristics: Digital representation of audio, music instruments digital interface Basic concepts, Devices, Messages, Timing Standards Speech generation, analysis and transmission.

1.1 Multimedia Systems

1.1.1 Properties of Multimedia Systems

Q1. What is Multimedia? Explain the properties of multimedia systems.

Ans :

A multimedia system is responsible for developing a multimedia application. A multimedia application is a bundle of different kinds of data. A multimedia computer system is one that can create, integrate, store, retrieve delete two or more types of media materials in digital form, such as audio, image, video, and text information.

Properties

The main properties of multimedia systems that differentiate it from other systems are as follows,

1. Combination of Media

Multimedia does not mean to be any random combination of media. But the combination should be justified. For example, a simple text program with images will be considered as a multimedia application because it is incorporated with two different media i.e., text data and image data. Note that, an application is said to be a multimedia application only when it uses the discrete media as well as continuous media.

Thus, a simple text program with images will not be a multimedia application as it did not incorporate with both the continuous and discrete media.

2. Media Independence

The level of independence of media distinguishes itself from other media. Thus, each media requests for different level of significance. In order to accomplish this, multimedia requires many levels of significance.

The video recorder incorporated in a computer store both the audio data and video data, but both the media have a close relation with each other that is common storage medium (tape). In contrast, the presentations consist of Digital Audio Tape (DAT) recorder signals and computer-available text which satisfies the media independence.

The multimedia systems enabled with communication should be approached. The various reasons for this are as follows,

- Generally, all the computers are interconnected with each other. Here, the restriction is assuming all the multimedia functions from one single viewpoint.
- Most of the distributed environments allows mainly the effective and efficient multimedia applications. Here, the information is not only be created but also be refined, stored, displayed and distributed out of the computer boundary.

3. Computer-supported Integration

The multimedia system should enable the computer-controlled media processing because the computers help in providing the

possibility of merging the media data in arbitrary forms. It must be reprogrammable by user as well as the system programmer. The computer integration will be satisfied only when there is a connection between different types of media data.

Q2. Explain about the global structure of multimedia system.

Ans : **(Imp)**

Global structure of Multimedia System

1. Device domain
2. System domain

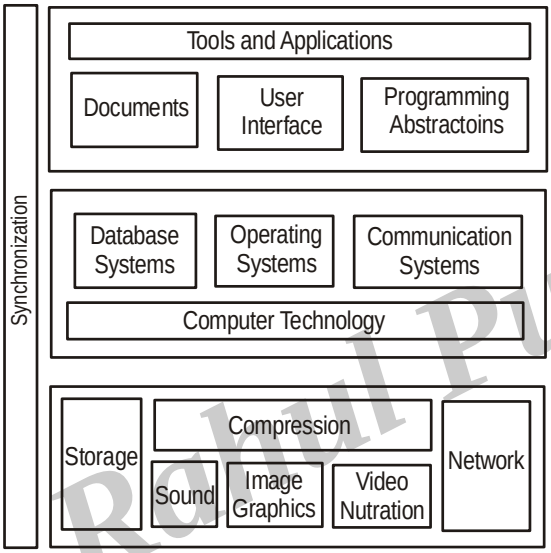


Fig.: Main topics in multimedia system with chapter number

3. Application domain
4. Cross domain

1. Device domain

It deals with interaction between multimedia application and multimedia devices such as AGP Card, Sound Card etc. Basic concepts for the processing of digital audio and video data are based on digital signal processing. Different methods for the processing of image, graphics and animation are described. The audio techniques section includes music (MIDI) and speech processing.

2. System Domain

The interface between the device domain and the system domain is specified by the computer technology. To utilize the device domain, several system services are needed. Basically, three services exits. These services are mostly implemented in software. The operating system, serves as an interface between computer hardware/system and all other software components. It provides the user with a programming and computational environment, which should be easy to operate. The database system allows a structured access to data and a management of large databases. The communication system is responsible for data transmission according to the timing and reliability requirements of the networked multimedia.

3. Application domain

Provides functions to the user to develop and present multimedia projects. This includes software tools, and multimedia projects development methodology. The services of the system domain are offered to the application domain through proper programming abstractions. Another topic embedded in the application domain is document handling.

4. Cross domain

It turns out that, some aspects such as synchronization aspects, are difficult to locate in one or two components or domains. The reason is that synchronization, being the temporal relationship among various media, relates to many components across all domains.

Q3. What is medium? What are the different types of Transmission media in Multimedia Systems.

Ans :

In general, one describes medium as information., Examples of a medium a means for distribution and presentation of are text, graphics, speech and music.

We classify media according to, perception, representation, presentation, storage, transmission information exchange.

1. Perception Media

Perception media refers to the nature of information perceived by humans, which is not strictly identical to the sense that is stimulated. For example, a still image and a movie convey information of a different nature, though stimulating the same sense. The question to ask here is: How do humans perceive information?

In this context, we distinguish primarily between what we see and what we hear. Auditory media include music, sound and voice. Visual media include text, graphics and still and moving pictures. This differentiation can be further refined. For example, a visual medium can consist of moving pictures, animation, and text. In turn, moving pictures normally consist of series of scenes that, in turn, are composed of single pictures.

2. Representation Media

The term representation media refers to how information is represented internally to the computer. The encoding used is of essential importance. The question to ask here is: How is information encoded in the computer? There are several option:

- Each character of a piece of text is encoded in ASCII.
- A picture is encoded by the CEPT or CAPTAIN standard, or the GKS graphics standard can serve as a basis.
- An audio data stream is available in simple PCM encoding and a linear quantization of 16 bits per sampling value.
- A single image is encoded as group-3 facsimile or in JPEG format.
- A combined audio-video sequence is stored in the computer in various TV standards (eg. PAL, SECAM or NTSC), in the CCIR-601 standard, or in MPEG format.

3. Presentation Media

The term presentation media refers to the physical means used by systems to reproduce information for humans. For example, a TV set uses a cathode-ray tube and loudspeaker. The question to ask here is: Which medium is used to output information from the computer or input in the computer?

We distinguish primarily between output and input. Media such as paper, computer monitors, and loudspeakers are output media, while keyboards, cameras, and microphones are input media.

4. Storage Media

The term storage media is often used in computing to refer to various physical means for storing computer data, such as magnetic tapes, magnetic disks, or digital optical disks. However, data storage is not limited to the components available in a computer, which means that paper is also considered a storage medium. The question to ask here is: Where is information stored?

5. Transmission Media

The term transmission media refers to the physical means-cables of various types, radio tower, satellite, or ether (the medium that transmit radio waves)-that allow the transmission of telecommunication signals. The question to ask here is: Which medium is used to transmit data?

6. Information Exchange Media

Information exchange media include all data media used to transport information, e.g., all storage and transmission media. The question to ask here is: Which data medium is used to exchange information between different locations?

For example, information can be exchanged by storing it on a removable medium and transporting the medium from one location to another. These storage media include microfilms, paper, and floppy disks. Information can also be exchanged directly, if transmission media such as coaxial cables, optical fibers, or radio waves are used.

1.2 Data Streams Characteristics

Q4. Explain Traditional Data Streams Characteristics.

Ans : **(Imp.)**

Transmission of information carrying different media leads to data streams with very different features. The attributes of asynchronous, synchronous, and. Isochronous.

Data transmission come from the fields of computer communication and switching. They are also used, for example, in FDDI (Fiber Distributed Data Interface) networks for the description of different data transmission modes with respect to end-to-end delay of individual packets.

1. Asynchronous Transmission Mode

There is no time restriction on the transmission of packets. Packets reach the receiver (client) as fast as possible. Transmission of discrete data typically requires only this transmission mode.

2. Synchronous Transmission Mode

A maximum end-to-end delay is specified, for the transmission of packets, and the maximal delay is never violated, although each packet may be received at any arbitrary earlier time. This is essential for multimedia applications which require that no packets are dropped due to network and server overheads. All packets sent will be received in a timely fashion. Note, however, that the sequence in which packets are received is not guaranteed. The client will still need to buffer packets which arrive out of sequence. There may also be a slight jitter in the playback of the stream as the playback system waits for data which has not yet arrived, but which is still within the bounds.

3. Isochronous Transmission Mode

As well as a maximum end-to-end delay, a minimum end-to-end delay is also specified and guaranteed. This reduces (but does not completely eliminate) the need for temporary storage in the client to buffer out-of-sequence packets, as well as jitter. However, the implication is that nodes on the network are reponsible for storing packets which have already been sent by the server, but which are not yet ready to be received by the client.

4. Data stream characteristics for continuous media

The data stream characteristics in transmission are associated with any audio and video data transfer. Also the info stream characteristics are influenced by the compression during the information transfer. Its characteristics apply for distributed furthermore as local environment. Hence the data stream characteristics can be discussed on the basic of three given factors or properties.

i) According to the time intervals between consecutive packets.

ii) According to the variation of the amount between consecutive packets.

iii) According to the continuity or connection between consecutive packets.

(i) According to time intervals between consecutive packets

On the basic of this factor we find out three properties they are:

a) Strongly periodic data stream

If time intervals are of the identical length between two consecutive to packets that's a continuing, then the stream is named strongly periodic and within the ideal case the jitter has the worth zero. For eg: PCM coded speech in traditional telephone switching.

b) Weakly periodic data stream

If time intervals between two consecutive packets is not constant but are of periodic nature with finite period then the data stream is called weakly periodic.

c) A-periodic data stream

If the sequence of time intervals is neither strongly nor weakly periodic, instead the time period or time gap various between packets to packets during transmission then such data stream is called A-periodic data stream.

(ii) According to variation of consecutive packet amounts

On the basic of these factors there are three types of data stream:

a) Strongly regular data stream

If the number of information stays constant during the life time of a knowledge stream, this feature is specially found in uncompressed digital data transmission, as an example audio stream of CD, video stream of camera in uncompres-sed form.

b) Weakley regular data stream

If the amount of data stream varies periodically with time and not shows the behaviors of strongly regular data stream then it is called Weakley regular data stream, For example compressed video stream.

c) Irregular data stream

If the number of information is neither constant nor changes in keeping with a periodic function, then the information streams are called irregular data stream. Transmission and processing of this category data stream is complicated. Since data stream includes a variable (bit) rate after applying compression methods.

(iii) According to continuity or connection between consecutive packets

On the basic of this factor their can be also 2 types or characters.

a) Continuous data stream

If consecutive packets are directly transmitted one after another without any time gap then such data streams are called continuous data stream, For example audio data use for B channel of Isdn with transmission rate for 64 kbps.

b) Unconnected data stream

A data stream with gaps between information units is named and unconnected data stream. The transmission of a connected data stream through a channel with the next capacity treads gaps between individual packets, as an example the information stream coded with JPEG method with 1.2 mbps on a FDDI network.

1.2.1 Digital Representation of Audio

Q5. What is Sound? Write about it.

(OR)

Explain briefly about sound.

Ans :

Meaning

Sound is a physical phenomenon produced by the vibration of matter. The matter can be almost anything: a violin string or a block of wood, for example. As the matter vibrates, pressure variations are

created in the air surrounding it. This alternation of high and low pressure is propagated through the air in a wave-like motion. When the wave reaches our ears, we hear a sound.

Figure graphs the oscillation of a pressure wave over time.

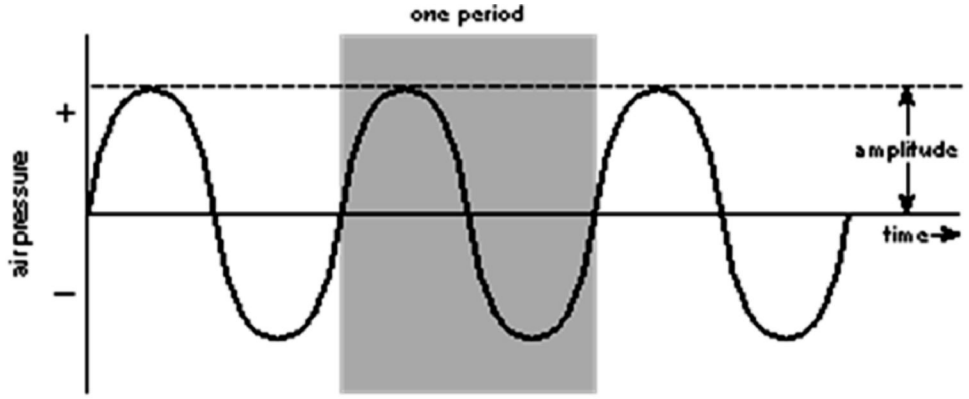


Fig.: Air Pressure Wave

The pattern of the pressure oscillation is called a waveform. Notice that the waveform in Figure repeats the same shape at regular intervals; the gray area shows one complete shape. This portion of the waveform is called a period. A waveform with a clearly defined period occurring at regular intervals is called aperiodic waveform.

Since they occur naturally, sound waveforms are never as perfectly smooth nor as uniformly periodic as the waveform shown in figure. However, sounds that display a recognizable periodicity tend to be more musical than those that are nonperiodic. Here are some sources of periodic and nonperiodic sounds:

(i) Periodic

- Musical instruments other than unpitched percussion
- Vowel sounds
- Bird songs
- Whistling wind

(ii) Nonperiodic

- Unpitched percussion instruments
- Consonants, such as “t,” “f,” and “s”
- Coughs and sneezes
- Rushing water

(iii) Frequency

The frequency of a sound the number of times the pressure rises and falls, or oscillates, in a second is measured in hertz (Hz). A frequency of 100 Hz means 100 oscillations per second. A convenient abbreviation, kHz for kilohertz, is used to indicate thousands of oscillations per second: 1 kHz equals 1000 Hz.

The frequency range of normal human hearing extends from around 20 Hz up to about 20 kHz.

The frequency axis is logarithmic, not linear: To traverse the audio range from low to high by equal-

sounding steps, each successive frequency increment must be greater than the last. For example, the frequency difference between the lowest note on a piano and the note an octave above it is about 27 Hz. Compare this to the piano's top octave, where the frequency difference is over 2000 Hz. Yet, subjectively, the two intervals sound the same.

(iv) Amplitude

A sound also has an amplitude, a property subjectively heard as loudness. The amplitude of a sound is the measure of the displacement of air pressure from its mean, or quiescent state. The greater the amplitude, the louder the sound.

Q6. How the Computer Represents Sound? Explain

Ans :

The smooth, continuous curve of a sound waveform isn't directly represented in a computer. A computer measures the amplitude of the waveform at regular time intervals to produce a series of numbers. Each of these measurements is called a sample. Figure 2-2 illustrates one period of a digitally sampled waveform.

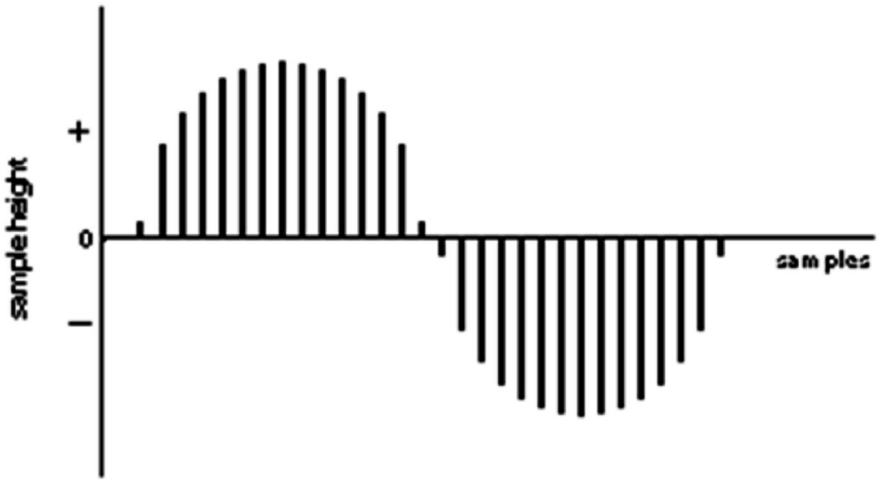


Fig.: Sampled Waveform

Each vertical bar in Figure represents a single sample. The height of a bar indicates the value of that sample.

The mechanism that converts an audio signal into digital samples is called an analog-to-digital converter, or ADC. To convert a digital signal back to analog, you need a digital-to-analog converter, or DAC (pronounced "dack").

Sampling Rate

The rate at which a waveform is sampled is called the sampling rate. Like frequencies, sampling rates are measured in hertz. The CD standard sampling rate of 44100 Hz means that the waveform is sampled 44100 times per second. This may seem a bit excessive, considering that we can't hear frequencies above 20 kHz; however, the highest frequency that a digitally sampled signal can represent is equal to half the sampling rate. So a sampling rate of 44100 Hz can only represent frequencies up to 22050 Hz, a boundary much closer to that of human hearing.

Quantization

Just as a waveform is sampled at discrete times, the value of the sample is also discrete. The quantization of a sample value depends on the number of bits used in measuring the height of the waveform. An 8-bit quantization yields 256 possible values; 16-bit CD-quality quantization results in over 65000 values. As an extreme example, Figure 2-3 shows the waveform used in the previous example sampled with a 3-bit quantization. This results in only eight possible values: .75, .5, .25, 0, -.25, -.5, -.75, and -1.

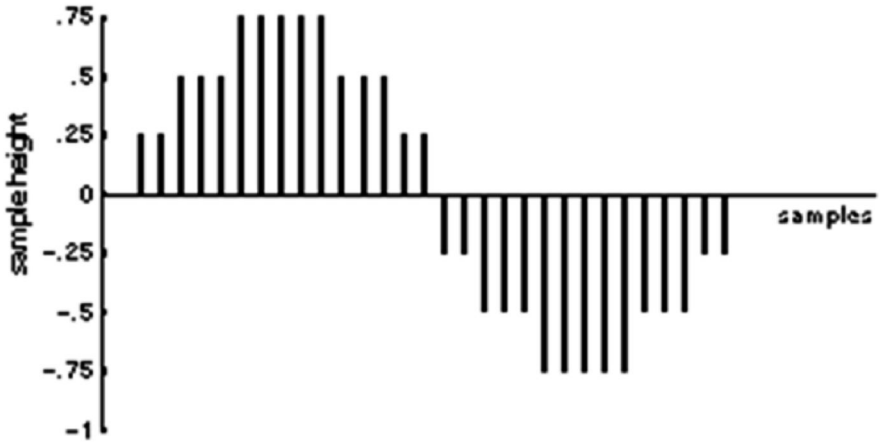


Fig.: Three-Bit Quantization

As you can see, the shape of the waveform becomes less discernible with a coarser quantization. The coarser the quantization, the "buzzier" the sound.

Storing Sampled Data

An increased sampling rate and refined quantization improves the fidelity of a digitally sampled waveform; however, the sound will also take up more storage space. Five seconds of sound sampled at 44.1 kHz with a 16-bit quantization uses more than 400,000 bytes of storage—a minute will consume more than five megabytes. A number of data compression schemes have been devised to decrease storage while sacrificing some fidelity.

1.3 Music Instruments Digital Interface

1.3.1 Basics concepts

Q7. Explain about the relationship between music and computer.

(OR)

Explain about MIDI.

Ans : **(Imp.)**

The relationship between music and computers has become more and more important, specially considering the development of MIDI (Music Instrument Digital Interface) and its important contributions in the music industry today. The MIDI interface between electronic musical instruments and computers is

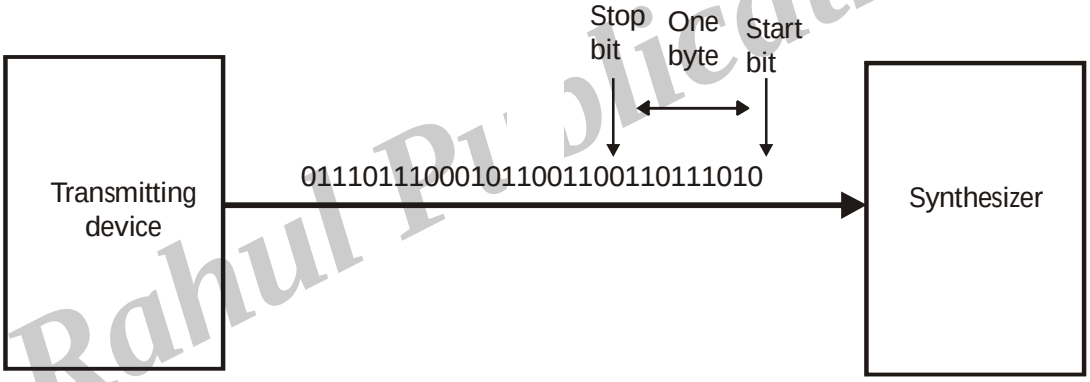
a small piece of equipment that plugs directly into the computer's serial port and allows the transmission of music signals. MIDI is considered to be the most compact interface that allows full-scale output.

MIDI Basic Concepts

Hardware connects the equipment. It specifies the physical connection between musical instruments, stipulated that a MIDI port is built into an instrument, specifies a MIDI cable which connects two instruments and deals with electronic signals that are sent over the cable.

A data format encodes the information traveling through the hardware. A MIDI data format does not include an encoding of individual samples as the audio format does. Instead of individual samples, an instrument connected data format is used. The encoding includes, besides the instrument specification, the notion of the beginning and end of a note, basic frequency and sound volume. The data in a MIDI status byte is between 128 and 255; each of the data bytes is between 0 and 127. Actual MIDI bytes are 8 bit, plus a 0 start and stop bit, making them 10 - bit "bytes". The above figure shows the MIDI data stream.

A MIDI device often is capable of programmability, which means it has filters available for changing the bass and treble response and can also change the "envelope" describing how the amplitude of a sound changes over time. The following figure shows a model of a digital instrument's response to Note On / Note Off messages.

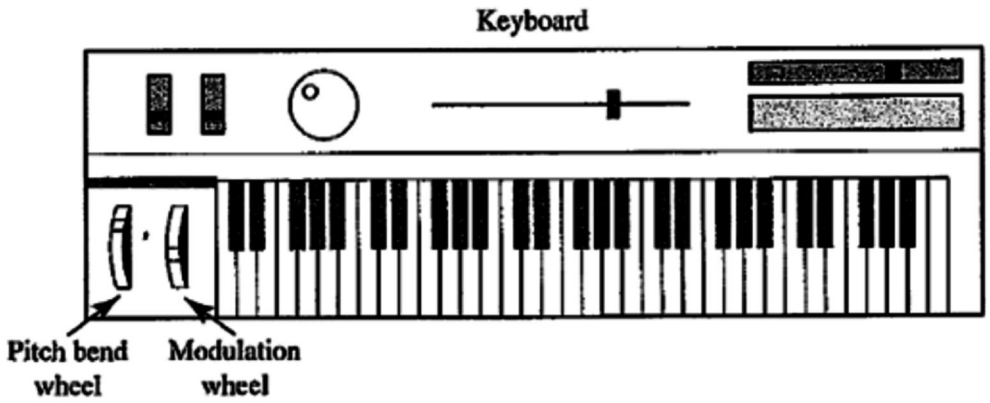


MIDI sequencers (editors) allow you to work with standard music notation or get right into the data, if desired. MIDI files can also store wave - table data. The advantage of wave - table data (WAV files) is that it much more precisely stores the exact sound of an instrument. A sampler is used to sample the audio data — for example, a "drum machine" always stores wave - table data of real drums.

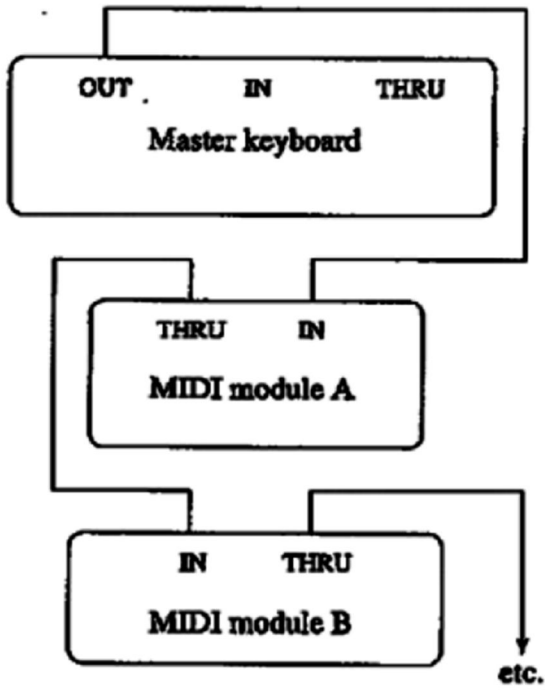
Sequencers employ several techniques for producing more music from what is actually available. For example, looping over (repeating) a few bars can be more or less convincing. Volume can be easily controlled over time — this is called time - varying amplitude modulation. More interestingly, sequencers can also accomplish time compression or expansion with no pitch change.

While it is possible to change the pitch of a sampled instrument, if the key change is large, the resulting sound begins to sound displeasing. For this reason, samplers employ multi - sampling. A sound is recorded using several band - pass filters, and the resulting recordings are assigned to different keyboard keys. This makes frequency shifting for a change of key more reliable, since less shift is involved for each note.

A MIDI Synthesizer



A typical MIDI setup



The above figure shows a typical MIDI sequencer setup. Here, the MIDI OUT of the keyboard is connected to the MIDI IN of a synthesizer and then THRU to each of the additional sound modules. During recording, a keyboard - equipped synthesizer sends MIDI messages to a sequencer, which records them. During playback, messages are sent from the sequencer to all the sound modules and the synthesizer, which play the music.

MIDI Software

Once a computer is connected to a MIDI system, a variety of MIDI applications can run on it. Digital computers afford the composer or sound designer unprecedented levels of control over the evolution and combination of sonic events.

The software applications generally fall into four major categories:

1. Music recording and performance applications

This Category of applications Provides funerals $1.191: as rec india 9f M191 messages as they enter the computer from other MIDI devices, and possibly editing and playing back the messages in performance.

2. Musical notations and printing applications

This category allows writing music using traditional musical notation. The user can then play back the music using a performance program or print the music on paper for live performance or publication.

a) Synthesizer patch editors and librarians

These programs allow information storage of different synthesizer patches in the computer's memory and disk drives, and editing of patches in the computer.

b) Music education applications

These software applications teach different aspects of music using the computer monitor, keyboard and other controllers of attached MIDI instruments.

The main issue in current MIDI—based computer music systems is interactivity. Music is a temporal art, and any computer program dealing with music must have sophisticated facilities for representing time and for scheduling processes to occur at a particular time, This capability of music applications became possible because of increased computational speeds therefore current computer music system are able to modify their behavior in response to input from other performing musicians.

1.3.2 Devices

Q8. Write about various devices used for MIDI.

Ans :

MIDI Devices

Through the MIDI interface, a computer can control output of individual instruments. On the other hand, the computer can receive, store or process coded musical data through the same interface. The data are generated with a keyboard 11d reproduced through a sound generator. A sequencer an store data. Further, it may also modify the musical data. In a multimedia system, the sequencer is a computer application.

The heart of any MIDI system is the MIDI synthesizer device. A typical synthesizer looks like a simple piano keyboard with a panel full of buttons, but it is far more (more detailed information on synthesizers can be found in [Bo087].).

Most synthesizers have the following common components:

(i) Sound Generators

Sound generators do the actual work of synthesizing sound; the purpose of the rest of the synthesizer is to control the sound generators. The principal purpose of the generator is to produce an audio signal that becomes when fed into a loudspeaker. By varying the voltage oscillation of the audio signal, a sound generator changes the quality of the sound — its pitch, loudness and tone color ,— to create a wide variety of sounds and notes.

Internally, sound generation can be done in different ways. One way is to store the acoustic as data in advance. Afterwards, MIDI data are transformed with a digital—analog adapter into acoustic signals. Individual notes are composed in a timely fashion. Another method is to create acoustic signals synthetically.

(ii) Microprocessor

The microprocessor communicates with the keyboard to know what notes the musician

is playing, and with the control panel, to know what commands the musician wants to send to the microprocessor. The micro processor then specifies note and sound commands to the sound generators; in other words, the microprocessor sends and receives MIDI messages.

(iii) Keyboard

The keyboard affords the musician's direct control of the synthesizer. Pressing keys on the keyboard signals the microprocessor what notes to play and how long to play them. Some synthesizer keyboards can also signal to the microprocessor how loud to play the notes and whether to add vibrato or other effects to the notes. The sound intensity of a tone depends on the speed and acceleration of the key pressure. The keyboard should have at least five octaves with 61 keys.

(iv) Control Panel

The control panel controls those functions that are not directly concerned with notes and durations (controlled by the keyboard). Panel controls include: a slider that sets the overall volume of the synthesizer, a button that turns the synthesizer on and off, and a menu that calls up different patches or the sound. generators to play.

(v) Auriliary Controllers

Auxiliary controllers are available to give more control over the notes played on the keyboard. Two very common variables on a synthesiser and pitch bend and modulation. Pitch bend controllers can bend pitch up and down, adding portamento to notes ; modulation controllers can increase or decrease effects such as vibrato.

(vi) Memory

Synthesizer memory is used to store patches for the sound generators and settings on the control panel. Many synthesizers also have a slot for external rmemory catridgess. By using several memory cartridges, the musician can plug in a different cartridge each time s / he wants a set of the synthesizer.

1.3.3 Messages

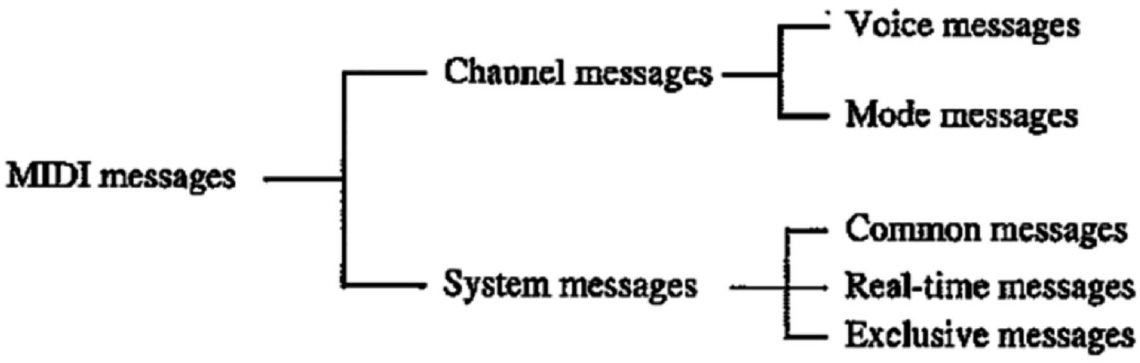
Q9. Explain various kinds of MIDI messages.

Ans : **(Imp.)**

MIDI Messages

MIDI messages can be classified into two types, as in the following figure — channel messages and system messages — and further classified as shown. Each type of message will be examined below.

MIDI message taxonomy



1. Channel Messages

Channel messages go only to specified devices. There are two types of channel messages:

- Channel voice messages send actual performance data between MIDI devices, describing keyboard action, controller action and control' panel changes. They describe music by defining pitch, amplitude, timbre , duration and other sound qualities. Examples of channel voice messages are Note On, Note Off, Channel Pressure, Control Change, etc.
- Channel mode messages determine the way that a receiving MIDI device responds to channel voice messages. They set the MIDI channel receiving modes for different MIDI devices, stop spurious notes from playing and affect local control of a device. Examples of such messages are Local Control, All Notes Off, omni Mode Off, etc.

2. System Messages

System messages go to all devices in a MIDI system because no channel numbers are specified. There are three types of system messages:

- System real-time messages are very short and simple, consisting of only one byte. They carry extra data with them. These messages synchronize 'the timing of MIDI devices in performance; therefore it is important that they be sent at precisely the time they are required. To avoid delays, these messages are sent in the middle of other messages, if necessary. Examples of such messages are System Reset, Timing Clock {MIDI clock), etc.
- System common messages are commands that prepare sequencers and synthesizers to play a song. 'The various messages enable you to select a song and find a common starting place in the sizers if they need tuning. Examples are Song Select, Tune Request, etc.
- System exclusive messages allow MIDI manufacturers to create customized MIDI messages to send between their MIDI devices. 'This coding' starts with a system-exclusive-message, where the manufacturer is speciûed, and ends with an end-of-exclusive message.

1.3.4 Timing Standards

Q10. What is SMPTE used for?

Ans : **(Imp.)**

SMPTE timecode is now used in film, TV, video and music production and "versions" of it appear embedded in digital video formats such as DV, and the MIDI language (MTC).

Here is a list of the most common uses of SMPTE...

- To aid the editing together of speech, sound effects and moving images during TV, film and video post production and editing
- To allow different sound recording technologies to be synchronized together during music production, such as ... tape recorders, computer sequencers, drum machine
- To aid composition and editing together of music and film, TV and video pictures
- To help automate the transfer of digital video clips between camcorders and computer editing systems
- To enable offline and online video editing work processes

Q11. What is SMPTE timecode?

Ans :

The SMPTE timecode is an hybrid computer/audio signal (more on this later) which can be recorded onto...

- a separate audio track of a portable (film location) audio recorder
- a separate audio track on film in a camera
- a separate audio track on a digital video cassette tape in a video camera
- a separate audio track (usually the highest) of a multitrack tape recorder

The SMPTE timecode signal carries Clock information in the form of hours, minutes, seconds and frames. Here is an example of a SMPTE display showing timing information ...

Q12. What are frame rates?

Ans :

SMPTE timecode allows each individual picture (or frame) of film or video to have its own unique identifying number. There are 4 SMPTE frame rates for the four worldwide TV broadcast and film frame rate standards each identified by its unique fps (frames per second) rate ...

- 30 fps The original black and white US TV rate. Unused since the advent of colour TV.
- 29.97 fps NTSC colour television and video. Also known as drop frame. Click here for a list of countries which use NTSC.
- 24 fps Worldwide film frame rate.
- 25 fps PAL colour television and video. Also known as EBU (European Broadcast Union), this rate was created for TV engineers working with the European PAL TV standard.

Q13. What is a SMPTE device?

Ans :

The SMPTE timecode signal is generated (and read back) by a SMPTE timecode generating device. Examples of SMPTE timecode devices include.

Stand alone devices Film production uses these to feed an identical SMPTE timecode signal simultaneously to both cameras and audio recorders on location

Devices built into cameras and camcorders

Interface devices (sometimes stand alone, sometimes PCI video cards) which are plugged into a computer system and controlled by software.

All SMPTE devices will generally have....

- A display showing timing and program parameter information.
- Controls to facilitate programming of frame rate, start time (for writing and reading), tempo, time signature etc.
- Output and input audio sockets for connection to a recorder.
- An output level control allowing the output (record) level of the SMPTE signal to be set
- MIDI In and Out ports if it is to designed for music production applications

A SMPTE timecode device must be able to both stripe (create/write/record) and read back a SMPTE timecode signal.

Q14. What physically is the SMPTE time code?

Ans :

Within a SMPTE device, the SMPTE timecode signal is in digital computer signal form (i.e. a very quiet pulse wave). Before it can be sent to be recorded by an audio recording device (camera, tape recorder etc) it must be amplified to the volume of an audible line level audio signal. In this form the SMPTE signal is unpleasant to listen to, making a sound that will be familiar to anyone who has loaded games programs from a cassette player into their Commodore 64 or Spectrum ZX80 home computer. Essentially the signal is a "digital" line level audio signal!

Once recorded, the SMPTE signal can be replayed and sent back to the SMPTE device, which will have been reset from "write" to "read" mode. The signal enters the SMPTE device at line level and is then "modulated" back down to digital computer signal level before re-entering the software environment.

Physically therefore, the SMPTE timecode signal will exist in 3 forms ...

1. digital computer signal
2. amplified "digital" line level audio signal
3. Magnetic pressure signal on magnetic tape

Q15. What are the other types of SMPTE.

Ans :

Other SMPTE "flavours"

SMPTE has been so successful in the TV and film industries that its principal concepts have been adopted for associated other industry technologies.

DV timecode

The DV (Digital Video) format includes an embedded timecode track that works exactly like SMPTE. There are 2 frame rates for NTSC (29.97fps) and PAL (25fps). DV timecode is used to automate batch capture (transfer) of clips from a DV camcorder to a computer editing system and to create EDL's for online/offline workflow's.

The process of using DV timecode is simple. The code is recorded along with pictures and sound to the DV tape during filming. Later during editing, when the camcorder is connected to a computer system, software on the computer reads the timecode and allows an editor to mark timecode "in" and "out" points for all the clips to be included in the edit. The software can then batch capture all the clips to the computer hard drive in preparation for editing. All through the editing process the timecode is used as a reference by the software to identify where edits, effects and transitions are made.

MIDI timecode (MTC)

SMPTE was adopted by the music business in the late 1970s to aid the synchronization of multitrack tape recorders, sequencers and drum machines and in 1989 a "flavour" of SMPTE called MIDI Time Code (MTC) was added to the MIDI specification to allow MIDI devices (computers, drum machines etc) to be inter grated seamlessly into the synchronisation process.

Q16. What timecode formats are used in music recording/production?

Ans :

In music production 3 timecode types are used...

- SMPTE (already discussed opposite)
- MTC (MIDI Time Code)
- MIDI Sync/Clock

MTC and MIDI Sync/Clock timecode's may be generated by a suitably equipped MIDI device such as a sequencer, drum machine or combined SMPTE/MIDI timecode device.

SMPTE is usually generated using a SMPTE device

1.4 Speech Generation

Q17. What is speech?

Ans :

Speech

Speech can be "perceived," "understood" and "generated" by humans and also by machines. A human adjusts himself/herself very efficiently to different speakers and their speech habits. The human speech signal comprises a subjective lowest spectral component known as tie pitch, which is not proportional to frequency.

Speech signals have two properties, which can be used in speech processing:

Voiced speech signals show during certain time intervals almost periodic behavior. Therefore, we can consider these signals as quasi-stationary signals for around 30_milliseconds.

The spectrum of audio signals shows characteristic maxima, which are mostly 3-5 frequency bands. These maxima, called formants, occur because of resonances of the voice tract.

Q18. What is speech generation?

Ans :

Speech generation and recognition are used to communicate between humans and machines. Rather than using your hands and eyes, you use your mouth and ears. This is very convenient when your hands and eyes should be doing something else, such as: driving a car, performing surgery, or (unfortunately) firing your weapons at the enemy. Two approaches are used for computer generated speech: digital recording and vocal tract simulation. In digital recording, the voice of a human speaker is digitized and stored, usually in a compressed form. During playback, the stored data are uncompressed and converted back into an analog signal.

