1

UNIT I

INTRODUCTION TO MULTIMEDIA

Unit Structure

1.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define Multimedia
2. Define the scope and applications of multimedia
3. Distinguish between Multimedia Presentation and Production
4. Understand how a multimedia presentation is created

1.2 INTRODUCTION

Multimedia refers to content that uses a combination of different content forms. This contrasts with media that use only rudimentary computer displays such as text-only or traditional forms of printed or hand-produced material. Multimedia includes a combination of text, audio, still images, animation, video, or interactivity content forms.

Multimedia is usually recorded and played, displayed, or accessed by information content processing devices, such as computerized and electronic devices, but can also be part of a live performance. Multimedia devices are electronic media devices used to store and experience multimedia content. Multimedia is distinguished from mixed media in fine art; by including audio, for example, it has a broader scope. The term "rich media" is synonymous for interactive multimedia. Hypermedia can be considered one particular multimedia application.

1.3 WHAT IS MULTIMEDIA?

The word “multimedia” comes from the Latin words *multus* which means “numerous” and *media* which means “middle” or centre. Multimedia therefore means “multiple intermediaries” or “multiple means”.

Multimedia is a combination of following elements:

- Text [E.G. books, letters, newspapers]
- Images and graphics [E.G. photographs, charts, maps, logos, sketches]
- Sound [E.G. radio, gramophone records and audio cassettes]
- Video and animation [E.G. TV, video cassettes and motion pictures]

1.3.1 CHARACTERISTICS OF MULTIMEDIA PRESENTATIONS

- Multiple media
- Non-linearity
- Scope of Interactivity
- Integrity
- Digital representation

1.3.1 MULTIPLE MEDIA:

Multiple Media may be described as transmission media using more than one type of transmission path (e.g., optical fiber, radio, and copper wire) to deliver information. In addition to text, pictures are also started being used to communicate ideas.

Pictures are sub-divided into two types.
I. A real-world picture captured by a camera is called images.
II. A hand-drawn picture like sketches, diagrams and portraits called graphics.

Text, images and graphics are together referred to as static elements, because they do not change overtime. With further improve in technology, time varying elements like sound and movies were used. Movies are again divided into two classes.

They are
- *Motion pictures*
- *Animation*

Legitimate multimedia presentation should contain at least one static media like text, images or graphics and at least one time varying media like audio, video or animation.

1.3.2 NON-LINEARITY:

Non-Linearity is the capability of jumping or navigating from within a presentation with one point without appreciable delay. Traditionally, TV shows and motion pictures are considered linear presentation because the user or viewer has to watch the information being prescribed. The user cannot modify the content. Note that the difference between linear and non-linear is blurring as time goes by. Now the DVD format allows you to have a non-linear experience by choosing scenes and going forward and backward.

In a multimedia presentation the user can instantly navigate to different parts of the presentation and display the frames in any way, without appreciable delay, due to which it is called a non-linear presentation.

Non-linear content offers user interactivity to control progress as used with a computer game or used in self-paced computer based training. Hypermedia is an example of non-linear content.

1.3.3 SCOPE OF INTERACTIVITY:

In a non-linear presentation user will have to specify the desire to watch the presentation. The presentation should be capable of user inputs and capable of change the content of the presentation. Interactivity is considered to be one of salient features on which next generation e-learning tools are expected to reply for greater effectively.

The term “rich media” is synonymous for interactive multimedia. Interactive media works with the user's participation. The media still has the same purpose but the user's input adds the interaction and brings interesting features to the system for a better enjoyment.

Interactive media is helpful in these four development dimensions in which young children learn: social and emotional, language development, cognitive and general knowledge, and approaches toward learning. Using computers and educational computer software in a learning environment helps children increase communication skills and their attitudes about learning.

Interactive media makes technology more intuitive to use. Interactive products such as smartphones, iPad's/iPod's, interactive whiteboards and websites are all easy to use. The easy usage of these products encourages consumers to experiment with their products rather than reading instruction manuals.

Interactive Media can be implemented in a wide variety of platforms and applications encompassing virtually all areas of technology. Some examples include mobile platforms such as touch screen smartphones and tablets, was well as other interactive mediums that are created exclusively to solve a unique problem or set of problems. Interactive media is not limited to a certain field of IT, it instead encompasses any technology that supplies for movie parts or feedback based on the users actions. This can include JavaScript and AJAX utilization in web pages, but can further be extended to any programming languages that share the same or similar functionality.

Scope of interactive media continues to advance today, with the advent of more and more powerful machines the limit to what can be input and manipulated on a display in real time is become virtually non-existent.

1.3.4 INTEGRITY:

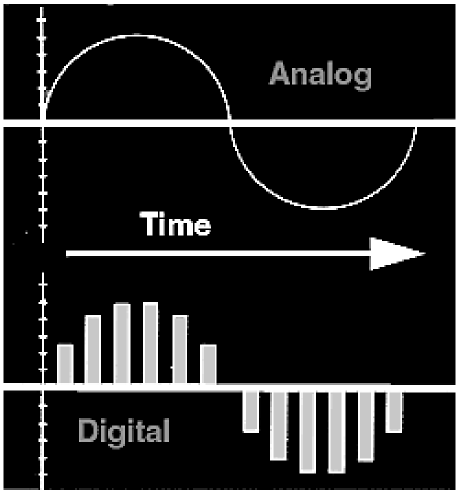
An important characteristic of a multimedia presentation is integrity. This means that although there may be several media types present and playing simultaneously, they need to be integrated or be part of a single entity which is the presentation. It should not be able to separate out the various media and control them independently; rather they should be controlled from within the frame work of the presentation. Moreover, the presentation should decide how the individual elements can be controlled.

1.3.5 DIGITAL REPRESENTATION:

Magnetic tapes are called the sequential access storage devices [i.e.] data is recorded sequentially along the length of the tape. When a specific potion of the data is required to be played back, the portion before that needs to be skipped. Multimedia requires instant access to different portion of the presentation. This is done by random access storage devices like hardware, floppy disks, and compact disks.

Digital multimedia are represented by a discrete set of values defined at specific instances of the input domain (time/space or both). For example, Digital sound – finite series of instantaneous pulses which are measured at distinct interval of time

Digital representations has other advantages, software based programs can be used to edit the digitized media in various ways to appearances and compress the file sizes to increase the performance efficiency.



DIGITAL REPRESENTATION

1.4 DEFINING THE SCOPE OF MULTIMEDIA

Today, much of the media content we consume is available in a variety of formats, intended to serve multiple purposes and audiences. For example, a book usually starts out as a print-only product. However, if the market demand is large enough, it may also be published in a spoken-word format and delivered via compact disc or MP3. With the right equipment, you can avoid paper altogether by downloading the e-book, a digital version of the text designed for reading on a computer screen or a handheld device such as the Kindle or iPad.

The website for a bestseller may offer bonus material or value-added content to online users through a gamut of multimedia channels—featuring audio excerpts, video interviews, background stories, pictures, and more. With such a vast sea of information and social networking potential, you can easily imagine many of the other possibilities that exist.

The opportunities for shaping content to meet the diverse needs and habits of different user groups are numerous, and they are changing rapidly, as the culture of multimedia continues to grow and permeate nearly every aspect of our personal and professional lives.

1.5 APPLICATIONS OF MULTIMEDIA

Multimedia has found extensive applications in various and varied fields. The following are some of the main areas where this technology is applied:

- Home entertainment
- Educational purposes
- Industrial training
- Information kiosk
- Corporate presentations
- Business
- Tourism and Travel industry
- E – Shopping
- Communication and networks
- Medicine
- Engineering Application
- Content based storage and retrieval [CBSR] systems.

Home Entertainment:

Application of Multimedia technology related to home entertainment includes computer based games for kids, interactive encyclopedia’s, storytelling, cartoons etc., Computer games are one of the best application of Multimedia because of the high amount of interactivity involved.