Vocal tract simulators are more complicated, trying to mimic the physical mechanisms by which humans create speech. The human vocal tract is an acoustic cavity with resonate frequencies determined by the size and shape of the chambers. Sound originates in the vocal tract in one of two basic ways, called voiced and fricative sounds. With voiced sounds, vocal cord vibration produces near periodic pulses of air into the vocal cavities.

Q19. what is speech recognition?

Ans :

Speech recognition

The automated recognition of human speech is immensely more difficult than speech generation. Speech recognition is a classic example of things that the human brain does well, but digital computers do poorly. Digital computers can store and recall vast amounts of data, perform mathematical calculations at blazing speeds, and do repetitive tasks without becoming bored or inefficient.

Digital Signal Processing generally approaches the problem of voice recognition in two steps: feature extraction followed by feature matching. Each word in the incoming audio signal is isolated and then analyzed to identify the type of excitation and resonate frequencies. These parameters are then compared with previous examples of spoken words to identify the closest match. Often, these systems are limited to only a few hundred words;

can only accept speech with distinct pauses between words; and must be retrained for each individual speaker.

Q20. Explain about speech generation system.

Ans :

Speech Generation

An important requirement for speech generation is real-time signal generation. With such a requirement net a speech output ; automatically without any lengthy pre-processing. Some applications only need a limited vocabulary; an example is the spoken time announcement of a telephone answering service. However, most applications need a large vocabulary, if not an unlimited vocabulary.

Generated speech must be understandable and must sound natural. The requirement of understandable speech is a fundamental assumption, and the natural sound of speech increases user acceptance.

Basic Notions

- The lowest periodic spectral component of the speech signal is called the fundamental frequency. It is present in a voiced sound.
- A phone is the smallest speech unit, that distinguish one utterance or word from another in a given language.
- Allophones mark the variants of a phone. For example, the aspirated p of pit and the unaspirated p of spit are allophones of the English phoneme p.
- The morph marks the smallest speech unit which carries a meaning itself. Therefore, consider is. a morph, but reconsideration is not.
- A voiced sound is generated through the vocal cords. in, v and l are examples of voiced sounds.. The pronunciation of a voiced sound depends strongly on each speaker.

During the generation of an unvoiced sound, the vocal cords are opened. f and s are unvoiced sounds. Unvoiced sounds are relatively independent from the speaker.

Exactly, there are:

- Vowels - a speech sound created by the relatively free passage of ,breath through the larynx and oral cavity, usually forming the most prominent and central sound of a syllable .
- Consonants - a speech sound produced by a partial or complete obstruction of the air stream by any ot he various constrictions of the speech

Reproduced Speech Output

The easiest method of speech generation/output is to use pre-recorded speech and play it back in a timely fashion. The speech can be stored as PCM samples. Further data compression methods, without using language typical properties can be applied to recorded speech.

Types of Speech generation

The important requirement of speech generation is the generation of the real time signals. The easiest method for speech generation is to use pre-coded speech and play it back in the timely fashion.

Time-dependent sound concatenation

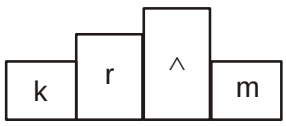


Fig.: Phone Sound Concatenation

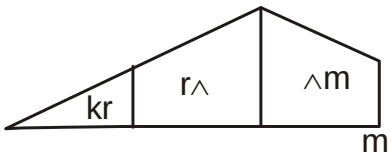


Fig.: Diphone Sound Concatenation

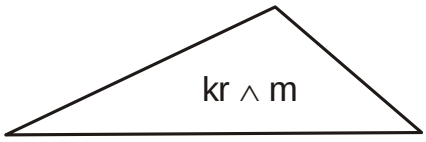


Fig.: Syllable Sound

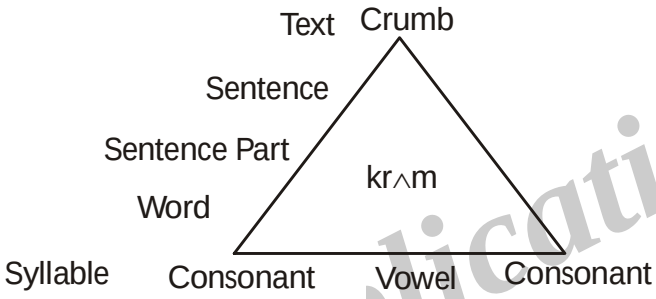


Fig.: Word Sound Concatenation

Speech generation can also be performed by sound concatenation in a timely fashion. Individual speech unit are composed like building blocks, where the composition can occur at different levels.

In the simplest case, the individual phones are understood as speech units. In the above given example individual phone of the word CRUM are shown. It is possible with just a few phones to create from an unlimited vocabulary. The phone sound concatenation shows the problem during the transition between individual phones. This problem is called co articulation which is mutual sound effect. To solve these problems, Di-phone sound concatenation is used. Two phones can constitute a Di-phone. In the above figure, Di-phone of word CRUM is shown. The transition problem is not solved sufficiently on this level.

To make the transition problem easier, syllabus can be created. The speech is generated through set of syllabus. The above given figure word sound concatenation and syllabus sound shows the syllabus sound of word CRUM. The best pronunciation of word is achieved is storage of whole word. This leads towards synthesize the speech sequence.

The transition between individual sound units create an essential problem called co articulation which is the mutual sound influence throughout the several sound.

Frequency dependent sound concatenation:

Speech generation can also be based on frequency dependent sound concatenation. E.g. through formant synthesizing. Formants are the frequency maxima in the spectrum of the speech. Formants synthesize

simulate the vocal track through filter. This characteristics value is filters middle frequency and their bandwidth. The method used for sound synthesize in order to simulate human speech is called the linear predictive coding (LPC) method.

Using speech synthesize a text can be transformed into acoustics signals. The typical component of this system is given in the figure below:

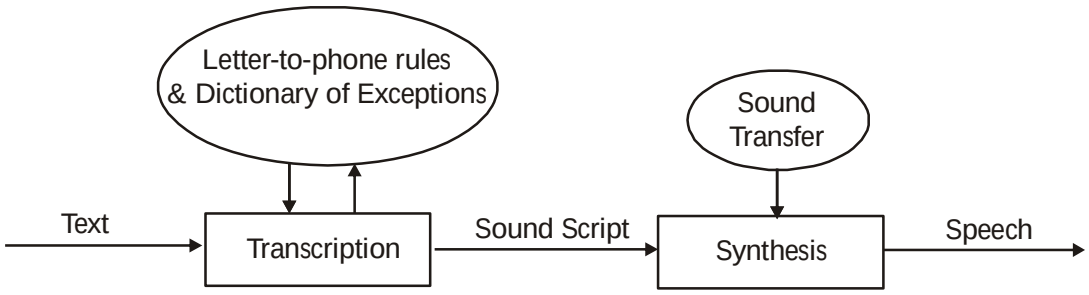


Fig.: Components of a speech synthesis system with time-dependent sound concatenation.

In the first step the transcription is performed where text is translated into sound script. Most transcription methods work with later to phone rules. A dictionary of exception stored in the library.

In the second step the sound script is translated into a speech signals. Besides the problem of co-articulation ambiguous pronunciation most be considered.

Q21. Explain about speech recognition system.

Ans :

Speech Recognition System

The automated recognition of human speech is immensely more difficult than speech generation. Speech recognition is a classic example of things that the human brain does well, but digital computers do poorly. Digital computers can store and recall vast amounts of data, perform mathematical calculations at blazing speeds, and do repetitive tasks without becoming bored or inefficient..

The system which provides the recognition and understanding of speech signal applies this principle several times as follows:

1. In the first step, the principle is applied to a sound pattern and/or word model. An acoustic and phonetic analysis is performed.
2. In the second step, certain speech units go through syntactical analysis. In this step the errors in the previous step can be recognized.
3. The third step deals with semantics of the previously recognized language. Here the decision error of previous step can be recognized and corrected.

There are still many problems into speech recognition and understanding research.

i) Specific problem is presented by Room Acoustic with existed environment noise.

ii) Word boundaries must be determined. Very often neighboring words follows into one another.

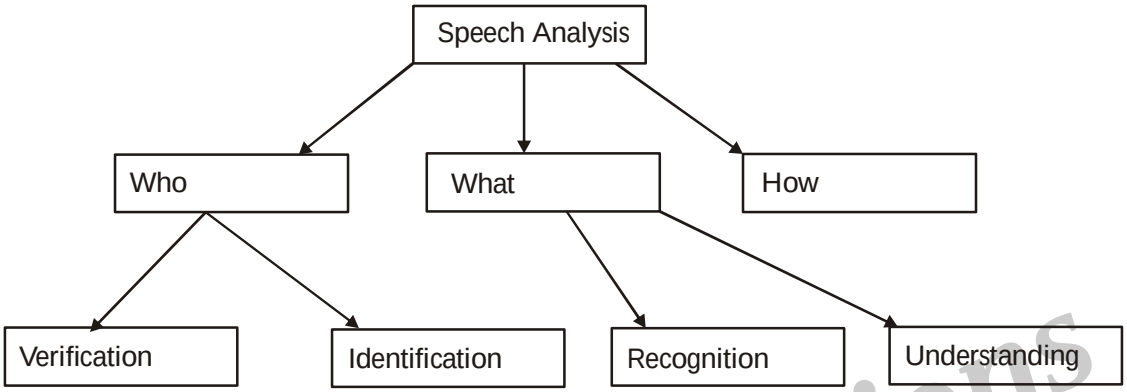
iii) For the comparison of the speech elements to the existing pattern, time normalization is necessary. The same word can be spoken quickly or slowly.

1.4.1 Analysis and Transmission

Q22. Write about the speech analysis areas.

Ans :

Speech analysis/input deals with the research areas shown in Figure:



Human speeches have certain characteristics determined by speaker. So speech analysis can server to analyze who is speaking. i.e. to recognize a speaker, for his identification and verification. The computer identifies and verifies the speaker using an acoustic finger print is digitally stored. It is digitally stored speech probe (certain statement of the speaker).

- Another main part of speech analysis is to analyze what has been said. To recognize and understand the speech itself.
- Another area of speech analysis tries to research speech pattern with respect to how the statement was said. E.g. a speaker sentence sounds differently if a person is angry or happy.

Q23. What is speech transmission? Explain about the components of speech transmission system.

Ans : **(Imp.)**

Speech Transmission

The area of speech transmission deals with efficient coding of the speech signal to allow speech/ sound transmission at low transmission rates over networks. The goal is to provide the receiver with the same speech/sound quality as was generated at the sender side.

Signal Form Coding

This kind of coding considers no speech-specific properties and parameters. Here, the goal is to achieve the most efficient coding of the audio signal. The data rate of a PCM-coded stereo-audio signal with CD-quality requirements is:

Rate = 2 * 44100/s * 16 bit / 8 bit/byte = 176400b bytes/s = 1411200 bits / s

Telephone quality, in comparison to CD-quality, needs only 64 Kbit/s. Using Difference Pulse Code Modulation (DPCM), the data rate can be lowered to 56 Kbits/s without loss of quality. Adaptive Pulse Code Modulation (ADPCM) allows a further rate reduction to 32 Kbits/s.

Source Coding

Parameterized systems work with source coding algorithms. Here, the specific speech characteristics are used for data rate reduction. Channel vo-coder is an example of such a parameterized system (Figure g.I2). The channel vo-coder is an extension of a sub-channel coding. The signal is divided into a set of frequency channels during speech analysis because only certain frequency maxima are relevant to speech.

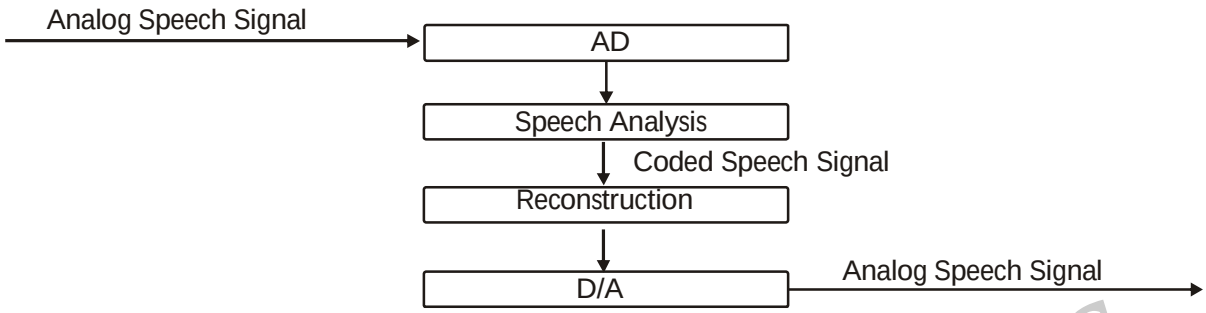


Fig.: Source coding in parameterized systems: components of a speech transmission system

- **Recognition / Synthesis Methods**

There have been attempts to reduce the transmission rate using pure recognition/ synthesis methods. Speech analysis (recognition) follows on the sender side of a speech transmission system and speech synthesis (generation) follows on the receiver side see Figure.

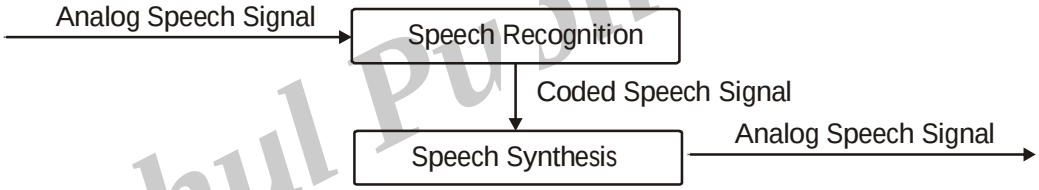
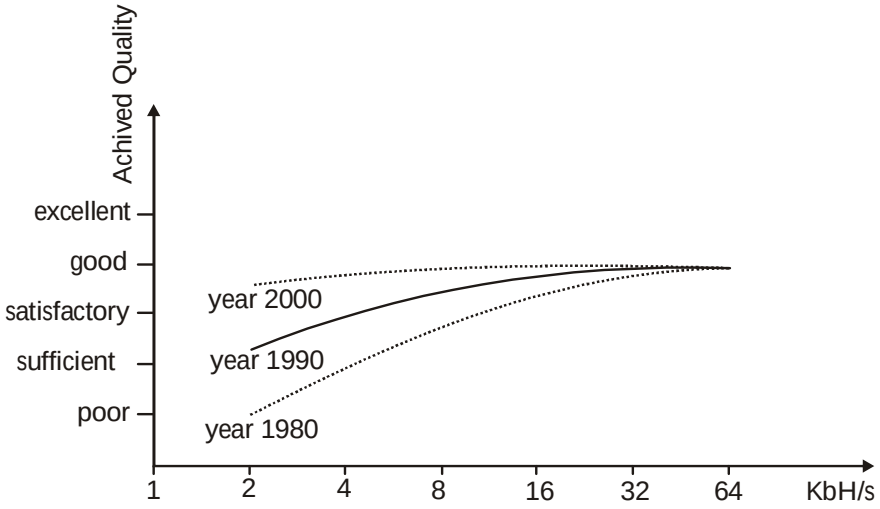


Fig.: Recognition/synthesis systems: components of a speech transmission system.

- **Achieved Quality**

The essential question regarding speech and audio transmission with respect to multimedia systems is how to achieve the minimal data rate for a given quality.



Short Question and Answers

1. What is Multimedia?

Ans :

A multimedia system is responsible for developing a multimedia application. A multimedia application is a bundle of different kinds of data. A multimedia computer system is one that can create, integrate, store, retrieve delete two or more types of media materials in digital form, such as audio, image, video, and text information.

2. What are the different types of Trans- mission media in Multimedia Systems.

Ans :

1. Perception Media

Perception media refers to the nature of information perceived by humans, which is not strictly identical to the sense that is stimulated. For example, a still image and a movie convey information of a different nature, though stimulating the same sense. The question to ask here is: How do humans perceive information?

In this context, we distinguish primarily between what we see and what we hear. Auditory media include music, sound and voice. Visual media include text, graphics and still and moving pictures. This differentiation can be further refined. For example, a visual medium can consist of moving pictures, animation, and text. In turn, moving pictures normally consist of series of scenes that, in turn, are composed of single pictures.

2. Representation Media

The term representation media refers to how information is represented internally to the computer. The encoding used is of essential importance. The question to ask here is: How is information encoded in the computer? There are several option:

- Each character of a piece of text is encoded in ASCII.
- A picture is encoded by the CEPT or CAPTAIN standard, or the GKS graphics standard can serve as a basis.
- An audio data stream is available in simple PCM encoding and a linear quantization of 16 bits per sampling value.
- A single image is encoded as group-3 facsimile or in JPEG format.
- A combined audio-video sequence is stored in the computer in various TV standards (eg. PAL, SECAM or NTSC), in the CCIR-601 standard, or in MPEG format.

3. Presentation Media

The term presentation media refers to the physical means used by systems to reproduce information for humans. For example, a TV set uses a cathode-ray tube and loudspeaker. The question to ask here is: Which medium is used to output information from the computer or input in the computer?

We distinguish primarily between output and input. Media such as paper, computer monitors, and loudspeakers are output media, while keyboards, cameras, and microphones are input media.

3. What is Sound?

Ans :

Meaning

Sound is a physical phenomenon produced by the vibration of matter. The matter can be almost anything: a violin string or a block of wood, for example. As the matter vibrates, pressure variations are created in the air surrounding it. This alternation of high and low pressure is propagated through the air in a wave-like motion. When the wave reaches our ears, we hear a sound.

4. MIDI.

Ans :

The relationship between music and computers has become more and more important, specially considering the development of MIDI (Music Instrument Digital Interface) and its important contributions in the music industry today. The MIDI interface between electronic musical instruments and computers is a small piece of equipment that plugs directly into the computer's serial port and allows the transmission of music signals. MIDI is considered to be the most compact interface that allows full-scale output.

5. MIDI Devices

Ans :

Through the MIDI interface, a computer can control output of individual instruments. On the other hand, the computer can receive, store or process coded musical data through the same interface. The data are generated with a keyboard 11d reproduced through a sound generator. A sequencer an store data. Further, it may also modify the musical data. In a multimedia system, the sequencer is a computer application.

The heart of any MIDI system is the MIDI synthesizer device. A typical synthesizer looks like a simple piano keyboard with a panel full of buttons, but it is far more (more detailed information on synthesizers can be found in [Bo087].).

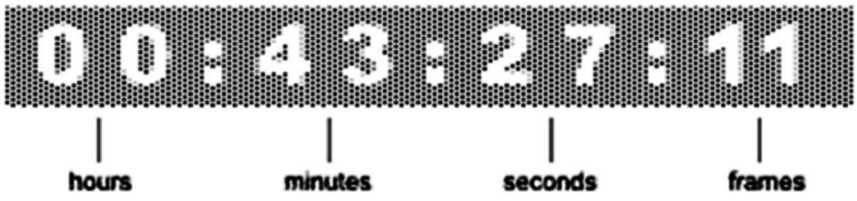
6. What is SMPTE timecode?

Ans :

The SMPTE timecode is an hybrid computer/audio signal (more on this later) which can be recorded onto...

- a separate audio track of a portable (film location) audio recorder
- a separate audio track on film in a camera
- a separate audio track on a digital video cassette tape in a video camera
- a separate audio track (usually the highest) of a multitrack tape recorder

The SMPTE timecode signal carries Clock information in the form of hours, minutes, seconds and frames. Here is an example of a SMPTE display showing timing information ...



7. What are frame rates?

Ans :

SMPTE timecode allows each individual picture (or frame) of film or video to have its own unique identifying number. There are 4 SMPTE frame rates for the four worldwide TV broadcast and film frame rate standards each identified by its unique fps (frames per second) rate ...

- 30 fps The original black and white US TV rate. Unused since the advent of colour TV.
- 29.97 fps NTSC colour television and video. Also known as drop frame. Click here for a list of countries which use NTSC.
- 24 fps Worldwide film frame rate.
- 25 fps PAL colour television and video. Also known as EBU (European Broadcast Union), this rate was created for TV engineers working with the European PAL TV standard.

8. What is speech?

Ans :

Speech

Speech can be "perceived," "understood" and "generated" by humans and also by machines. A human adjusts himself/herself very efficiently to different speakers and their speech habits. The human speech signal comprises a subjective lowest spectral component known as tie pitch, which is not proportional to frequency.

Speech signals have two properties, which can be used in speech processing:

Voiced speech signals show during certain time intervals almost periodic behavior. Therefore, we can consider these signals as quasi-stationary signals for around 30_milliseconds.

The spectrum of audio signals shows characteristic maxima, which are mostly 3-5 frequency bands. These maxima, called formants, occur because of resonances of the voice tract.

9. What is speech generation?

Ans :

Speech generation and recognition are used to communicate between humans and machines. Rather than using your hands and eyes, you use your mouth and ears. This is very convenient when your hands and eyes should be doing something else, such as: driving a car, performing surgery, or (unfortunately) firing your weapons at the enemy. Two approaches are used for computer generated speech: digital recording and vocal tract simulation. In digital recording, the voice of a human speaker is digitized and stored, usually in a compressed form. During playback, the stored data are uncompressed and converted back into an analog signal.

10. What is speech recognition?

Ans :

Speech recognition

The automated recognition of human speech is immensely more difficult than speech generation. Speech recognition is a classic example of things that the human brain does well, but digital computers do poorly. Digital computers can store and recall vast amounts of data, perform mathematical calculations at blazing speeds, and do repetitive tasks without becoming bored or inefficient.

Digital Signal Processing generally approaches the problem of voice recognition in two steps: feature extraction followed by feature matching. Each word in the incoming audio signal is isolated and then analyzed to identify the type of excitation and resonate frequencies. These parameters are then compared with previous examples of spoken words to identify the closest match. Often, these systems are limited to only a few hundred words; can only accept speech with distinct pauses between words; and must be retrained for each individual speaker.

Choose the Correct Answers

1. _____ refers to the nature of information perceived by humans [b]

   (a) Representation Media (b) Perception media

   (c) Presentation Media (d) Storage Media

2. What is multimedia file ? [b]

   (a) Is same as any other regular file

   (b) Must be accessed at specific rate

   (c) Stored on remote server can not be delivered to its client

   (d) None of the mentioned

3. The major difference between a multimedia file and a regular file is _____ [d]

   (a) The size (b) The attributes

   (c) The ownership (d) The rate at which the file must be accessed

4. Multimedia files stored on a remote server are delivered to a client across the network using a technique known as _____ [b]

   (a) Download (b) Streaming

   (c) Flowing (d) Leaking

5. What are the two types of streaming techniques? [a]

   (a) Progressive download & real time streaming

   (b) Regular download & real time streaming

   (c) Real time & virtual time streaming

   (d) Virtual time streaming

6. A media file containing audio or video is downloaded and stored on the client's local file system in _____ [a]

   (a) Progressive download (b) Regular download

   (c) Real time streaming (d) Virtual time streaming

7. The full form of MIDI [c]

   (a) Music Instrument Digital Instructions

   (b) Music Instrument Data Interface

   (c) Music Instrument Digital Interface

   (d) Multimedia Instruction Digital Interface

8. The SMPTE time code is an hybrid computer/audio signal (more on this later) which can be recorded onto _____ [d]

(a) a separate audio track of a portable (film location) audio recorder

(b) a separate audio track on film in a camera

(c) a separate audio track on a digital video cassette tape in a video camera

(d) all the above

9. Among the following which is not the element of multimedia [d]

(a) Text (b) image,

(c) audio (d) time

10. Among the following which domain deals with interaction between multimedia application and multimedia devices [a]

(a) Device domain (b) System domain

(c) Application domain (d) Cross domain

Fill in the blanks

1. The _____ of a sound the number of times the pressure rises and falls, or oscillates, in a second is measured in hertz (Hz).
2. The _____ refers to the physical means-cables of various types, radio tower, satellite etc.
3. _____ Provides functions to the user to develop and present multimedia projects
4. An important aspect of different media is their level of _____ from each other
5. _____ mode of transmission specified a maximum end-to-end delay transmission of packets,
6. If consecutive packets are directly transmitted one after another without any time gap then such data streams are called _____
7. The mechanism that converts an audio signal into digital samples is called an _____
8. The _____ of a sample value depends on the number of bits used in measuring the height of the waveform
9. _____ are available to give more control over the notes played on the keyboard
10. _____ and _____ are used to communicate between humans and machines

ANSWERS

1. Frequency
2. Transmission media
3. Application domain
4. Independence
5. Synchronous Transmission mode
6. Continuous data stream
7. Analog-to-digital converter
8. Quantization
9. Auxiliary controllers
10. Speech generation and recognition

UNIT II

DIGITAL IMAGE & ANIMATIONS:

Digital Image: Analysis, recognition, transmission, **Video**: Representation, Digitalization, transmission.

Animations: Basic concepts, animation languages, animations control transmission.

2.1 DIGITAL IMAGE

2.1.1 Image Analysis

Q1. What is image? How to represent digital image?

Ans :

Image

An image may be defined as two dimensional function f(x, y) where, x y are the spatial co-ordinate and the amplitude of 'f' at any pair of co-ordinates x, y is called the intensity or gray level of image at that point x, y and amplitude values of f are all finite, discrete quantities we all the image is digital image.

Digital Image Representation

For computer representation, function (e.g. intensity) must be sampled at discrete intervals.

Sampling quantizes the intensity values into discrete intervals.

- Point at which an image is sampled are called picture elements or pixels.
- Resolution specifies the distance between points accuracy.

A digital image is represented by a matrix of numeric values each representing a quantized intensity value.

A digital image is a numeric representation (normal binary) of two dimensional images.

When I is a two-dimensional matrix, then I (r, c) is the intensity value at the position corresponding to row r and column c of the matrix. Intensity value can be represented by bits for black and white images (binary valued images), 8 bits for monochrome imagery to encode color or grayscale levels, 24 bit (RGB).

Q2. What is image format? Discuss its types.

Ans : **(Imp.)**

Image Format

There are different kinds of image formats in the literature. We shall consider the image format that comes out of an image frame grabber, i.e., the captured image format, and the format when images are stored, i.e., the stored image format.

(i) Captured Image Format

(ii) Stored Image Format

(i) Captured Image Format

The image format is specified by two main parameters: spatial resolution, which is specified as pixels x pixels (e.g. 640x480) and color encoding, which is specified by bits per pixel. Both parameter values depend on hardware and software for input/output of images.

(ii) Stored Image Format

When we store an image, we are storing a two-dimensional array of values, in which each value represents the data associated with a pixel in the image. For a bitmap, this value is a binary digit. A bitmap is a simple information matrix describing the individual dots that are the smallest elements of resolution on a computer screen or other display or printing device.

Image file format include:

- GIF (Graphic Interchange Format)
- X11 bitmap
- Postscript

- JPEG (Joint Picture Expert Group)
- TIFF (Tagged Image File Format) etc.

There are many file formats used to store bitmaps and vectored drawing. Following is a list of few image file formats.

Format	Extension
Microsoft Windows DIB	.bmp .dib .rle
Microsoft Palette	.pal
Autocad format 2D	.dxf
JPEG	.jpg
Windows Meta file	.wmf
Portable network graphic	.png
Compuserve gif	.gif
Apple Macintosh	.pict .pic .pct

Graphics Format

Graphic image formats are specified through graphics primitives and their attributes.

- Graphic primitive – line, rectangle, circle, ellipses, specification 2D and 3D objects.
- Graphic attribute – line style, line width, color.

Graphics formats represent a higher level of image representation,

i.e., they are not represented by a pixel matrix initially.

- PHIGS (Programmer's Hierarchical Interactive Graphics)
- GKS (Graphical Kernel System)

Q3. What is digital image processing?

Ans :

- Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image.
- Image processing usually refers to digital image processing, but optional and analog image processing also are possible.
- Computer graphics concern the pictorial synthesis of real or imaginary objects from their computer-based models.

The related field of image processing treats the converse process: the analysis of scenes, or the reconstruction of model from pictures of 2D or 3D objects.

Q4. Explain the frame work of interactive graphics system.

Ans :

The Framework of Interactive Graphics System

Graphics are not confined to static pictures. Picture can be dynamically varied; for example, a user can control animation by adjusting the speed, portion of the total scene inn view, amount of detail shown, etc.

Image can be generated by video digitizer cards that capture NTSC (PAL) analog signals and create a digital image.

Graphical images are generated using interactive graphics systems.

The high-level conceptual framework of almost any interactive graphics system consists of three software components: *an application model, an application program and a graphics system*, and a hardware component: graphics hardware.

1. Application Model

The application model represents the data or objects to be picture on the screen; it is stored in an application database. The model is an application-specific and is created independency of any particular display system.

2. Application Program

The application program handles user input. It produces views by sending to the third component, the graphics system, a series of graphics output commands that contain both a detailed geometric description of what is to be viewed and the attributes describing how the objects should appear.

3. Graphics System

The graphics system is responsible for actually producing the picture from the detailed descriptions and for passing the user's input to the application program for processing.

The graphics system is an intermediary component between the application program and the display hardware.

Q5. Explain graphics hardware for digital image processing.

Ans :

Graphics Hardware

At the hardware level, a computer receives input from interaction devices and output images to display devices.

Input

Current input technology provide us with the ubiquitous mouse, the data tablet and transparent, touch sensitive panel mounted on the screen. The other graphics input are track-balls, space-balls or the data glove. Track-ball can be made to sense rotation about the vertical axis in addition to the about two horizontal axes.

A space-ball is a rigid sphere containing strain gauges. The user pushes or pulls the sphere in any direction, providing 3D translation and orientation. The data glove records hand position and orientation as well as finger movements. It is a glove covered with small, lightweight sensors.

Output: Raster Display

- Most common type of graphic monitors using raster scan display type CRT
- Point plotting device
- Based on TV technology
- Electron beam is swept across the screen, one row at a time from top to bottom,starting at the upper left corner of the display
- Process is repeated until the entire screen is covered, and the beam is then
- Returned to the upper left corner to start a new scan
- Beam intensity is turned on and off to create a pattern of illuminated spots
- Pictures are dynamically stored in a piece of memory known as frame buffer or refresh buffer
- This buffer holds the set of intensity values all the screen points (pixels)
- Requirement to control the intensity of the screen positions:
- Simple black and white system:
  - o 1 bit per pixel (bitmap)
  - o Color system:
    - 24 bits/pixel (maximum no. of color representation, pixmap)

Frame buffer or refresh buffer (storage) require-ments:

- Large storage e.g. 24 bits/pel, screen resolution of 1024x1024 requires 3mb of RAM
- Refresh rate: 60 to 80 frames per second

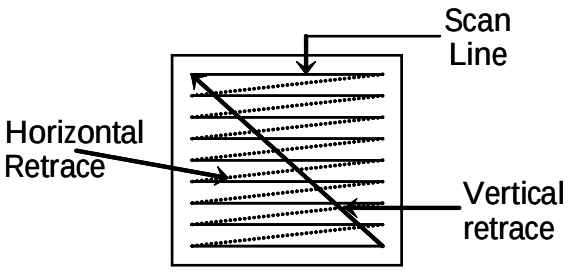


Fig. : Raster Scan

Advantages of Raster Scan Display:

- Capable of presenting bright pictures
- Unaffected by picture complexity
- Suitable for showing dynamic motion
- Lower cost
- Ability to display areas filled with solid colors or patterns

Disadvantages of Raster Scan Display:

- Requires large amount of memory (RAM)
- Produced "stair stepped" appearance of diagonal lines on the image (known as aliasing effect)
- True line cannot be represented exactly due to the discretization of the display surface (discrete nature of pixel representation)

Q6. Explain analysis process of image.

Ans :

- Image analysis is concerned with techniques for extracting descriptions from images that are necessary for high-level scene analysis methods.
- Image analysis techniques include computation of perceived brightness and color, partial or complete recovery of three-dimensional data in the scene, location of discontinuities corresponding to objects in the scene and characterization of the properties of uniform regions in the image.
- Image processing includes image enhancement, pattern detection and recognition and scene analysis and computer vision.
- Image enhancement deals with improving image quality by eliminating noise or by
- enhancing contrast.
- Pattern detection and recognition deal with detecting and clarifying standard patterns and finding distortions from these patterns.
- Scene analysis and computer vision deal with recognizing and reconstructing 3D models of a scene from several 2D images.

2.1.2 Image Recognition

Q7. Define Image Recognition. Explain image recognition steps with diagram.

Ans :

Image Recognition

Image recognition can be defined as the capability of a computer to determine the objects or features of an image.

Steps

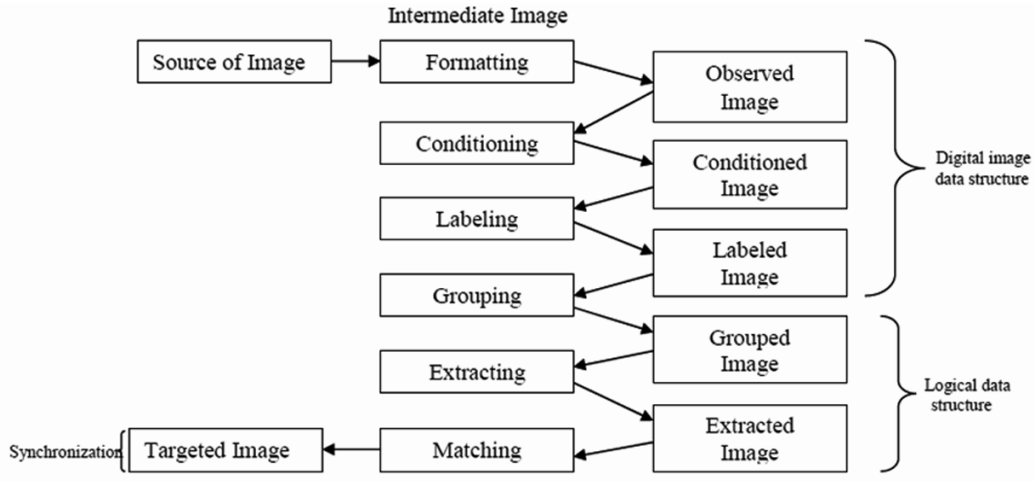


Fig. : Image Recognition Steps

1. Formatting

Capturing an image from a camera and bringing it into a digital form. It means that we will have a digital representation of an image in the form of pixels.

2. Conditioning

In image, there is usually uninteresting object introduced during digitize as noise. In conditioning, interesting objects are highlighted by suppressing or analyzing uninteresting in systematic or patterned variations. Conditioning is typically applied uniformly and is context-independent.

3. Labeling

The informative pattern has structure as a spatial arrangement of events, each spatial event being a set of connected pixels. Labeling determines in what kinds of spatial events each pixel participates.

E.g. edge detection technique

Edge detection technique determines continuous adjacent pairs which differ in intensity or color. Another labeling operation must occur after edge detection, namely thresholding.

Thresholding specifies which edges should be accepted and which should not; the thresholding operation filters only the significant edges from the image and labels them.

4. Grouping

It can turn edges into line by determining edges belongs to same spatial event. A grouping operation, where edges are grouped into lines, is called line filtering. The grouping operation involves a change of logical data structure.

5. Extracting

Generating list of properties from set of pixel in spatial event. Extraction can also measure topological or spatial relationship between two or more grouping.

6. Matching

After the completion of the extracting operation, the events occurring on the image have been identified and measured but the events in and of themselves have no meaning.

It is the matching operation that determines the interpretation of some related set of image events, associating these events with some given three dimensional object or two dimensional shape.

The classic example is template matching, which compares the examined pattern with stored models (templates) of known patterns and chooses the best match.

2.1.3 Image Transmission

Q8. What is image transmission? Explain various methods of raw image transmission?

Ans :

Image transmission takes into account transmission of digital images through computer networks. There are several requirements on the networks when images are transmitted:

- The network must accommodate bursty data transport because image transmission is bursty (The burst is caused by the large size of the image).
- Image transmission requires reliable transport.
- Time-dependence is not a dominant characteristic of the image in contrast to audio/video transmission.

Image size depends on the image representation format used for transmission. There are several possibilities.

(i) Raw Image Data Transmission

The image is generated through a video digitizer and transmitted in its digital format.

Size = Spatial_resolution x Pixel_quantization

For example, the transmission of an image with a resolution of 640 x 480 pixels and pixel quantization of 8 bit per pixel requires transmission of 307,200 bytes through the network.

(ii) Compressed Image Data Transmission

- The image is generated through a video digitizer and compressed before transmission.
- The reduction of image size depends on the compression method and compression rate.
- JPEG (Joint Picture Expert Group) & MPEG (Motion Picture Expert Group)

(iii) Symbolic Image Data Transmission

The image is represented through symbolic data representation as image primitives (e.g. 2D or 3D geometric representation), attributes and other control information.

2.2 VIDEO REPRESENTATION

Q9. Explain about various measures taken for visual representation of a video signal.

Ans : **(Imp.)**

Visual Representation

The objective is to offer the viewer a sense of presence in the scene and of participation in the events portrayed. To meet the objective, the televised images should convey spatial and temporal content of the scene.

Important measures taken for the purpose include:

1. Vertical Detail and the Viewing Distance

Geometry of the field occupied by the television image is based on the ratio of picture width and height, called as aspect ratio (W/L = 4/3 = 1.33)

Viewing distance D determines the angle h subtended by the picture height. The angle is usually measured by the ration of the viewing distance to the picture height (D/H). The smallest detail that can be reproduced in the image is pixel

Both the number of pixels per scanned line and the number of lines per frame vary; the actual numbers used being determined by the aspect ratio. This is the ratio of the screen width to the screen height. The aspect ratio of current television tubes is 4/3 with older tubes – on which PC monitors are based – and 16/9 with wide screen television tubes.

2. Horizontal Detail and Picture Width

Picture width chosen for the conventional television services is 4/3 * picture height. Using the aspect ratio, we can determine the horizontal field of view from the horizontal angle.

The horizontal detail or the picture width is again dependent upon the aspect ratio and is given by aspect ratio *height of picture.

3. Total Detail Content of the Image

The vertical resolution is equal to the number of picture elements separately presented in the picture height, while the number of elements in the picture width is equal to the horizontal resolution times the aspect ratio.

The product of horizontal and vertical elements gives the total picture elements in the image. The total detail content of the image is given by the number of pixels used to represent the image and is given by the resolution. Thus total detail content of the image is the product of Vertical Detail * Horizontal Detail.

4. Perception of Depth

In natural vision, perception of the third spatial dimension, depth, depends primarily on the angular separation of the images received by the two eyes if the viewer.

In the flat image television, a considerable degree of depth perception is inferred from the perspective appearance of the subject matter.

As the screen is two dimensional a special measure is taken to give the sense of depth. This involves the perspective representation of the image content. This is again governed by the focal length of the lenses and changes in depth of focus in camera.

5. Luminance and Chrominance

Color vision is achieved through three signals, proportional to the relative intensities of Red, green and blue lights (RGB) in each portion of the scene. During the transmission of the signals from the camera to the television (display), a different division of signals in comparison to the RGB division is often used. This color encoding during transmission uses luminance and chrominance.

The term luminance is used to refer to the brightness of a source, and the color information is referred by the chrominance. Human eye however is more sensitive to luminance than to chrominance. The color encoding during transmission uses luminance and two chrominance signals (for hue and saturation).

Thus luminance (Ys) and two chrominance blue chrominance (Cb), and the red chrominance (Cr) are then used to represent the video content where

$$Y_s = 0.299\ Rs + 0.587\ G_s + 0.144\ B_S$$

$$C_b = B_s - Y_s \text{ and } Cr = R_s - Y_s$$

6. Temporal Aspects of Illumination

In contrast to continuous pressure waves of an acoustic signal, a discrete sequence of individual pictures can be perceived as a continuous sequence. This property is used in television and motion pictures. The rate of repetition of the images must be high enough to guarantee smooth motion from frame to frame and the persistence of vision extends over the interval between flashes.

One of the major characteristics of human eye is that if the still images are shown in rapid succession, human eye perceive the motion i.e. it does not notice the brief cut off of the light.

Thus if the rate of succession is sufficiently high still images can be used to represent motion as the persistence of vision extends over the interval between flashes.

7. Continuity of Motion

NTSC specified the frame rate for maintaining the visual-aural carrier separation at 4.5 MHz was 30 frames/sec but has been changed to 29.97 HZ. PAL adopted the frame repetition rate of 25 Hz, and the frame rate is 25frames/sec. Digital motion picture uses frame rate of 50 frames/sec.

As stated above, motion can be represented by showing the still images in rapid succession. This can be done by showing the still images 15 frames per second. Video motion seems smooth and is achieved at only 30 frames per second.

8. Flicker

Through a slow motion, a periodic fluctuation of brightness perception, a flicker effect arises. The marginal value to avoid flicker is at least 50 refresh cycles per second. To achieve continuous flicker-free motion, we need relatively high refresh frequency. Movies and television apply some technical measures to work with lower motion frequencies.

When the refresh rate of frames is smaller than the eye will notice the cutoff of light between the frames and there arises the flickering rate. The marginal value to avoid flicker is at least 50 refresh cycles per second.

9. Temporal Aspect of Video Bandwidth

The most important factor to determine at which bandwidth the video can be transmitted is its temporal specification. Temporal specification depends on the rate of the visual system to scan pixels, as well as on the human eye's scanning capabilities. In a HDTV (High Definition TV) device, a pixel

can be scanned in less than a tenth of a millionth of a second. From the human visual perspective, the eye requires that a video frame to be scanned every 1/25 second, this time is equivalent to the time during which human eye does not see the flicker effect.

The choice of bandwidth for transmission of the video signal depends on the rate of the visual system to scan pixels, as well as on the human eye's scanning capabilities. From the human visual perspective, the eye requires that a video frame be scanned every 1/25 second.

Q10. Explain about the transmission process of video signal.

(OR)

Discuss in deail about video transmission.

Ans :

Standards for transmission of video are: NTSC (National Television System Committee) and PAL (Phase Alternating Line). Approaches of color encoding are: RGB, YUV (luminance Y, two chrominance channels U and V), YIQ, Composite, etc.

Different color encoding techniques can be employed while transmitting the video signal and they are:

1. RGB Signal

In this case separate signal coding is done for the individual R, G and B components of the image and these components are transmitted separately.

2. YUV Signal

This model exploits the fact that the human eye is more sensitive to brightness i.e. luminance than to color information i.e. chrominance. Thus instead of separating the color component the brightness and the coloration of the image is separated. Thus luminance (Ys) and two chrominance blue chrominance (Cb), and the red chrominance (Cr) are then used to represent the video content where

$$Y = 0.30\ R_s + 0.59\ G_s + 0.11\ B_s$$

$$U = (B_s - Y_s)*0.493$$

$$V = (R_s - Y_s)*0.877$$

As the bandwidth for color broadcasts must be the same as monochrome, in order to fit the Y, Cb, and Cr signals in the same bandwidth, the three signals must be combined together for transmission. The resulting signal is then known as composite video signal.

3. YIQ Signal

This model is very much similar to YUV model and is used in NTSC system. The image information is broken down into three components a luminance component and two chrominance component given by following relation.

Q11. What is known as composite signal?

Ans :

The alternative to component coding composes all information into one signal; consequently, the individual components (RGB, YUV, and YIQ) must be combined into one signal. The basic information consists of luminance information and chrominance difference signals.

2.2.1 Video Digitalization

Q12. What is digitalization of a video signal?

Ans :

Digitalization is a process of changing the continuously varying signal into discrete components. This basically uses mathematical process like Fourier analysis or a series of step consisting of sampling and quantizing. Sampling is the process that actually digitizes the spatial position of the image while quantizing digitizes its color information. Sampling involves dividing the picture at M*N array of points while quantizing involves dividing the signal into a range of gray level values. Finally the digital motion video is created by digitizing the pictures temporally.

Q13. Explain about computer video format and its standards.

Ans : **(Imp.)**

In order to standardize the process of digitization the international body for television standards, the International Telecommunications Union – Radio communications Branch (ITU-R) defined a standard for the digitization of video pictures known as Recommendation CCIR-601. In addition, a number of variants of this standard have been defined for use in other application domains such as television broadcasting, video telephony, and video conferencing. The video digitizers differ in digital image resolution, quantization and frame rate. For e.g. VINO's resolution for NTSC is 640 by 480 with 8 bits/pixel and frame rate of 4 frames/second.

Computer video format depends on the input and output devices for the motion video medium

- Current video digitizers differ in digital image resolution, quantization and frame rate
- Most often used display is raster display
- The raster display architecture (as shown below)

Typically, the graphics program is used to create the high-level version of the image interactively and the Display controller part of the program interprets sequences of display commands and converts them into displayed objects by writing the appropriate pixel values into the refresh buffer.

Normally the video controller is a hardware subsystem that reads the pixel values stored in the refresh buffer in time-synchronism with the scanning process and, for each set of pixel values, converts these into the equivalent set of red, green, and blue analog signal for output of the display.

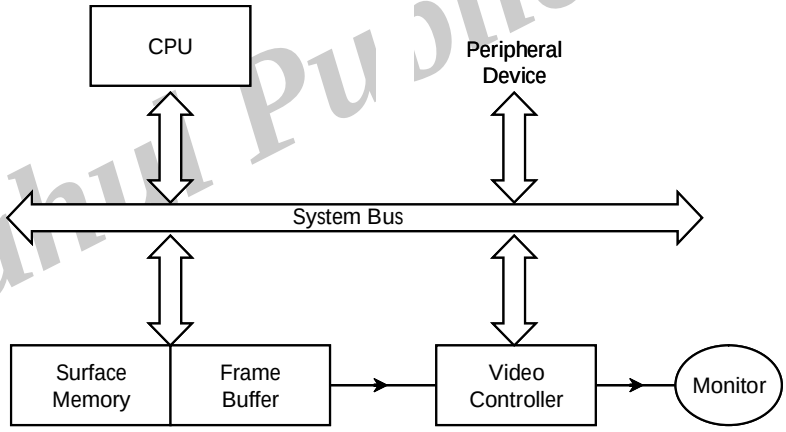


Fig. : A common raster display system

A range of color values are stored in a table and each pixel value is used as an address to a location within the table which contains the corresponding three color values. The table is known as color look up table.

Some computer video controller standards are given below:

Color Graphics Adapter (CGA)

Resolution: 320*200

Color depth: 2 bits/pixel

Image size: 16 KB

No. of colors: 4

Storage Capacity: 320*200*4/8 = 16,000 bytes

Enhanced Graphics Adapter (EGA)

Resolution: 640*480

Color depth: 4 bits/pixel

Image size: 112 KB

No. of colors: 16

Storage Capacity: 640*480*4/8 = 112,000 bytes

Video Graphics Array (VGA)

Resolution: 640*480

Color depth: 8 bits/pixel

Image size: 307.2 KB

No. of colors: 256

Storage Capacity: 640*480*8/8 = 307,200 bytes

8514/A Display Adapter Mode

Resolution: 1024*768

No. of colors: 256

Storage Capacity: 1024*768*8/8 = 786,432 byte

Extended Graphics Array (XGA)

Resolution: 640*480 / 1024*768

Color depth: 65,000 colors / 256 colors

No. of colors: 256

Storage Capacity: 1024*768*8/8 = 786,432 bytes

Super VGA (SVGA)

Resolution: 1024*768

Color depth: 24 bits/pixel

Image size: 2.35 MB

No. of colors: 224

Storage Capacity: 1024*768*24/8 = 2,359,296 bytes.

Q14. Explain various standards of television conventional system.

Ans : **(Imp.)**

Conventional television systems employ the following standards:

1. NTSC
2. SECAM
3. PAL

1. NTSC

- National Television Systems Committee

- Developed in the U.S.
- Oldest and most widely used television standard
- Color carrier is used with approximately 4.429 MHz or with approximately 3.57 MHz.
- Uses a quadrature amplitude modulation with a suppressed color carrier
- Works with a motion frequency of approximately 30 Hz
- A picture consists of 525 lines

- 4.2 MHz is used for the luminance and 1.5 MHz is used for each of the two chrominance channels.

NTSC **National Television System Committee**	
Lines/Field	525/60
Horizontal Frequency	15.734 kHz
Vertical Frequency	60 Hz
Color Subcarrier Frequency	3.579545 MHz
Video Bandwidth	4.2 MHz
Sound Carrier	4.5 MHz

2. SECAM

- Sequential Couleur Avec Memoire
- Used in France and Eastern Europe
- Unlike NTSC and PAL, it is based on frequency modulation
- Uses a motion frequency of 25 Hz
- Each picture has 625 lines

SECAM **Sequential Couleur Avec Memoire** ***or* Sequential Color with Memory**		
SYSTEM	**SECAM B,G,H**	**SECAM D,K,K1,L**
Line/Field	625/50	625/50
Horizontal Frequency	15.625 kHz	15.625 kHz
Vertical Frequency	50 Hz	50 Hz
Video Bandwidth	5.0 MHz	6.0 MHz
Sound Carrier	5.5 MHz	6.5 MHz

3. PAL

- Phase Alternating Line
- Invented in 1963 by W. Bruch
- Used in parts of Western Europe
- Uses a quadrature amplitude modulation similar to NTSC, but the color carrier is not Suppressed

PAL Phase Alternating Line			
SYSTEM	**PAL**	**PAL N**	**PAL M**
Line/Field	625/50	625/50	525/60
Horizontal Freq.	15.625 kHz	15.625 kHz	15.750 kHz
Vertical Freq.	50 Hz	50 Hz	60 Hz
Color Sub Carrier	4.433618 MHz	3.582056 MHz	3.575611 MHz
Video Bandwidth	5.0 MHz	4.2 MHz	4.2 MHz
Sound Carrier	5.5 MHz	4.5 MHz	4.5 MHz

2.2.2 Video Transmission

Q15. Write about different sub-standards and data reduction methods of transmission.

Ans : **(Imp.)**

For multimedia transmission, the data rates created by motion video are important. The data rate for U.S. HDTV is 1.0368 × 109 bits/second in general. The data rate for European HDTV is 1.152 × 109 bits/second in general.

In digital television, different substandards and data reductions can be applied, and hence the data rate can be calculated as follows.

Substandard 1

- Works with a luminance sampling frequency of 11.25 MHz and with a chrominance sampling frequency of 5.625 MHz
- Data rate is 190 × 106 bits/second.

Substandard 2

- Works with a luminance sampling frequency of 10.125 MHz and with a chrominance sampling frequency of 3.375 MHz
- Data rate is 125 × 106 bits/second.

Substandard 3

- Works with a luminance sampling frequency of 9.0 MHz and with a chrominance sampling frequency of 2.25 MHz
- Data rate is 108 × 106 bits/second.

Data Reduction Methods

Further reduction of the data rates can be achieved by following three steps:

Removing the sampling gaps; this means that only visible lines are coded. For example, the luminance consists of 648 samples out of which only 540 are visible. The chrominance consists of 216 samples out of which only 180 are visible. Hence, if the picture consists of 575 lines, then the data rate equals 103.5 × 106 bits/second, which is calculated as shown below.

(540 + 180 + 180) sample values per line × 575 visible lies/picture = 517,500 sample values /picture

517,500 sample values/picture × 8 bits/ sample value × 25 pictures/second = 103.5 × 106 bits/second.

Reduction of vertical chrominance resolution; Not all but alternating lines of both the chrominance components are digitized. Hence, if the picture consists of 575 lines, then the data rate equals 82.8 × 106 bits/second, which is calculated as shown below.

(540 + 90 + 90) sample values per line × 575 visible lies/picture = 414,000 sample values / picture 414,000 sample values/picture × 8 bits/ sample value × 25 pictures/second = 82.8 × 106 bits/second.

Different kinds of source coding can be applied to the components;

For example, with an intra frame working ADPCM, a data rate reduction of 3 bits/sample (instead of 8 bits/sample) for luminance and chrominance can be achieved. Hence, if the picture consists of 575 lines, then the data rate equals 31.050 × 106 bits/second, which is calculated as shown below.

(540 + 90 + 90) sample values per line × 575 visible lies/picture = 414,000 sample values / picture 414,000 sample values/picture × 3 bits/ sample value × 25 pictures/second = 31.050 × 10 6 bits/second.

2.3 ANIMATIONS

2.3.1 Basic Concepts

Q16. What is animation? Explain different categories of animation.

Ans :

Meaning

Animation is the processes of making something look and feel live. Animation can broadly be categorized into two categories:

(i) Motion Dynamics

Induction of life by changing the position of the object Ex: animation of a bouncing ball.

(ii) Update Dynamics

Induction of life by changing the shape, size and/or structure of the object Ex: animation of flower blooming from bud.

Q17. Write about computer based animation.

Ans :

Computer based animation used computer based graphical tools to provide visual effects.

Advantage of computer based animation includes efficiency, time and simplicity. In several applications the video can be generated by a computer program rather than a video camera. This type of video content is normally referred to as computer animation or sometimes, because of the way it is generated, animated graphics.

A range of special programming languages is available for creating computer animation. To animate something means to bring it life. An animation covers all changes that have a visual effect. Visual effect can be of different nature. They might include time varying positions (motion dynamics), shape, color, transparency, structure and texture of an object (update dynamics), and changes is lighting camera position, orientation and focus. Computer based animation is an animation performed by a computer using graphical tools to provide visual effects. Animation covers all changes that have a visual effect. Visual effect can be of different nature.

They might include:

- Time-varying positions (motion dynamics),
- Shape, color, transparency, structure and texture of an object (update dynamics), and
- Changes in lighting, camera position, orientation and focus.

A computer-based animation is an animation performed by a computer using graphical tools to provide visual effects.

Q18. Mention the basic concepts in computer based animations.

Ans : **(Imp.)**

(i) Input Process

Before computer can be used, drawing must be digitized because key frames, in which the entities being animated are at extreme or characteristic positions, must be drawn. This can be done through optical scanning, tracing the drawings with a data tablet or producing the original drawings with a drawing program. Drawings may need to be post-processed (e.g. filtered) to clean up any glitches arising from the input process

(ii) Composition Stage

In this stage, foreground and background figures are combined to generate individual frames for the final animation. This is done by image-composition techniques.

(iii) In between Process

Animation of movement from one position (starting frame) to another (ending frame) needs a composition of frames with intermediate positions (intermediate frames) in between the key frames. This is called in between process.

This process is performed in computer-based animation through interpolation. The easiest interpolation is linear interpolation (sometimes called lerping-Linear intERPolation). Splines are an alternative to lerping for removing the drawbacks of lerping, but they, too, do not entirely solve the inbetweening problem. Inbetweening also involves interpolating the shapes of objects in intermediate frames.

(iv) Changing Colors

For changing colors, computer based animators use CLUT (lut) in a frame buffer and the process of double buffering. The lut animation is generated by manipulating the lut. The simplest method is to cycle the colors in the lut, thus changing the colors of various pieces of the image. Using lut animation is faster than sending an entire new pixmap to the frame buffer for each frame.

2.3.2 Animation Languages

Q19. Explain various types of animation languages.

Ans : **(Imp.)**

Animation languages fall into three categories:

(i) Linear-list Notations

Each event in the animation is described by a starting and ending frame number and an action that is to take place (event). The actions typically take parameters, such as degree of rotation, axis of rotation, etc.

E.g. 42, 53, B, ROTATE "PALM", 1, 30 means "between 42 and 53 rotate object called PALM about axis 1 by 30 degrees, determining the amount of rotation at each frame from table B.

E.g.: Scefo (SCEneFOrmat)

(ii) General-purpose Languages

Another way to describe animation is to embed an animation capability within a general purpose programming language. The values of variables in the language can be used as parameters to the routines, which perform the animation.

E.g.: ASAS.

Values in the language can be used as parameters to the routine, which perform the animation

Ex.: (grasp my-cube); /* cube becomes the current object

(cw 0.05); /* spin clockwise by a small amount

(grasp camera); /* make the camera the current object

(right 5); /*move right by 5 units

(iii) Graphical Languages

Graphical animation languages describe animation in a more visual way than textual animation languages. These languages are used for expressing, editing and comprehending the simultaneous changes taking place in an animation. Rather than writing out description of actions, the animator provides a picture of the action.

E.g.: GENESYSTM, S-Dynamics System.

Graphical languages provide user the ability to visualize the action clearly. Graphical language describes animation in a more visual way. It can be used for expressing, editing and comprehending simultaneous changes.

2.3.3 Animations Control Transmission

Q20. Explain various methods of controlling animation.

Ans : **(Imp.)**

Methods of Controlling Animation

Controlling animation is independent of the language used for describing it. Animation control mechanisms can employ different techniques.

1. Full Explicit Control

It is a simplest type of animation control. Animator provides a description of everything that occurs in the animation, either by specifying simple changes, such as scaling, translation, and rotation, or by providing key frame information and interpolation methods to use between key frames.

This interpolation may be given explicitly or by direct manipulation with a mouse, joystick, data glove or other input device (in an interactive system). E.g. BBOP system

2. Procedural Control

It is based on communication between various objects to determine their properties. It is a significant part of several other control mechanisms. In physically-based systems, the position of one object may influence the motion of another (eg: balls cannot pass through walls). In actorbased systems, the individual actors may pass their positions to other actors to affect the other actors' behaviors.

3. Constraint-based Systems

Many objects move in a manner determined by other objects with which they are in contact, and this compound motion may not be linear at all. Such motion can be modeled by constraints. Specifying an animated sequence using constraints is often much easier to do than using explicit control. E.g. Sutherland's Sketchpad

4. Tracking Live Action

Trajectories of objects in the course of an animation can be generated by tracking live action. Traditional animation uses rotoscoping. A film is made in which people/ animals act out the parts of the characters in the animation. Then animators draw over the film, enhancing the background and replacing the human actors with their animated equivalents.

Another live-action technique is to attach some sort of indicator to key points on a person's body. By tracking the positions of the indicators, one can get location for corresponding key points in an animated model.

Eg: data glove, which measures the position and orientation of the wearer's hand as well as the flexion and hyperextension of each finger point.

5. Kinematics and Dynamics

Kinematics refers to the position and velocity of points. E.g. a kinematic description of a scene might say, "The cube is at origin at time t = 0. It moves with a constant acceleration in the direction (1, 1, 5) thereafter".

Dynamics takes into account physical laws that govern kinematics (Eg: Newton's

laws of motion for large bodies etc.). A particle moves with an acceleration proportional to the forces acting on it, and the proportionality constant is the mass of the particle.

Thus, a dynamic description of a scene might be, "At time t=0 seconds, the cube is at position (0 meters, 100 meters, 0 meters). The cube has a mass of 100 grams. The force of gravity acts on the cube.

Having the knowledge of kinematics of kinematics and dynamics of the object at a point, we candetermine the next point in motion.

6. Display of Animation

Animated objects must be scan-converted into their pixmap in frame buffer. To show a rotating object, we can scan-convert into the pixmap successive views from slightly different locations, one after another. This scan-conversion must be done at least 10 (preferably 15 to 20) times per sec to give a reasonably smooth effect. Hence, a new image must be created in no more than 100 milliseconds.

To remove distracting effect (due to slow scan-conversion), double buffering is used. To display animations with raster systems, animated objects must be scan-converted into their pixmap in the frame buffer. To show a rotating object, we can scan-convert into the pixmap successive views from slightly different locations, one after another. This scan conversion must be done at least 10 times per second to give reasonably smooth effect.

Hence, a new image must be created in no more than 100 ms. Let us assume that the two halves of the pixmap are image0 and image1.

7. Transmission of Animation

One of the following two approaches may be used for the transmission of animation over computer networks.

8. Symbolic Representation

The symbolic representation (e.g. circle) of animation objects (e.g. ball) is transmitted together with the operation commands (e.g. roll the ball) performed on the object, and at the receiver side the animation is displayed. The transmission time is short because the symbolic representation of an animated object is smaller in byte size than its pixmap representation.

However, the display time at the receiver takes longer because the scan-converting operation has to be performed at the receiver side. The transmission rate (bits/sec or bytes/sec) of animated objects depends on:

- The size of the symbolic representation structure, where the animated object is encoded,
- The size of the structure, where the operation command is encoded, and
- The number of animated objects and operation commands sent per second.

9. Pixmap Representation

The pixmap representation of the animated objects is transmitted and displayed on the receiverside. The transmission time is longer in comparison to the previous because of the size of thepixmap representation. However, the display time is shorter because the scan-conversion of the animated objects is avoided at the receiver side. It is performed at the sender side where animation objects and operation commands are generated. The transmission rate of the animation is equal to the size of the pixmap representation of an animated object (graphicalimage) multiplied by the number of graphical images per second.

Short Question and Answers

1. What is image?

Ans :

An image may be defined as two dimensional function f(x, y) where, x y are the spatial co-ordinate and the amplitude of 'f' at any pair of co-ordinates x, y is called the intensity or gray level of image at that point x, y and amplitude values of f are all finite, discrete quantities we all the image is digital image.

2. Digital Image Representation

Ans :

For computer representation, function (e.g. intensity) must be sampled at discrete intervals.

Sampling quantizes the intensity values into discrete intervals.

- Point at which an image is sampled are called picture elements or pixels.
- Resolution specifies the distance between points accuracy.

A digital image is represented by a matrix of numeric values each representing a quantized intensity value.

A digital image is a numeric representation (normal binary) of two dimensional images.

When I is a two-dimensional matrix, then I (r, c) is the intensity value at the position corresponding to row r and column c of the matrix. Intensity value can be represented by bits for black and white images (binary valued images), 8 bits for monochrome imagery to encode color or grayscale levels, 24 bit (RGB).

3. What is digital image processing?

Ans :

- Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image.
- Image processing usually refers to digital image processing, but optional and analog image processing also are possible.
- Computer graphics concern the pictorial synthesis of real or imaginary objects from their computer-based models.

4. What is image synthesis?

Ans :

Image synthesis is an integral part of all computer user interfaces is indispensable for visualizing 2D, 3D and higher dimensional objects. Areas as diverse as education, science, engineering, medicine, advertising and entertainment all rely on graphics.

5. Process of Image.

Ans :

- Image analysis is concerned with techniques for extracting descriptions from images that are necessary for high-level scene analysis methods.
- Image analysis techniques include computation of perceived brightness and color, partial or complete recovery of three-dimensional data in the scene, location of discontinuities corresponding to objects in the scene and characterization of the properties of uniform regions in the image.
- Image processing includes image enhancement, pattern detection and recognition and scene analysis and computer vision.
- Image enhancement deals with improving image quality by eliminating noise or by
- enhancing contrast.
- Pattern detection and recognition deal with detecting and clarifying standard patterns and finding distortions from these patterns.
- Scene analysis and computer vision deal with recognizing and reconstructing 3D models of a scene from several 2D images.

6. Image Recognition

Ans :

Image recognition can be defined as the capability of a computer to determine the objects or features of an image.

7. What is image transmission?

Ans :

Image transmission takes into account transmission of digital images through computer networks. There are several requirements on the networks when images are transmitted:

- The network must accommodate bursty data transport because image transmission is bursty (The burst is caused by the large size of the image).
- Image transmission requires reliable transport.
- Time-dependence is not a dominant characteristic of the image in contrast to audio/video transmission.

Image size depends on the image representation format used for transmission. There are several possibilities.

8. What is image enhancement?

Ans :

Enhancement is the process an image so that the result is more suitable than the original image for a specific application.

Enhancement Approaches:

(i) Spatial Domain

Spatial domain techniques are techniques that operate directly on pixels.

(ii) Frequency Domain

Frequency domain techniques are based on modifying the Fourier transform of an image.

9. Video Digitalization

Ans :

Digitalization is a process of changing the continuously varying signal into discrete components. This basically uses mathematical process like Fourier analysis or a series of step consisting of sampling and quantizing. Sampling is the process that actually digitizes the spatial position of the image while quantizing digitizes its color information. Sampling involves dividing the picture at M*N array of points while quantizing involves dividing the signal into a range of gray level values. Finally the digital motion video is created by digitizing the pictures temporally.

10. Define EDTV and HDTV.

Ans :

Enhanced Definition System

EDTV are conventional systems modified to offer improved vertical and/or horizontal resolution. This stands for Enhanced-Definition TV, and usually it describes a television that can display HDTV signals but doesn't have enough resolution to really do them justice. Most often it applies to plasma TVs and denotes 852x480 pixels.

High Definition Television

HDTV is high-resolution digital television (DTV) combined with Dolby Digital surround sound (AC- 3). HDTV is the highest DTV resolution in the new set of standards. This combination creates a stunning image with stunning sound. HDTV requires new production and transmission equipment at the HDTV stations, as well as new equipment for reception by the consumer. The higher resolution picture is the main selling point for HDTV.

11. Animations

Ans :

Meaning

Animation is the processes of making something look and feel live. Animation can broadly be categorized into two categories:

(i) Motion Dynamics

Induction of life by changing the position of the object Ex: animation of a bouncing ball.

(ii) Update Dynamics

Induction of life by changing the shape, size and/or structure of the object Ex: animation of flower blooming from bud.

Choose the Correct Answers

1. Video is represented as a series of images formally known as _____. [c]
   (a) pics (b) shots
   (c) frames (d) snaps
2. Progressive download is most useful for _____. [a]
   (a) short video clips (b) long video clips
   (c) extremely long and high quality videos (d) none of the mentioned
3. Real time streaming is most useful for _____. [b]
   (a) short video clips (b) long video clips
   (c) extremely short and low quality videos (d) none of the mentioned
4. The faster the frames are displayed, _____. [b]
   (a) the rougher the video appears (b) the smoother the video appears
   (c) it gets blurry (d) none of the mentioned
5. The characteristic of the eye to retain the image for a short time after it has been presented is known as _____. [a]
   (a) persistence of vision (b) learning power
   (c) memory mapped input (d) none of the mentioned
6. In which type of streaming multimedia file is delivered to the client, but not shared? [a]
   (a) real-time streaming (b) progressive download
   (c) compression (d) none of the mentioned
7. Which of the method is correct for the Moving Picture Experts Group (MPEG) is used to compress _____. [b]
   (a) Audio (b) Video
   (c) Frames (d) Images
8. During Compression Audio and Video, each frame is divided into small grids, called picture elements or _____. [a]
   (a) Pixels (b) Mega Pixels
   (c) Frame (d) Packets
9. We get _____ impression, If frames are displayed on screen fast enough. [c]
   (a) Bits (b) Signals
   (c) Motions (d) Packets
10. Whenever the system receives the signal, a translator is needed to decode the signal and encode it _____ quality. [c]
   (a) Same (b) Bad
   (c) High (d) Lower

Fill in the Blanks

1. A _____ is represented by a matrix of numeric values each representing aquantized intensity value.
2. Full form of JPEG _____.
3. _____ is an integral part of all computer user interfaces for visualizing 2D, 3D and higher dimensional objects.
4. Pictures are dynamically stored in a piece of memory known as _____.
5. The appearance of diagonal lines on the image isknown as _____ effect.
6. _____ deal with recognizing 3D models of a scene from several 2D images.
7. _____ is the process an image so that the result is more suitable than the original image for a specific application.
8. _____ is a process of changing the continuously varying signal into discrete components.
9. The Full form of CGA _____.
10. The _____ representation of the animated objects is transmitted and displayed on the receiver side.

ANSWERS

1. digital image
2. Joint Picture Expert Group
3. Image synthesis
4. refresh buffer
5. aliasing
6. Scene analysis
7. Image Enhancement
8. Digitalization
9. Color Graphics Adapter
10. pixmap

UNIT III

DATA COMPRESSION STANDARDS&STORAGE:

Data Compression Standards: JPEG, H.261, MPEG, DVI

Optical storage devices and Standards: WORHS, CDDA, CDROM, CDWO, CDMO. Real Time Multimedia, Multimedia file System.

3.1 DATA COMPRESSION STANDARDS

Q1. Explain various types of coding techniques in data compression.

Ans : **(Imp.)**

Source Coding

Source coding takes into account the semantics and the characteristics of the data. Thus the degree of compression that can be achieved depends on the data contents. Source coding is a lossy coding process in which there is some loss of information content.

For e.g. in case of speech the speech is transformed from the time domain to frequency domain. In the psychoacoustic the encoder analyses the incoming audio signals to identify perceptually important information by incorporating several psychoacoustic principles of the human ear.

One is the critical-band spectral analysis, which accounts for the ear's poorer discrimination in higher frequency regions than in lower-frequency regions. The encoder performs the psychoacoustic analysis based on either a side-chain FFT analysis or the output of the filter bank.

E.g. Differential Pulse Code Modulation, Delta Modulation, Fast Fourier Transform, Discrete Fourier Transform, Sub-band coding etc

Entropy Coding

Entropy coding is used regardless of the media's specific characteristics. The data stream to be compressed is considered to be a simple digital sequence and the semantics of the data is ignored.

It is concerned solely with how the information is represented.

Run Length Coding

Typical applications of this type of encoding are when the source information comprises long substrings of the same character or binary digit. Instead of transmitting the source string in the form of independent code-words or bits, it is transmitted in the form of a different set of code-words which indicate not only the particular character or bit being transmitted but also indication of the number of characters / bits in the substring.

For e.g. if the string is AAAAABBBTTTTTTT MMMMMMMM, it is encoded as A!5BBBT!6M!8 (In this case there is no point in encoding characters that repeats itself less than 4 times).

Diatomic encoding is a variation of run-length encoding based on a combination of two data bytes. This technique determines the most frequently occurring pairs of bytes.

For e.g. in English language "E", "T", "TH", "A", "S", "RE", "IN" and "HE" occurs most frequently.

Huffman Coding

Huffman coding is an example of variable length coding. It is based in the concept that the probability of occurrence of the characters is not same so different number of bits is assigned for different character. Basically in variable length coding the characters that occur most frequently are assigned fewer numbers of bits.

Arithmetic Coding

Unlike Huffman coding which used a separate code word for each character, arithmetic coding yields a single code word for each encoded string of characters.

The first step is to divide the numeric range from 0 to 1 into a number of different characters present in the message to be sent – including the termination character – and the size of each segment by the probability of the related character.

Hybrid Coding

This type of coding mechanism involve the combine use of both the source coding and the entropy coding for enhancing the compression ratio still preserving the quality of information content. The example of Hybrid Coding includes MPEG, JPEG, H.261, DVI techniques.

Q2. Explain the need of Data Compression.

Ans : **(Imp.)**

- Multimedia object consists of color image, photographic or video image, audio data etc. These data object needs to be stored, retrieved, transmitted and displayed. This large data objects present two problems viz, storage and transmission.
- Transfer of data in the Internet age is very time-dependent. Take for example an audio file, which is nothing but changes in the intensity of sound over a fixed period. This audio is transferred across networks in the form of sound files. If the size of the sound files is too large, the time taken to transfer the files increases. The time taken to transfer a file can be decreased using compression.
- Compression in computer terms means reducing the physical size of data such that it occupies less storage space and memory, Compressed files are, therefore, easier to transfer because there is a sizable amount of reduction in the size of data to be transferred.
- This results in a reduction in the time needed for file transfer as well as a reduction in the bandwidth utilization even for a slow network.
- Thus, the process of reducing the amount of data required to represent information is called compression.

There are two types of compression

Lossless Compression

- Lossless data compression is a class of data compression algorithms that allows the exact original data to be reconstructed from the compressed data. i.e. In lossless data compression, the compressed-then-decompressed data is an exact replication of the original data.
- Lossless data compression is used in many applications. For example, the zip file format produced by the WinZip program on a PC produces exact copies of the original material that was encoded within the zip file. It is also used in the Unix tool gzip.
- Lossless compression is , used in cases where it is important that the original and the decompressed data be identical. Typical examples are executable programs, text documents and source code. Some image file formats, like PNG or GIF, use only lossless compression.

Lossless compression techniques

- Packbits encoding (Run Length Encoding) RLE
- CCITT Group 3 1D
- CCITT Group 3 2D
- CCITT Group 4
- LZW

Lossy Compression

- Lossy data compression, which only allows an approximation of the original data to be reconstructed i.e. Inlossy data compression, the decompressed data may be different from the original data. Typically, there is some distortion between the original and reproduced data.
- Lossy compression is a data encoding method which discards (loses) some of the data, in order to achieve its goal, with the result that

decompressing the data yields content that is different from the original, though similar enough to be useful in some way.

- Lossy compression is most commonly used to compress multimedia data (audio, video). This method is used where absolute data accuracy is not essential. Lossy compression is the most commonly used compression type. This compression technique is used for image documents, audio, and video objects.

Lossy Compression techniques

- Joint Photographic Experts Group (JPEG)
- Motion Picture Experts Group (MPEG)
- Adaptive Differential Pulse Code Modulation (ADPCM)
- CCITT H.261 (Px64) Video Coding Algorithm
- Intel DVI (Digital Video Interactive)

Q3. Explain major steps in data compression.

Ans :

(i) Preparation

Preparation involves analog to digital conversion of the picture where the image is divided into blocks of 4*4 or 8*8 pixels.

(ii) Processing

This involves the conversion of the information from the time domain to the frequency domain by using DCT.

(iii) Quantization

It defines discrete level or values that the information is allowed to take. This process involves the reduction of precision. The quantization process may be uniform or it may be differential depending upon the characteristics of the picture.

Manor Steps of Data Compression

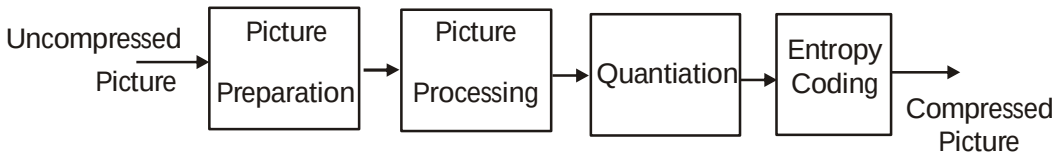


Fig.: Major steps of data compression

(iv) Entropy Encoding

This is the lossless compression method where the semantics of data is ignored but only its characteristics are considered. It may be run length coding or entropy coding.

After compression, the compressed video stream contains the specification of the image starting point and an identification of the compression technique may be the part of the data stream. The error correction code may also be added to the stream. Decompression is the inverse process of compression.

3.1.1 JPEG

Q4. Explain about JPEG standards and requirements.

Ans :

JPEG (Joint Photographic Experts Group)

The JPEG standard for compressing continuous –tone still pictures (e.g. photographs) was developed by photographic experts working under the joint auspices of ITU, ISO and IEC. JPEG is significant in compression because MPEG or the standard for motion picture compression is just the JPEG encoding applied to each frame separately.

There are some requirements of JPEG standard and they are:

- The JPEG implementation should be independent of image size.
- The JPEG implementation should be applicable to any image and pixel aspect ratio.
- Color representation itself should be independent of the special implementation.
- Image content may be of any complexity, with any statistical characteristics.
- The JPEG standard specification should be state of art (or near) regarding the compression factor and achieved image quality.
- Processing complexity must permit a software solution to run on as many available standard processors as possible. Additionally, the use of specialization hardware should substantially enhance image quality.

Q5. Explain the steps in JPEG compression.

Ans :

Steps in JPEG Compression

Step 1: (Block Preparation)

This step involves the block preparation. For e.g. let us assume the input to be 640*480 RGB image with 24 bits/pixel. The luminance and chrominance component of the image is calculated using the YIQ model for NTSC system.

Y = 0.30R + 0.59G + 0.11B

I = 0.60R – 0.28G – 0.32B

Q = 0.21R – 0.52G + 0.31B

For PAL system YUV model is used. Separate matrices are constructed for Y, I and Q each elements inthe range of 0 and 255. The square blocks of four pixels are averaged in the I and Q matrices to reducethem to 320*240. Thus the data is compressed by a factor of two. Now, 128 is subtracted from eachelement of all three matrices to put 0 in the middle of the range. Each image is divided up into 8*8blocks. The Y matrix has 4800 blocks; the other two have 1200 blocks.

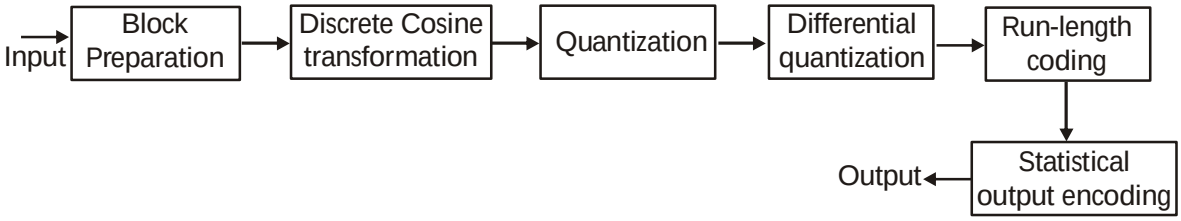


Fig.: The operation of JPEG in lossy-sequential mode

Step 2: (Discrete Cosine Transformation)

Discrete Cosine Transformation is applied to each 7200 blocks separately. The output of each DCT is an8*8 matrix of DCT coefficients. DCT element (0,0)is the average value of the block. The other element tells how much spectral power is present at each spatial frequency.