Educational purposes:

These applications include learning packages and simulation of lab experiments [especially those which cannot be easily performed]. <u>The multisensory perceptions of such study material are expected to provide a good grasp of the subject matter and interactivity elements to provide for better retention</u>.

Industrial Training:

These applications involve Computer Aided Instruction [CAI] and Computer Based Training [CBT] for employee both technical and marketing. Successful organizations are required to maintain a high level of staff training and development.

Information kiosk:

These are devices where information is accessed through a touch screen and viewed on a monitor. Examples may include multi-lingual product catalogues for placing orders or for dispensing important information. Kiosks can also be used to capture statistical data for an in-depth marketing research to be carried out on customer trends.

Corporate presentations:

Corporate presentations are emphasizing the salient features and activities of a company, its products, business partners like suppliers and retailers can be built by incorporate multimedia elements along with textual descriptions.

Business:

Items like glass utensils are difficult to stock; industrial equipment can be displayed through perspectives buyers by company sales people through multimedia presentations.

Tourism and Travel industries:

Travel companies can market packaged tools by showing prospective customers, glimpses of the places they would like to visit, details on lodging, special attractions. A multimedia system implementing intelligent travel agent software will enable the user to their travel need and budget.

E-shopping:

Like the travel industry, customized presentations for consumer and industrial products can be created and distributed to prospective customers. Customers can compare different products in relation to their quality, price, and appearances without leaving their homes and offices.

1.6 HARDWARE AND SOFTWARE REQUIREMENTS

Hardware and software requirements of a multimedia personal computer can be classified into 2 classes.

They are:
a. Multimedia playback [usually requires lesser amount of resources for viewing an existing presentation]
b. Multimedia production [generally requires greater and more powerful resources fulfilling designing requirements]

Multimedia playback:

- **Processor –** At least Pentium class and minimum of 512 MB RAM-to-4GB RAM.

- **Hard disk drive [HDD] –** At least 40GB having 15M/s. access time.
- **The monitor and video display adapter** should confirm through SVGA standards and support 800x600 display modes with true color.

- **CD-ROM drives** having a speed of at least 4X but highest speed like 36X are recommended.

- PC should have a **sound card** with attached speakers standard 101 keys keyboard and mouse.

- Multimedia PC system software should be compatible with Windows XP or higher, with standard software with playback of media files in standard formats. E.G. Windows Media Player, Apple QuickTime Player, etc.

Multimedia production:

- **Processor - Pentium** IV or higher, memory should be at least 2GB RAM-to-8GB RAM.

- **Hard disk drive [HDD] –** Typical requirements would be around 80GB with 540GB recommended.

- **The monitor and video display adapter** should confirm through SVGA standards and should be able to support 800x600 display mode with true color, RAM should be 4MB to 8MB.

- **CD-ROM drive** having a speed of at least 4X to 36X, PC should have a CD/DVD writer.

- PC should have a **sound card** with attached speakers standard 101 keys keyboard and mouse.

- Multimedia PC system software should be compatible with windows or higher, with standard software with playback of media files in standard formats. E.G. Windows Media Player, Apple QuickTime Player, etc.

- **Editing software** is used to manipulate media components to suit the developers, requirements. [E.G.] Adobe Photoshop, Flash, Cool Edit, and sound Forge.

- **Authoring softwares** are used to integrate all the edited media into single presentations and build navigational pathways for accessing the media.

- To display the web content **web browsers** will be required. [E.G.] Mozilla Firefox, MS Internet Explorer, to create web content HTML, and JavaScript editors might be required [E.G.] Macromedia, Dream Weaver, etc..

1.7 SUMMARY

- The word "multimedia" comes from the Latin words *multus* which means "numerous" and *media* which means "middle" or centre. Multimedia therefore means "multiple intermediaries" or "multiple means".

- Characteristics of multimedia are: Multiple media, Non-linearity, Scope of Interactivity, Integrity, Digital representation

- The multisensory perceptions of such study material are expected to provide a good grasp of the subject matter and interactivity elements to provide for better retention.

1.8 UNIT END EXERCISES

1. What is meant by a multimedia presentation? Describe some of its important characteristics.
2. What distinguishes a multimedia presentation from a video clip or a movie show? In this context, explain the concept of 'non-linearity'.
3. Distinguish between multimedia 'production' and 'playback'. How would hardware and software requirements vary in each case?
4. Mention some of the uses of multimedia presentations giving examples.
5. Describe some of the important criteria required for promoting multimedia technology.
6. Clearly distinguish between a story, a script and a storyboard, stating their purposes.
7. What is the role of testing and feedback stage in improving the quality of the presentation?

Reference:

1. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

2. Wikipedia:http://en.wikipedia.org/wiki/Interactive_media

❖❖❖❖

2

INTRODUCTION TO MULTIMEDIA

Unit Structure

2.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand how a multimedia presentation is promoted.
2. Understand how a multimedia presentation is created.

2.2 INTRODUCTION

World Wide Web is the name given to the vast system of interconnected servers used in the transmission of digital documents formatted using Hypertext Markup Language (HTML) via the Internet.

Timothy Berners-Lee, the inventor of World Wide Web, coined the term *Semantic Web* to describe Web 3.0, which many see as the next significant iteration of the Web. The Semantic Web is defined as “a web of data that can be processed directly and indirectly by machines.” Web 3.0 will likely involve the creation of new standards and protocols for unifying the way that content is coded, organized, and analysed by computers monitoring the Web. This may involve transforming the Web into a massive unified database akin to a virtual library.

Such a vast library will require more sophisticated computers and search engines to categorize and make sense of its holdings. As computers become “smarter” and the engineering behind the Web grows more sophisticated, many believe that we will see a significant increase in computer generated content.

Web 3.0 signals, in part, a larger role for computers as creators of intellectual content. Is a day coming when you will no longer be able to distinguish between content written by humans and machines? Has it already arrived?

2.3 MULTIMEDIA DATABASE

Automated retrieval of syntactically and semantically useful contents from multimedia databases is crucial, especially when the contents have become so rich and the size ofthe databases has grown so rapidly. This unit looks at a particular promotion of multimedia database systems, examining the variables involved in content-based multimedia.

A Multimedia Database (MMDB) hosts one or more multimedia data types (i.e. text, images, graphic objects, audio, video, animation sequences.

These data types are broadly categorized into three classes:

- Static media (time-independent: image and graphic object).
- Dynamic media (time-dependent: audio, video and animation).
- Dimensional media (3D game and computer aided drafting programs).

A Multimedia Database Management System (MMDBMS) is a framework that manages different types of data potentially represented in a wide diversity of formats on a wide array of media sources. It provides support for multimedia data types, and facilitate for creation, storage, access, query and control of a multimedia database.

Like the traditional databases, Multimedia databases should address the following requirements:

- **Integration**
  Data items do not need to be duplicated for different programs invocations

- **Data independence**
  Separate the database and the management from the application programs

- **Concurrency control**
  Allows concurrent transactions

- **Persistence**
  Data objects can be saved and re-used by different transactions and program invocations

- **Privacy**
  Access and authorization control

- **Integrity control**
  Ensures database consistency between transactions

- **Recovery**
  Failures of transactions should not affect the persistent data storage

- **Query support**
  Allows easy querying of multimedia data

Multimedia databases should have the ability to uniformly query data (media data, textual data) represented in different formats and have the ability to simultaneously query different media sources and conduct classical database operations across them. (Query support)

They should have the ability to retrieve media objects from a local storage device in a good manner. (Storage support)

They should have the ability to take the response generated by a queryand develop a presentation of that response in terms of audio-visualmedia and have the ability to deliver this presentation. (Presentation and delivery support)

2.3.1 PROMOTION OF MULTIMEDIA BASED CONTENT:

For promotion of multimedia based applications and presentations, so as to capture an appreciable sector of the IT marketplace, the following should be present:

- Demand from Customer
- Compression Techniques
- Processing power
- Standards
- Bandwidth
- Distribution mechanism

Demand from customers:

If the customer is to invest additional amount is acquiring multimedia based content, then the customer will see how much they can fulfil their requirements. There should be appreciable value addition factor provided by the multimedia application, something that the normal print media cannot fulfil.