Step 3: (Quantization)

In this step the less important DCT coefficients are wiped out. This transformation is done by dividing each of the coefficients in the 8*8 DCT matrix by a weight taken from a table. If all the weights are 1 the transformation does nothing however, if the weights increase sharply from the origin, higher spatial frequencies are dropped quickly.

Step 4: (Differential Quantization)

This step reduces the (0,0) value of each block by replacing it with the amount it differs from the corresponding element in the previous block. Since these elements are the averages of their respective blocks, they should change slowly, so taking the differential values should reduce most of them to small values. The (0,0) values are referred to as the DC components; the other values are the AC components.

Step 5: (Run length Encoding)

This step linearizes the 64 elements and applies run-length encoding to the list. In order to concentrate zeros together, a zigzag scanning pattern is used. Finally run length coding is used to compress the elements.

Step 6: (Statistical Encoding)

Huffman encodes the numbers for storage or transmission, assigning common numbers shorter codes than uncommon ones.

JPEG produces a 20:1 or even better compression ratio. Decoding a JPEG image requires running the algorithm backward and thus it is roughly symmetric: decoding takes as long as encoding.

Lossy Sequential DCT-based Mode of JPEG

Image Processing

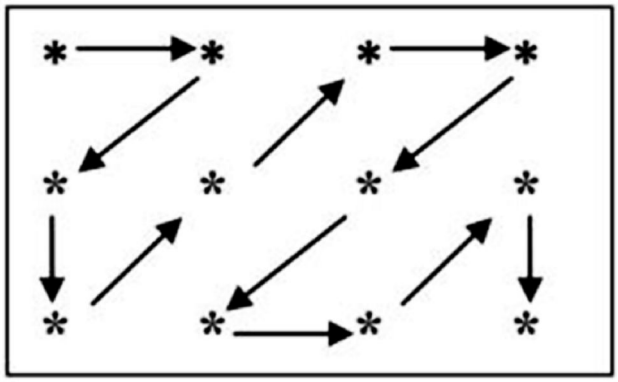
It basically involves the block preparation where the image samples are grouped into 8*8 pixels and passed to the encoder. Then Discrete Cosine Transformation is applied to the blocks where the pixel values are shifted into the range [-128,127] with zero as the center. Each of these values is then transformed using Forward DCT (FDCT). DCT is similar to Discrete Fourier Transformation as it maps the values from the time to the frequency domain.

Quantization

The JPEG application provides a table of 64 entries. Each entry will be used for the quantization of one of the 64 DCT-coefficients.

Entropy Encoding

During the initial step of entropy encoding, the quantized DC-coefficients are treated separately from the quantized AC-coefficients.



- The DC-coefficient determines the basic color of the data units.
- The DCT processing order of the AC coefficients involves the zigzag sequence to concentrate the number of zeros.

JPEG specifies Huffman and arithmetic encoding as entropy encoding methods. However, as this is lossy sequential DCT-based mode, only Huffman encoding is allowed. In lossy sequential mode the framework of the whole picture is not formed but parts of it are drawn i.e. sequentially done.

3.1.2 H.261

Q6. Explain the Encoder and Decoder diagrams of H.261.

Ans :

- H.261 is an algorithm that determines how to encode and compress the data electronically.
- It is a video coding standard published by the ITU (International Telecommunication Union) in 1990.
- It is the most widely used international compression technique for encoding videos.
- H.261 encoding technique can encode only video part of an audiovisual service.
- H.261 is a two-way communication over ISDN lines (Video conferencing and Video calling) and supports data rate in multiples of 64 KBPS.
- H.261 defines a video encoder that is intended to be used to compress video data that will be sent over Integrated Services Digital Network (ISDN) lines.
- The H.261 codec is intended primarily for use in video telephony and video conferencing applications.
- H.261 was the first practical digital video coding standard. The H.261 design was a pioneering effort, and all subsequent international video coding standards (MPEG-1, MPEG-2/H.262, H.263, and even H.264 have been based closely on its design). Additionally, the methods used by the H.261 development committee to collaboratively develop the standard have remained the basic operating process for subsequent standardization work in the field.
- The images supplied as input to an H.261 compressor must meet both color space and size (width and height) requirements.

- In terms of color space, the images must be YCbCr images.
- In terms of size, the images must adhere to either the Common Interchange Format (CIF) or the Quarter-CIF (QCIF) format. Table below indicates the widths and heights defined by these formats.

	Width	Height
CIF images	352	288
QCIF images	176	144

The three main elements in an H.261 encoder are as follows:

- Prediction
- Block Transformation
- Quantization and Entropy Coding

Two types of image frames are defined

- Intra-frames (I-frames) and Inter-frames (P-frames)
- I-frames are treated as independent images. Transform coding method similar to JPEG is applied within each I-frame, hence "Intra".

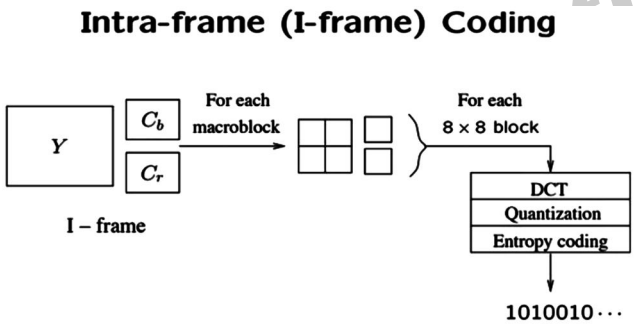


Fig.: I-frame Coding

- Macroblocks are of size 16 × 16 pixels for the Y frame, and 8 × 8 for Cb and Cr frames, since 4:2:0 chroma Subsampling is employed.
- A macroblock consists of four Y, one Cb, and one Cr 8 × 8 blocks.
- For each 8 × 8 block a DCT transform is applied, the DCT coefficients then go through quantization zigzag scan and entropy coding.

Inter-frame (P-frame) Predictive Coding:

- P-frame coding scheme is based on motion compensation. In motion compensation a search area is constructed/predicted in previous frame to determine the best reference macroblock.
- After the prediction, a difference macroblock is derived to measure the prediction error.
- Each of these 8×8 blocks go through DCT, quantization, zigzag scan and entropy coding procedures.
- The P-frame coding encodes the difference macroblock (not the Target macroblock itself).
- Sometimes, a good match cannot be found, i.e., the prediction error exceeds a certain acceptable level.

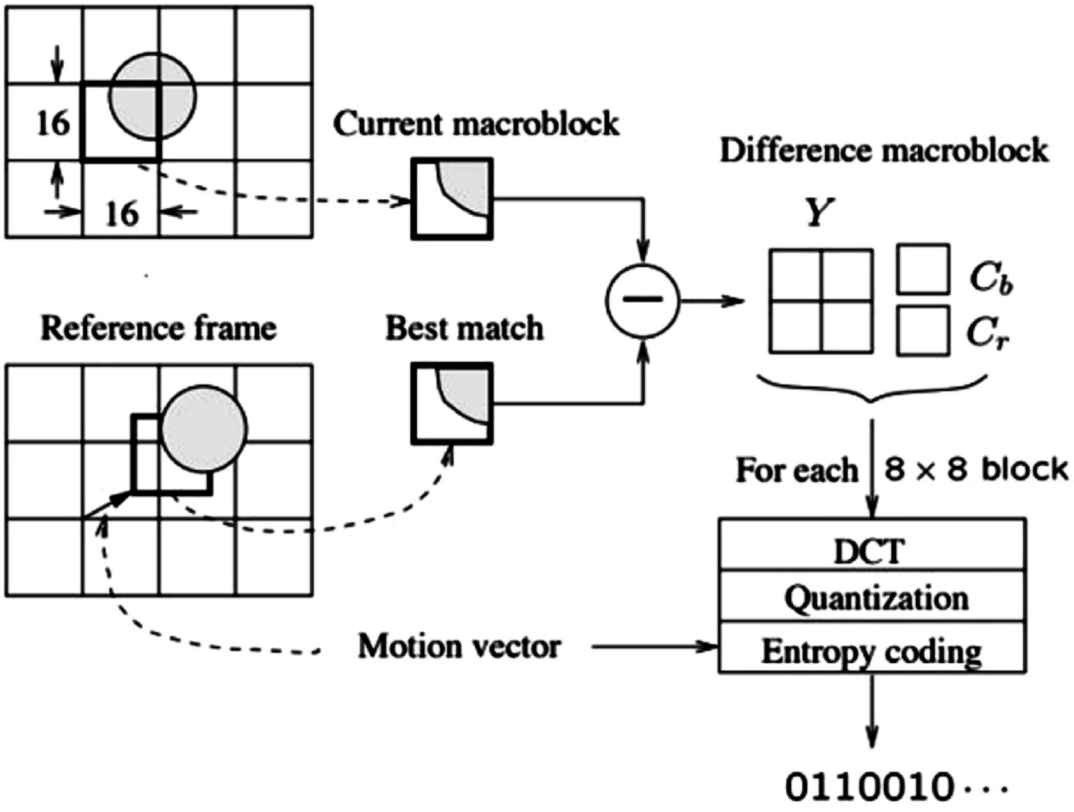
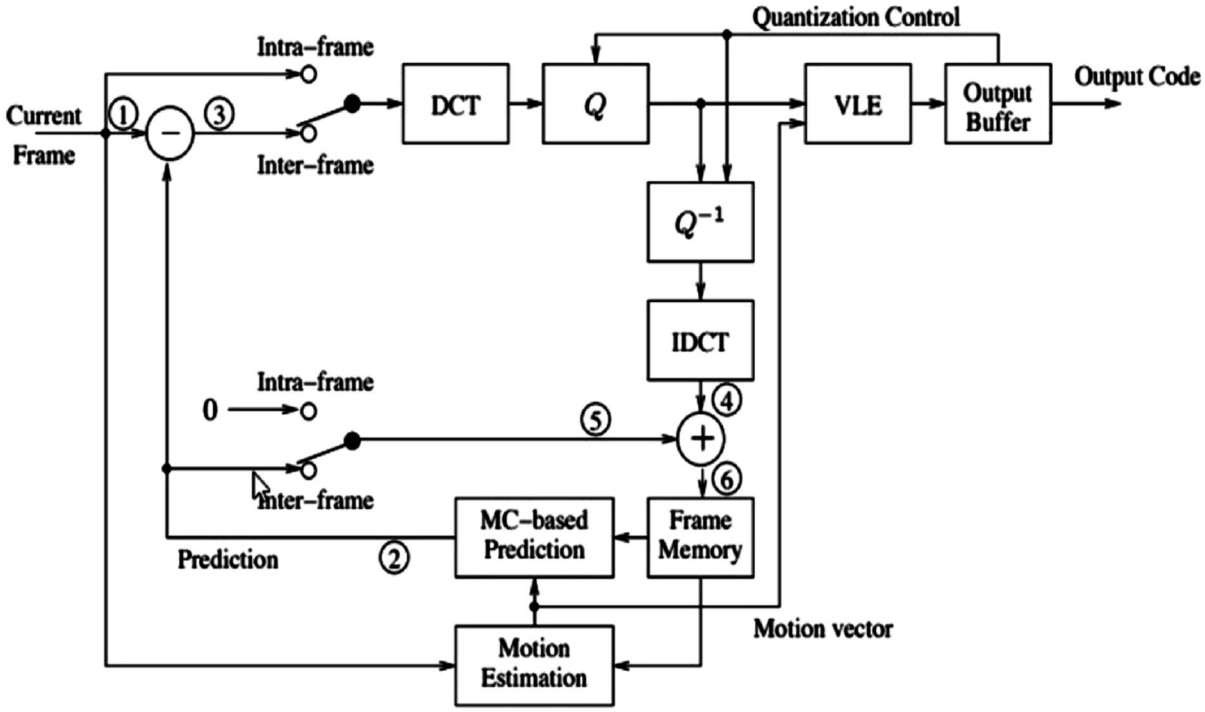
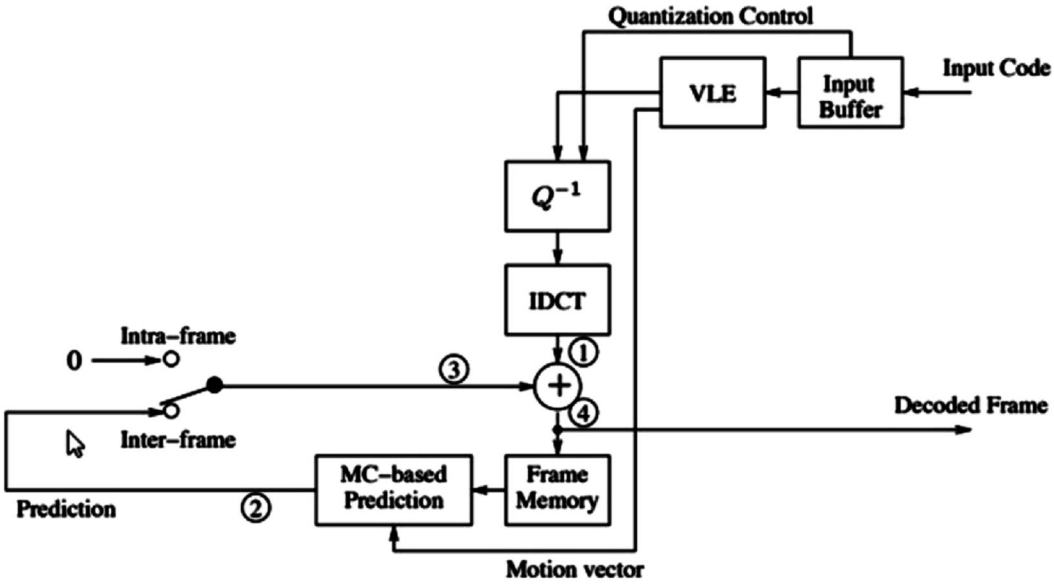


Fig.: H.261 P-frame Coding Based on Motion Compensation

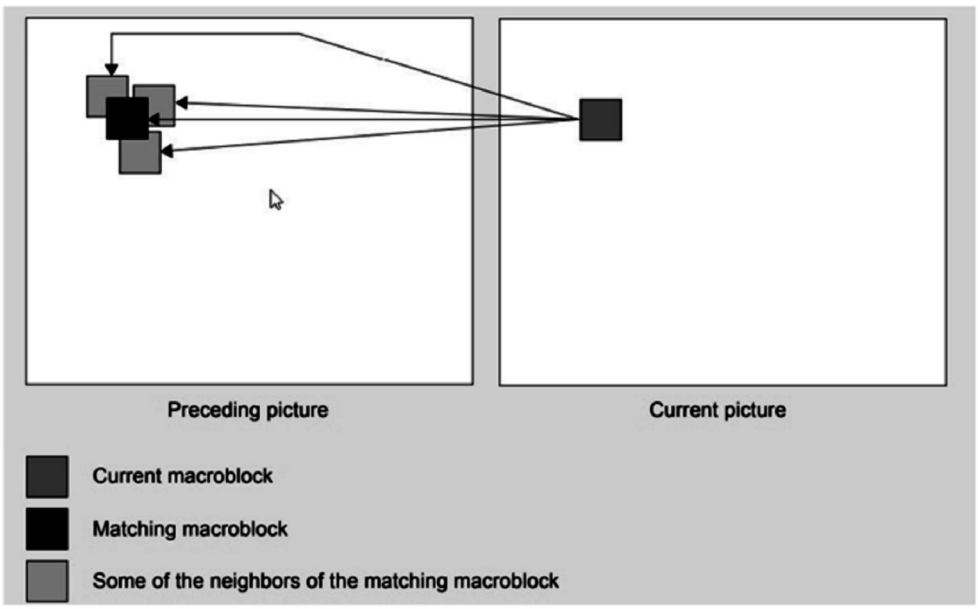


(a) Encoder

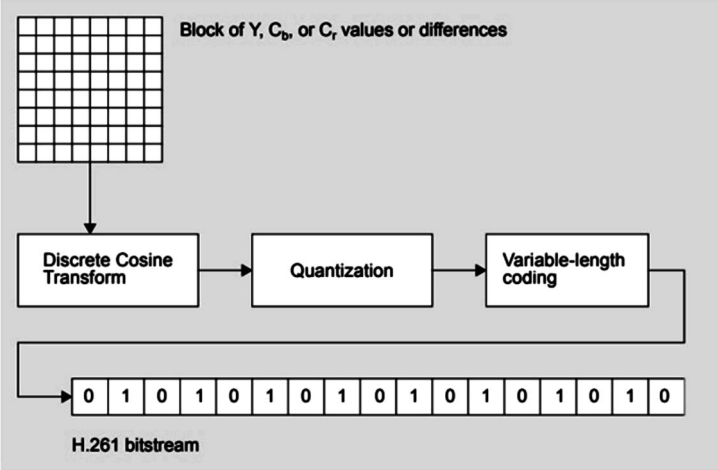


IntraframeVersusInterframe Encoding

- For each macroblock that it encodes, the H.261 encoder can perform Intraframe or Interframe compression. In intraframe mode, the compressor encodes the actual YCbCr values in the macroblock.
- In Interframe mode, the compressor:
- Looks at the YCbCr values in the macroblock it is encoding.
- Calculates the difference between the predicted values for the macroblock and the actual values in the macroblock. The predicted values are taken from the most recently compressed image, which is stored in a history buffer.
- Encodes the difference values if they are significant.



- Encoding YCbCr Values
- This encoding scheme is very similar to the one used in the JPEG still-image compression standard.



3.1.3 MPEG

Q7. What is MPEG?

(OR)

Write short notes on MPEG.

Ans : **(Imp.)**

MPEG (Motion Picture Expert Group)

The Motion Pictures Expert Group was formed by the ISO to formulate a set of standards relating to a range of multimedia applications that involve the use of video with sound. The coders associated with the audio compression part of these standards are known MPEG audio coders and a number of these use perceptual coding. MPEG can deliver a data rate of at most 1856000 bits/second, which should not be exceeded. Data rates for audio are between 32 and 448 Kbits/second; this data rate enables video and audio compression of acceptable quality.

MPEG video uses video compression algorithms called Motion-Compensated Discrete Cosine Transform algorithms.

The algorithms use the following basic algorithm

- Temporal Prediction: It exploits the temporal redundancy between video pictures.
- Frequency Domain Decomposition: It uses DCT to decompose spatial blocks of image data to exploit statistical and perceptual spatial redundancy.
- Quantization: It reduces bit rate while minimizing loss of perceptual quality.
- Variable-length Coding: It exploits the statistical redundancy in the symbol sequence resulting from quantization as well as in various types of side information.

As far as audio compression is concerned the time-varying audio input signal is first sampled and quantized using PCM, the sampling rate and number of bits per sample being determined by the specific application. The bandwidth that is available for transmission is divided into a number of frequency sub b and s using a bank of analysis function which because of their role, are also known ascritical-band filters.

Q8. Explain MPEG video compression standard.

Ans : **(Imp.)**

The name MPEG is an acronym for Moving Pictures Experts Group. MPEG is a method for video compression, which involves the compression of digital images and sound, as well as synchronization of the two.

There currently are several MPEG standards.

- **MPEG-1** is intended for intermediate data rates, on the order of 1.5 Mbit/sec.
- **MPEG-2** is intended for high data rates of at least 10 Mbit/sec.
- **MPEG-3** was intended for HDTV compression but was found to be redundant and was merged with MPEG-2.
- **MPEG-4** is intended for very low data rates of less than 64 Kbit/sec.

i) In principle, a motion picture is a rapid flow of a set of frames, where each frame is an image. In other words, a frame is a spatial combination of pixels, and a video is a temporal combination of frames that are sent one after another.

ii) Compressing video, then, means spatially compressing each frame and temporally compressing a set off names.

iii) **Spatial Compression:** The spatial compression of each frame is done with JPEG (ora modification of it). Each frame is a picture that can be independently compressed.

iv) **Temporal Compression:** In temporal compression, redundant frames are removed.

v) To temporally compress data, the MPEG method first divides frames into three categories:

vi) I-frames, P-frames, and B-frames. Figure 1 shows a sample sequence off names.

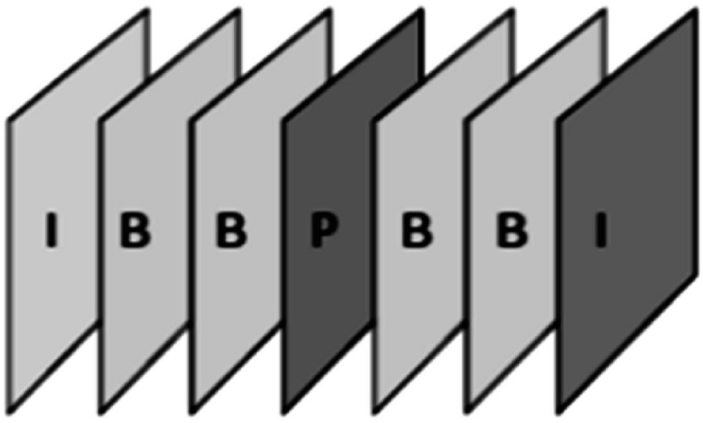


Fig.: MPEG frames

vii) Figure 2 shows how I-, P-, and B-frames are constructed from a series of seven frames.

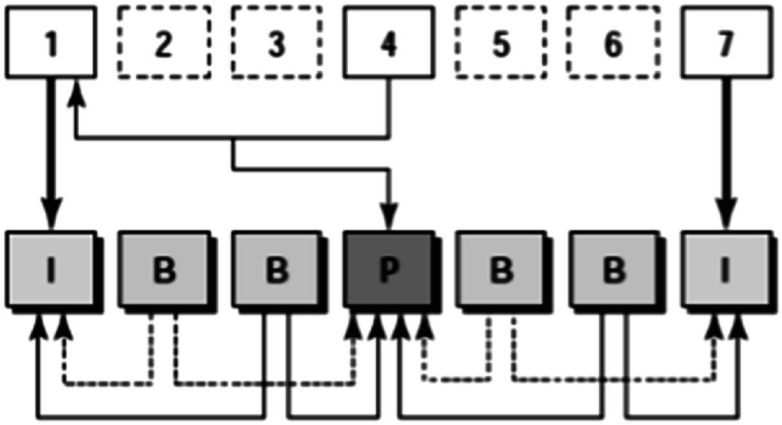


Fig.: MPEG frame construction

- **I-frames**: An intracoded frame (I-frame) is an independent frame that is not related to any other frame.

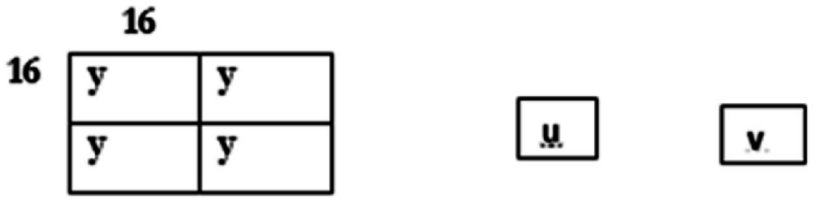
  They are present at regular intervals. An I-frame must appear periodically to handle some sudden change in the frame that the previous and following frames cannot show. Also, when a video is broadcast, a viewer may tune at any time. If there is only one I-frame at the beginning of the broadcast, the viewer who tunes in late will not receive a complete picture. I-frames are independent of other frames and cannot be constructed from other frames.

- **P-frames:** A predicted frame (P-frame) is related to the preceding I-frame or P-frame. In other words, each P-frame contains only the changes from the preceding frame. The changes, however, cannot cover a big segment. For example, for a fast-moving object, the new changes may not be recorded in a P-frame. P-frames can be constructed only from previous I- or P-frames. P-frames carry much less information than other frame types and carry even fewer bits after compression.

- **B-frames:** A bidirectional frame (B-frame) is related to the preceding and following I-frame or P-frame. In other words, each B-frame is relative to the past and the future. Note that a B-frame is never related to another B-frame.

  - According to the MPEG standard the entire movie is considered as a video sequence which consist of pictures each having three components, one luminance component and two chrominance components (y, u & v).
  - The luminance component contains the gray scale picture & the chrominance components provide the color, hue & saturation.
  - Each component is a rectangular array of samples & each row of the array is called the raster line.
  - The eye is more sensitive to spatial variations of luminance but less sensitive to similar variations in chrominance. Hence MPEG – 1 standard samples the chrominance components at half the resolution of luminance components.
  - The input to MPEG encoder is called the resource data and the output of the MPEG decoder is called the reconstructed data.
  - The MPEG decoder has three parts, audio layer, video layer, system layer.
  - The system layer reads and interprets the various headers in the source data and transmits this data to either audio or video layer.

- The basic building block of an MPEG picture is the macro block as shown:



- The macro block consist of 16×16 block of luminance gray scale samples divided into four 8×8 blocks of chrominance samples.
- The MPEG compression of a macro block consist of passing each of the °6 blocks their DCT quantization and entropy encoding similar to JPEG.
- A picture in MPEG is made up of slices where each slice is continuous set of macro blocks having a similar gray scale component.
- The concept of slice is important when a picture contains uniform areas.
- The MPEG standard defines a quantization stage having values (1, 31). Quantization for intra coding is:

$$Q_{DCT} = \frac{(16 \times DCT) + sign(DCT) \times quantizer\ scale)}{2 \times quantizer - scale \times \theta}$$

Where,

DCT = Discrete cosine transform of the coefficienting encoded

Q = Quantization coefficient from quantization table

$$Sign\ (DCT) = \begin{cases} +1 & DCT > 0 \\ 0 & DCT = 0 \\ -1 & DCT < 0 \end{cases}$$

Quantization rule for encoding,

$$Q_{DCT} = \frac{16 \times DCT}{2 \times quantizer - scale \times \theta}$$

- The quantized numbers Q_(DCT)are encoded using non adaptive Haffman method and the standard defines specific Haffman code tables which are calculated by collecting statistics.

3.1.4 DVI

Q9. Explain about DVI technology.

Ans :

DVI is a technology that includes coding algorithms. The fundamental components are a VLSI chip set for the video subsystem, a well specified data format for audio and video files, an application user interface to the audio-visual kernel and compression, as well as decompression, algorithms. Forencoding audio standard signal processor is used. Processing of images and video is performed by a video processor.

Audio and Still Image encoding:

Audio signals are digitized using 16-bits per sample. Audio signals may be PCM-encoded or compressed using the adaptive differential pulse coded modulation (ADPCM) technique. Supported sampling frequencies are: 11025Hz, 22050Hz and 44100 Hz for one or two PCM-coded channels. And 8268Hz,31129Hz, 33075Hz for ADPCM.

For Still Images, DVI assumes an internal digital YUV format for image preparation. Any video input signal must first be transformed into this format. The color of each pixel is split into luminance component and the two chrominance components (U and V). The luminance represents the gray scale image. With RGB, DVI computes the YUV signal using the following relationship.

Y = 0.30R + 0.59G + 0.11B

U = B-Y

V = R-Y

It leads to:

U = -30R-0.59G + 0.89B

V = 0.70R-059G-0.11B

DVI Determines the components YUV according to the following:

Y = 0.299R + 0.587G + 0.144B + 16

U = 0.577B-0.577Y + 137.23

V = 0.730R-0.730Y + 139.67

DVI is able to process image in the 16-bit YUV format and the 24-bit YUV format. The 24-bit YUV format uses 8 bits for each component. The 16-bit YUV format coded the Y components of each pixel with 6bits and the color difference components with 5 bits each.

There are 2 bitmap formats: Planer and Packed.

Planer:

All data of the Y component are stored first, followed by the U component values and then all V values.

Packed:

For the packed bitmap format, the Y, U, and V information of each pixel is stored together by the data of the next pixel.

3.2 Optical Storage Devices and Standards

Q10. What is Optical Storage? Why to use optical storage?

Ans :

Optical storage is also known as "Optical Media" or "Optical Memory" or "Optical Medium", and it allows all read and write activities which are performed by laser beam

In Optical Memory, all recording information is stored at an optical disk. As per the opinions of data scientist that compact space is most useful for huge data storage. Their big advantages are not more costly, light weight, and easy to transport because it is removable device unlike hard drive.

Use Optical Storage

In the optical storage devices, all data is saved like as patterns of dots which can be easily read with using of LIGHT. Laser Beam is used like as "Light Source".

The data is read while bouncing laser beam on the surface of storage medium. Laser beam creates the all Dots while reading process, but it is used with high power mode to mark the surface of storage device, and make a dot. This entire process is also called the "Burning" data onto Disc.

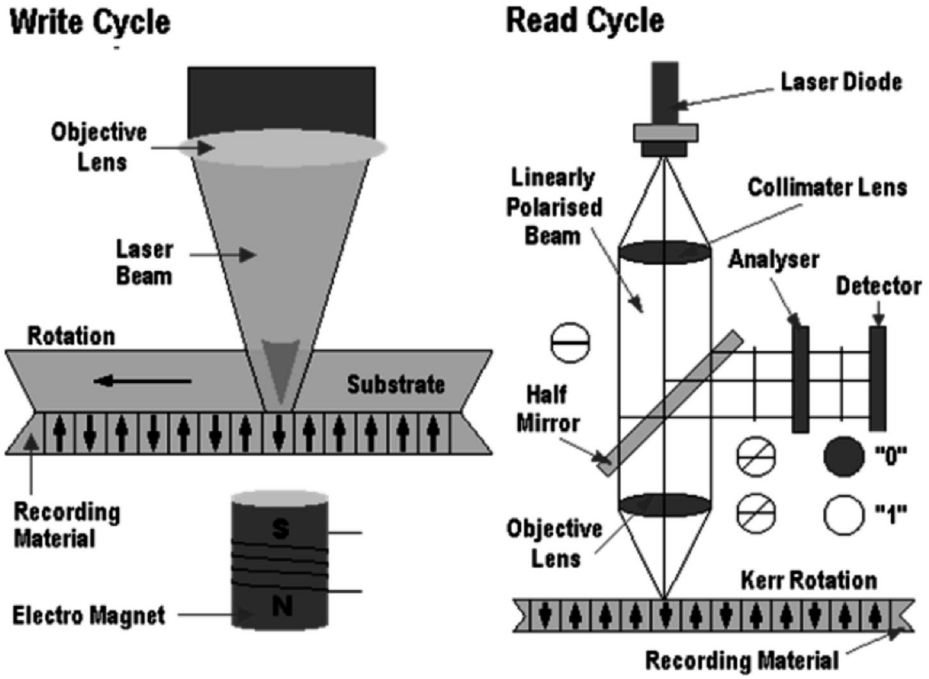
3.2.1 WORMS

Q11. Explain about Write Once, Read Many storage device.

Ans : **(Imp.)**

WORM (Write Once, Read Many) storage had emerged in the late 1980s and was popular with large institutions for the archiving of high volume, sensitive data. When data is written to a WORM drive, physical marks are made on the media surface by a low-powered laser and since these marks are permanent, they cannot be erased.

Rewritable, or erasable, optical disk drives followed, providing the same high capacities as those provided by WORM or CD-ROM devices. However, despite the significant improvements made by recent optical technologies, performance continued to lag that of hard disk devices. On the plus side optical drives offered several advantages. Their storage medium is rugged, easily transportable and immune from head crashes and the kind of data loss caused by adverse environmental factors.



The result is that the relative advantages of the two types of system make them complementary rather than competitive – optical drives offering security, magnetic drives real-time performance. The development of the CD/DVD technologies to include recordable and rewritable formats has had a dramatic impact in the removable storage arena and compatibility is an important and unique advantage of the resulting family of products. Today's market is large enough to accommodate a number of different technologies offering a wide range of storage capacities.

The predominant are:

- magnetic disk
- magneto-optical
- phase-change

At time of writing, the resulting range of capacities can be categorized as follows:

- floppy replacements (100MB to 150MB)
- super-floppies (200MB to 300MB)
- hard disk complement (500MB to 1GB)
- removable hard disks (1GB plus)

3.2.2 CDDA

Q12. Explain about CDDA technology.

Ans :

Compact Disc Digital Audio(CDDA or CD-DA)

The original Compact Disc format developed by Sony and Philips in the 1980s, which was designed for audio only. Since "CD" is used loosely for all Compact Disc formats, the CD-DA designation differentiates a music disc or player from its data counterparts, such as CD-ROM, CD-R and CD-RW.

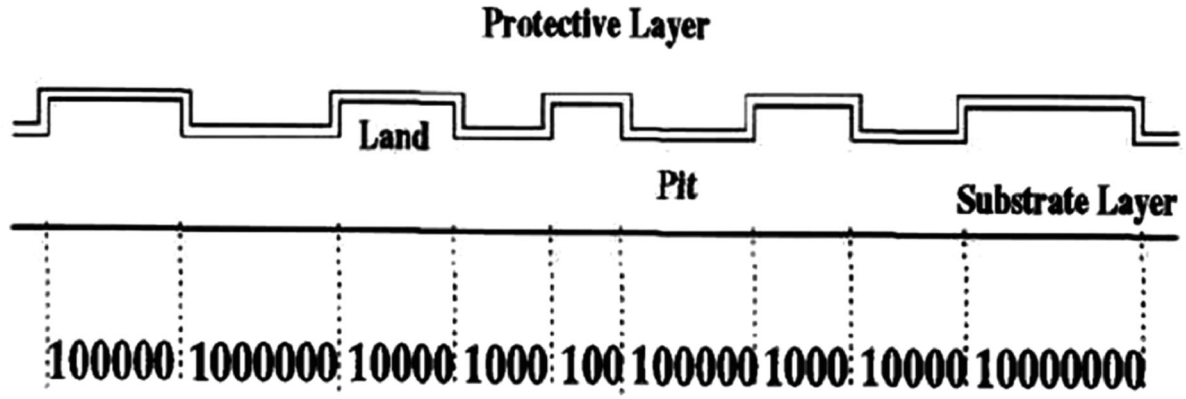
Physical characteristics:

- **Diameter:** 120 mm
- Constant linear velocity (CLV), i.e. number of rotations/s depends on position of head relative to disc center
- **Track shape:** One spiral with appr. 20000 turns (LP: 850 turns)

Audio data rate

- **Sampling frequency:** 44100Hz
- 16 bit quantization
- Pulse code modulation (PCM)
- Audio data rate = 1411200 bit/s = 176,4 Kbyte/s Quality:
- **Signal to noise ratio (S/N):** ~ 6dB/bit, 16 bit quantization ©' S/N exactly 98 dB
- **LP, tape: S/N 50-60 dB Capacity:** (without error correction data)
- **Playback time:** maximal 74 min
- Capacity =74 min "1411200 bit/s = 6265728000 bit ~ 747 MByte

CD-DA: Pits and Lands



Length of pits: multiples of 0.3μm Coding:

- Transition from pit to land / from land to pit: '1'
- Between transitions: sequence of '0's

Audio Data Rate

The audio data rate can be easily derived from the given sample frequency of 44.I kHz and the 16-bit linear quantization. The stereo-audio signals obey the pulse-code modulation rules and the following audio data rate is derived:

$$\text{Audio data rate CD} - \text{DA} = 16\frac{\text{bits}}{\text{sample}} \times 2 \text{ channels} \times 44100\frac{\text{samples}}{\text{s} \times \text{channel}}$$

$$= 1{,}411{,}200\frac{\text{bits}}{\text{s}} = 1{,}411{,}200\frac{\text{bits / s}}{8 \text{ bits / byte}}$$

$$= 176.4\frac{\text{k bytes}}{8} \cong 172.3\frac{\text{k bytes}}{8}$$

Capacity

A CD-DA play time is at least 74 minutes. With this value, the capacity of a CD-DA can be easily determined. The following example shows the computation of a capacity for pure audio data without taking into consideration additional information.

$$\text{Capacity CD} - \text{DA} = 74 \text{ min} \times 1{,}411{,}200\frac{\text{bits}}{\text{s}} = 6{,}265{,}728{,}000 \text{ bits}$$

$$= 6{,}265{,}728{,}000 \text{ bits} \times \frac{1}{8\frac{\text{bits}}{\text{byte}}} \times \frac{1}{\frac{\text{K bytes}}{\text{M byte}}} \cong 747 \text{ M bytes}$$

CD-DA: Eight-to-Fourteen Modulation

Restricted laser resolution:

- Minimal distance between transitions (pit to land, land to pit)

- At least two “0” between two “1” Generation of clock signal:
- Maximal distance between transitions (pit to land, land to pit)
- Not more than 10 consecutive “0” $\Rightarrow$ Eight-to-Fourteen Modulation:
- 8 bit value is encoded using 14 bits
- 267 combinations possible
- 256 are used (criterion: efficient implementation with small number of gates)

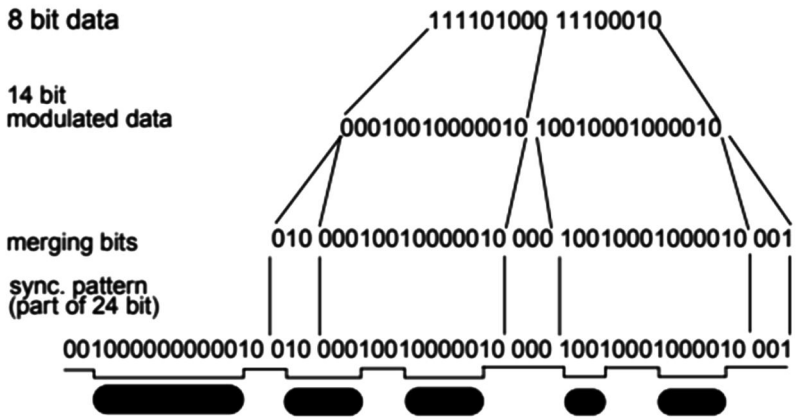
Example for a code conversion table:

data bits	channel bits
00000000	01001000100000
00000001	10000100000000
....	....

Concatenation of independent 14 bit values Í potential violation of:

- Min. distance of 2 bits
- Max. distance of 10 bits ±’ Three additional merging bits

CD-DA: Eight-to-Fourteen Modulation Example



CD-DA: Error Handling

Typical Errors:

- Scratches, dust, fingerprints
- “Burst errors”
- To be detected and corrected Two-level Reed-Solomon Code with frame interleaving:
- **First level:** byte level, EDC and ECC
- **two groups:** each with 4 correction bytes for 24 data bytes
- **first group:** correction of single byte errors
- **second group:** correction of double byte errors, detection of further errors
- **Second level:** Frame interleaving
- **frame:** 588 channel bits = 24 audio data bytes
- distribution of consecutive data bytes and corresponding ECC bytes over adjacent frames

Frames, Tracks, Areas and Blocks of a CD-DA

Frame consists of:

- Data:
- 2 groups of 12 audio data bytes each (actual data)
- Error detection and correction code:
- 2 groups of 4 parity bytes
- According to Reed-Solomon
- Control & display byte:
- Together with c&d bytes of other frames it forms subchannel stream
- E.g., subchannel byte for track start identification
- Synchronization pattern:
- Start of a frame
- 12 x “1” + 12 x “0” + 3 merging bits = 27 bits

CD-DA: Data Streams

Audio bit stream ~ 1.41 x 106 bit/s:

- 44,1 kHz sampling frequency ~ 1411200 bit/s
- 16-bit stereo PCM
- Uniform quantization Data bit stream ~ 1.94 x 106 bit/s:
- Audio bit stream + parity bytes + control & display byte Channel bit stream ~ 4.32 x 106 bit/s:
- Data bit stream + EFM + merging bits + synchronization pattern

CD-DA: Areas

Areas:

- Lead-in area:
- List of contents
- Indication to start of each track
- Program area:
- Up to 99 tracks of different lengths
- Typically one track relates to one song
- Lead-out area Random Access:
- Tracks
- Index points:
- IP0: start of track
- **IP1:** start of audio data
- **Track pregap:** part between IP0 and IP1

3.2.3 CDROM

Q13. Write about CD-ROM?

Ans :

CD-ROM, abbreviation of compact disc read-only memory, type of computer memory in the form of a compact disc that is read by optical means. A CD-ROM drive uses a low-power laser beam to read digitized (binary) data that has been encoded in the form of tiny pits on an optical disk. The drive then feeds the data to a computer for processing.

The standard compact disc was introduced in 1982 for digital audio reproduction. But, because any type of information can be represented digitally, the standard CD was adapted by the computer industry, beginning in the mid-1980s, as a low-cost storage-and-distribution medium for large computer programs, graphics, and databases. With a storage capacity of 680 megabytes, the CD-ROM found rapid commercial acceptance as an alternative to so-called floppy disks (with a maximum capacity of 1.4 megabytes).

Q14. Explain Block Format and Disk Layout on CD-ROM?

Ans :

A typical block format is displayed in Figure (a) It comprises the subsequent fields:

- **Sync:** Sync field identifies beginning of a block. It comprises a byte of all 0s, 10 bytes of all 1s and bytes of all 0s.
- **Header:** The header comprises block address and mode byte. Mode 0 specifies a blank data field and mode 1 specifies use of an error-correcting code and 2048 bytes of data as well as mode 2 specifies 2336 bytes of user data with no error correcting code.
- **Data:** User data.
- **Auxiliary:** Extra user data in mode 2. In mode 1 this is a 288-byte error correcting code.

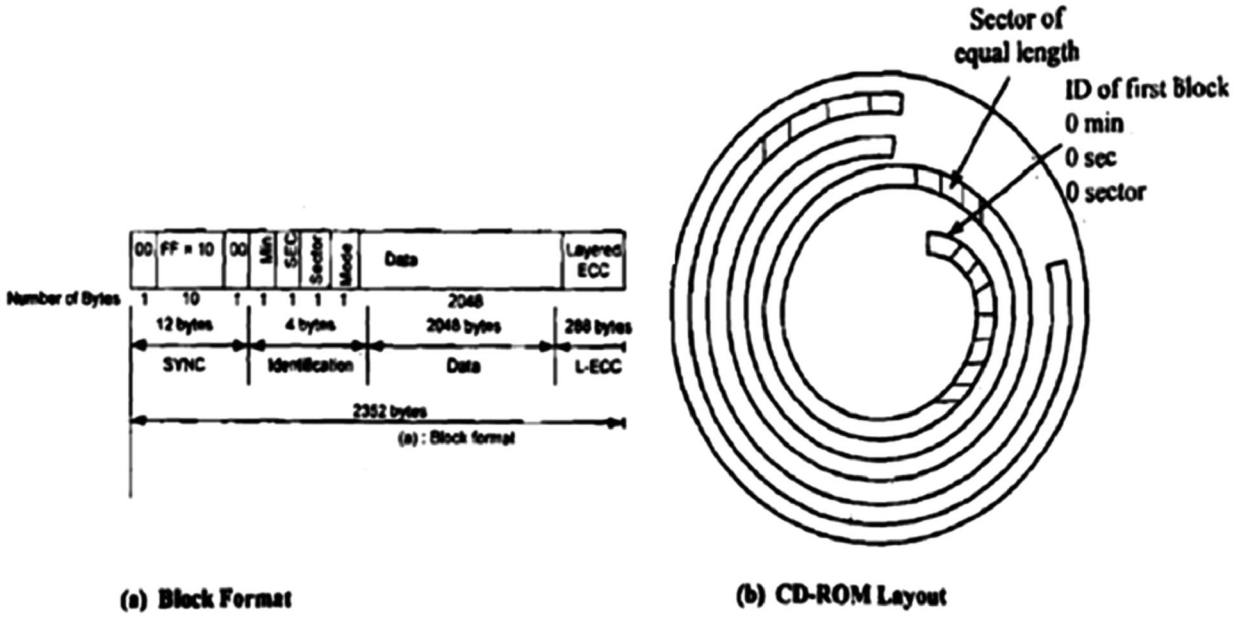


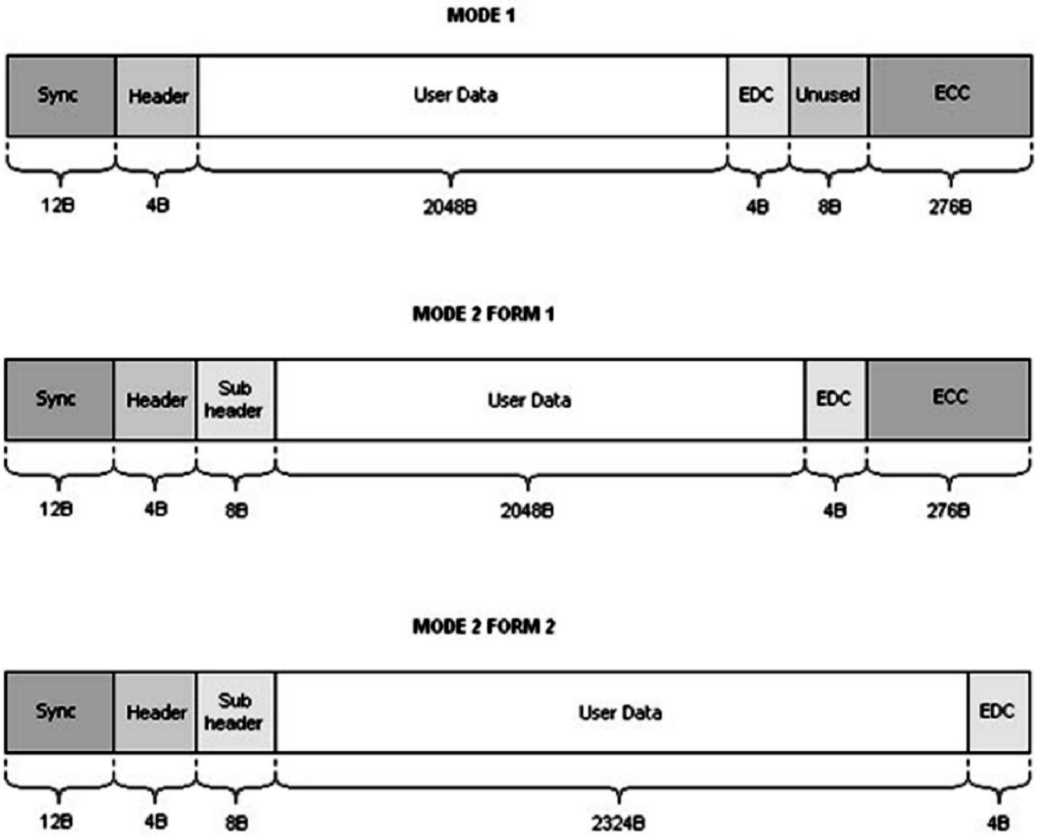
Fig.: Block Format and Disk Layout on CD-ROM

However what is Min (Minute), Sec (Second) and Sector fields in Header field? The sectors of CD-ROM aren't arranged like sectors in hard disks (Please see Figure (b)). Rather they are all equal length segments. If we rotate CD drive at constant speed the linear velocity of disk surface movement will be higher at the outer side than that of centre portions. To offset this liner speed gap either we store less data on outer sectors or we decrease speed of rotation while reading outer tracks. CD follows later approach which is instead of moving CD drive at constant velocity. It's rotated at variable velocity. Speed or rotation of disk decreases as we move away from centre such that sector's can be read in constant time. This method of reading is known as Constant Liner Velocity (CLV).

Modes

The CD-ROM was specified with the following goal: it should serve to hold uncompressed CD-DA data and computer data. This goal is achieved by introducing two modes: mode I and mode 2.

The CD-ROM specification (Yellow Book) defines the use of two types of sectors, mode 1 and mode 2. There are two different types of sectors defined in the CD-ROM specification, Mode 1 and Mode 2 (the latter being used for CD-ROM XA discs).



Mode 1 and Mode2 sector fields description:

- Sync (12 bytes) which is used to enable the player to identify the start of each sector.
- Header (4 bytes) consisting of Minutes, Seconds, Sectors and Mode (= 1).
- Subheader (8 bytes) contains content related parameters eg data type.
- EDC (Error Detection Code - 4 bytes) for detecting errors to be corrected.
- ECC (Error Correction Code - 276 bytes), which comprises an additional level of CIRC error protection.

Mode 2 sectors are used for formats based on CD-ROM XA and can be either Form 1 or 2:

- Mode 2 Form 1 sectors contain 2048 bytes with the same ECC as Mode 1 sectors.
- Mode 2 Form 2 sectors contain 2324 bytes of user data per sector, with no ECC are are suitable only for data where errors can be concealed (eg audio or video data).

Note: That Mode 1 and Mode 2 Form 1 use the same error correction.

The capacity of a CD-ROM depends on whether it is a Mode 1 or Mode 2. Assuming the maximum size is 76 minutes 30 seconds (as recommended) this means that there are 336,300 sectors on a CD-ROM. From this must be subtracted 166 sectors at the start of track 1 plus a few sectors for the file system, amounting to, say, 200 sectors leaving 336,100 sectors for user data.

- **Mode 1** sectors contain 2048 bytes per sector giving a total capacity of 688,332,800 bytes or 656MB (where 1 MB = 1024 * 1024).
- **Mode 2** sectors contain either 2048 or 2324 bytes per sector so will have a somewhat higher data capacity depending on the mix of the two types of sector.

The above assumes a CD-ROM comprising a single track in a single session. For multiple track/ session discs the data capacity will be reduced.

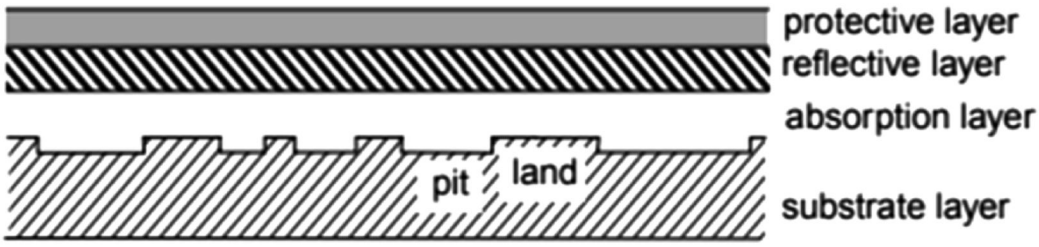
3.2.4 CDWO

Q15. Explain CDWO technology.

Ans : **(Imp.)**

The compact Disk write once (cD-wo), like woRM (write once Read Many), allows the user to write once to a CD and afterwards to read it many times .cd-wo is specified in the second part of the orange Book

Figure shows a cross-section of a CD-WO, vertical to the disk surface and data track. In the case of read-only CDs, the substrate (a polycarbonate) lies directly next to the reflection layer.



Principles:

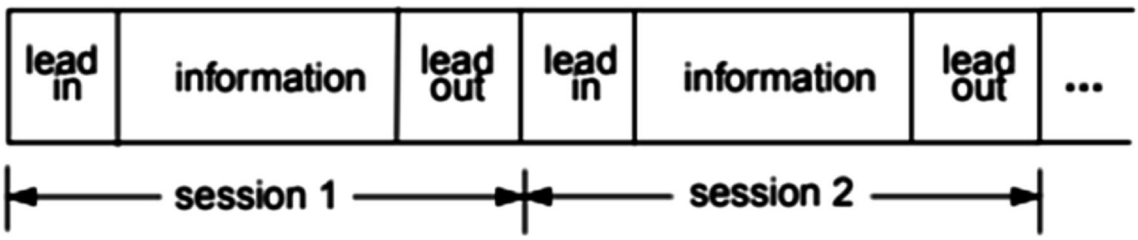
- Pre-grooven track
- Absorption layer between substrate and reflective layer
- Irreversible change of the reflection characteristics by:
- Heating up the absorption layer
- CD-WO can be played in CD-DA players

Sessions

- CD player needs table of content
- Several sessions each with:

- Lead-in part
- Data part
- Lead-out part
- New sessions can be added
- Maximum of 99 sessions

Disc layout with several sessions:



- CD players before 1992 can only read first session
- Regular CD-WO: only one session
- Hybrid CD-WO: several sessions

3.2.5 CDMO

Q16. Explain about CDMO technology.

Ans : **(Imp.)**

The Compact Disk Magneto Optical (CD-MO)has a high storage capacity and allows one to write multiple times to the CD.

Principle of the Magnetic-Optical Method

The magnetic-optical method is based on the polarization of the magnetic field where the polarization is caused by a heat. To be written, the block (sector) is heated to above 1500C. Simultaneously, a magnetic field approximately 10 times the strength of the earth's magnetic field is created.

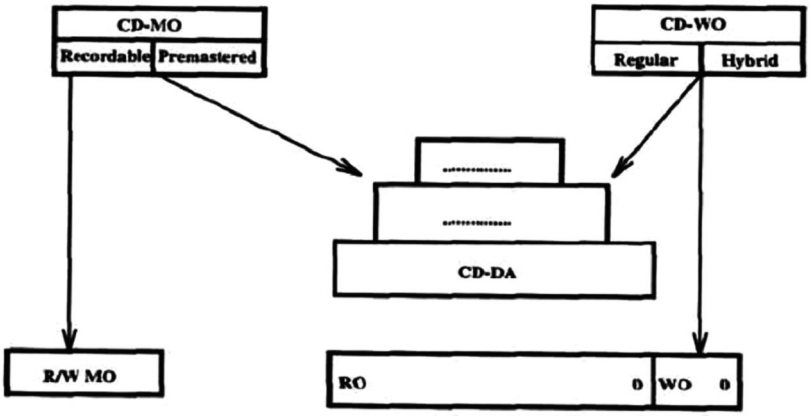
The individual dipoles in the material are then polarized according to this magnetic field. Hereby, a pit corresponds to a low value of the magnetic field. A land is coded through a high value of the magnetic field',

After the CD is irradiated with a laser beam, the polarization of the light changes corresponding to the existing magnetization. Using this process, the read operation is executed.

For a delete activity, a constant magnetic field is created in the area of a block and the sector is simultaneously heated.

Areas of the CD-MO

A CD-MO consists of an option al read,-only area and a write-many (recordable) area. The read-only area includes data which were written in a specified format onto the disk. Figure shows the relationship between the premastered area of a CD-MO and read-only technologies. Therefore, only the CD-MO read-only alea can be read by available playback devices.

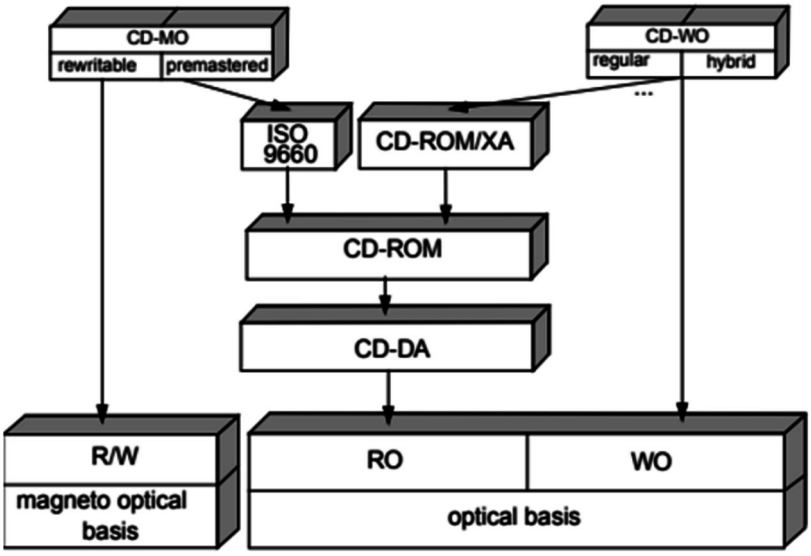


The recordable area of a CD-MO cannot be played because the CD-MO read/write technology is fundamentally different from any CD-DA, CD-ROM, CD-ROM/XA or CD-WO device.

The Prospects of CD Technologies

Compact disk technology will remain the optical storage technology for all kinds of media. The relationship between different standards, shown in Figure allows for a broad field of applications. Except for the CD-MO, the CD-DA with its optical technology still serves as the basis,

A closer view and comparison of the formats show how the specification of CD technology has progressed. CD-ROM mode 1 defines improved error handiing for computer data. In CD-ROM/XA form 1, based on CD-ROM mode 2, the same improved error handling is offered as in CD-RQM mode 1. This may mean that, for example, CD-RoM mode 1 could be omitted if so many applications did not already exist.



- The disadvantage of all CDs is still the relatively high duration of the approximately 200 ms mean access.
- The storage capacity of CDs is sufficient for many current systems. In research labs there are already optical storage media which work with a blue - instead of red - laser.
- Existing data rates require and determine efrcient compression methods for audio, images and especially video

3.2.6 Real Time Multimedia

Q17. What is the use of real time scheduling in multimedia?

Ans :

Audio and video data streams consist of single, periodically changing values of continuous media data, e.g., audio samples or video frames. To fulfill the timing requirements of continuous media, the operating system must use real-time scheduling techniques.

These techniques must be applied to all system resources involved in the continuous media data processing, i.e., the entire end-to-end data path is involved.

The real-time requirements of traditional real-time scheduling techniques have a high demand for security and fault-tolerance.

- The fault-tolerance requirements of multimedia systems are usually less strict than those of real-time systems that have a direct physical impact.
- For many multimedia system applications, missing a deadline is not a severe failure, although it should be avoided.
- A sequence of digital continuous media data is the result of periodically sampling a sound or image signal.
- The bandwidth demand of continuous media is not always that stringent; it must not be a priori fixed, but it may eventually be lowered.

Q18. Explain about resource management in real time multimedia.

Ans :

Resource Management

Current computers do not allow processing of data according to their deadlines without any resource reservation and real-time process management Processing in this context refers to any kind of manipulation and communication of data. This stage of development is known as the window of insufficient resource.

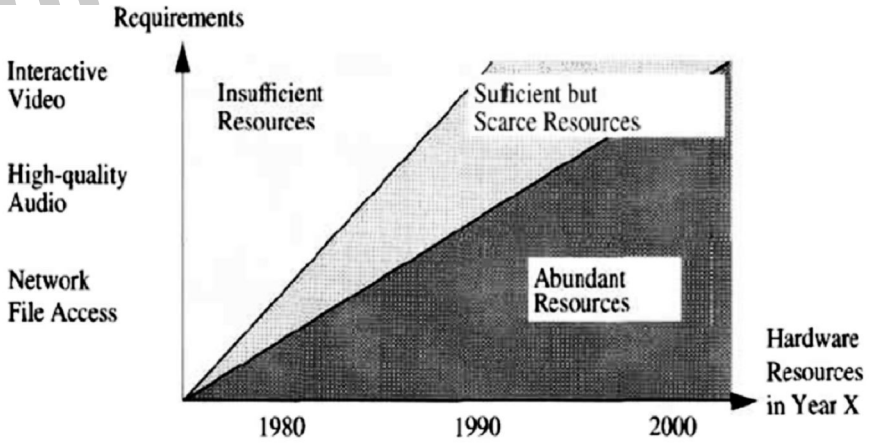


Fig.: Window of insufficient resources

Thus, in an integrated distributed multimedia system, several applications compete for system resources. This shortage of resources requires careful allocation. The system management must employ adequate scheduling algorithms to serve the requirements of the applications. Thereby, the resource is first allocated and then managed.

Resource management in distributed multimedia systems covers several computers and the involved communication networks. It allocates all resources involved in the data transfer process between sources and sinks.

Resources

A resource is a system entity required by tasks for manipulating data. Each resource has a set of distinguishing characteristics classified using the following scheme:

- A resource can be active or passive. An active resource is the CPU or a network adapter for protocol processing; it provides a service. A passive resource is the main memory, communication bandwidth or a file system .it denotes some system capability required by active resources.
- A resource can be either used exclusively by one process at a time or shared between various processes. Active resources are often exclusive passive resources can usually be shared among processes.
- A resource that exists only once in the system is known as a single, otherwise
- it is a multiple resource. In a transputer-based multiprocessor system, the individual CPU is a multiple resource.

Each resource has a capacity which results from the ability of a certain task to perform using the resource in a given time-span.

Requirements

The requirements of multimedia applications and data streams must be served by the single components of a multimedia system. The resource management maps these requirements onto the respective capacity. The transmission and processing requirements of local and distributed multimedia applications can be specified according to the following characteristics:

1. The throughput is determined by the needed data rate of a connection to satisfy the application requirements. It also depends on the size of the data units.
2. We distinguish between local and global (end-to-end) delay:
   (a) The delay "at the resource" is the maximum time span for the completion of a certain task at this resource.
   (b) The end-to-end delay is the total delay for a data unit to be transmitted from the source to its destination. For example, the source of a video telephone is the camera, the destination is the video window on the screen of the partner.
3. The jitter (or delay jitter) determines the maximum allowed variance in the arrival of data at the destination.
4. The reliability defines error detection and correction mechanisms used for the transmission and processing of multimedia tasks. Errors can be ignored, indicated and/or corrected.

In accordance with communication systems, these requirements are also known as Quality of Service parameters (QoS).

Components and Phases

One possible realization of resource allocation and management is based on the interaction between clients and their respective resource managers. The client selects the resource and requests a resource allocation by specifying its requirements through a QoS specification.

This is equivalent to a workload request. First, the resource manager checks its own resource utilization and decides if the reservation request can be served or not.

All existing reservations are stored. This way, their share in terms of the respective resource capacity is guaranteed.

A more elaborate method is to optimize single parameters. In this case, two parameters are determined by the application , and the resource manager calculates the best achievable value for the third parameter (e.g., delay).

In the case shown in Figure 9.2, two computers are connected over a LAN. The transmission of video data between a camera connected to a computer server and the screen of the computer user involves, for all depicted components, a resource manager.

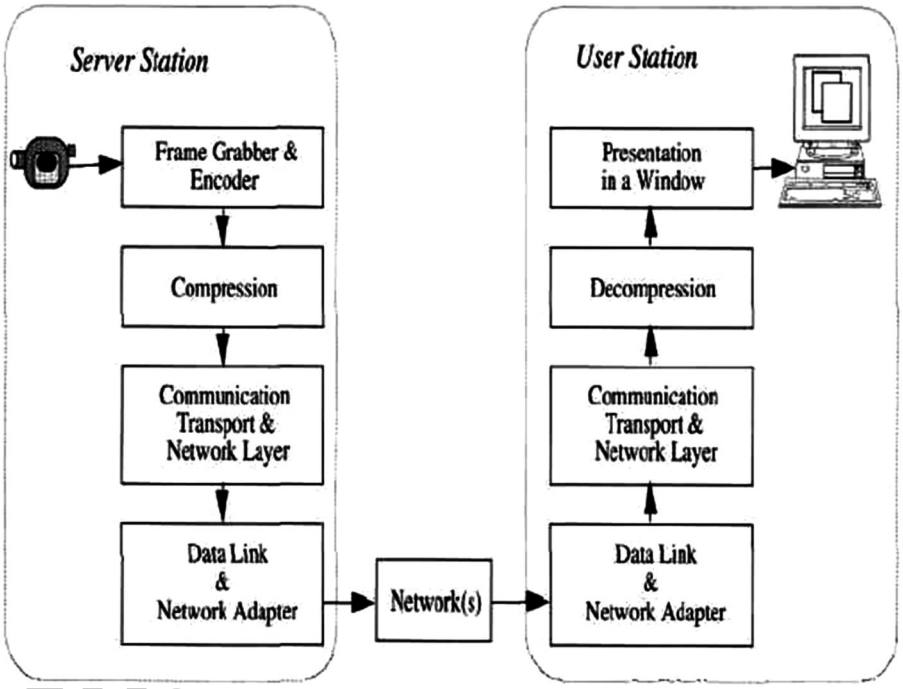


Fig.: Components grouped for the purpose of video data transmission

Phases of the Resource Reservation and Management Process

A resource manager provides components for the different phases of the allocation and management process:

1. Schedulability Test

The resource manager checks with the given QoS parameters (e.g., throughput and reliability) to determine if there is enough remaining tesource capacity available to handle this additional request.

2. Quality of Service Calculation

After the schedulability test, the resource manager calculates the best possible performance (e.g., delay) the resource can guarantee for the new request.

3. Resource Reservation

The resource manager allocates the required capacity to meet the QoS guarantees for each request.

4. Resource Scheduling

Incoming messages from connections are scheduled according to the given QoS guarantees. For process management, for instance, the allocation of the resource is done by the scheduler at the moment the data arrive for processing.

With respect to the last phase, for each resource a scheduling algorithm is defined. The schedulability test, QoS calculation and resource reservation depend on this algorithm used by the scheduler.

Allocation Scheme

Reservation of resources can be made either in a pessimistic or optimistic way:

- The pessimistic approach avoids resource conflicts by making reservations for the worst case, i.e., resource bandwidth for the longest processing time and the highest rate which might ever be needed by a task is reserved. Resource conflicts are therefore avoided. This leads potentially to an underutilization of resources. In a multimedia system, the remaining processor time can be used by discrete media tasks. This method results in a guaranteed QoS.
- With the optimistic approach, resources are reserved according to an average workload only. This means that the CPU is only reserved for the average processing time. This approach may overbook resources with the possibility of unpredictable packet delays. QoS parameters are met as far as possible. Resources are highly utilized, though an overload situation may result in failure.

To detect an overload situation and to handle it accordingly a monitor can be implemented. The monitor may, for instance, pre-empt processes according to their importance

Q19. Explain about continuous media resource model of real time scheduling.

Ans :

Continuous Media Resource Model

- This specifies a model frequently adopted to define QoS parameters and hence, the characteristics of the data stream.
- It is based on the model of Linear Bounded Arrival Processes(I BAP),
- In this model a distributedsystem is decomposed into a chain of resources traversed by the messages on their end-to-end path.
  - **Examples:** CPU, networks.
- The data stream consists of LDUs. In this context, we call them messages
- Various data streams are independent of each other.
- This variance of the data rate results in an accumulation of messages (burst), where the maximal range is defined by the maximum allowed number of messages.

In the LBAP model, a burst of messages consists of messages that arrived ahead of schedule. IBAP is a message arrival plocess at a resource defined by three parameters:

- M = Maximum message size (byte/message).
- R= Maximummessagerate (message/second).
- B= Maximum Burstiness (message).

Q20. Explain about real time scheduling system model.

Ans :

Real-time Scheduling: System Model

All scheduling algorithms to be introduced are based on the following system model for the scheduling of real-time tasks. Their essential components are the resources (as discussed previously), tasks and scheduling goals.

A task is a schedulable entity of the system, and it corresponds to the notion of a thread in the previous description. In a hard real-time system, a task is characterized by its timing constraints, as well as by its resource requirements. In the considered case, only periodic tasks without precedence constraints are discussed, i.e., the processing of two tasks is mutually independent. For multimedia systems, this can be assumed without any major restriction. Synchronized data, for example, can be processed by a single process.

The time constraints of the periodic task 7 are characterized by the following parameters (s, e, d,, p)

o s: Starting point

o e: Processing time of 7

o d: Deadline of 7

p: Period of ?

o r: Rate of T(r = 1/p,)

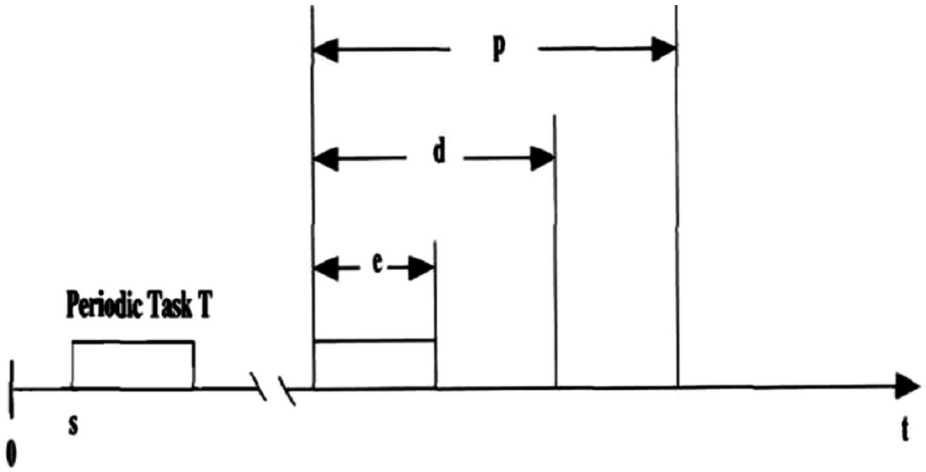


Fig.: Characterization of periodic tasks.

whereby 0 d" e d" d d" p).

- The starting point s is the first time when the periodic task requires processing.
- Afterwards, it requires processing in every period with a processing time of e.
- At s + (k – 1) *p, the task T is ready for k-processing.
- The processing of 7 in period k must be finished at s + (k – 1) *p +d.
- For continuous media tasks, it is assumed that the deadline of the period (k - 1) is the ready time of period k.
- This is known as congestion auoiding deadlines: The deadline for each message (d) coincides with the period of the respective periodic task (p).

Tasks can be pre-emptive or non-pre-emptive.

A pre-emptive task can be interrupted by the request of any task with a higher priority. Processing is continued in the same state later on.

A non-pre-emptive task cannot be interrupted until it voluntarily yields the processor. Any high-priority task must wait until the low-priority task is finished. The high-priority task is then subject to priority inversion.

In the following, all tasks processed on the CPU are considered as preemptive unless otherwise stated.

In a real-time system, the scheduling algorithm must determine a schedule for an exclusive, limited resource that is used by different processes concurrently such that all of them can be processed without violating any deadlines.

This notion can be extended to a model with multiple resources (e.g., CPU) of the same type.

A major performance metric for a real-time scheduling algorithm is the guarantee ratio. The guarantee ratio is the total number of guaranteed tasks versus the number of tasks which could be processed.

Another performance metric is the processor utilization. This is the amount of processing time used by guaranteed tasks versus the total amount of processing time.

$$U = \sum_{i=1}^{n} \frac{e_i}{p_i}$$

Q21. Explain briefly EDF algorithm with an example.

Ans :

Earliest Deadline First Algorithm

Earliest Deadline First (EDF) is an optimal dynamic priority scheduling algorithm mainly used in real-time operating systems. It can be described through the following points:

a) **Priority Driven:** Each process is assigned a priority and the scheduler selects the process to run according to it. Hence, the process with the highest priority is carried out first. In the case of EDF, the priority is set according to the absolute deadline of each process.

b) **Dynamic:** Priorities are calculated and can change while the processes are running; unlike static scheduling in which the priorities are already assigned before the processes are being carried out.

c) **Executes in preemptive mode:** This means that time slots of the CPU can be divided for a process. In other words, it is not required that the same process holds the resource given to it throughout its life cycle. Instead, it can be interrupted by another process if that other process has a higher priority. EDF also runs on non-preemptive uniprocessors, but it does not allow inserted idle time.

How EDF works

Each process has an absolute deadline assigned to it and the scheduler runs the processes based on those deadlines. The process with the closest deadline is assigned the highest priority. The priorities are assigned and changed dynamically.

To ensure that a set of n processes is schedulable using EDF, the following formula is used:

$$U = \sum_{i=1}^{n} \frac{C_i}{T_i} \leq 1 - p,$$

where U is the CPU Utilization, Ci is the execution time, and Ti is the period for each process.

EDF can guarantee that all deadlines are met given that their utilization is less than or equal to 100%

Advantages and Limitations

EDF has some advantages compared to fixed-priority scheduling algorithms. for instance, it has less context switches, less idle times (CPU can be fully utilized), and it can schedule all sets of processes that fixed-priority algorithms can, while EDF schedulable process are not all schedulable with fixed-priority algorithms. However, there are a few limitations. For example, it is less predictable because of the variable priorities of the tasks, which might affect the system when it is overloaded. it is also less controllable in terms of priorities and execution. In addition, EDF has high overhead and is difficult to implement in hardware.

Example 1:

Consider the 3 following processes:

Execution Time (C)	Deadline (D)	Period (T)
3	2	2
7	4	8
20	5	10

1) First, we will check whether the processes are schedulable by calculating the utilization:

   U = C1/P1 + C2/P2 + C3/P3 = (3/20) + (2/5) + (2/10) = 0.75 = 75%

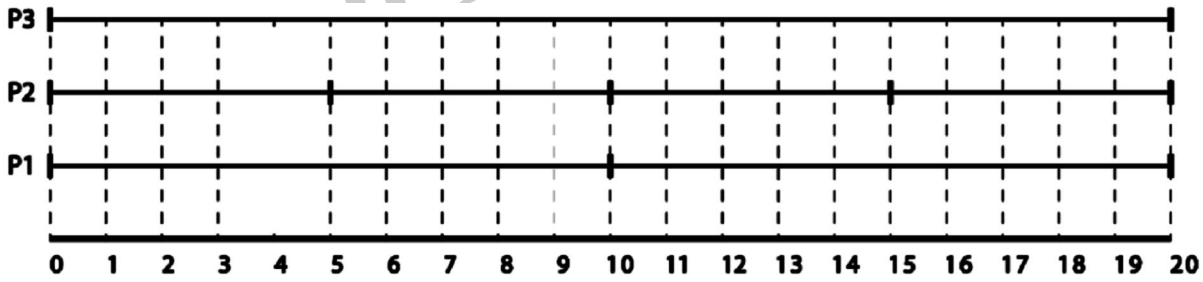
   U = 75 < 100%, which indicates that the processes are schedulable.

2) We take the Least common multiple of the periods of the processes to know how many units we need to execute the three processes:

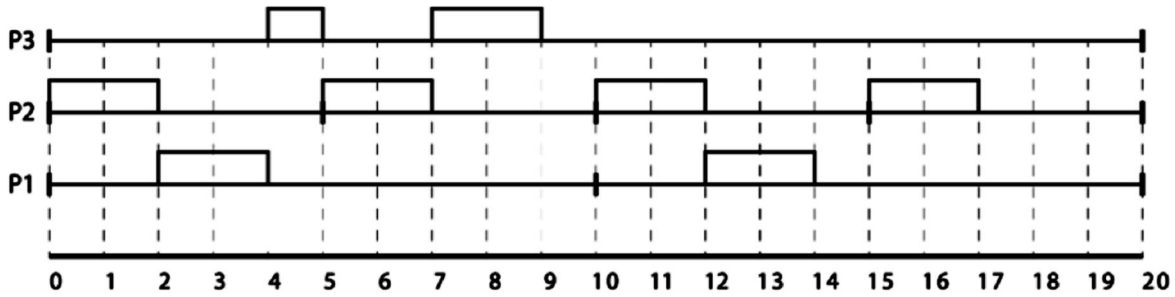
   Lcm(20, 5, 10) = 20

   We need 20 units at most to execute the processes.

3) because the period of P1 = 10, we give it 20/10 = 2 time slices, each is 10 units long. Similarly, with P2 we get 20/5 = 4 time slices, each with a length of 5 units, and P3 will be given 20/20 = 1 time slice with a length of 20 units. The dividers are marked in green and shown in the figure below:



4) the process with the nearest deadline is given the priority to run. When its Execution time is completed for that time slice, the next process with the nearest deadline will run. In this case, the deadlines are equal to the periods. this is illustrated below:



Q22. Explain Rate Monotonic Algorithm with example.

Ans : **(Imp.)**

Rate Monotonic Algorithm

ate monotonic scheduling Algorithm works on the principle of preemption. Preemption occurs on a given processor when higher priority task blocked lower priority task from execution. This blocking occurs due to priority level of different tasks in a given task set. rate monotonic is a preemptive algorithm which means if a task with shorter period comes during execution it will gain a higher priority and can block or preemptive currently running tasks. In RM priorities are assigned according to time period. Priority of a task is inversely proportional to its timer period. Task with lowest time period has highest priority and the task with highest period will have lowest priority. A given task is scheduled under rate monotonic scheduling Algorithm, if its satisfies the following equation:

$$\sum_{k=1}^{n} \frac{C_k}{T_k} \leq \text{URM} = n\left(2^{1/2} - 1\right)$$

where n is the number of tasks in a task set.

Example of RATE MONOTONIC (RM) SCHEDULING ALGORITHM

For example, we have a task set that consists of three tasks as follows

Tasks	Release time(ri)	Execution time(Ci)	Deadline(Di)	Time period(Ti)
T1	0	0.5	3	3
T2	0	1	4	4
T3	0	2	6	6

Table 1. Task set

U= 0.5/3 +1/4 +2/6 = 0.167+ 0.25 + 0.333 = 0.75

As processor utilization is less than 1 or 100% so task set is schedulable and it also satisfies the above equation of rate monotonic scheduling algorithm.