Compression Techniques:

Non-Textual digital media, usually occupy a lot of disk space. While several text pages of a document occupy few kilobytes [kb], full screen images have sizes in the range of few megabytes, audio content occupy tens' of megabytes while video content can span Gigabytes. It is difficult to manipulate these large files in their original form, so they need to be compressed to reduce their sizes. Efficient compression techniques should be present to make this possible.

Compression techniques can be classified as either lossless or lossy. If the compression and decompression processes induce no information loss, the compression scheme is lossless; otherwise, it is lossy. We will study various image, video, and audio compression techniques in later units.

Processing power:

Dynamic media like audio, video, and animation require the central processor and screen display system to process a large quantity of data, every second. 3D animation requires a large number of floating point calculation to be performed every second.

Synchronization markersareeasy to recognize and are particularly well suited to devices with limited processing power, such as cellphones and mobile devices.

Standards:

The need for standard cannot be overemphasized that standard guarantee the interoperability. These include both hardware and software standards like buses, cables, connectors, signals, file formats, data transmission protocols, compression and de-compression techniques.

Most companies concentrate on developing products using standards, in the interest of interoperability of mobiles and networks. The standard has evolved to allow integration with script languages, dynamic manipulation of almost all elements and properties after display on the client side.

Bandwidth:

The term bandwidth, in analog devices, refers to the part of the response or transfer function of a device that is approximately constant, or flat, with the x-axis being the frequency and the y-axis equal to the transfer function. So for analog devices, the bandwidth is expressed in frequency units, called Hertz (Hz), which is cycles per second. For digital devices, on the other hand, the amount of data that can be transmitted in a fixed bandwidthis usually expressedinbits per second (bps) orbytes per amount of time.

A powerful processing machine and large files sizes also means that a large amount of data need to transferred at high speed between devices or components. This requires high bandwidth and data rates between internal and external components of a system.

Electronic delivery depends on network bandwidth at the user side. A streaming option may be available, depending on the presentation.

Distribution mechanism:

After creation of multimedia content need to distribute effortlessly to the customer. Portable storage media supporting high storage volumes are essential for this distributed mechanism. A majority of the multimedia content are distributed via CD-ROM, which additionally provide high durability of the content.

2.3.2 STEPS FOR CREATING A MULTIMEDIA PRESENTATION:

Choosing a Topic ➔Writing a Story ➔Writing a Script ➔ Preparing a Storyboard➔Preparing a flow line ➔Implementation ➔ Testing and Feedback ➔Final Delivery

Choosing a Topic:

The first topic/task is to choose a topic on which to create the presentation. In principle, one can select any topic/topics which can be explained or demonstrated using various media types are more conductive to multimedia presentation. Use of text should be kept at a minimum. For example, not more than a few lines per page of the presentation when choosing a topic one should make a mental note of how the subject matter should be divided and what entry points should give access to which module. The author shall also decide who should be the target audience further deciding the objectives of the presentation [i.e.] what the audience is expected to learn after going through presentation.

Writing a story:

A story describes the overall content of the presentation. The writing style should be textual and should resemble an essay. Within the story, the author can divide the matter into logical divisions like chapters, sections, topics, etc. for better readability and modularization.

Writing a script:

Once the overall subject matter has been finalized, the next step is to create a script. A script emphasizes how the subject matter unfolds. While writing a script, the author visualizes the content in terms of frames. For example, what is to be displayed on the first screen? This requires the subject matter of the story be divided into small modules one for each screen. The script may also include other accessory information like how the elements are displayed on the screen.

Preparing a Storyboard:

Once the script has been prepared, the author needs to prepare the storyboard. The storyboard depicts what should be the layer of each screen within the presentation. The screen should have an aesthetic feel about them and should be pleasant to look. These are like "key frames" in a video - the story hangs from these "stopping places".

Preparing a flow line:

Along-side a storyboard, the author should also prepare a flow line. A flow line at a glance tells us how the user can access different pages of the presentation.

A flowchart organizes the storyboards by inserting navigation information - the multimedia concept's structure and user interaction. The most reliable approach for planning navigation is to pick a traditional data structure. A hierarchical system is perhaps one of the simplest organizational strategies.

Multimedia is not really like other presentations, in that careful thought must be given to organization of movement between the "rooms" in the production. For example, suppose we are navigating an African safari, but we also need to bring specimens back to our museum for close examination – just how do we effect the transition from one locale to the other?

Implementation:

Implementation means actually creating the physical presentation using required hardware and software. Implementation has a number of sub steps. The first step is the collection of media items. The author can use software to create their own items.

There are two types of implementation software.

[i] The first type is the editing software, which are used to edit the digitized items.

[ii] The second type is the authoring software; which are used to integrate all the editor media into a single presentation. The output of the authoring software is usually an executable file [exe] which contains its own runtime engine and therefore can be played without the help of any other software.

Testing and feedback:

After the implementation phase is completed, an important step of testing and feedback should be done for improving the quality of the presentation. This step involves distributing whole [or] part of the presentation to sections of the target audience and heading the feedback from them about the possible areas which need improvement. Developers always work under various constraints and do not have indefinite time on their hands.

Some multimedia designers use an authoring tool at this stage already, even if the intermediate prototype will not be used in the final product or continued in another tool. User testing is, of course, extremely important before the final development phase.

Final delivery:

The final phase in the production schedule is the delivery of the application to be intended client. Usually the runtime version of the application files are copied into a CD-ROM / flash drive and physically handed over to the customer. It is also important for the author to state clearly the hardware and software requirements which should be present on the client machine to run the application smoothly.

2.4 SUMMARY

- For promotion of multimedia based applications and presentations, so as to capture an appreciable sector of the IT marketplace, the following should be present: Demand from Customer, Compression Techniques, Processing power, Standards, Bandwidth, Distribution mechanism.

- Steps for creating a multimedia presentation include: Choosing a Topic ➔ Writing a Story ➔ Writing a Script ➔ Preparing a Storyboard ➔ Preparing a flow line ➔ Implementation ➔Testing and Feedback ➔ Final Delivery

2.5 UNIT END EXERCISES

1. Describe some of the important criteria required for promoting multimedia technology.
2. Clearly distinguish between a story, a script and a storyboard, stating their purposes.
3. What are the requirements that Multimedia databases should address?
4. What is the role of testing and feedback stage in improving the quality of the presentation?

Reference:

1. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

2. Wikipedia : http://en.wikipedia.org/wiki/Multimedia_database

❖❖❖❖

3

UNIT II

DIGITAL REPRESENTATION

Unit Structure

3.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define and distinguish analog and digital waves
2. Convert analog to digital
3. Convert Digital to analog
4. Describe fundamental properties of waves

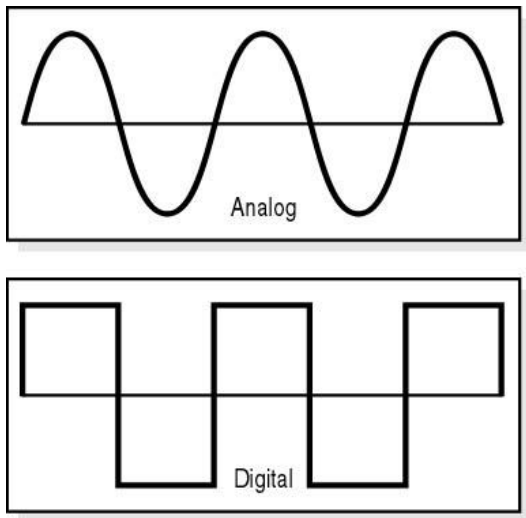
3.2 INTRODUCTION

Vibrations in air create waves of pressure that are perceived as sound. Multimedia system sound is digitally recorded audio or MIDI (Musical Instrumental Digital Interface) music. Digital audio data is the actual representation of a sound, stored in the form of samples. Digital video has replaced analog as the method of choice for making and delivering video for multimedia. Digital video device produces excellent finished products at a fraction of the cost of analog. Digital video editing and playback require a large amount of free disk space. When working with digital video, it is necessary to defragment and optimize the disk before recording and playing back movie files. It is necessary to compress movie files before delivery. Special hardware (video capture boards) are often needed to convert analog signal (camera) to digital (computer) – more recently digital cameras and camcorders have become available

(Adobe Premiere) allows you to edit video clips assembled from camera videotape, scanned images, digitized audio or MIDI files.