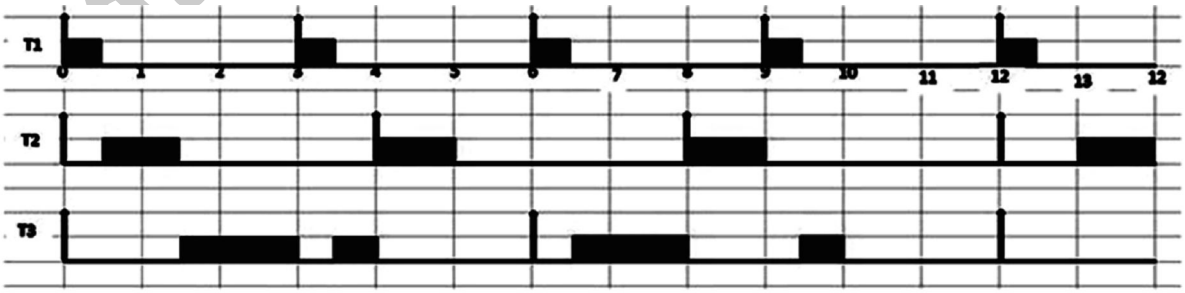


Fig.: 1 RM scheduling of Task set in table 1.

A task set given in table 1 it RM scheduling is given in figure 1. The explanation of above is as follows

1. According to RM scheduling algorithm task with shorter period has higher priority so T1 has high priority, T2 has intermediate priority and T3 has lowest priority. At t=0 all the tasks are released. Now T1 has highest priority so it executes first till t=0.5.
2. At t=0.5 task T2 has higher priority than T3 so it executes first for one-time units till t=1.5. After its completion only one task is remained in the system that is T3, so it starts its execution and executes till t=3.

3. At t=3 T1 releases, as it has higher priority than T3 so it preempts or blocks T3 and starts it execution till t=3.5. After that the remaining part of T3 executes.
4. At t=4 T2 releases and completes it execution as there is no task running in the system at this time.
5. At t=6 both T1 and T3 are released at the same time but T1 has higher priority due to shorter period so it preempts T3 and executes till t=6.5, after that T3 starts running and executes till t=8.
6. At t=8 T2 with higher priority than T3 releases so it preempts T3 and starts its execution.
7. At t=9 T1 is released again and it preempts T3 and executes first and at t=9.5 T3 executes its remaining part. Similarly, the execution goes on.

3.2.7 Multimedia File System

Q23. Why do we need multimedia file systems? Explain.

Ans : **(Imp.)**

Multimedia file systems

Multimedia Applications and Systems are getting more and more involved in our everyday lives. Their main purpose is to deal with various media types like pictures, video data, audio data and text. Video and audio belong to continuous media data. Pictures and text belong to discrete media data.

When most people refer to multimedia, they generally mean the combination of two or more continuous media. In practice, the two media are normally audio and video, that is, sound plus moving pictures.

The challenge on multimedia systems are media types that need to be played continuosly. That means that the data that should be played has to arrive in real time (or at least until a certain strict deadline).

Continuous media data differs from discrete data but not only in its real time characteristics. A challenge for these systems is also the synchronization of pictures and the according sound. Hence these can be two different data streams, it is important to synchronize these before showing them on the monitor.

Another difference to discrete data is the file size. Video and audio need much more storage space than text data and the multimedia file system has to organize this data on disk in a way that efficiently uses the limited storage.

Characteristics of Multimedia Files

- Multimedia files can be quite large.
- Continuous media may require very high data rate.
- Multimedia files/applications are very critical with timing delays during the playback.

Multimedia and File System

For computer systems to deliver continuous media data, it must provide the specific rate and timing requirements associated with such data known as quality of services. For the computing system to be able to deliver such requirements, several components with such system will be affected such as CPU scheduling, network managements, and disk scheduling; and some of such examples are:

- Decoding and compression may require significant CPU processing.
- Processes scheduling algorithm must change for multimedia tasks to meet the deadline requirements of continuous media.
- File systems must be efficient to meet the rate requirements of continuous media.
- Network protocols must support bandwidth requirements to support the streaming technique while minimizing delay in data transformation.

As a result of the characteristics described above, multimedia applications/files require levels of services from the operating system that differ from the traditional applications, such as spreadsheets or word processor. The most important fact of such requirements are the timing, and the rate where the playback of audio and video data demand that the data should be delivered within a certain deadline and at a continuous fixed rate. Such case with the traditional application is not valid since the time and the rate is not considered as constraints for such traditional files.

To provide the quality of service to the multimedia data, operating system often uses an admission control, which is a practice of admitting a request of service only if the system has sufficient resources to satisfy the request. Multimedia data (also known as continuous-media files) have two constraints that are not valid in the conventional data file: timing deadlines and rate requirements. To satisfy both constraints required by the multimedia data, disk scheduling algorithm must be optimized for such constraints. Also, these two constraints are usually in conflict with one another i.e. continuous-media files (multimedia files) require very high-disk bandwidth rates to satisfy such requirements, and disks usually have relatively low transfer rates and high latency rates, and to satisfy the requirements for multimedia data, disk schedulers must reduce the latency time to ensure high bandwidth.

Some of the following features are required in the modern computing systems:

- **High processing power:** to deal with large data in the multimedia files, and deliver the data within the time constraints required to deliver in a real time.
- **Multimedia capable file system:** needed to deliver the data within the time constraints.
- **Data representation/ file formats that support multimedia:** multimedia data should have the format that can be easily handled, and also be able to be compressed-decompressed.
- **Efficient and high I/O:** the file subsystems within the file system implemented within the operating system must be efficient and fast.
- **Special operating system:** to be able to handle the file manipulation with multimedia data operating systems must provide the capabilities to access the file system and process data efficiently and quickly by providing the support for direct transfers to disk, fast interrupt processing, real-time scheduling, and I/O streaming.
- **Storage and memory:** to handle multimedia data, large storage and large memory is required.

Short Question and Answers

1. JPEG

Ans :

The JPEG standard for compressing continuous –tone still pictures (e.g. photographs) was developed by photographic experts working under the joint auspices of ITU, ISO and IEC. JPEG is significant in compression because MPEG or the standard for motion picture compression is just the JPEG encoding applied to each frame separately.

2. H.261.

Ans :

- H.261 is an algorithm that determines how to encode and compress the data electronically.
- It is a video coding standard published by the ITU (International Telecommunication Union) in 1990.
- It is the most widely used international compression technique for encoding videos.
- H.261 encoding technique can encode only video part of an audiovisual service.
- H.261 is a two-way communication over ISDN lines (Video conferencing and Video calling) and supports data rate in multiples of 64 KBPS.
- H.261 defines a video encoder that is intended to be used to compress video data that will be sent over Integrated Services Digital Network (ISDN) lines.
- The H.261 codec is intended primarily for use in video telephony and video conferencing applications.
- H.261 was the first practical digital video coding standard. The H.261 design was a pioneering effort, and all subsequent international video coding standards (MPEG-1, MPEG-2/H.262, H.263, and even H.264 have been based closely on its design). Additionally, the methods used by the H.261 development committee to collaboratively develop the standard have remained the basic operating process for subsequent standardization work in the field.

3. MPEG

Ans :

The Motion Pictures Expert Group was formed by the ISO to formulate a set of standards relating to a range of multimedia applications that involve the use of video with sound. The coders associated with the audio compression part of these standards are known MPEG audio coders and a number of these use perceptual coding. MPEG can deliver a data rate of at most 1856000 bits/second, which should not be exceeded. Data rates for audio are between 32 and 448 Kbits/second; this data rate enables video and audio compression of acceptable quality.

4. DVI

Ans :

DVI is a technology that includes coding algorithms. The fundamental components are a VLSI chip set for the video subsystem, a well specified data format for audio and video files, an application user interface to the audio-visual kernel and compression, as well as decompression, algorithms. Forencoding audio standard signal processor is used. Processing of images and video is performed by a video processor.

5. What is Optical Storage?

Ans :

Optical storage is also known as "Optical Media" or "Optical Memory" or "Optical Medium", and it allows all read and write activities which are performed by laser beam

In Optical Memory, all recording information is stored at an optical disk. As per the opinions of data scientist that compact space is most useful for huge data storage. Their big advantages are not more costly, light weight, and easy to transport because it is removable device unlike hard drive.

6. WORMS

Ans :

WORM (Write Once, Read Many) storage had emerged in the late 1980s and was popular with large institutions for the archiving of high volume, sensitive data. When data is written to a WORM drive, physical marks are made on the media surface by a low-powered laser and since these marks are permanent, they cannot be erased.

Rewritable, or erasable, optical disk drives followed, providing the same high capacities as those provided by WORM or CD-ROM devices. However, despite the significant improvements made by recent optical technologies, performance continued to lag that of hard disk devices. On the plus side optical drives offered several advantages. Their storage medium is rugged, easily transportable and immune from head crashes and the kind of data loss caused by adverse environmental factors.

7. CDROM

Ans :

CD-ROM, abbreviation of compact disc read-only memory, type of computer memory in the form of a compact disc that is read by optical means. A CD-ROM drive uses a low power laser beam to read digitized (binary) data that has been encoded in the form of tiny pits on an optical disk.The drive then feeds the data to a computer for processing.

The standard compact disc was introduced in 1982 for digital audio reproduction. But, because any type of information can be represented digitally, the standard CD was adapted by the computer industry, beginning in the mid-1980s, as a low-cost storage-and-distribution medium for large computer programs, graphics, and databases. With a storage capacity of 680 megabytes, the CD-ROM found rapid commercial acceptance as an alternative to so-called floppy disks (with a maximum capacity of 1.4 megabytes).

8. CDWO technology

Ans :

The compact Disk write once (cD-wo), like woRM (write once Read Many), allows the user to write once to a CD and afterwards to read it many times .cd-wo is specified in the second part of the orange Book

Figure shows a cross-section of a CD-WO, vertical to the disk surface and data track. In the case of read-only CDs, the substrate (a polycarbonate) lies directly next to the reflection layer.

9. CDMO

Ans :

The Compact Disk Magneto Optical (CD-MO)has a high storage capacity and allows one to write multiple times to the CD.

Principle of the Magnetic-Optical Method

The magnetic-optical method is based on the polarization of the magnetic field where the polarization is caused by a heat. To be written, the block (sector) is heated to above 1500C. Simultaneously, a magnetic field approximately 10 times the strength of the earth's magnetic field is created.

The individual dipoles in the material are then polarized according to this magnetic field. Hereby, a pit corresponds to a low value of the magnetic field. A land is coded through a high value of the magnetic field',

After the CD is irradiated with a laser beam, the polarization of the light changes corresponding to the existing magnetization. Using this process, the read operation is executed.

For a delete activity, a constant magnetic field is created in the area of a block and the sector is simultaneously heated.

10. Real time multimedia.

Ans :

Audio and video data streams consist of single, periodically changing values of continuous media data, e.g., audio samples or video frames. To fulfill the timing requirements of continuous media, the operating system must use real-time scheduling techniques.

These techniques must be applied to all system resources involved in the continuous media data processing, i.e., the entire end-to-end data path is involved.

The real-time requirements of traditional real-time scheduling techniques have a high demand for security and fault-tolerance.

- The fault-tolerance requirements of multimedia systems are usually less strict than those of real-time systems that have a direct physical impact.
- For many multimedia system applications, missing a deadline is not a severe failure, although it should be avoided.
- A sequence of digital continuous media data is the result of periodically sampling a sound or image signal.

Choose the Correct Answer

1. Among the following which is an example of variable length coding [c]
   (a) Arithmetic Coding (b) Entropy Coding
   (c) Huffman Coding (d) Hybrid Coding
2. Which of the following is correct in the multimedia system require hard real-time scheduling to [b]
   (a) Deliver the media file to the client
   (b) Ensure critical tasks will be serviced within timing deadlines
   (c) Security
   (d) Minimize the delay
3. Lossy and lossless are classifications of _____. [c]
   (a) Multimedia storage systems (b) files
   (c) compression algorithms (d) All of the mentioned
4. Where this range occurs H.323 uses G.71 or G.723.1… [b]
   (a) Communication (b) Compression
   (c) Conferencing (d) Controlling
5. Lossy techniques provide _____ when compared to lossless techniques. [b]
   (a) Lower compression ratios (b) Much higher compression ratios
   (c) Similar compression ratios (d) None of the mentioned
6. Full form of MPEG is [b]
   (a) Movie Picture Expert Group (b) Motion Picture Expert Group
   (c) Motion Pixel Express Group (d) Motion Picture Express Group
7. Among the following which frame is an independent frame that is not related to any other frame. [a]
   (a) I-frame (b) P-frame
   (c) B-frame (d) T-frame
8. DVI is able to process image in _____ and _____ bit YUV format [c]
   (a) 8 bit and 16 bit (b) 16 bit and 32 bit
   (c) 16 bit and 24 bit (d) 8 bit and 24 bit

9. Full form of WORM _____. [c]

(a) Write Once, Read More (b) Write Only, Read More

(c) Write Once, Read Many (d) Write Only, Read Many

10. In resource reservation and management process in which phase, the resource manager allocates the required capacity to meet the QoS guarantees for each request. [c]

(a) Schedulability Test (b) Resource Scheduling

(c) Resource Reservation (d) Allocation Scheme

Fill in the blanks

1. In _____ coding the data stream to be compressed is considered to be a simple digital sequence and the semantics of the data is ignored.
2. In computer terms _____ means reducing the physical size of data such that it occupies less storage space and memory
3. _____ compression is a data encoding method which discards some of the data, in order to achieve its goal.
4. _____ is the most widely used international compression technique for encoding videos.
5. The _____ coding encodes the difference macroblock.
6. MPEG video uses video compression algorithms called _____.
7. Audio signals are digitized using _____ per sample
8. _____ allows all read and write activities which are performed by laser beam.
9. When data is written to a _____ drive, physical marks are made on the media surface by a low-powered laser and since these marks are permanent, they cannot be erased.
10. The _____ method is based on the polarization of the magnetic field where the polarization is caused by a heat.

ANSWERS

1. Entropy Coding
2. Compression
3. Lossy
4. H.261
5. P-frame
6. Motion-Compensated Discrete Cosine Transform algorithms.
7. 16-bits
8. Optical storage
9. WORM
10. Magnetic-optical

UNIT IV

Multimedia Communication System, Databases & Synchronization

Multimedia Communication System: Collaborative computing session management, transport subsystem, QOS, resource management.
Multimedia Databases: Characteristics, data structures, operation, integration in a database model.
Synchronization: Issues, presentation requirements, reference to multimedia synchronization, MHEG.

4.1 MULTIMEDIA COMMUNICATION SYSTEM

4.1.1 Collaborative Computing

Q1. What is Collaborative Computing? Explain the dimensions of Collaborative Computing.

Ans : **(Imp.)**

Meaning

Collaborative computing is the computer supported cooperative work supported by the networks, PCs and the software that facilitates the cooperation. The examples of collaborative computing tools are electronic mail, bulletin boards, screen sharing tools, text-based conferencing systems, telephone conference systems, conference rooms and video conference systems.

Dimensions

The collaboration dimensions are

1. Time
2. User Scale
3. Control

1. Time

According to time there can be two types of collaborative work and they are a synchronous where the cooperative works do not happen at the same time; while the other is the synchronous where the cooperative works happen at the same time.

2. User Scale

The user of the application can be a single user interacting to the other user or to a group of users. Email between two individuals is a user to user interaction where as Email to a group is auser to group interaction. Video conferencing is also a user to group interaction application.

There can be following types of groups.

- **Dynamic or Static Group**

  When the new users can join the group for cooperative work in the real time the group is said to be dynamic where as when the number of members and their membership is predefined it is static group.

- The members of the group may simply be a participant of the collaborative work or he/she may be the co-coordinator, conference initiator, conference chairman, a token holder or just an observer.

- The members of the group may have homogenous or heter ogeneous characteristics. Fore.g. they may belong to different ethnic group or they may differ in the level of intelligence.

3. Control

Control during the collaboration can be centralized or distributed. Centralized control means there is a chairman who controls the collaborative work and every group member reports to him. Distributed control means every group member has control over his/her own tasks in the collaborative works.

Q2. Explain group communication architecture with the diagram.

(OR)

Discuss in detail about group Communi cation Architecture.

Ans : **(Imp.)**

Group Communication Architecture

Group communication is a cooperative activity which may be synchronous or asynchronous which may be central control or a distributed control. A group communication architecture consists of:

1. Support model
2. System model
3. Interface model

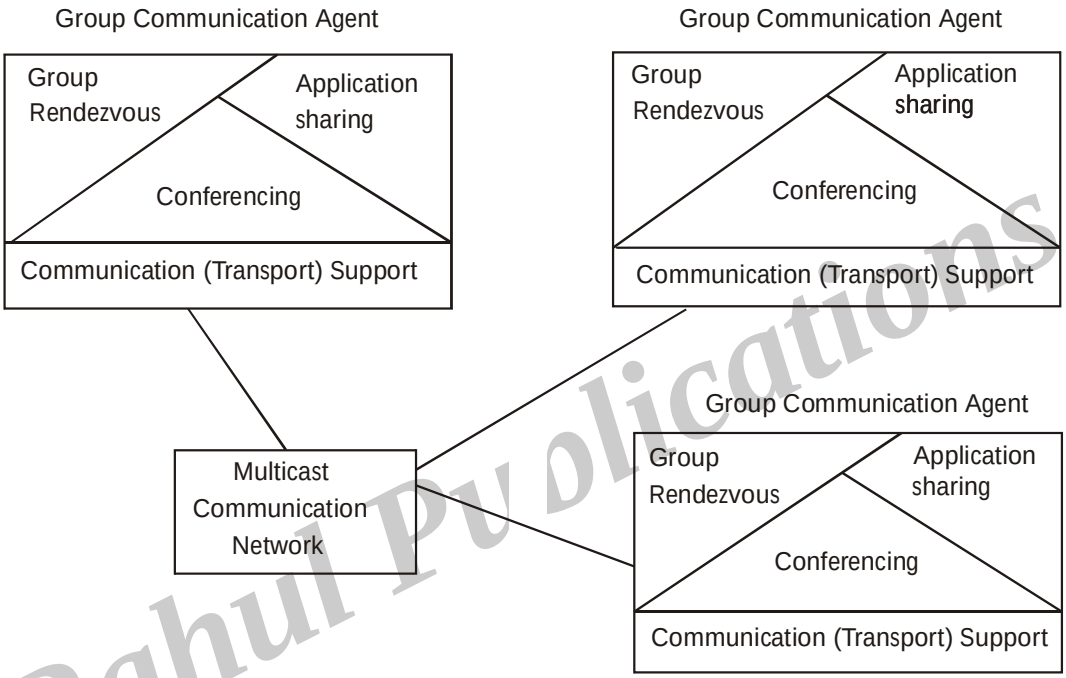


Fig.: Communication Support Model

1. Support Model:

It includes group communication agents that communicate via a multi-point multicast communication network as shown in the figure above. Group communication agents may use the following for their collaboration:

(i) Group Rendezvous

It represents methods which allows one to organize meetings, and to get information about the group, ongoing meetings and other static and dynamic information. The rendezvous can be Synchronous where the Directory services allows the access to information stored in a knowledge base about the conference, registered participants, authorized users and name and role of the participants. Explicit invitations method sends invitations either point-to-point or point-to-multipoint to conference participants. Asynchronous rendezvous methods may be implemented through email or bulletin boards.

(ii) Shared Applications

It refers to the techniques which allow the replication of the information which can be delivered to all participants of the collaborative work. It may use the Centralized Architecture or Replicated

Architecture. In the former a single copy of the application runs at one site say server.

All participants post their information to the central application which again distributes the information to the end users. This architecture is easy to manage and is less bandwidth hungry.

In the replicated architecture, a copy of the shared application runs locally at each site. Input events to each application are distributed to all sites and each copy of the shared application is executed locally at each site.

(iii) Conferencing

It represents the service which manages the multiple users to communicate and interact with each other by the use of multimedia data. Thus conferencing is basically a management service that controls the communication among multiple users via multiple media, such as video and audio, to achieve simultaneous face-to-face communication. The conferencing services control a conference by the implementation of following functions.

- Establishing a conference
- Closing a conference
- Adding new users and removing users who leave the conference

The conference may enforce its control centrally where the state of the conference is located on a central machine or it may enforce the control in the distributed manner.

2. System Model

It is based on a client-server model where the clients are applications that provides interface to the users who interact with the system while servers refers to function which makes it possible for the clients to communicate with each other and manage the communication.

3. Interface Model

It includes the user presentation protocols and group work management protocols. User presentation protocol is the interface available to the end users from which they can initiate, join, manage, communicate and terminate the conference. Group work management protocols specify the communication between the client and the servers for services like registration and querying the status of the conference.

4.1.2 Session Management

Q3. Explain about session management architecture.

Ans : **(Imp.)**

A session is the total logged in time of a user or it can be the entire conference from its commencement to its termination. The management of session is a very important task and it should consider several issues like allowing users to join and leave the conferencing, selection of the coordinator, distributing information between users.

Architecture of Session Management

It consists of following components:

1. Session manager
2. Media agent
3. Shared workspace agent

1. Session Manager

It includes local and remote functionalities. The local functionalities includes the

- **Membership Control management**

Authenticating the users or allowing members to join the session.

- **Control Management**

It may involve floor management which involves the distribution and sharing of the information and resources available to the conference.

- **Media Control Management**

It is required for the synchronization between the different media.

- **Configuration Management**

  It refers to the exchange and optimization of the QoS parameters.

- **Conference Control Management**

  It consists of functionalities for starting, changing and closing the conference.

2. Media Agents

These are responsible for decisions specific to each type of media. Each agent performs its own control mechanism over the particular medium, such as mute, unmute, change video quality, start sending, stop sending etc.

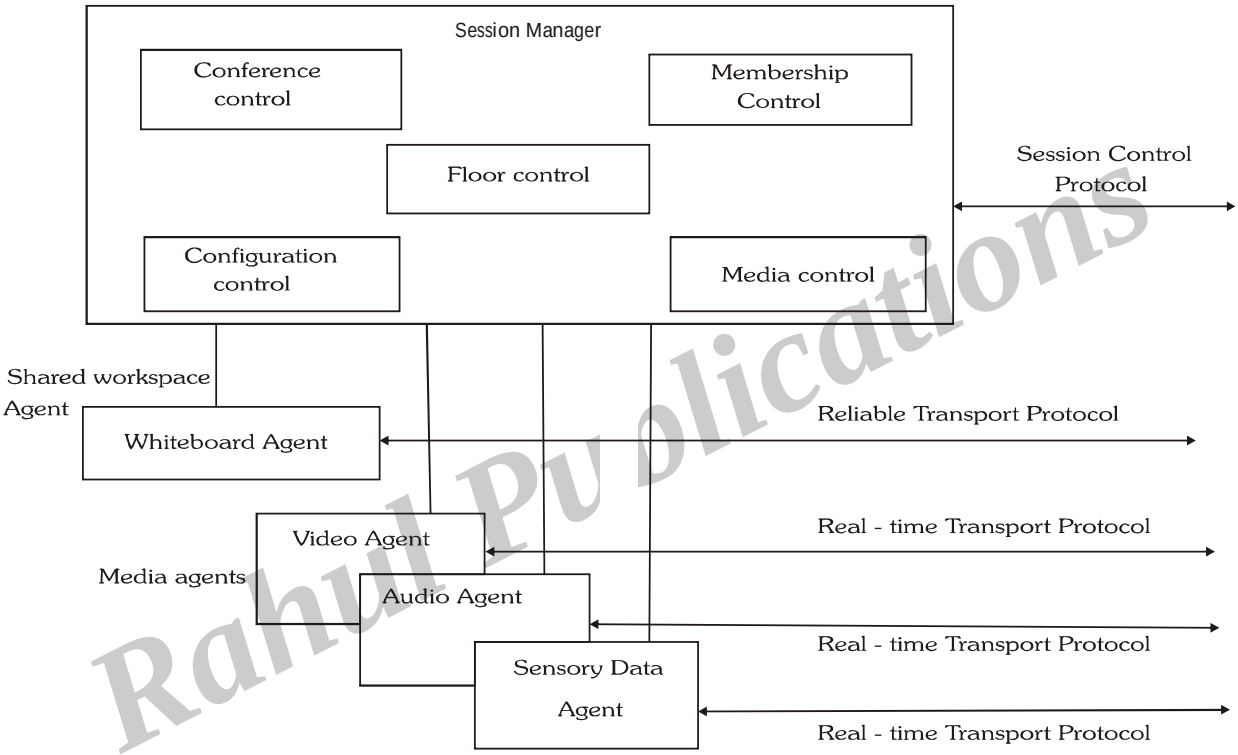


Fig.: Session Control Architecture

3. Shared Workspace Agent

The shared workspace agent transmits shared objects (e.g., telepointer coordinate, graphical or textual object) among the shared applications.

Q4. Explain various Control Mechanism Embedded in session Management.

Ans :

Each session is represented by its state which may the number of current users, media types, and the policies that are implemented.

Session Management includes two steps to process session state: an establishment and a modification of the session. During the establishment, the session manager negotiates, agrees, and sets the logical state of its own session. Modification means redefining the session variables.

The control mechanisms implemented for the management of session includes.

(i) Floor Control

Floor control is required to control and manage the access to the shared workspace, the resources etc. It is also required for maintaining the data consistency. In a typical floor passing mechanism at any time, only one participant has the floor. The floor is handed off to another participant when requested.

(ii) Conference Control

The conferencing services control a conference by the implementation of following functions

- Establishing a conference
- Closing a conference
- Adding new users and removing users who leave the conference

The conference may enforce its control centrally where the state of the conference is located on a central machine or it may enforce the control in the distributed manner.

(iii) Media Control

It is required for the synchronization of the different media streams or allocating them bandwidth.

(iv) Configuration Control

It refers to the exchange and optimization of the QoS parameters, managing the media quality and users' requirement.

(v) Membership Control

It refers to the sending invitation to the members, registration of users and modification of the Media Control:

It is required for the synchronization of the different media streams or allocating them bandwidth.

4.1.3 Transport Subsystem

Q5. What are the requirements of the transport sub system.

Ans : **(Imp.)**

Networked multimedia applications by themselves impose new requirements onto data handling in computing and communication because they need: Sustainable data throughput, fast data forwarding, service guarantees, and multicasting.

1. Data Throughput

This requirement wants the processing of the system to be fast and effective.

2. Fast Data Forwarding

The users or application wants very low end to end delay and jitter when communicating multimedia data. The holding time should be very less due to the real time requirement.

3. Service Guarantees

The loss of the information is undesired and the system or protocol used must ensure that the information is delivered to the intended destination.

3. Multicasting

It is important for sharing the bandwidth and the communication protocol processing at end systems.

Processing and Protocol Constraints

Processing system and protocols have constraints which need to be considered while processing and transmitting multimedia information.

- Following the "shortest possible path" for quicker delivery
- Buffer management
- Segmentation and reassembly
- Retransmission on error
- Error-recovery
- Asynchronous transfer

Q6. Write about TCP

Ans :

TCP (Transmission Control Protocol)

It was designed to provide a reliable end-to-end byte stream over an unreliable inter-network. Each machine supporting TCP has a TCP transport entity, either a library procedure, a user process, or part of the kernel. In all cases, it manages TCP streams and interfaces to the IP layer.

A TCP entity accepts user data streams from local processes, breaks them up into pieces not exceeding 64KB, and sends each piece as a separate IP datagram.

The IP layer gives no guarantee that datagrams will be delivered properly, so it is up to TCP to time out and retransmit them as needed. Datagrams that do arrive may well do so in the wrong order; it is also up to TCP to reassemble them into messages in the proper sequence. A key feature of TCP is that every byte on a TCP connection has its own 32-bit sequence number. The sending and receiving TCP entities exchange data in the form of segments. A TCP segment consists of a fixed 20-byte header followed by zero or more data bytes. However TCP is not suitable for real-time video and audio transmission because its retransmission mechanism may cause a violation of deadlines which disrupt the continuity of the continuous media streams.

Q7. Write a note on UDP protocol and its features.

Ans :

UDP Stands for (User Datagram Protocol). The Internet protocol suite supports a connection less transport protocol, UDP (User Datagram Protocol). UDP provides a way for applications to send encapsulated IP datagrams and send them without having to establish a connection. UDP transmits segments consisting of an 8-byte header followed by the payload. UDP does not do flow control, error control or retransmission upon receipt of a bad segment. All of that is up to the user processes. Many multimedia applications use this protocol because it provides to some degree the real-time transport property, although loss of PDUs may occur.

Features of UDP

The two important features of UDP are,

(i) Simple

It is a simple protocol that uses a very straightforward messaging structure that is similar to the message format used by many other TCP/IP protocols. The real goal of this protocol is to serve as an interface between networking application processes running at the higher layer and the inter networking capacities of IP. Like TCP, UDP layer is on the top of IP which is a method of addressing through the use of UDP port numbers. It does include an optional checksum capability for an error-detection but adds virtually no other functionality.

(ii) Fast

UDP is a fast protocol as it doesn't require error-detection mechanism. This make it unsuitable for use in many networking applications.

Q8. Describe about Real-time Transport Protocol (RTP).

Ans :

Real-time Transport Protocol (RTP)

RTP is a UDP protocol used in the client server environment and in the real-time multimedia applications. The multimedia application consists of multiple audio, video, text, and possibly other streams. These are fed into the RTP library, which is in the user space along with the application. This library then multiplexes the streams and encodes them in RTP packets, which it then stuffs into a socket. At the other end of the socket, UDP packets are generated and embedded in IP packets.

The basic function of RTP is to multiplex several real-time data streams onto single stream of UDP packets. The UDP stream can be sent to a single destination or to multiple destinations by unicasting and multicasting respectively.

Each RTP stream has a sequence number. If a packet is missing, the best action for the destination to take is to approximate the missing value by interpolation. It has no flow control, no error control, no acknowledgements, and no mechanism to request transmissions.

Q9. Write about Xpress transport protocol.

Ans :

Xpress Transport Protocol (XTP)

XTP integrates transport and network protocol functionalities to have more control over the environment in which it operates. XTP is intended to be useful in a wide variety of environments, from real-time control systems to remote procedure calls in distributed operating systems and distributed databases to bulk data transfer.

It defines for this purpose six service types:

connection, transaction, unacknowledged data gram, acknowledged datagram, isochronous stream and bulk data.

In XTP, the end-user is represented by a context becoming active within an XTP implementation. Two contexts are joined together to form an association. The path between two XTP sites is called a route. There are two types of XTP packets: information packets which carry user data, and control packets which are used for protocol management. It provides flow control and retransmission.

There are some features which meet the requirements for multimedia communication, such as

- XTP provides a connection-oriented transport and network transmission; hence it gives the benefit to map XTP on ATM networks and to use the possibilities of bandwidth reservation of ATM networks.
- Different transport services are provided: connection-mode, connectionless-mode and transaction-mode.
- It has error management and retransmission mechanism.
- It has rate-based flow control which allows it to provide a convenient mechanism for throughput and bandwidth reservation when QoS request is issued.

There are few problems though

- XTP requires VLSI while most current implementations are done in software and hence are slower.
- If the round rotation time of the network fluctuates it enters into handshake stage continuously which is undesirable.
- It has large header (44 bytes) which is an overhead.
- Source identification and discrimination is missing in XTP.
- Internetworking with other protocols is not worked out to provide QoS handling and resource reservation.

Q10. Explain how routing is done in network layer and write the supportive protocols to perform routing in network layer.

Ans : **(Imp.)**

1. Internet Protocol

In the TCP/IP protocol stack the network layer protocol is the Internet Protocol (IP) and, in order to transfer packets of information from one host to another, it is the IP in the two hosts, together with the IP in each Internet gateway and router involved that perform the routing and other harmonization functions necessary.

The IP in each host has a unique Internet-wide address assigned to it. This is known as the host's Internet address or, more usually, its IP address. Each IP address has two parts: a network identifier and a host identifier.

IP provides a connectionless best-effort service to the transport layer above it which is either the transmission control protocol (TCP) or the user datagram protocol (UDP). The source IP first adds the destination and source IP addresses to the head of the block, together with an indication of the source protocol (TCP or UDP)., to form the IP datagram. The IP then forwards the datagram to its local gateway. At this point the datagram is often referred to as a packet. An IP packet consists of header part and a text part. The header has a 20-byte fixed part and a variable length optional part.

Routers act as the guide to these packets. According to the information they store in their routing table they guide the datagrams to their destination. The routing is often determined according to congestion, shortest possible path etc. The recipient of the IP packets strips off the header from the packet and passes the block of information contained within it- known as the payload- to the peer transport layer protocol indicated in the packet header.

If the length of the packet is greater than maximum transfer unit (MTU) of the receiver

the packet is broken into segments. There is a Fragment offset field in the IP header which tells where in the current datagram this fragment belongs. All fragments except the last one in a datagram must be a multiple of 8 bytes.

The Time to live field is a counter used to limit packet lifetimes while the Protocol field specifies the transport layer protocol the packet is to be handed to. All IPv4 addresses are 32 bits long and are used in the Source address and Destination address fields of IP packet.

Inter connectivity between Internet Protocol and Underlying Networks The protocol used in Internet, data used and the services vary for e.g. there could be datagram transmission or multimedia traffic and the communication may take place through twisted pair telephone network, an Ethernet or an ATM B-ISDN. Hence, the mapping between the Internet protocol and the underlying layers is very important. More than that each and every machine is identified with its Ethernet address which is 48 bit in length so mapping of IP address to its corresponding Ethernet address is very important and which is done through ARP (Address Resolution Protocol). A related protocol, reverse ARP, can be used to map the Ethernet address to IP address.

2. Routing

Routing is used for guiding the packets from its source to the destination. Routers are dedicated for this purpose. Routing is based upon the congestion information, shortest path method.

These routers may be administered by a common authority and are called Autonomous Systems (AS) for which protocol- Interior Gateway Protocol (IGP). ASs of gateways exchange reachability information by means of Exterior Gateway Protocol (EGP).

3. Internet Group Management Protocol (IGMP)

Multicasting in Internet is done with the help of multicast routers. About once a minute, each multicast router sends a hardware multicast to the hosts on its LAN asking them to report back on the groups their processes currently belongs to. Each host sends back responses for all the class D addresses it is interested in. These query and response packets use a protocol called IGMP (Internet Group Management Protocol).

It has only two kinds of packets: query and response, each with a simple, fixed format containing some control information in the first word of the payload field and a class D address in the second word.

4. Resource Reservation Protocol (RSVP)

In order to ensure that the real-time traffic flows does not exceed that which is allocated for it, the resources required for each flow are reserved in advance of each packet flow starting. The resources can be bandwidth and buffer capacity. The protocol used to do this is called Resource Reservation Protocol. Because many of the new real-time applications involve multiple participants, RSVP is used to reserve resources in each router along either a unicast or a multicast path. When making a reservation, a receiver can specify one or more sources that it wants to receive from. It can also specify whether these choices are fixed for the duration of the reservation or whether the receiver wants to keep open the option of changing sources later.

The routers use this information to optimize bandwidth planning. Once a receiver has reserved bandwidth, it can switch to another source and keep that portion of the existing path that is valid for the new source.

4.1.4 QOS

Q11. Explain several issues which need to be addressed in QoS.

Ans : **(Imp.)**

During a multimedia communication, the services in the multimedia systems need to be parameterized. Parameterization of the services is defined in ISO standards through the notion of Quality of Service (QoS). Each service can be

characterized by a quality of service. As a simple example some services are reliable i.e. they do not loose data while some are unreliable as they may loose data. The parameters can be bandwidth, maximum and minimum end to end delay, jitter, buffer allocation etc.

There are several issues that need to be addressed and they are:

1. QoS Layering

The QoS requirement is associated with each layer of the OSI model, or the TCP/IP model.

However the QoS for multimedia communi- cation system (MCS) consists of three layers:

- application,
- system
- devices.

Application means the software and the program parameters, where as system refers to the overall system of communication and then in the network.

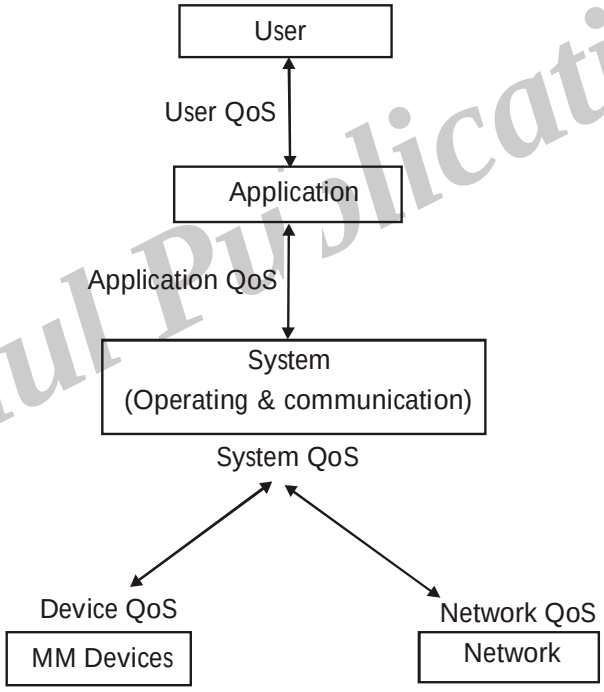


Fig.: QoS - layered model for the MCS

Services are performed on different objects, for example, media sources, media sinks, connections and Virtual Circuits (VCs), hence QoS parameterization species these service objects.

2. QoS Description

The QoS is described in terms of the required parameters by the end systems. The application QoS parameters may include media quality, transmission delay, jitter, synchronization.

The system QoS parameters describe requirements on the communication services and OS services resulting from the application QoS. They may be specified in terms of both quantitative (bandwidth, PDU size, buffer size) and qualitative criteria (level of synchronization, order of data delivery, recovery).

The network QoS parameters describe requirements on network services. They may be specified in terms of network load (packet size, service time) and network performance (congestion, delay).

The device QoS parameters typically specify timing and throughput demands for media data units.

3. QoS Parameter Values and Types of Service

There are three major type of service and they are:

- Guaranteed Services
- Predictive Services
- Best-effort services

In the Guaranteed services the QoS parameter values are deterministic or statistical in nature. It may ensure the lossless transmission of data i.e. reliable transmission. In the Predictive Services the QoS parameter are predictable with the help of the past parameters. Though the exact value of the parameters may not be known a rough estimate can be made.