3.3 ANALOG REPRESENTATION

An analog or analogue signal is any continuous signal for which the time varying feature (variable) of the signal is a representation of some other time varying quantity, i.e., analogous to another time varying signal. For example, in an analog audio signal, the instantaneous voltage of the signal varies continuously with the pressure of the sound waves.



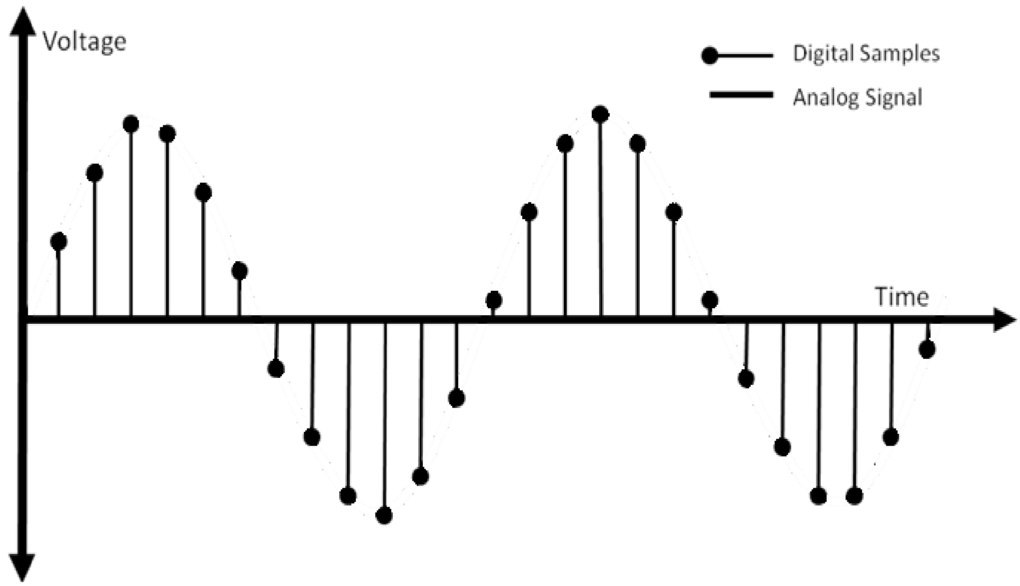
Analog v/s Digital Representation of Wave

An analog signal uses some property of the medium to convey the signal's information. In an electrical signal, the voltage, current, or frequency of the signal may be varied to represent the information. Any information may be conveyed by an analog signal; often such a signal is a measured response to changes in physical phenomena, such as sound, light, temperature, position, or pressure.

Analog signals have two essential properties:

1. The signal delivered by the capturing instrument may take any possible value within the limits of the instrument (sensors). Thus, the value can be expressed by any real number in the available range. Analog signals are thus said to be amplitude continuous.

2. The value of the analog signal can be determined for any possible value of time or space variable. Analog signals are therefore also said to be time continuous or space continuous.



An Analog Signal

3.4 WAVES

Waves can be conceived of energy propagating from one place to another. A wave represents a graph or plot of the motion of a particle (or a set of particles) in the path of a wave over space or time. Three fundamental properties can be identified in waves that determine the characteristics of the physical quantity that the wave represents. These are: amplitude, frequency and waveform.

3.4.1 FUNDAMENTAL PROPERTIES OF WAVE:-

PEAK AMPLITUDE

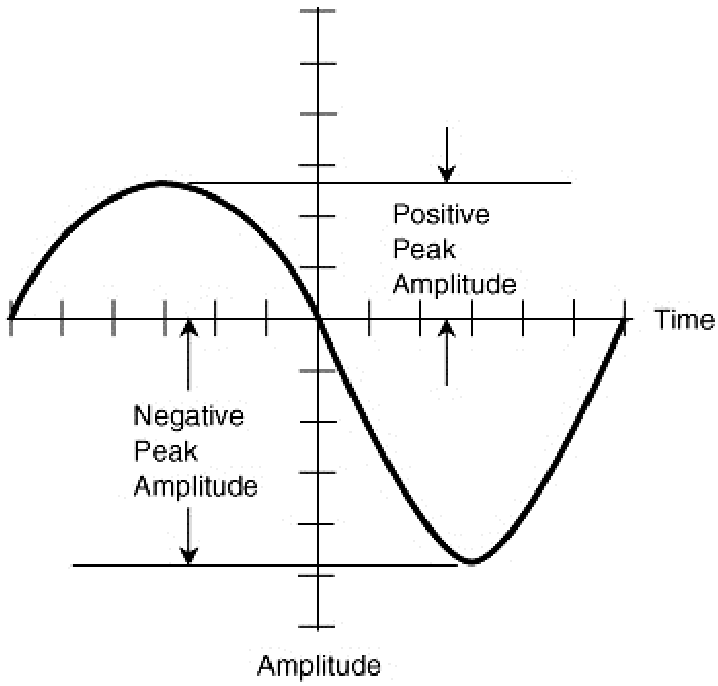
It is the maximum displacement of the oscillating particle from its mean position. The amplitude actually represents the intensity of the physical quality. For e.g.:-Brightness of light, loudness of sound, voltage or current level of electricity etc..

In audio system measurements, telecommunications and other areas where the measurand is a signal that swings above and below a zero value but is not sinusoidal, peak amplitude is often used. This is the maximum absolute value of the signal.

Peak-to-peak amplitude is the change between peak (highest amplitude value) and trough (lowest amplitude value, which can be negative). With appropriate circuitry, peak-to-peak amplitudes of electric oscillations can be measured by meters or by viewing the waveform on an oscilloscope.

Peak-to-peak is a straightforward measurement on an oscilloscope, the peaks of the waveform being easily identified and measured against the graticule. This remains a common way of specifying amplitude, but sometimes other measures of amplitude are more appropriate.

A wave is a pictorial representation for a physical quantity, therefore changing a fundamental property of a wave usually brings a change in some physical property of the quantity.



Representation of Wave

FREQUENCY

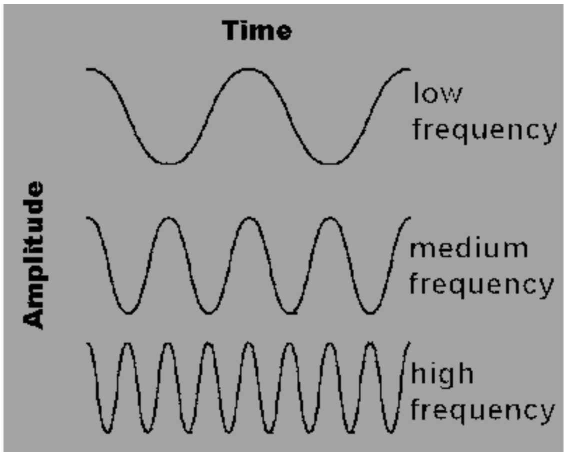
It refers to how fast a particle in the path of a wave is oscillating and is defined as the number of oscillations completed by the particle in unit time. Frequency is measured in a unit called Hertz represented as Hz. It is also referred to as temporal frequency, which emphasizes the contrast to spatial frequency and angular frequency.

A wave is said to have a frequency of 100 Hz when it completes oscillations (or complete cycles) in one second. The time period T of a wave is the time required to complete 1 cycle of the wave and hence is the reciprocal of the frequency f, i.e. $f=1/T$.

The frequencies an ear can hear are limited to a specific range of frequencies. The audible frequency range for humans is typically given as being between about 20 Hz and 20,000 Hz (20 kHz), though the high frequency limit usually reduces with age.

Other species have different hearing ranges. For example, some dog breeds can perceive vibrations up to 60,000 Hz.

In many media, such as air, the speed of sound is approximately independent of frequency, so the wavelength of the sound waves (distance between repetitions) is approximately inversely proportional to frequency.



Different frequencies of Wave

WAVEFORM

It represents the shape of a wave and signifies the change of the amplitude and frequency over time. The sinusoidal shape represents the elementary wave but waves can have any shape like triangular, square, etc..

The waveform determines the quality of the physical quantity, e.g. timbre of sound or texture of light.

An instrument called an oscilloscope can be used to pictorially represent a wave as a repeating image on a screen. By extension, the term "waveform" also describes the shape of the graph of any varying quantity against time.

Common periodic waveforms include (t is time):

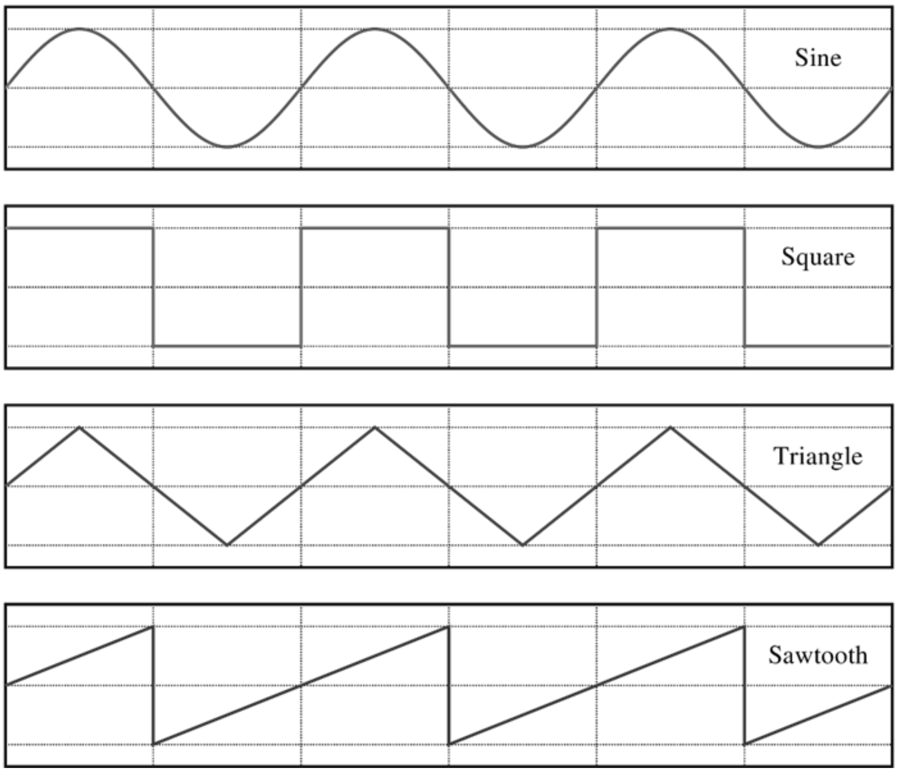
- **Sine wave**: sin (2 π t). The amplitude of the waveform follows atrigonometric sine function with respect to time.

- **Square wave**: This waveform is commonly used to represent digital information. A square wave of constant period contains odd harmonics that fall off at −6 dB/octave.

- **Triangle wave**: It contains odd harmonics that fall off at −12 dB/octave.

- **Sawtooth wave**: This looks like the teeth of a saw. Found often in time bases for display scanning. It is used as the starting point for subtractive synthesis, as a saw tooth wave of constant period contains odd and even harmonics that fall off at −6dB/octave.

Other waveforms are often called composite waveforms and can often be described as a combination of a number of sinusoidal waves or other basis functions added together.



Different Waveforms

3.5 DIGITAL REPRESENTATION

Digital data, in information theory and information systems, are discrete, discontinuous representations of information or works, as contrasted with continuous, or analog signals which behave in a continuous manner, or represent information using a continuous function.

Although digital representations are discrete mathematics, the information represented can be either discrete, such as numbers and letters or continuous, such as sounds, images, and other measurements.

The word digital comes from the same source as the words digit and *digitus* (the Latin word for finger), as fingers are often used for discrete counting. Mathematician George

Stibitz of Bell Telephone Laboratories used the word digital in reference to the fast electric pulses emitted by a device designed to aim and fire anti-aircraft guns in 1942.It is most commonly used in computing and electronics, especially where real-world information is converted to binary numeric form as in digital audio and digital photography.

We therefore do not represent digital quantities as equations but rather as sets of values to numbers along the time instants when the values are occurring. E.g. $\{x_1{:}t_1, x_2{:}t_2, \ldots x_n{:}t_n\}$. The state of the digital quantity between these points is unknown and therefore cannot be represented graphically.

For example, the state of an electric switch can be represented by two numbers: 1 when it is in the ON position and 0 when it is in the OFF position. In such a case, the number 1.5 does not carry any significance as it does not represent a valid state of the switch. Even if the switch is half-closed, either a contact exists between the opposite poles of the switch in which it behaves as ON (1) or there is not contact in which it behaves as OFF (0).

3.6 NEED FOR DIGITAL REPRESENTATION

The fundamental property of digital representations is that they are based on the use of a finite number of discrete symbols to represent information. Because finite systems can represent only a finite number of symbols, in any such system there is only a finite number of possible meaningful states; this is a fundamental difference between digital and analog representations of information. (In practice, the number of states distinguishable in a digital system is large enough that it often simplifies reasoning to pretend that it is infinite.)

In digital systems, the physical similarity of two representations of information is no guide to the similarity of the information they represent. (For example, the bit sequences 0000 0000 and 1000 0000 differ only in the value of a single bit, but if they are taken as unsigned integers, they denote 0 and 128; many numbers much closer to zero than 128 have representations very different from either.) Small errors occurring in the physical representation of information (e.g. the accidental flipping of a single bit) can and often do lead to wildly erratic results.

Typically, digital representations provide better access to information of interest than do analog representations. Full-text search of digital representations is a straightforward operation, while full-text search of images representing pages of text is possibly only via a detour through a textual representation (often created automatically by optical character recognition, with the high error rates entailed by that operation). Because purely digital representations can omit extraneous information, they tend to be more compact than analog representations.

An image of a page typically requires many times the storage space needed for a character-based transcription of the page. Conversely, because analog representations do not discriminate between relevant and extraneous information, they will typically convey information omitted from a purely digital representation of the same thing. Sometimes this extra information will prove useful or important.

3.7 SUMMARY

- An analog or analogue signal is any continuous signal for which the time varying feature (variable) of the signal is a representation of some other time varying quantity.
- A wave represents a graph or plot of the motion of a particle (or a set of particles) in the path of a wave over space or time.
- Three fundamental properties can be identified in waves that determine the characteristics of the physical quantity that the wave represents. These are: amplitude, frequency and waveform.
- The amplitude actually represents the intensity of the physical quality.
- Frequency refers to how fast a particle in the path of a wave is oscillating and is defined as the number of oscillations completed by the particle in unit time.
- Waveform represents the shape of a wave and signifies the change of the amplitude and frequency over time.

3.8 UNIT END EXERCISES

1. Write are analog signals and their essential properties?
2. How can analog signals be represented as waves?
3. Explain the fundamental properties of waves.
4. Why is there a need to represent data digitally?

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.
II. Wikipedia

❖❖❖❖

4

SIGNAL CONVERSIONS

Unit Structure

4.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Convert analog signal to digital signal
2. Convert Digital signal to analog signal
3. Understand Nyquist Theorem.