In the Best-effort services, the QoS parameters depend on the load of the network. It ensures that the best possible service is provided to the multimedia data.

4.1.5 Resource Management

Q12. What is meant by resource?

Ans :

Resource

Resource is a system used for processing, storing, manipulating data.

- The resource can be active and passive. The active resource can be a CPU which processes data or manipulates data where as the passive resource is a bandwidth which only serves a particular purpose.
- A resource can be exclusive i.e. used by a single process or it may be shared where it is shared between various processes.
- The resource may be single or it may be multiple.

Q13. Explain resource management architecture with a neat diagram.

Ans : **(Imp.)**

Resource Management Architecture

The main goal of resource management is to offer guaranteed quality of service. It addresses three main actions

- Reserve and allocate resource
- Provide resources according to QoS specification
- Adapt to resource changes during on-going multimedia data processing.

Relation between QoS and Resources

The QoS parameters and their corresponding resources are mapped for the management of resources. For e.g. the end to end delay QoS parameter determines the behavior of transmission services along the path between source and sink with respect to packet scheduling, queuing and task scheduling. Description of a possible realization of resource allocation and management shows the QoS and resource relation.

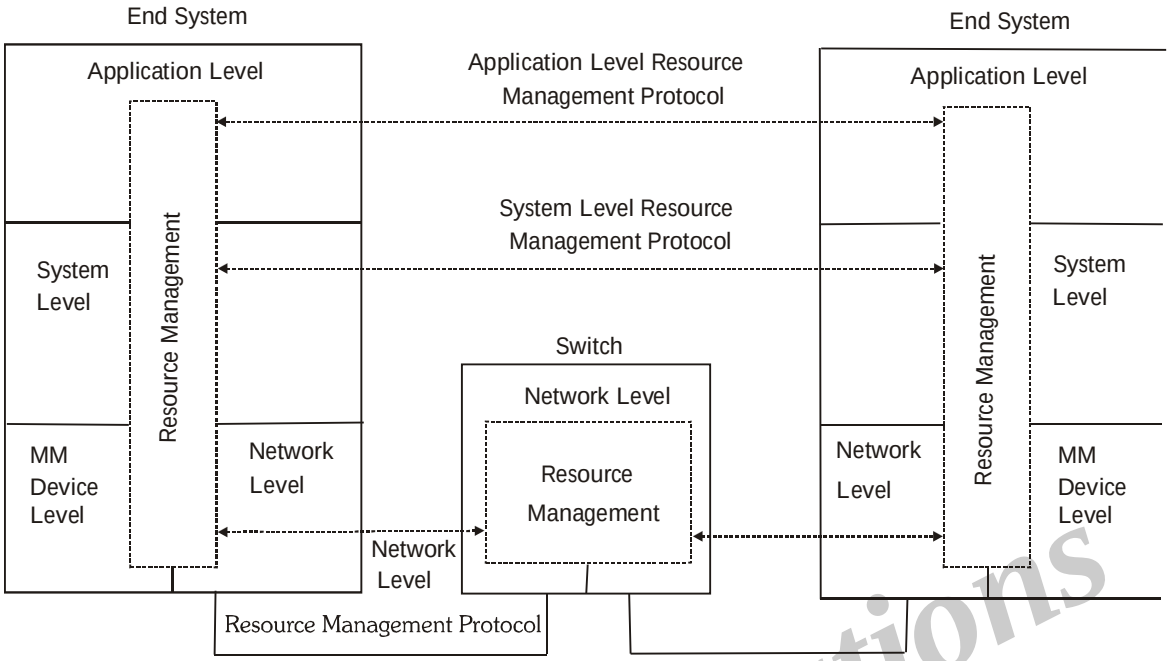


Fig.: Resource Management in MCSs

Establishment and Closing of the Multimedia Call

QoS Negotiation

It can be:

Bilateral Peer-to-Peer Negotiation

Negotiation occurs between the two service users and the service provider is not involved.

Bilateral Layer-to-Layer Negotiation

It occurs between the service user and the service provider.

Unilateral Negotiation

The user and the provider cannot modify the QoS parameter. It is based on “take it or leave it” model.

Hybrid Negotiation

The negotiation between host and sender is bilateral layer-to-layer negotiation and negotiation between network and host-receiver is unilateral.

Triangular Negotiation for Information Exchange

The user specifies its required QoS parameters while the provider may change it according to the possibility and availability before confirming that the caller agrees upon it.

Triangular Negotiation for a Bounded Target

It is similar to the above method but in this type of negotiation the caller specifies both the target value and the minimum required value. If the parameter that the provider does not meet the minimum required value the caller rejects the provider.

Triangular Negotiation for a Contractual Value

In this case, the QoS parameters are specified through a minimal requested value and bound of strengthening. The goal of this negotiation is to agree on a contractual value, which in this case is the minimal request QoS parameter value.

Translation

The QoS parameters for the users and application, system and network are different. However it must always be possible to derive all QoS values from the user and application QoS values. This derivation is known as translation which may require "additional knowledge" stored together with the specific component.

The translation functions used for the purpose are as follows:

- **Human Interface – Application QoS**

  The user must be able to tune the QoS parameters that are provided by the application.

  For e.g. in case of video he/she must be able to specify the resolution, and the compression ratio.

- **Application QoS – System QoS**

  The application level parameters are mapped to the System parameters for e.g. frame size specification at application level can be mapped as segmentation and reassembly.

- **System QoS – Network QoS**

  It maps the System QoS parameters like transport packet, end-to-end delay into the underlying network QoS parameters like in ATM, the end-to-end delay of cells and vice versa.

Q14. What is scaling? Write about it.

Ans :

Scaling

Scaling is a process of sub-sampling the data stream and only present a fraction of its original contents which can be done either at the source or at the receiver. Scaling can be

- Transparent Scaling
- In the transparent scaling the scaling done by the lower levels is not visible to the application level or the higher level.
- Non-Transparent Scaling
- Here the higher level notices that the lower levels have performed scaling and it may compensate the loss of data stream accordingly.

For audio, transparent scaling is a difficult task because even a small drop in data stream would be noticeable to the listeners.

For video the scaling can be Temporal Scaling which reduces the resolution of the video stream in the time domain i.e. the number of video frames transmitted within a time interval decreases.

In Spatial Scaling the number of pixels of each image in a video stream. Frequency Scaling reduces the number of DCT coefficients applied to the compression of an image. Amplitude Scaling reduces the color depth of the image pixel. Color space scaling reduces the number of entries in the color space.

Q15. Write about Resource Admission Procesure.

Ans :

Resource Admission

It is a procedure in which the source specifies the resources and their parameters that can be available to the user. The user tests for the availability of resource are called admission tests. Based on the results of the admission tests, the reservation protocol creates either a "reserve" message with admitted QoS values or a "reject" message when the minimal bound of QoS values cannot be satisfied.

The tests can be

- Schedulability test : e.g. CPU schedulability or packet schedulability
- Spatial Test : for buffer allocation for delay and reliability guarantees
- Link bandwidth Test: for throughput guarantees

Q16. Explain the terms which are used to manage resources during multimedia transmission.

Ans :

Managing Resources during Multimedia Transmission

Rate Control

A rate-based service discipline is one that provides a client with a minimum service rate independent of the traffic characteristics of other clients. Several rate-based scheduling disciplines have been developed.

Fair Queuing

If N channels share an output trunk, then each one should get 1/Nth of the bandwidth. The users are usually determined by first come first serve approach combined with round robin.

Virtual Clock

A virtual transmission time is allocated to each packet. It is the time at which the packet would have been transmitted, if the server would actually be doing Time Division Multiplexing.

Delay Earliest-Due-Date

If the source obeys a peak and average sending rate, then the server provides bounded delay. The server sets a packet's deadline to the time at which it should be sent, if it had been received according to the agreed contract.

Jitter Earliest-Due-Date

Here the server sets the jitter boundary and at the entrance of router and switch it is ensured the maximum value of allowable jitter is not exceeded.

Stop – and – Go

Here if a packet arrives at a frame n, the packet is sent only at frame n+1. It ensures that the packets that would arrive quickly are instead being delayed.

Hierarchical Round Robin

The server sets a priority and discriminates between the packets according to the priority. The packets at the same level get equal share of bandwidth. However packets at lower level get lesser bandwidth than the higher level.

End – to – End Error Control

Error Detection

Error Detection can be done by parity check, cyclic redundancy check or by including the checksum values.

Error Correction

Go-back-N Retransmission

If a packet arrives with an error, all the subsequent packets are also rejected unless the packet with error is retransmitted.

Selective Retransmission

If a packet arrives with error, it is discarded but the subsequent packets are not rejected but are buffered. The error packet is then asked to be retransmitted and subsequent processing occurs.

Forward Error Correction (FEC)

FEC uses Huffman codes, Checksum mechanism so that the packets can be corrected at the receiver and there is no need of retransmission.

Slack Automatic Repeat Request (S-ARQ)

For voice when there is jitter the receiver observes gaps, which result in interruptions of continuous playback of the voice stream. With this approach the first packet is artificially delayed at the receiver so that the buffer becomes full and the playback would exhibit no gaps. The slack time of a packet is defined as the difference between its arrival time at the receiver and its playback time. S-ARQ allows the control time to extend so that this time can be used to lengthen the slack time of arriving packets.

Resource monitoring

Resource monitoring function is responsible for checking the proper utilization of the resource and the shared resource is fairly used by the servers. For this the different parameter related to resource is exchanged with the help of agents.

Monitoring can add overhead during multimedia transmission, which should not cause violation of QoS guarantees. Monitoring at end-systems includes a supervisor function to continuously observe that the processed QoS parameters do not exceed their negotiated values. Different protocols like Resource Administration Protocols, Simple Network Management Protocol are used for the purpose.

Resource Adaptation

In the network it should be possible to dynamically change the QoS parameters and hence the use of Resources. This should be supported by the protocols for reporting the QoS changes in the existing connections and adaptive resource schemes to respond to and accommodate the changes either in the network, the hosts or both.

User Request for Renegotiation

If the user-receiver required change of QoS for the receiving media, first the resource manager checks the local resource and reserves it. Then, the sender is notified via a resource administration protocol and the same admission procedure follows as in the case of a user sender requiring QoS changes.

4.2 MULTIMEDIA DATABASES

4.2.1 Characteristics

Q17. Write about multimedia database and its characterstics.

Ans :

- Multimedia database is database containing multimedia collections.
- Multimedia database management system is essential to manage multimedia data like text, graphics, animation, music, etc.
- Multimedia database management system can be defined as a software system that manages a collection of multimedia data and provides access to users to query and retrieve multimedia objects .
- Generally, multimedia database contains text, image, animation, video, audio, movie, sound etc. which is stored in binary form.
- SQL query language is used for query and retrieval of data.
- There are generally two types of multimedia databases
- Linked Multimedia Databases and Embedded Multimedia Databases.
- Linked multimedia databases In this database, multimedia elements are organized as image, audio/ MP3, video etc. All the data may be stored either on off-line sources (CD-ROM, Hard Disc, DVD etc.) or on Online sources. One great advantage of this type of database is that the size of database will be small due to the reason that multimedia elements are not embedded in the database, but only linked to it.
- Embedded multimedia database Embedded Multimedia Database implies that the database itself contains the multimedia objects as in the binary form in the database. The main advantage of such kind of database is that retrieval of data will be faster because of the reduced data access time. However, the size of the database will be very large.

Characteristics

- A MDBMS (Multimedia Database Management System) can be characterized based on its objectives at the time of handling multimedia objects.
- **Corresponding storage media**

  Multimedia data must be stored and managed according to the specific characteristics of the available storage media.
- **Comprehensive search methods**

  During a search in the database, an entry, given in the form of text or a graphical image, is found using different search queries and the corresponding search methods.
- **Format independent interface**

  Database queries should be independent of media format. MDBMS should provide information in formats requested by the application.
- **Simultaneous data access**

  The same multimedia data can be accessed (even simultaneously) through different queries by several applications. Hence, consistent access to shared data can be implemented.

- **Management of large amount of data**

  The MDBMS must be capable of handling and managing large amounts of data.

- **Long Transaction**

  The performance of a transaction in a MDBMS means that transfer of a large amount of data will take a long time and must be done in a reliable manner.

- **Real-time Data**

  The read and write operations of continuous data must be done in real-time. The data transfer of continuous data has a higher priority than other database management actions.

4.2.2 Data Structures

Q18. Explain about the data structures used for multimedia database with an examples.

Ans : **(Imp.)**

In general, data can be stored in databases either in unformatted (unstructured) form or in formatted (structured) form. Unformatted, or unstructured data are presented in a unit where the content cannot be retrieved by accessing any structural detail.

Formatted or structured data are stored in variables, fields or attributes with corresponding values. Here, the particular data parts are characterized according to their properties. For example, a data description such as:

A.Student.Surname = Engler

A.Student.GivenNane = Clemens

A.Term = 8

can be accessed by structural details (student's given name, surname or term). Additionally, multimedia data can be stored in database s as ra,u)) registering and, descriptive data types.

1. Raw Data

An uncompressed image consists of a set of individual pixels. The pixels represent raw data in the form of bytes and bits. They create the unformatted information units, which represent a long sequence or set of symbols, pixels, sample values, etc.

2. Registering Data

These registering data are necessary to provide a correct interpretation of the raw data. Traditional DBMSs usually know only numbers and characters, which have fixed semantics therefore, no additional description is required.

3. Descriptive Data

Today, the search for textual and numerical content is very effective. However, the search for image, audio or video information is much more difficult. Therefore, optional description (descriptive data) should be assigned to each multimedia unit.

In the case of an image, the particular scene could be described in the form of text. These descriptive data provide additional redundant information and ease data retrieval during later searches. Descriptive data could be presented in unstructured or structured form.

Examples

We present examples of raw, registered, and descriptive data for different media such as text, image, video and audio. In the case of tert, the individual forms are:

1. Characters represent raw data.
2. The registering data describe the coding (e.g., ASCII). Additionally, a length entry must follow or an end symbol must be defined.
3. The descriptive data may include information for layout and logical structuring of the text or keywords.

Images can be stored in databases using the following forms:

Pixels (pixel matrix) represent raw data. A compressed image may also consist of a transformed pixel set. For example, the coefficients of the discrete cosine transformation in two-dimensional frequency presentation represent a transformed pirel set. A further compression of the raw data can include a set of entropy-encoded data.

The registering data include the height and width of the picture. Additionally, the details of coding are stored here. For example, in the case of a JPEG compressed image, the mode is entered first. This may be a specific JPEG mode based on a discrete cosine transformation. Additionally, for example, the eight bits per sample value for image processing, the sequential image structure and the entropy encoding scheme are defined. The tables for quantizing process and entropy encoding must also be specified.

Examples of descriptive data are individual lines, surfaces and subjects, or situations as a whole scene (e.g., "Birthday and 1995 New Year's Eve celebration.

at Lisa's favorite restaurant", or

B.Reason = New Year's Eve / Birthday

B.Date = 72/37/95

B.Place = Favorite Restaurant

B.Name = Lisa

B.Keywordl = New Year,s Eve Celebration

A motion video sequence can have a very different set of characteristics. It consists of the following information:

Raw data are defined in the simplest case through a sequence of pixel matrices. Mostly through motion video coding, the redundancies over several images are used for data reduction (intra-frame coding), so that each image does not carryall the necessary data for decompression. Also, a variable-rate data stream can be created.

The registering data provide, in addition to other information, the number of images per second .A data stream, coded according to the CCITT H.26l standard, is described as being QCIF with a resolution of I77 × 144 pixels for the luminance component and 88 X 72 pixels for the color difference components. The motion video, coded according to MPEG-2, is described by the relation between consecutive images; types are coded (1 I-frame, 2 B-frame, 1 P-frame, 2 P-frame,1 I-frame, etc.). Random access to each individual image of the motion video must follow.

The descriptive data provide a scene description (e.g., "Jan's birthday party with his friends from kindergarten")

Individual audio sequences can be classified according to the following scheme:

Raw data may be the digital sample values created by a simple PCM coding. The compressed values may also be considered as raw data.

The registering data represent the properties of the audio coding. Using a PCM coding, the sample rate, the quantization line and the resolution of the individual samples are the registering data. Compressed audio data can also carry additional information used by a parameterized decoder. Often, the coding information is already included in the raw data (e.g., ADPCM coding).

The descriptive data represent the content of the audio passages in a short form. In the case of a music composition, the name of the composition, the composer name and the name of the player can be entered. In the case of speech, a short content description, or the whole text can be written down.

4.2.3 Operation

Q19. What are the Operations that can be performed on MMDBMS?

Ans : **(Imp.)**

An MDBMS must offer, for all data types corresponding operations for:

- archival and
- retrieval

The media related operations will be handled as part of or an extension of query languages, e.g. SQL

Different classes of operations are needed:

- input
- output
- modification
- deletion
- comparison
- evaluation

1. Input

The input operation is used to write the data into the database. In this operation, the raw and registered data must be required where as the descriptive data can be provided later. While inputting a video or audio data, the size of the files must be specified. Otherwise, the MDBMS faces problem in selecting the required storage disk.

2. Output

The output operation is used to read the data from database as per the registering data. Therefore, the decoded data will be transmitted first followed by raw data.

3. Modification

The modification operation is used to modify the data using the respective editors like image editor, video editor, text editor etc. Usually, this operation performs on the raw data. The registering data is stored in modification attributes. Modification can be done in two different types. They are,

(i) Data conversion from one format to another format - Here, both the registering data as well as raw data will be modified.

(ii) Transformation from one medium to another medium - These types of conversions should be performed outside the MDBMS.

4. Deletion

The deletion operation is used to delete the data from databases. During this operation, the raw data should be protected. This is because, if the raw data is deleted, then the data types respective to the raw data will also be deleted.

5. Comparison

The comparison operation is used to compare the data. Searching and retrieving data from a database works on the comparison operation. It can be performed in two ways. They are,

(i) Here, the individual patterns of a specific medium will be compared with the raw data.

(ii) Here, a pattern from raw data will be stored as registered data. Then, the comparison will be performed based on this pattern.

6. Evaluation

The evaluation operation on raw and registering data is used to produce the respective descriptive data.

4.2.4 Integration in a Database Model

Q20. Explain the Integration of MMDBMS?

Ans : **(Imp.)**

Design of multimedia database management system is based on two different kinds of DBMS:

1. Extensible Relational MMDBMS
2. Object Oriented Database Management System

1. Extensible Relational MMDBMS

ERDBMS (Extensible Relational Database Management System):

- Definition of additional, application-dependent data types as domains for attributes.
- Definition of new functions to control behaviour and access to the data.
- Embedding new types and functions into existing RDMBS.

OODBMS (Object-Oriented Database Management System):

- Different media are represented by classes, whose instance variables include the data as internal state.
- Class hierarchy allows objects relations, offer well information navigation and flexible presentation possibilities.

Extensible Relational MMDBMS

Simplest possibility to implement a multimedia database is to use the relational database model.

The attributes of different media in relational databases are defined

Attributes can specify

- text
- audio
- video

Advantage

- compatibility with existent database applications

Example of ERDBMS:

A relation "student" is given

```
Student (
  Admission_Number          Integer,
  Name                      String,
  Picture                   Image,
  Exercise_Device_1         Video,
  Exercise_Device_2         Video
)
```

A relation's attributes can be specified through different media types

- picture
- exercise
- video

Other entries are "athletics", "swimming" and "analysis"

```
Athletics (
  Admission_Number          Integer,
  Qualification             Integer,
  The_High_Jump             Video,
  The_Mile_Run              Video
)

Swimming(
  Admission_Number          Integer,
  Crawl                     Video
)

Analysis (
  Qualification             Integer,
  Error_Pattern             String,
  Comment                   Audio
)
```

Classification of Extensible Relational Model:

Type 1 Relational Model

- Value of a certain attribute can be fixed over the particular set of the corresponding attribute types, e g. the frame rate of the video can be fixed
- In the example, the videos from the exercise devices 1 and 2 will play at the fixed rate defined by the type 1 specification

Type 2 Relational Model

- A varible number of entries can be defined through the type 2 relational model

- In the example, the individual disciplines of each admitted student are identified through their admission numbers

Type 3 Relational Model

- Additionally, an entry can simultaneously belong to several relations
- In the example, a video entry of a student can be assigned to the relation "athletics" as well as to the relation "analysis"

2. Object Oriented Model

In object-oriented databases

- classes with objects are defined
- objects can be put in relations via a class hierarchy
- a semantic specialization of classes and objects can **follow**

Example

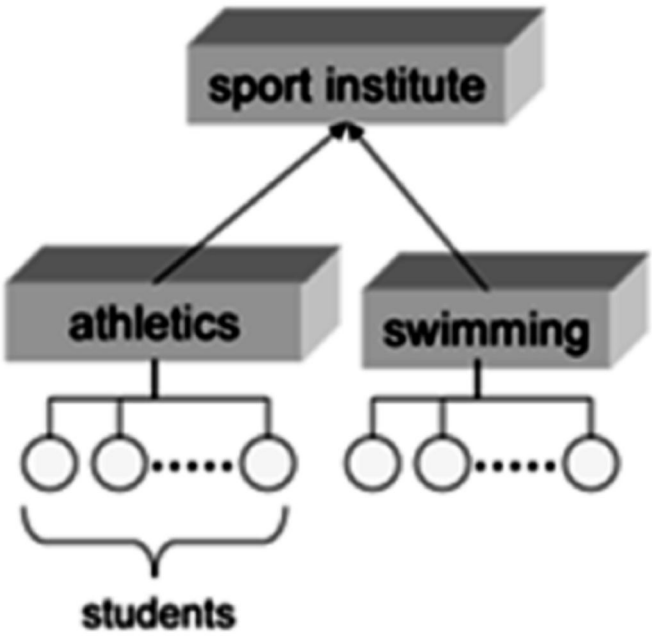
- Main class: sport institute
- Subclass: athletics, swimming
- Objects: students

Advantage:

- These system offer good information navigation and flexible presentation possibility

Disadvantages

- Query operations are incompletely supported



4.3 SYNCHRONIZATION

4.3.1 Issues

Q21. What is synchronization?

Ans :

The word synchronization refers to time. Synchronization in multimedia refers to the temporal relation between media objects in the multimedia system.

Synchronization between media object comprises relation between time dependent media object and time independent media object. A common example is synchronization between visual and the acoustical information in television.

Q22. Explain about the basic issues in synchronization.

Ans : **(Imp.)**

Basic Synchronization issues

1. Content Relations

Content relations define a dependency of media objects from some data. An example of content relation is the dependency between a filled spreadsheet and a graphics that represent the data filled in the spreadsheet.

2. Spatial Relations

The spatial relations that are usually known as layout relationships define the space used for the presentation of a media object on an output device at a certain point of time in a multi media presentation. If an output device is 2-D, the layout specifies the 2-D area to be used.

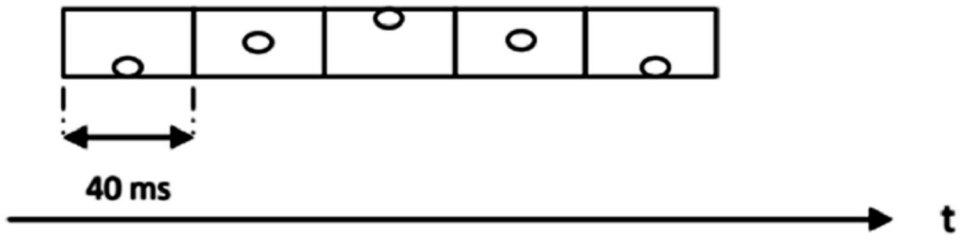
3. Temporal Relations

The temporal relations define the temporal dependencies between media objects. They are of interest whenever time-dependent media objects exist. E.g., the relation between a video and an audio object that are recorded during a concert.

4. Intra and inter-object synchronization

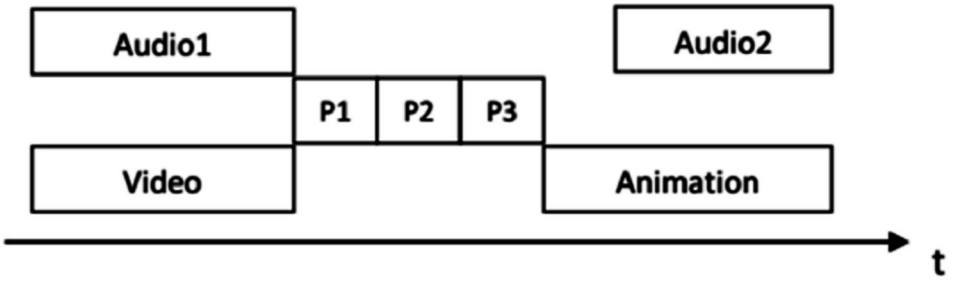
(i) Intra-object synchronization

Intra-object synchronization refers to the time relation between various presentation units of onetime-dependent media object. An e.g. is the time relation between the single frames of a video sequence of a bouncing ball. For a video with a rate of 25 frames per second, each of the frames must be displayed for 40 ms.



(ii) Inter-object synchronization

Inter-object synchronization refers to the synchronization between media objects. An e.g. is the time relations of a multimedia synchronization that starts with an audio/video sequence, followed by several pictures and an animation that is commented by an audio sequence.



5. Live and Synthetic synchronization

(i) Live synchronization

The goal of the live synchronization is to exactly reproduce a presentation the temporal relations as they existed during the capturing process. E.g. video conferencing.

(ii) Synthetic synchronization

In the synthetic synchronization, the temporal relations are artificially specified. E.g. Moving through the path of car race.

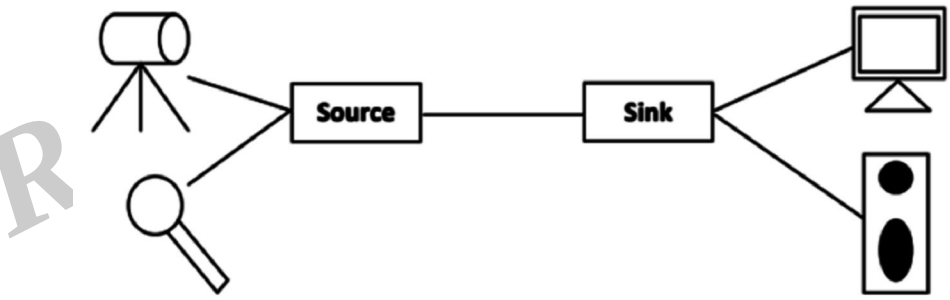


Fig.: Live Synchronization without intermediate long-term storage

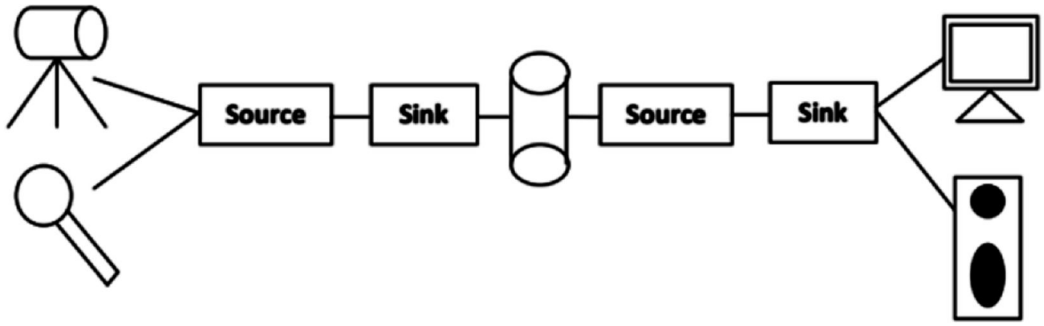


Fig.: Live Synchronization with intermediate long-term storage and delayed presentation.

4.3.2 Presentation Requirements

Q23. What are the presentation requirements in synchronization?

Ans :

For delivering multimedia data correctly at the user interface, synchronization is essential. A presentation requirement comprises, for intra-object synchronization, the accuracy concerning delays in the presentation of LDUs and, for inter-object synchronization, the accuracy in the parallel presentation of media objects.

Lip synchronization requirements

Lip synchronization refers to the temporal relationship between an audio and video stream for the particular case of human speaking. The time difference between related audio and video LDUs is known as *skew.* The streams which are perfectly in synchronization have no skew i.e. 0ms. The lip synchronization can be tolerated within skew of -80ms (audio behind video) and +80ms (audio ahead of video).

Pointer Synchronization requirements

In a computer-supported co-operative work (CSCW) environment, cameras and microphones are usually attached to the user's workstation. Using the pointer the speaker pointed out individual elements of the graphics. This obviously required synchronization of the audio and pointer. From the experiments, the synchronization area related to audio ahead of pointing is 750ms and for pointing ahead of audio is 500ms

4.3.3 Reference to Multimedia Synchronization

Q24. Explain the reference model of multimedia synchronization.

Ans : **(Imp.)**

A Reference Model for Multimedia Synchronization

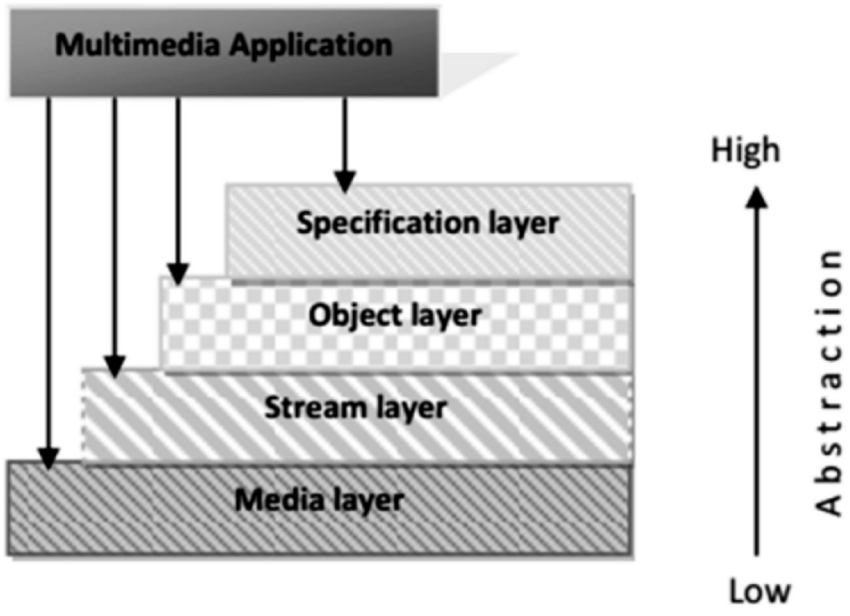


Fig.: Four-layer reference model

Each layer provides synchronization mechanism, which is provided by an appropriate interface. These interfaces can be used to specify or enforce the temporal relationships each interface defines services, i.e. offering the user a mean to define his/her requirements. Each layer can be used by an application directly, or by the next high or layer to implement an interface.

1. Media layer

At the media layer, an application operates on a single continuous media stream, which is treated as a sequence of LDU. The abstraction offered at this layer is a device independent interface with operation like read and write. Media layer implementation can be classified into simple implementation and implementation that provide across two interleaved media stream.

2. Stream layer

Stream layer operates on continuous media stream as well as on groups of media streams. In a group, all streams are presented in parallel by using mechanisms for inter stream synchronization. Continuous media is seen in the stream layers as a data flow with implicit time constraint. Individual LDUs are not visible. The streams are executed in real time environment, where all processing is constrained by well defined time specification. An application using the stream layer is responsible for starting, stooping and grouping the stream for the definition of the required QoS in terms of timing parameters supported by the stream layer. It is also responsible for the synchronization with time independent media object.

3. Object layer

Object layer operates on all types of media and hides the differences between discrete and continuous media. The abstraction offered to the application is that of a complete, synchronized presentation. This layer takes a synchronization specification as input and is responsible for the correct schedule of the overall presentation. The task of this layer is to close the gap between the needs for the execution of a synchronized presentation and the stream-oriented services. The functions located at the object layer are to compute and execute complete presentation schedule that include the presentation of the noncontinuous media object and the calls to the stream layer.

4. Specification layer

The specification layer is an open layer. It doesn't offer an explicit interface. This layer contains application and tools are located that allow to create synchronization specification. Such tools are synchronization editors, multimedia document editors, and authoring systems. It also contains tools for converting specification to an object layer format. It is also responsible for mapping QoS requirements of the user level to the qualities offered at the object layer interface

Q25. Write about Synchronization in Distributed Environment

Ans :

Synchronization in Distributed environment is more complex than in a local environment. Elements of synchronization in distributed environment are:

1. Transport of the synchronization specification

Three main approaches for the delivery of the synchronization information:

i) Delivery of the complete synchronized information before the start of the presentation

ii) Use of an additional synchronization channel

iii) Use of multiplexed data stream

2. Location of synchronization

To synchronize the media objects by combining the objects into a new media object.

3. Clock Synchronization

In distributed system the synchronization accuracy between the clocks of the source and sink nodes must be considered.

4. Multiple communication synchronization/Relation

Possible communication pattern with multiple sinks demand at run time.

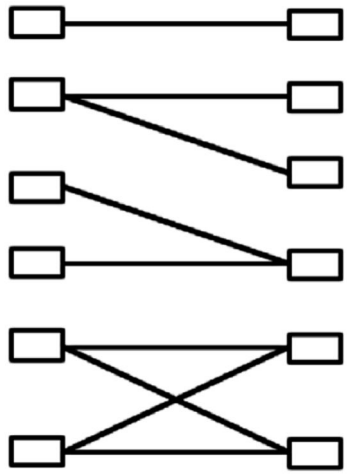


Fig.: Communication Pattern

5. Multi step synchronization

The synchronization must be maintained in a way that enabled the sink to perform the final synchronization.

The steps are:

i) synchronization during object acquisition (i.e. Digitizing video frame)

ii) synchronization of retrieval (i.e. synchronized access to frames of stored video)

iii) synchronization during the delivery of the logical data units (LDUs) to network

iv) synchronization during transport

v) synchronization at the sink

vi) Synchronization within the output device.

4.3.4 MHEG

Q26. What is MHEG? Explain Armedia client architecture.

Ans : **(Imp.)**

The generic space in MHEG provides a virtual coordinate system that is used to specify the layout and relation of content objects in space and time according to the virtual axes-based specification method.

MHEG Engine

Several components may be requires in implementing and MHEG systems:

Run time Engine (RTE)

MHEG-5 run time engines generally run accross a client-server architecture system referenced below for an example application).

Armedia Client Architecture

Armedia is a client-server based interactive multimedia application retrieval system.

A preceding Start-up Module may be used to perform general initialization etc.:

- The client can be launched either as an autonomous Windows application or
- As a plug-in by an HTML browser, allowing seamless navigation between the World Wide Web and the webs of MHEG-5 applications. (See Armida system for more details).

The MHEG-5 RTE is the kernel of the client's architecture. It performs the pure interpretation of MHEG-5 objects and, as a platform-independent module, issues I/O and data access requests to other components that are optimized for the specific run time platform.

The RTE performs two main tasks. First, it prepares the presentation and handles accessing, decoding, and managing MHEG-5 objects in their internal format. The second task is the actual presentation, which is based on an event loop where events trigger actions. These actions then become requests to the Presentation layer along with other actions that internally affect the engine.

Presentation layer

The presentation layer (PL) manages windowing resources, deals with low-level events, and performs decoding and rendering of contents from different media to the user. This functionality is available to the RTE via an object-oriented API that encapsulates all I/O platform specific aspects.

The basic MHEG-5 Presentable classes have

counterparts at this API level, which makes provisions for initialization/termination, data access and decoding, setting specific attributes (such as text font, color, and so on), and performing spatial and temporal controls. In addition, an informative flow exists from the PL back to the RTE, which notifies user interaction and stream events.

Access module

This module provides a consistent API for accessing information from different sources. It's used by the RTE to get objects and the PL to access content data (either downloaded or streamed). Note that the selection of a particular delivery strategy is out of MHEG-5's scope, and hence remains an implementation issue. Typical applications should support:

- bulk download for bitmaps, text, and MHEG-5 objects; and
- progressive download for audio and audiovisual streams

Q27. Write about MHEG engine.

Ans :

MHEG Engine

MHEG engine is considered as the implementation of object layer. It was developed at the European Networking Centre in Heidelberg. The basic architecture of an MHEG engine is as shown in the below figure,

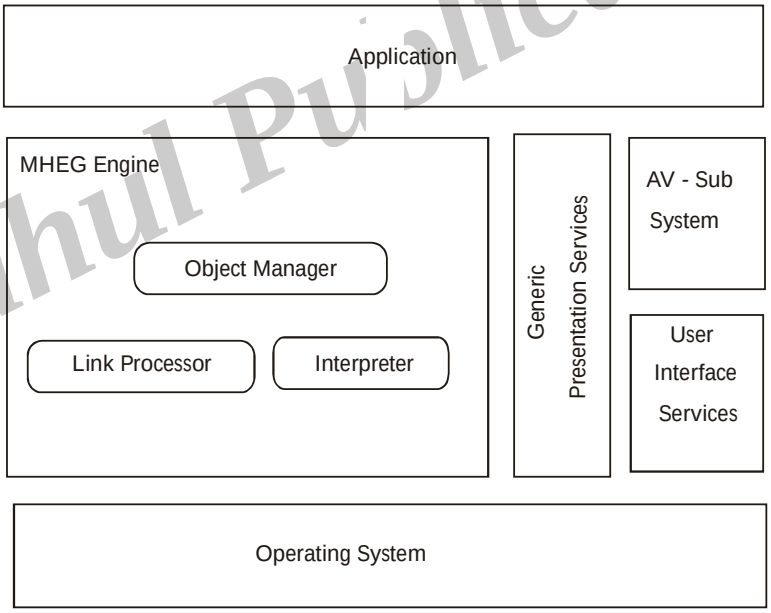


Fig: Architecture of an MHEG Engine

1. Components

The various components in MHEG engine are as follows,

2. Generic Presentation Services

This component is used to provide abstractions from presentation modules that are used to present the content objects.

3. Audio/Video Subsystems

This component is used to provide presentations of continuous media streams like audio stream and video stream. It is considered as the stream layer implementation.

4. User Interface Services

This component is used to provide the presentations of time-dependent media stream like text and graphics.

MHEG engine receives the MHEG objects sent from the application.

5. Object Manager

This component is used to handle all the MHEC i objects in run-time environment.