4.2 INTRODUCTION

Some multimedia formats are restricted to particular hardware/operating system platforms, while others are platform independent, or cross-platform, formats. Even if some formats are not cross-platform, conversion applications can recognize and translate formats from one system to another thus generating the need to digitize.

Digitization means conversion to a stream of numbers - preferably integers for efficiency.

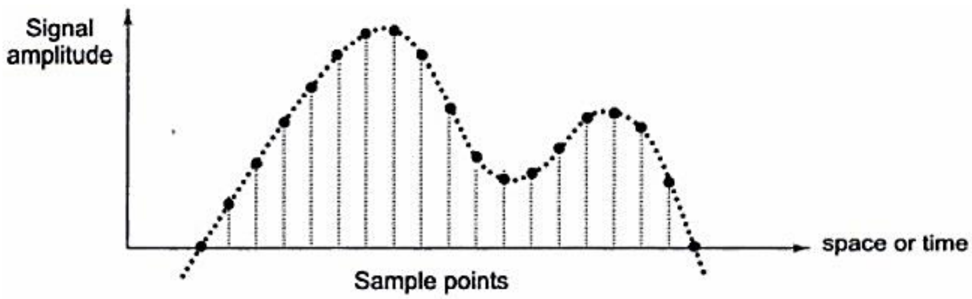
4.3 A TO D CONVERSION

4.3.1 Sampling

The first step in converting an analog quantity to digital form is called sampling. In signal processing, sampling is the reduction of a continuous signal to a discrete signal. A common example is the conversion of a sound wave (a continuous signal) to a sequence of samples (a discrete-time signal).

A sample refers to a value or set of values at a point in time and/or space. A sampler is a subsystem or operation that extracts samples from a continuous signal. A theoretical ideal sampler produces samples equivalent to the instantaneous value of the continuous signal at the desired points.

The wave can depict the changing loudness of sound from a musical instrument, or the changing brightness of light reflected from a photograph. For quantities like sound, sampling is done at specific intervals of time (eg.10 times per second) and is said to create time-discretization of the signal. For quantities like a static image, sampling is done at regular space intervals (along the length and breadth of the image) and is said to create space-discretization of the signal (e.g. 10 times per inch).

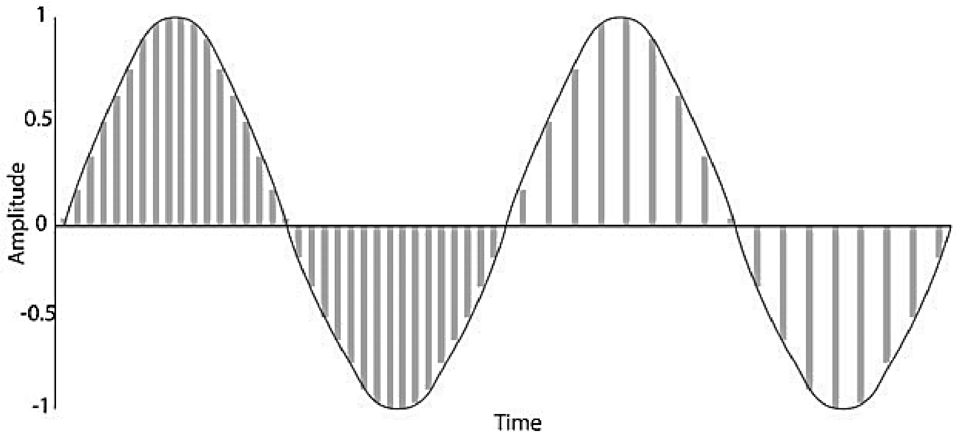


Sampling Rate

The rate or frequency at which the sampling occurs is called sampling frequency, and denotes the number of samples taken per second or per inch. This determines the quality of the output signal, higher the sampling rate more is the number of samples of the original wave that is obtained and better is the representation after digitization.

The digitized version will always be degraded version of the original analog wave, because no matter how large the sample rate is, we are bound to discard and lose some data in between two samples values. Only for an infinite sampling rate we get an exact representation of the analog quantity, which is practically impossible.

For audio, typical sampling rates are from 8 kHz (8,000 samples per second) to 48 kHz. The human ear can hear from about 20Hz (a very deep rumble) to as much as 20 kHz; above this level, we enter the range of ultra sound. The human voice can reach approximately 4 kHz and we need to bind our sampling rate from below by at least double this frequency. Thus we arrive at the useful range about 8 to 40 or so kHz.

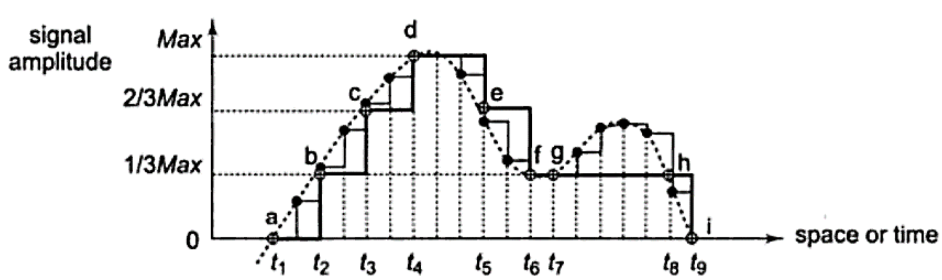


Higher and Lower Sampling Frequencies

4.3.2 Quantization

The amplitude values obtained after sampling may be long real numbers which are usually rounded to the nearest predefined discrete values. This process of converting the real numbers to the predefined discrete numbers is called quantisation. Quantization corresponds to a discretization of the intensity values. That is, of the co-domain of the function.

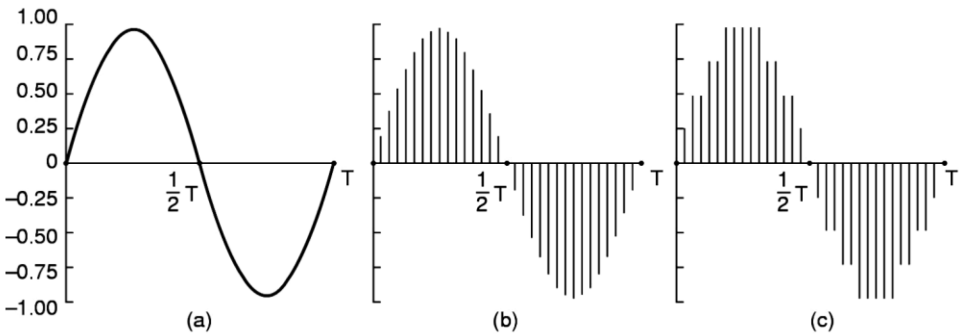
In other words, after the sampling we have a sequence of numbers which can theoretically still take on any value on a continuous range of values. Because this range in continuous, there are infinitely many possible values for each number, in fact even unaccountably infinitely many. In order to be able to represent each number from such a continuous range, we would need an infinite number of digits - something we don't have. Instead, we must represent our numbers with a finite number of digits, that is: after discretizing the time-variable, we now have to discretize the amplitude-variable as well. This discretization of the amplitude values is called quantization.



Consider the above diagram. The intersection points **a, b, c, d, e, f, g, h, i** represent the new sample values occurring at instant t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9 respectively, in time (or space). Due to the amplitude discretization process, some of the existing sample values are approximated to slightly changed values (a,b) while some other points remain almost at their earlier values (d, f). Joining these points together and remembering to continue the value of an existing sample value until the next sample value is obtained, we get the 4-level quantized output curve, shown in the black bold line.

It would be obvious that larger the number of quantization levels, the better would be the approximation of the output curve to the original analog wave.

<u>Practical Scenario</u>: Audio Encoding



(a) a sine wave
(b) sampling the sine wave
(c) quantizing the samples to 4 bits

4.3.3 Code-word Generation

The output of quantization stage contains set of sample values quantized to a specific number of levels, occurring at specific instants of time (or space). The amplitude levels of the samples are expressed as fractions of the difference between the minimum and maximum values. Code-word generation involves expressing these amplitude levels in terms of binary codes or numbers.

BIT DEPTH

In digital audio, bit depth is the number of bits of information in each sample, and it directly corresponds to the resolution of each sample. A string of bits can be divided into 8-bit words. A 16-bit sample is two words. A 24-bit sample is three words. The professional sample rate of 48 kHz is slightly higher than that used for CD quality.

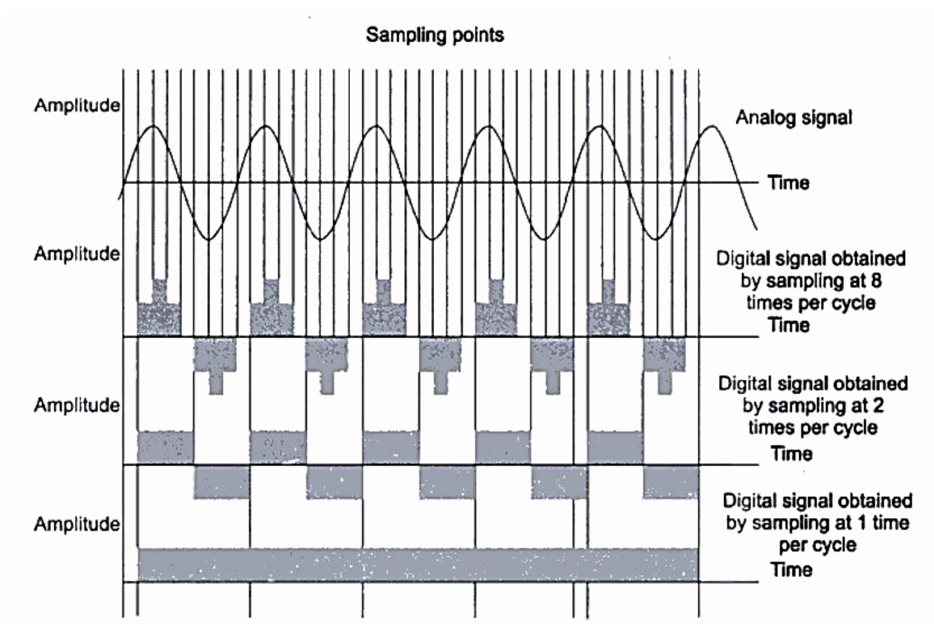
Doubling the sample rate adds one musical octave to the audio bandwidth. This is the major reason for 96 kHz and 192 kHz sample rates, the latter producing an audio bandwidth approaching 100 kHz. Increasing the sample rate also reduces the minimum delay step (1 sample) available in digital signal processors, which is an additional motivation for increasing the sample rate beyond what is needed to reproduce the audible spectrum.

A bit depth of 24 bits and a 48 kHz sample rate is commonly presented as 24/48k. This is the default resolution for most professional ADCs. This yields a theoretical dynamic range of approximately 224, or 144dB. This is a mathematical resolution. In practice the actual resolution will be far less. A respectable system dynamic range, resulting from stringing together a chain of digital audio devices, is about 100 dB—roughly CD quality.

NYQUIST's SAMPLING THEOREM

The origin of the sampling theorem is the American engineer Harry Nyquist. According to the theorem:

"When converting an analog signal into digital form the sampling frequency must be greater than twice the bandwidth of the input signal to be able to reconstruct the original signal accurately from the sampled version."



The low frequency distortion produced due to lowest sampling rate, known as under-sampling, is referred to as aliasing, because the input wave is not truly represented at the output.

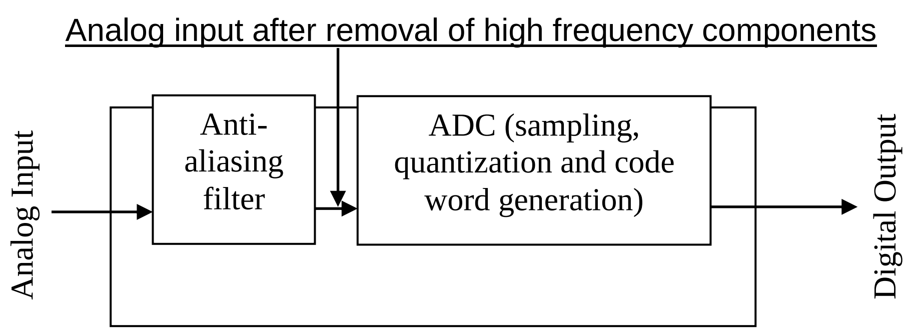
The sampling frequency that is half of the input wave frequency, is referred as Nyquist frequency. When sampling is done at much higher rate, it is called over-sampling. Both under and over sampling, in general, should be avoided.

ANALOG TO DIGITAL CONVERTER: ENCODER

An analog-to-digital converter (ADC, A/D, or A to D) is a device that converts a continuous physical quantity (usually voltage) to a digital number that represents the quantity's amplitude.

The conversion involves quantization of the input, so it necessarily introduces a small amount of error. Instead of doing a single conversion, an ADC often performs the conversions ("samples" the input) periodically. The result is a sequence of digital values that have been converted from a continuous-time and continuous-amplitude analog signal to a discrete-time and discrete-amplitude digital signal.

ELECTRICAL SYMBOL OF Analog to Digital Converter



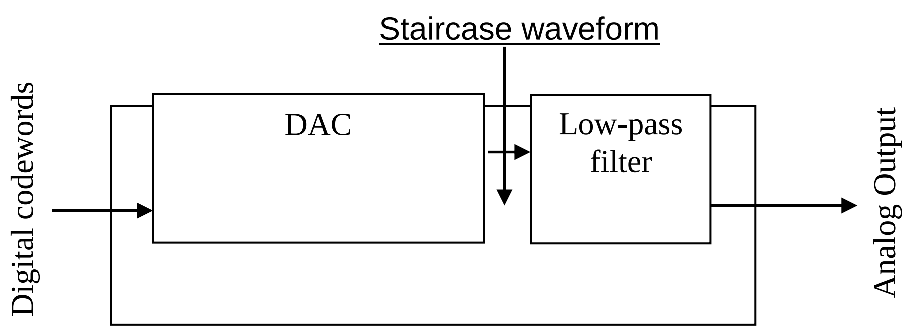
ENCODER: Analog to Digital Conversion

4.4 D TO A CONVERSION

So far we have only investigated generation of digitally represented signals that are sampled in time and quantized in amplitude to a specific number of bits *b*. Accordingly, in an ideal convertor there are 2*b*equally spaced quantisation levels that follow an exact linear relationship with the input code. We call this process digital to analogue conversion, and it comprises several distinct sequentially connected processing such as:

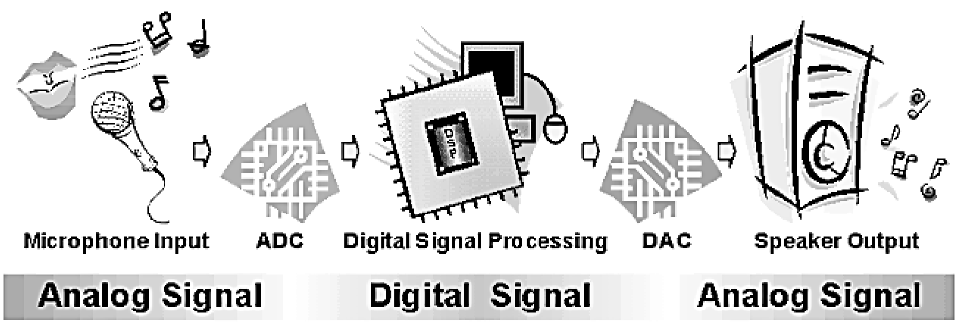
- the digital to analogue convertor (DAC);
- a typically low-pass filter;
- analogue post-processing

In effect, a DAC multiplies the reference voltage by a factor determined from the input word as a fraction of full-scale. Accordingly, the often overlooked DAC reference is a critical design consideration and for our present discussion we subsume it into the DAC function. There are many DAC architectures reported in the literature and available in 'single chip form' from commercial suppliers such as Analog Devices, Linear Technology and Texas Instruments.