6. Link Processor

This component is used to observe the states of trigger links and objects. It also detennines whether the trigger conditions of a link are fulfilled or not.

7. Interpreter

This component is used to initiate the making and presentations of the objects. It also processes the action objects and events.

Short Questions and Answers

1. What are the layers in multimedia communication system

Ans :

From the communication perspective, we can divide the higher layers of the Multimedia Communication System (MCS) into two architectural subsystems:

- Application Subsystem
- Transport Subsystem

2. What is session?

Ans :

A session is the total logged in time of a user or it can be the entire conference from its commencement to its termination. The management of session is a very important task and it should consider several issues like allowing users to join and leave the conferencing, selection of the coordinator, distributing information between users.

3. What are the features and functions provide by transport layer

Ans :

Transport protocols, to support multimedia transmission, need to have new features and provide the following functions:

- Timing information
- Semi-reliability
- Multicasting
- NAK (none-acknowledgement)-based error recovery mechanism
- Rate control

4. What is resource allocation?

Ans :

Resource Reservation/Allocation

It can be pessimistic or optimistic. In the former the user system doubts that the resources will be available when communication will begin so, it reserves the required resources in advance. In the later, the user system reserves the required system only when required.

The different data structures and functions that are used are: Resource Table, Reservation Table, and Reservation function.

5. What is Collaborative Computing?

Ans :

Meaning

Collaborative computing is the computer supported cooperative work supported by the networks, PCs and the software that facilitates the cooperation. The examples of collaborative computing tools are electronic mail, bulletin boards, screen sharing tools, text-based conferencing systems, telephone conference systems, conference rooms and video conference systems.

6. Session Management Architecture.

Ans :

A session is the total logged in time of a user or it can be the entire conference from its commencement to its termination. The management of session is a very important task and it should consider several issues like allowing users to join and leave the conferencing, selection of the coordinator, distributing information between users.

7. TCP

Ans :

It was designed to provide a reliable end-to-end byte stream over an unreliable inter-network. Each machine supporting TCP has a TCP transport entity, either a library procedure, a user process, or part of the kernel. In all cases, it manages TCP streams and interfaces to the IP layer.

A TCP entity accepts user data streams from local processes, breaks them up into pieces not exceeding 64KB, and sends each piece as a separate IP datagram.

The IP layer gives no guarantee that datagrams will be delivered properly, so it is up to TCP to time out and retransmit them as needed. Datagrams that do arrive may well do so in the wrong order; it is also up to TCP to reassemble them into messages in the proper sequence. A key feature of TCP is that every byte on a TCP connection has its own 32-bit sequence number. The sending and receiving TCP entities exchange data in the form of segments. A TCP segment consists of a fixed 20-byte header followed by zero or more data bytes. However TCP is not suitable for real-time video and audio transmission because its retransmission mechanism may cause a violation of deadlines which disrupt the continuity of the continuous media streams.

8. Real-time Transport Protocol.

Ans :

RTP is a UDP protocol used in the client server environment and in the real-time multimedia applications. The multimedia application consists of multiple audio, video, text, and possibly other streams. These are fed into the RTP library, which is in the user space along with the application. This library then multiplexes the streams and encodes them in RTP packets, which it then stuffs into a socket. At the other end of the socket, UDP packets are generated and embedded in IP packets.

The basic function of RTP is to multiplex several real-time data streams onto single stream of UDP packets. The UDP stream can be sent to a single destination or to multiple destinations by unicasting and multicasting respectively.

9. Xpress Transport Protocol (XTP)

Ans :

XTP integrates transport and network protocol functionalities to have more control over the environment in which it operates. XTP is intended to be useful in a wide variety of environments, from real-time control systems to remote procedure calls in distributed operating systems and distributed databases to bulk data transfer.

It defines for this purpose six service types:

connection, transaction, unacknowledged data gram, acknowledged datagram, isochronous stream and bulk data.

In XTP, the end-user is represented by a context becoming active within an XTP implementation. Two contexts are joined together to form an association. The path between two XTP sites is called a route. There are two types of XTP packets: information packets which carry user data, and control packets which are used for protocol management. It provides flow control and retransmission.

10. What is scaling? Write about it.

Ans :

Scaling

Scaling is a process of sub-sampling the data stream and only present a fraction of its original contents which can be done either at the source or at the receiver. Scaling can be

- Transparent Scaling
- In the transparent scaling the scaling done by the lower levels is not visible to the application level or the higher level.
- Non-Transparent Scaling

- Here the higher level notices that the lower levels have performed scaling and it may compensate the loss of data stream accordingly.

For audio, transparent scaling is a difficult task because even a small drop in data stream would be noticeable to the listeners.

For video the scaling can be Temporal Scaling which reduces the resolution of the video stream in the time domain i.e. the number of video frames transmitted within a time interval decreases.

11. What is synchronization?

Ans :

The word synchronization refers to time. Synchronization in multimedia refers to the temporal relation between media objects in the multimedia system.

Synchronization between media object comprises relation between time dependent media object and time independent media object. A common example is synchronization between visual and the acoustical information in television.

12. MHEG.

Ans :

The generic space in MHEG provides a virtual coordinate system that is used to specify the layout and relation of content objects in space and time according to the virtual axes-based specification method.

MHEG Engine

Several components may be requires in implementing and MHEG systems:

Run time Engine (RTE)

MHEG-5 run time engines generally run accross a client-server architecture system referenced below for an example application).

Choose the Correct Answers

1. Among the following which is required to control and manage the access to the shared workspace, the resources etc. [a]

   (a) Floor Control (b) Conference Control

   (c) Media Control (d) Configuration Control

2. In file server which of the following resource is not required.? [c]

   (a) Processor (b) Secondary storage

   (c) Monitor (d) Network

3. When a delay occurs during the playback of the stream is known as _____ [a]

   (a) Jitter (b) Event delay

   (c) Stream delay (d) Playback delay

4. Which algorithm can be optimized to meet the timing deadlines and rate requirements of continuous media? [c]

   (a) Earliest-Deadline-First Scheduling

   (b) SCAN-EDF Scheduling

   (c) Both Earliest-Deadline-First scheduling & SCAN-EDF scheduling

   (d) None of the Mentioned

5. What is MPEG Compression? [b]

   (a) Stores the compression values of each frame

   (b) Stores the differences between successive frames

   (c) Stores multiple frames' values together

   (d) None of the mentioned

6. What are the levels in QoS? [d]

   (a) Best effort service (b) Soft QoS

   (c) Hard QoS (d) All of the mentioned

7. The level that treats different types of traffics in different ways, giving certain traffic streams higher priority than other streams and with best efforts, but no guarantees are made _____ [b]

   (a) Best effort service (b) Soft QoS

   (c) Worst effort service (d) Hard QoS

8. What are the factors that define QoS? [d]

(a) Throughput (b) Jitter

(c) Delay (d) All of the mentioned

9. Full form of RTP [a]

(a) Real-time Transport Protocol (b) Real-time Transmission Protocol

(c) Random Transport Protocol (d) Random Transmission Protocol

10. Among the following which layer in multimedia synchronization reference model operates on all types of media and hides the differences between discrete and continuous media. [c]

(a) Media layer (b) Stream layer

(c) Object layer (d) Specification layer

Fill in the blanks

1. _____ includes group communication agents that communicate via a multi-point multi cast communication network
2. A _____is the total logged in time of a user or it can be the entire conference from its commencement to its termination.
3. _____ is required for the synchronization of the different media streams or allocating them bandwidth.
4. _____ are responsible for decisions specific to each type of media.
5. _____ is required to control and manage the access to the shared workspace, the resources etc.
6. _____ protocol suite supports a connectionless transport protocol,
7. XTP stands for _____
8. _____ is used for guiding the packets from its source to the destination. Routers are dedicated for this purpose.
9. _____ is a process of sub-sampling the data stream and only present a fraction of its original contents which can be done either at the source or at the receiver.
10. In _____ multimedia elements are organized as image, audio/ MP3, video etc

ANSWERS

1. The Support Model
2. Session
3. Media Control
4. Media Agents
5. Floor Control
6. UDP
7. Xpress Transport Protocol
8. Routing
9. Scaling
10. Linked Multimedia Databases

UNIT V

MULTIMEDIA APPLICATION:

Media preparation, Composition, integration communication, consumption, entertainment.

5.1 MULTIMEDIA APPLICATION

5.1.1 Media Preparation

Q1. Write about multimedia preparation programs in US and EUROPE

Ans : **(Imp.)**

The availability of multimedia hardware and software components has driven the enhancement of existing applications towards being more user-friendly (known as re-engineering).

It has also initiated the continuous development of new multimedia applications. Applications are crucial for the whole domain of multimedia computing and communications because they are the only reason why any body would invest in this area.

Programs

Several programs for the development of multimedia applications have been established during the last few years, some well known from the US and Europe are:

US

HPCC (High performance computing and communication) program accelerates the development of scalable, high performance computers, advanced high-speed computer communication network and advanced software. One of the most significant program components of the HPCC program is IITA program (Information infrastructure Technology and application). IITA research and development effortare directed towards National Challenge problems such as civil infrastructure, digital libraries, education and lifelong learning, energy management, environment, healthcare, manufacturing processes and products, national security and public access to government information. IITA technology will support advanced applications such as:

- Tele medicine
- Remote education and training
- Tele operation
- Information access

Europe

ESPRIT (European strategic program to research in Information technology) is a well known scientific program of the European community). Its primary goal is to support development of technology and science similar to the HPCC program.

The smaller program RACE (Research in Advanced Communication in Europe) is similar to RACE but focuses on communication issues. In the second phase, RACE-II program focused on the residential and small business user market to use multimedia communication applications.

The RACE program covers applications such as

- Tele-interaction
- Tele-shopping
- Interactive TV and electronic newspaper
- Tele-working

Q2. What is the structure of multimedia applications.

Ans :

- There are many views on how multimedia applications are classified.
- For example, a market view may divide the current multimedia applications into kiosk applications, educational applications and applications in the area of cooperative work.
- Another view would be a communication oriented view dividing multimedia applications into interactive or distributive oriented applications.
- The third possibility is some view derived from the hypertext/hypermedia area.

Q3. Explain, how multimedia preparation can be done on various hardware Technologies

Ans :

Media Preparation

Media preparation is performed by multimedia I/O hardware and its supporting software. Therefore, hardware and software are the basic components for introducing media into the digital world of the computer.

New hardware technology is needed for multimedia applications and their interactive experience. Here we want to expand briefly on other devices also available for media preparation.

Audio Support

Some audio support with multiple channel digital sound tracks is already available. For e.g. a six channel digital sound track (front-left, center, front-right, surrounded-left, surrounded-right and subwoofer) has been developed. In the area of virtual reality entertainment, sound interaction occurs via a helmet.

Video Support

Video bards and digitizers aim towards a high-resolution picture presentation. The ultimate goal is high resolution and rate of 60 frames per second. Currently, several basic kinds of displays are used in virtual reality applications.

- Head mounted displays (HDD)
- Surrounded displays
- Digital Holography

Scanner Devices

- Photo scanner
- Image scanner
- Photo CD devices

Recognition devices

- Object oriented character recognition engine (Example: AQUIRE)
- Image recognition
- Speech recognition

Tracking devices

Trackers report information about position, orientation, acceleration, pressure or joint angles or tracked objects. There are several technologies which have been deployed:

- Electromagnetic trackers
- Ultrasound
- Optical tracking system
- Position only tracking
- Eye tracking technologies

5.2 MEDIA COMPOSITION

Q4. What are the different editors used for media composition?

Ans : **(Imp.)**

Media composition involves editing single media. i.e., changing its objects, such as characters, audio, sentences, video frames and attribute such as the font of the character, recording speed of an audio sentence of color of an image.

Different media editors are:

1. Text Editors

Text editors are used to write, arrange and modify the text in a document by providing various features. Some of the examples of text editors are as follows,

- CorelDRAW
- MS Visio
- Notepad, Notepad++ etc.

Some of the features need to consider while editing a text document are as follows,

(i) Fonts

Fonts determine the characteristics of the text. The font files include these descriptions either in vector form or in bitmap form.

(ii) Text Styles

Text styles determine the different style of text. The various text styles include italic, bold, underlined etc.

(iii) Text Effects

Text effects determine the different visual effects added to a particular text in the document. These effects include shadowing, textured fills, text-on-curve and extrusion etc.

Now-a-days the text editors are updated in such a way that they also can include the graphical representations such as images, tables etc. Microsoft word is the best example for these types of editors.

2. Graphics Editors

Graphics editors provide various features to edit the structural representation (i.e., structure-level editing) as well as to change the high-level operations (object-level editing) of graphical objects. The object and structure of an object will be modified because the primitive values of objects are stored in structural representations.

Example

X fig is an X windows drawing program that runs on UNIX systems.

3. Image Editors

Image editors are used by the applications when the application and software package stores the details of primitives.

The implementation of scaling cannot be done by providing the primitives with scaled endpoint coordinates. But, it can be done by editing the contents of image with the help of read-pixel and write-pixel operations. The two different techniques used for scaling are as follows,

(i) Pixel Replication

Pixel replication provides an easy and fast way to scale up an image. As a result, the image will become large as well as rough because no additional information is provided.

(ii) Sampling and Filtering

Sampling can be defined as a process of choosing a set of pixels from a continuous signal. Using this set, the pixels will be reconstructed and recreates an original signal.

Filtering can be defined as a process of obtaining a clear image. It is done before sampling or after reconstruction. It removes the noise from the image.

Aliasing

Aliasing is used to convert the high frequency components into low frequency components. Here, the resulted image consists of rough edges i.e., staircasing. This can be removed either by increasing the resolution or by using any anti-aliasing techniques.

Examples

- Adobe's photoshop™
- PhotoStyler™.

4. Animation Editors

Animation editing is performed on the frames generated from the graphical editors with respect to 2D or 3D spatial graphic objects. It includes the time as an extra component which can also be edited and is called as 4D editing. Some of the various functionalities of animation editor are as follows,

- Cutting one frame from an animation clip.
- Adding one frame to an animation clip.

The two different advanced animation techniques include,

(i) Tweening

Tweening is a method used in computer

animation to generate in-between frames between two images. It is also referred as in-betweening. The term tweening is mostly used while animating the digital context and the output set of frames are referred as the tweens.

(ii) Morphing

Morphing is a unique animation technique which changes/converts the shape of on object from one form to other form.

Examples

The various examples of animation editing tools are as follows,

- Animation Works Interactive (AWI) - It is used for the assembly of complete animations. It is mostly used by movie editors.
- Cel Editor - It is used to build cel and actors.
- Background Editor - It is used to build the background.

5. Sound Editors

Sound editors allow the user to perform various operations inorder to access, change and play the sound data. These operations are classified into four types. They are as follows,

Locating and Storing Sounds

The sound signals can be located and stored in four different ways. They are as follows,

1. Recording sound using the analog to digital converter.
2. Reading sound from the sound file.
3. Accessing sound from pasteboard.
4. Creating sound data by using the algorithms.

- **Recording**

This operation is used to record the sounds through microphones until it is stopped.

- **Playback**

This operation is used to play the sound using digital to analog converter device.

- **Editing**

This operation allows the user to copy/paste, cut, delete, insert or replace the sampled sound signal. Whenever the sound data is edited, it composes into fragments where the fragment sounds are considered as less efficient. Thus, the sound editing should be done in such a way the generated sound data should be contiguous.

5.3 MEDIA INTEGRATION COMMUNICATION

Q5. Define media integration. Write about media editors used for integration.

Ans : **(Imp.)**

Media integration specifies relationships between various media elements to represent and manipulate a media object. Integration is still very much dependent on technology i.e. platform-specific and format specific.

Multimedia editors

Multimedia editors support the ability to manipulate multimedia document that include structuredtext, multi-font text, bitmap image, graphics, video, digitized voice and other modifiable objects.

Several design issues need to be considered when implementing editors for multimedia documents.

1. Document Structure

The document structure determines the functionality of the editor. Thus, the structure of the multimedia document should achieve the international standards like ODA or SGML.

2. Media Editor Integration

Multimedia editors consist of the composition of individual editors for respective data. The various issues involved in media editor integration are as follows,

(a) Data Levels: The individual data should transmit from one type to another type.

(b) User Interface: The user interface of individual media editors should be consistent.

(c) Processes: The design of individual processes should be done using the respective editors and then serves as a parent process for management.

(d) Display Surface: The individual media should be viewed in an integrated way even though separate editors are used for editing.

3. Multiple Buffers and Multiple Panes

- Multiple buffers are used to manage multiple documents simultaneously while editing the documents.
- Multiple Panes are used to divide the document into some parts. Thus, the user can view the individual parts of a document.

4. Large Documents

The editor should be in such a way that they should be able to manage the large documents stored in distributed databases.

5. External Representation

The media should be represented in standardized format.

5.4 MEDIAN COMMUNICATION

Q6. Write about various sources for media communication.

Ans : **(Imp.)**

Media communication denotes applications which exchange different media over a network via Teleservices(e.g. video conferencing, cooperative works, mailing etc) to multimedia application end users.

The advantage of Tele-services in multimedia is that the end users can be located in different places,and

- Still internet closely in a quite natural way or
- Operate on remote data and resources in the same way as with local data and resources.

The disadvantage is that the delivery time of the Tele-services is longer than the processing time of local multimedia application.

Tele-services

Tele-services are services provided by communication system which are based on and make use of audio and video data. It can be Interactive services or distribution services.

1. Interactive services

Interactive services include an exchange of control data between remote sites to influence the presentation of continuous media data. Communication of between the sender and receiver can be performed either synchronously or asynchronously.

Depending on the task, the interactive services are again classified into five services. They are as follows,

(i) Conversational Service

This service provides the two-way communication/transfer of data between two end users in a synchronous mode. In this service, the data delivers to the destination quickly.

(ii) Messaging Service

This service provides communication/transfer of data between two users in asynchronous mode. Here, the users are human being.

(iii) Retrieval Service

This service provides communication/transfer of data between sender (client) and receiver (server). Here, human being is considered as client and computer with database provision is considered as server.

When the client requests for information stored in the server. The server retrieves the required information from the database and sends back to the client.

(iv) Tele-action Service

Tele-action service establishes a connection between two users located remotely. In this, the user writes or reads the data to the destination by remotely logging into other system. Tele-action services are again divided into six different services on the basis of their industry.

(v) Tele-operation Service

It provides a bi-directional communication. This allows the user to perform a specific task by manipulating a master controller.

2. Distribution services

Distributive services are services for the distribution of information to different remote sites. For e.g.TV or radio broadcasting use distributive services. There are two kinds of sub-process of distributive services: without individual user presentation control and with individual user presentation control. E.g. video on demand, Pay-per-view, etc

The distribution services are again classified into three types. They are as follows,

(i) Pay-per-view

It is a special type of service provided by cable operators. In this service, in order to watch particular programs on TV, a user has to pay for it.

(ii) Near Video-on-Demand

Near video on demand is considered a "pay-per-view consumer video delivery service" that offers more flexible viewing schedule than video-on-demand. In near video on demand, the service provider broadcasts a number of copies of a same video at short intervals of about 10-20 minutes. Such broadcasting allows the user's to view the video without tuning at every regular intervals though it is incapable of offering the interactive video services (such as pause/ resume) as in video on demand, it however ensures that no part of the video is missed. This type of video service is offered at a relatively low cost, because it requires less equipment and bandwidth to deliver the video. This is because video generated from the video server is transmitted multiple users simultaneously.

(iii) True Video-on-Demand

It offers the functions of VCR and allows the user to modify the view temporally.

Q7. Write about,

(i) Books, Proceedings and Newspapers

(ii) Kiosks

(iii) Tele-shopping/Mobiie Shopping.

Ans :

(i) Books, Proceedings and Newspapers

Books, proceeding and newspapers come under the category of interactive multimedia. This type of media can also be available electronically through the Internet. The user can either print the given data or can view it in the computer itself.

The e-paper will be updated frequently depending on the news. Using e-paper saves the trees and environment.

(ii) Kiosks

The Video Kiosks are the interactive, freestanding, electronic computer terminals. These are usually placed at store aisles and hotel lobbies.

It displays various products on a video screen and provides touch screen to select the products with ease. Customers can place order, perform the transactions and can make payment through credit/ debit card.

Around 2 million Kiosks are being used by US customers at present. Over 80 percent of video kiosks are performing retail related transactions across the world.

Examples

- Ticket Counters: The various activities like booking a ticket, cancelling a ticket etc., will occur at ticket counters.
- Bank Teller: Bank teller helps in the sale of life-insurance, fund transfers and investment tracking.
- Education System: Education system allows an online-student to interact with the lessons and quick review or feedback.
- Co-operative Work System: Co-operative work system helps in developing the team and status reports.

(iii) Tele-shopping/Mobile Shopping

Tele-shopping/Mobile shopping is a new form of shopping that enables a customer to locate the required product from intended seller using an internet connected mobile phone. It involves in searching a product in various online stores and comparing the features and prices of a product.

Example

When user wants to buy a 'SIA spectrum'. The user will search for the book in various online stores like 'Amazon', 'Flipkart' etc. After locating the required book, he/she compares the features and prices of the book and buys the book from the store which sells the book at low price. This process can be done through a mobile device having a good internet connection.

5.5 MEDIA CONSUMPTION

Q8. Define media consumption. What are the various media consumption modes?

Ans :

Media Consumption

Media consumption is the act of viewing, listening or feeling multimedia information. Viewing and listening are the most common ways user consume media. Feeling multimedia information can be experienced in motion-based entertainment for e.g. through virtual reality.

Viewing multimedia documents

Multimedia document can be viewed (consumed) in two modes:

1. Browsing
2. Detailed media consumption

1. Browsing

Browsing means that the user goes quickly through the document to get an overview of what the document includes. Example: user reading only the title of articles in the newspaper, table of contents of book, etc.

2. Detailed media consumption

Detailed media consumption means a detailed reading, viewing, or listening of the multimedia entity.

- **Kiosks :** Kiosk systems are the public information services located in the public areas, accessible to the visitors or customers. Response time is short and user interface is simple and easy to handle. Example: airport or train station kiosks with maps of terminals, arrival/ departure times and gate numbers; cinema information kiosks with information of upcoming movie preview clips, trivia, etc.

5.6 MEDIAN ENTERTAINMENT

Q9. Write about media entertainment technologies.

Ans : **(Imp.)**

Media Entertainment

Virtual reality entertainment, location-based entertainment, motion-based simulators, large screen films and games are applications that used multimedia for entertainment and bring a different and more involved entertainment experience that what is available with a standard TV or movie theatre.

Virtual reality (VR)

The term VR promises far more than our technology can currently deliver. It has been variously used to describe user interfaces ranging from synthesized physical environments presented or head-mounted displays to or ordinary CRTs. Computer based VR system are 3-dimensional interactive as opposed to passive and use one or more devices in an attempt to provide the user with a sense of presence of real situation.

Interactive video

- Interactive TV
- Video on demand (VOD)

Interactive audio (Example: CD-on-demand, thematic audio channel, etc) Games (Example: tele-games)

Q10. Explain the latest trends in multimedia application.

Ans :

- Applications are going from reengineering of existing applications to establish new application domain. The new applications may require reengineering of user-interface, new integration technologies etc.
- Multimedia applications are moving from a single PC environment to either a multi user
- environment or to a personalized user environment.
- Multimedia applications are developed less and less for local environment only and more and
- more for distributive environment.
- The solutions of current application are often platform specific and system dependent. The
- trend is going toward open solution so that, applications are portable across various platforms.
- Media consumption is going from the passive mode of user-computer interaction to an active mode of interaction.
- Technical improvements and changes in multimedia application improve productivity through better collaboration opportunities', visualization of different manufacturing processes etc.

Q11. Explain the concept of interactive audio and games.

Ans : **(Imp.)**

Interactive Audio

Interactive audio services represent a CD- on-Demand where the audio server stores the music libraries and allows the user to select the required audio.

Example

LyricTime

LyricTime is a music system consists of music libraries. User can select the audio from the server and can listen to it. It also displays the cover image of that particular audio. It allows the user to perform start, stop, forward, rewind on the audio.

Game

Computer game is an audio visual system that is able to maintain the internal model of a dynamic system. It abstracts the information regarding the internal state and creates a challenge to the user.

Games are categorized on the basis of storage location, environment sophistication and the number of players.

(i) Storage Location

The electronic games i.e., tele-games are stored on a remote or local computer.

(ii) Environment Sophistication

The games will be stored in,

- Interactive environment with advanced technology components like VR groupware, movies.
- Interactive environment with audio-visual components.

(iii) Number of Players

The one-on-one games, terminal-sharing games with two players will come under this category. In this, individual users have input devices and play simultaneously. Ludo King is an example of this category.

Games in an Interactive Environment

Branching is considered as one of the interactive game techniques. It helps the user in experiencing short, linear story segments. Each of these segments consists of many of choices at the end. These ends lead to new linear segment.

Branching techniques is not used by all interactive games. Some of the interactive games use clever algorithms.

Example:

SimCity

SimCity game imitates the growth and development of an urban metropolis having a player in a city mayor role. The player performs multiple operations on the landscape like zoning land, demolishing buildings and layering down roads, electric power lines and water mains. As it consists of large map and a player with capability of performing various operations, each and every moment in time is clearly huge.

Tele-games

Tele-games are mainly used in VR games and video games.

- **Video Games**

  The menu of video game is connected with a central system from which a game will be selected. Then a game scenario will be sent to the home participant by the central system. As the game starts, the home participant will have control over some objects. When he/she moves or plays with the objects, instructions will be sent to the machine via network and the central system responds with the targets. Also, the home participant can play game with another home participant instead of the central machine.

- **VR Games**

  The virtual reality games are similar to the video games but these games provide a realistic environment by using a helmet with HMD, gloves with motion sensors and headphones.

Short Question and Answers

1. Media Preparation.

Ans :

Media preparation is performed by multimedia I/O hardware and its supporting software. Therefore, hardware and software are the basic components for introducing media into the digital world of the computer.

New hardware technology is needed for multimedia applications and their interactive experience. Here we want to expand briefly on other devices also available for media preparation.

2. Media Composition.

Ans :

Media composition involves editing single media. i.e., changing its objects, such as characters, audio, sentences, video frames and attribute such as the font of the character, recording speed of an audio sentence of color of an image.

Different media editors are:

1. Text Editors

Text editors are used to write, arrange and modify the text in a document by providing various features. Some of the examples of text editors are as follows,

- CorelDRAW
- MS Visio
- Notepad, Notepad + + etc.

Some of the features need to consider while editing a text document are as follows,

(i) Fonts

Fonts determine the characteristics of the text. The font files include these descriptions either in vector form or in bitmap form.

(ii) Text Styles

Text styles determine the different style of text. The various text styles include italic, bold, underlined etc.

(iii) Text Effects

Text effects determine the different visual effects added to a particular text in the document. These effects include shadowing, textured fills, text-on-curve and extrusion etc.

Now-a-days the text editors are updated in such a way that they also can include the graphical representations such as images, tables etc. Microsoft word is the best example for these types of editors.

2. Graphics Editors

Graphics editors provide various features to edit the structural representation (i.e., structure-level editing) as well as to change the high-level operations (object-level editing) of graphical objects. The object and structure of an object will be modified because the primitive values of objects are stored in structural representations.

3. Define media integration.

Ans :

Media integration specifies relationships between various media elements to represent and manipulate a media object. Integration is still very much dependent on technology i.e. platform-specific and format specific.

4. Define Kiosks.

Ans :

The Video Kiosks are the interactive, freestanding, electronic computer terminals. These are usually placed at store aisles and hotel lobbies.

It displays various products on a video screen and provides touch screen to select the products with ease. Customers can place order, perform the transactions and can make payment through credit/ debit card.

Around 2 million Kiosks are being used by US customers at present. Over 80 percent of video kiosks are performing retail related transactions across the world.

Examples

- Ticket Counters: The various activities like booking a ticket, cancelling a ticket etc., will occur at ticket counters.
- Bank Teller: Bank teller helps in the sale of life-insurance, fund transfers and investment tracking.
- Education System: Education system allows an online-student to interact with the lessons and quick review or feedback.
- Co-operative Work System: Co-operative work system helps in developing the team and status reports.

5. Mobile Shoping

Ans :

Mobile shopping is a new form of shopping that enables a customer to locate the required product from intended seller using an internet connected mobile phone. It involves in searching a product in various online stores and comparing the features and prices of a product.

Example

When user wants to buy a 'SIA spectrum'. The user will search for the book in various online stores like 'Amazon', 'Flipkart' etc. After locating the required book, he/she compares the features and prices of the book and buys the book from the store which sells the book at low price. This process can be done through a mobile device having a good internet connection.

6. Define media consumption.

Ans :

Media consumption is the act of viewing, listening or feeling multimedia information. Viewing and listening are the most common ways user consume media. Feeling multimedia information can be experienced in motion-based entertainment for e.g. through virtual reality.

7. Media Entertainment

Ans :

Virtual reality entertainment, location-based entertainment, motion-based simulators, large screen films and games are applications that used multimedia for entertainment and bring a different and more involved entertainment experience that what is available with a standard TV or movie theatre.

8. Interactive Audio.

Ans :

Interactive audio services represent a CD- on-Demand where the audio server stores the music libraries and allows the user to select the required audio.

Example

LyricTime

LyricTime is a music system consists of music libraries. User can select the audio from the server and can listen to it. It also displays the cover image of that particular audio. It allows the user to perform start, stop, forward, rewind on the audio.

9. Write about internet gopher.

Ans :

Gopher is a search and retrieval tool that allows users to search for some information on the internet. It can also be called a document delivery tool that delivers documents and indexes. This simple tool delivers information both in textual format and structural format. This can be done by organizing the lists of resources that are browsed on internet.

10. Write short notes on tele-shopping.

Ans :

Tele-shopping

Tele-shopping/Mobile shopping is a new form of shopping that enables a customer to locate the required product from intended seller using an internet connected mobile phone. It involves in searching a product in various online stores and comparing the features and prices of a product.

Choose the Correct Answer

1. The RACE program covers applications such as [d]

   (a) Tele-interaction (b) Tele-shopping

   (c) Tele-working (d) All the above

2. Full form of HPCC _____. [c]

   (a) High performance commercial computers

   (b) High level Programing computing and communication

   (c) High performance computing and communication

   (d) High level Programming commercial Computers

3. Among the following which is not a kind of display in video support [d]

   (a) Head mounted displays (HDD) (b) Surrounded displays

   (c) Digital Holography (d) Kiosk

4. AQUIRE is an example for _____. [a]

   (a) Object oriented character recognition engine

   (b) Image recognition

   (c) Speech recognition

   (d) Image scanner

5. _____ is an example of virtual reality applications. [c]

   (a) Photo scanner (b) Image scanner

   (c) Head mounted displays (HDD) (d) Photo CD devices

6. Among the following which include an exchange of control data between remote sites to influence thepresentation of continuous media data [b]

   (a) Tele-services (b) Interactive services

   (c) Distribution services (d) Media services

7. _____ are the public information services located in the public areas, accessible to the visitors or customers [a]

   (a) Kiosk systems (b) Display boards

   (c) Servers (d) All the above

8. Among the following which applications that used multimedia for entertainment and bring a different and more involved entertainment experience. [d]

(a) Virtual reality entertainment (b) location-based entertainment,

(c) motion-based simulators (d) All the above

9. Among the following which is not the example of interactive service [c]

(a) Conversational services (b) Messaging services

(c) Distributive services (d) Retrieval service

10. Among the following which is the example of Hyper text editors [d]

(a) Hypercard (b) DynaText

(c) Hyperbole (d) All the above

Fill in the blanks

1. Full form of ESPRIT_____.
2. _____ program accelerates the development of scalable, high performance computers, advanced high-speed computer communication network and advanced software.
3. _____ type of applications are used in markets to guide the customers
4. _____ has been developed. In the area of virtual reality entertainment, sound interaction occurs via a helmet.
5. _____ involves editing single media. i.e., changing its objects, such as characters, audio, sentences, video frames and attribute such as the font of the character, recording speed of an audio sentence of color of an image.
6. _____ specifies relationships between various media elements to represent and manipulate a media object.
7. _____ denotes applications which exchange different media over a network via Teleservices to multimedia application end users.
8. _____ is the act of viewing, listening or feeling multimedia information.
9. _____ has been variously used to describe user interfaces ranging from synthesized physical environments presented or head-mounted displays to or ordinary CRTs.
10. _____ is going from the passive mode of user-computer interaction to an active mode of interaction.

ANSWERS

1. European strategic program to research in Information technology
2. HPCC
3. Kiosk
4. Sixchannel digital sound track
5. Media composition
6. Media integration
7. Media communication
8. Media consumption
9. Virtual reality (VR)
10. Media consumption

FACULTY OF COMMERCE

B.Com. III Year VI - Semester(CBCS) Examination

Model Paper - I

MULTIMEDIA SYSTEMS

(Only for Computer Applications Courses)

Time: 1½ Hours] [Max. Marks : 50

PART – A (5 × 2 = 10 Marks)

Note: Answer any five of the following questions not exceeding 20 lines each.

ANSWERS

1. MIDI. **(Unit-I, SQA.4)**
2. What is speech? **(Unit-I, SQA.8)**
3. Digital Image Representation **(Unit-II, SQA.2)**
4. What is image synthesis? **(Unit-II, SQA.4)**
5. CDROM **(Unit-III, SQA.7)**
6. Real-time Transport Protocol. **(Unit-IV, SQA.8)**
7. MHEG. **(Unit-V, SQA.12)**
8. Media Preparation. **(Unit-V, SQA.1)**

PART – B (5 × 8 = 40 Marks)

Note: Answer all the questions in not exceeding four pages each.

9. (a) Explain various kinds of MIDI messages. **(Unit-I, Q.No.9)**

OR

(b) What is speech transmission? Explain about the components of speech transmission system. **(Unit-I, Q.No.23)**

10. (a) What is image format? Discuss its types. **(Unit-II, Q.No.2)**

OR

(b) Explain various methods of controlling animation. **(Unit-II, Q.No.20)**

11. (a) Explain various types of coding techniques in data compression. **(Unit-III, Q.No.1)**

OR

(b) Explain about Write Once, Read Many storage device. **(Unit-III, Q.No.11)**

12. (a) Explain resource management architecture with a neat diagram. **(Unit-IV, Q.No.13)**

OR

(b) What are the Operations that can be performed on MMDBMS? **(Unit-IV, Q.No.19)**

13. (a) What are the different editors used for media composition? **(Unit-V, Q.No.4)**

OR

(b) Explain the concept of interactive audio and games. **(Unit-V, Q.No.11)**

FACULTY OF COMMERCE

B.Com. III Year VI - Semester(CBCS) Examination

Model Paper - II

MULTIMEDIA SYSTEMS

(Only for Computer Applications Courses)

Time: 1½ Hours] [Max. Marks : 50

PART – A (5 × 2 = 10 Marks)

Note: Answer any five of the following questions not exceeding 20 lines each.

ANSWERS

1. What is speech recognition? **(Unit-I, SQA.10)**
2. What is Multimedia? **(Unit-I, SQA.1)**
3. What is image? **(Unit-II, SQA.1)**
4. What is digital image processing? **(Unit-II, SQA.3)**
5. CDWO technology **(Unit-III, SQA.8)**
6. What is synchronization? **(Unit-IV, SQA.11)**
7. What is session? **(Unit-IV, SQA.2)**
8. Define Kiosks. **(Unit-V, SQA.4)**

PART – B (5 × 8 = 40 Marks)

Note: Answer all the questions in not exceeding four pages each.

9. (a) Explain about the global structure of multimedia system. **(Unit-I, Q.No.2)**

OR

(b) Explain about speech generation system. **(Unit-I, Q.No.20)**

10. (a) Define Image Recognition. Explain image recognition steps with diagram. **(Unit-II, Q.No.7)**

OR

(b) Explain various types of animation languages. **(Unit-II, Q.No.19)**

11. (a) Explain about CDMO technology. **(Unit-III, Q.No.16)**

OR

(b) Why do we need multimedia file systems? Explain. **(Unit-III, Q.No.23)**

12. (a) Explain how routing is done in network layer and write the supportive protocols to perform routing in network layer. **(Unit-IV, Q.No.10)**

OR

(b) Explain about the basic issues in synchronization. **(Unit-IV, Q.No.22)**

13. (a) Write about various sources for media communication. **(Unit-V, Q.No.6)**

OR

(b) Write about multimedia preparation programs in US and EUROPE. **(Unit-V, Q.No.1)**

FACULTY OF COMMERCE

B.Com. III Year VI - Semester(CBCS) Examination

Model Paper - III

MULTIMEDIA SYSTEMS

(Only for Computer Applications Courses)

Time: 1½ Hours] [Max. Marks : 50

PART – A (5 × 2 = 10 Marks)

Note: Answer any five of the following questions not exceeding 20 lines each.

ANSWERS

1. What is Sound? **(Unit-I, SQA.3)**
2. What is SMPTE timecode? **(Unit-I, SQA.6)**
3. What is image transmission? **(Unit-II, SQA.7)**
4. Animations **(Unit-II, SQA.11)**
5. JPEG **(Unit-III, SQA.1)**
6. What is Collaborative Computing? **(Unit-IV, SQA.5)**
7. TCP **(Unit-iV, SQA.7)**
8. Mobile Shoping **(Unit-V, SQA.5)**

PART – B (5 × 8 = 40 Marks)

Note: Answer all the questions in not exceeding four pages each.

9. (a) Explain about the relationship between music and computer. **(Unit-I, Q.No.7)**

OR

(b) (i) What is a SMPTE device? **(Unit-I, Q.No.13)**

(ii) What physically is the SMPTE time code? **(Unit-I, Q.No.14)**

10. (a) Explain about various measures taken for visual representation of a video signal. **(Unit-II, Q.No.9)**

OR

(b) Explain various standards of television conventional system. **(Unit-II, Q.No.14)**

11. (a) Explain MPEG video compression standard. **(Unit-III, Q.No.8)**

OR

(b) Explain about CDDA technology. **(Unit-III, Q.No.12)**

12. (a) What is MHEG? Explain Armedia client architecture. **(Unit-IV, Q.No.26)**

OR

(b) Explain about session management architecture. **(Unit-IV, Q.No.3)**

13. (a) Write about media entertainment technologies. **(Unit-V, Q.No.9)**

OR

(b) Define media integration. Write about media editors used for integration. **(Unit-V, Q.No.5)**