DECODER: Digital to Analog Conversion

ADC and DAC are generally combined as a single unit called an encoder-decoder or ADC-DAC set. These are typically implemented in the form of a single or a group of integrated circuits on a printed circuit board, e.g. sound card.



ENCODING DECODING CYCLE

4.5 SUMMARY

- The first step in converting an analog quantity to digital form is called sampling.
- The rate or frequency at which the sampling occurs is called sampling frequency, and denotes the number of samples taken per seconds or per inch.

- Quantization corresponds to a discretization of the intensity values. That is, of the co-domain of the function.

- Code-word generation involves expressing these amplitude levels in terms of binary codes or numbers.

- In digital audio, bit depth is the number of bits of information in each sample, and it directly corresponds to the resolution of each sample.

- According to Nyquist theorem: “When converting an analog signal into digital form the sampling frequency must be greater than twice the bandwidth of the input signal to be able to reconstruct the original signal accurately from the sampled version.”

- ADC and DAC are generally combined as a single unit called an encoder-decoder or ADC-DAC set.

- In effect, a DAC multiplies the reference voltage by a factor determined from the input word as a fraction of full-scale.

- The low frequency distortion produced due to lowest sampling rate is known as under-sampling.
- When sampling is done at much higher rate, it is called over-sampling.

4.6 UNIT END EXERCISES

1. Define Sampling. What is meant by sampling rate?
2. Why is quantization necessary? Distinguish between analog and digital waves.
3. Write a note on Nyquist Theorem.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.
II. Wikipedia

❖❖❖❖

5

CONVERSIONS, RELATION AND REPRESENTATION

Unit Structure

5.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define and distinguish Sampling Rate and Bit Depth
2. Understand Quantization error
3. Represent signals in Fourier Series
4. Understand importance and Drawback of Digital Representation

5.2 INTRODUCTION

Although discussions of digital audio conversion have filled several books, a fundamental understanding of two terms is particularly important to correctly using your computer-based recording system: sample rate and bit depth.

The conversion process is complex, and there are multiple ways to accomplish it. But no worries: We’re just going to discuss sample rate and bit depth at a basic level, as applied to linear pulse-code modulation (PCM), one of the most common conversion technologies.

Finally, one often encounters the term “high-resolution audio” but it is rarely defined. That’s because there is no agreed-upon definition. For many years, “resolution” referred to bit depth, but in recent years, the term has been used more broadly to refer to both sample rate and bit depth. And “high resolution,” in particular, is a relative term. When 8-bit audio was in common use, 16-bit was “high resolution.” Today, 24-bit, 96 kHz audio is considered “high resolution.” In the future, it might be 32-bit, 192 kHz and beyond.

5.3 RELATION BETWEEN SAMPLING RATE AND BIT-DEPTH

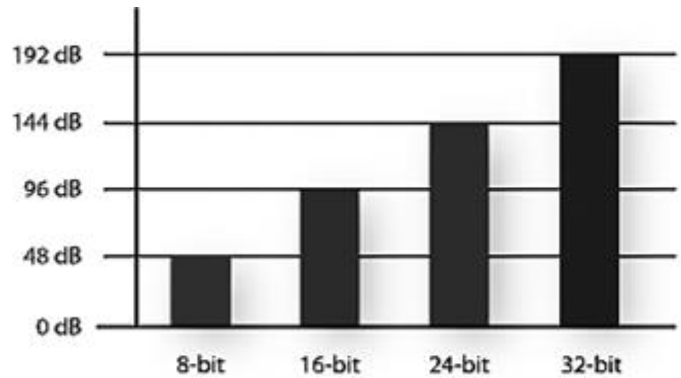
Two factors affect the potential quality of a digital audio system: its sample rate and bit depth. The sample rate defines how often the input voltage is measured, or sampled, with faster rates allowing higher frequencies to be captured. The bit depth refers to the size of the numbers used to store the digitized data, with larger numbers giving a lower noise floor. The higher the sampling rate of a digital file (measured in samples per second), the more frequencies that file can reproduce. The frequency of a sound determines its pitch; low frequencies produce low-pitch sounds, and high frequencies produce high-pitch sounds. The human ear can hear frequencies ranging from 20 Hz to about 20 kHz.

According to Nyquist's theorem, which says you have to be able to take two complete samples per second to reconstruct an accurate waveform, for a particular frequency to be captured in a digital file, the sampling rate must be at least double that frequency's value. So a recording containing a peak frequency of 12,000 Hz, for example, would require a sampling rate of 24,000 samples per second.

Bit depth	Quality level	Amplitude values	Dynamic range
8-bit	Telephony	256	48 dB
16-bit	CD	65,536	96 dB
24-bit	DVD	16,777,216	144 dB
32-bit	Best	4,294,967,296	192 dB

Comparing Bit Depths

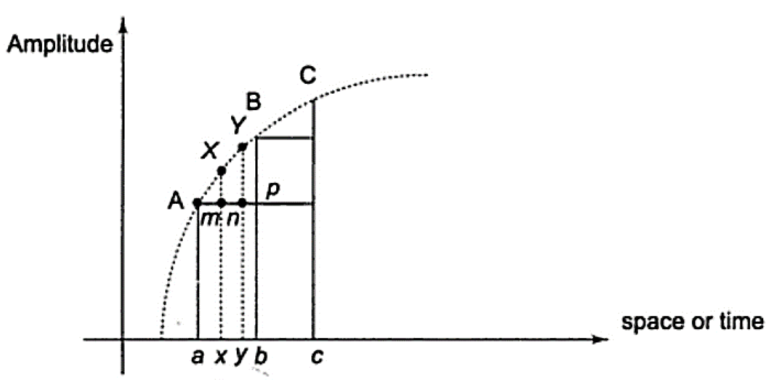
A file's bit depth affects the range of amplitude, or volume, that the file can capture. A file with a larger bit depth will have more dynamic range than one with a smaller bit depth. When audio is captured in a digital format, each sample is assigned the most appropriate amplitude value among those available. With a 16-bit depth (the standard for audio compact discs), 65,536 amplitude values are available. In contrast, millions of amplitude values are available with higher bit depths such as 24- and 32-bit (which is the highest bit depth that Audition supports). For this reason, it is recommended to record the digital audio and edit at the highest bit depth possible, even if you have to down-sample the resulting file to a lower bit depth later in the editing process.



HIGHER BIT DEPTHS PROVIDE GREATER DYNAMIC RANGE

5.4 QUANTIZATION ERROR

By definition, because all voltages in a certain quantization interval are represented by the voltage at the centre of this interval, the process of quantization is a non-linear process and creates an error, called quantization error (or, sometimes, round-off error). A quantized signal takes only discrete, predetermined levels: Compared to the original continuous signal, quantization error has been introduced. This error is correlated with the signal, and is properly called distortion. However, classical signal theory deals with the addition of noise to signals. Providing each quantizer step is small compared to signal amplitude, we can consider the loss of signal in a quantizer as addition of an equivalent amount of noise instead: Quantization diminishes signal-to-noise ratio.



Quantization Error

5.5 FOURIER REPRESENTATION

Named after French mathematician Jean Baptiste Joseph Fourier, it is mathematical technique which shows that any periodic signal is made up of an infinite series of sinusoidal frequency components. The lowest frequency component is known as fundamental frequency and the other components are called Harmonics and are multiplies of the fundamental.

There are four types of Fourier representations depending on the nature of the input wave:

(a) ***Fourier Transform***: resulting from continuous and a-periodic input signals
(b) ***Fourier Series***: resulting from continuous and periodic input signals
(c) ***Discrete Time Fourier Transform***: resulting from discrete and a-periodic input signals
(d) ***Discrete Fourier Transform***: resulting from discrete periodic input signals

All these classes of signals extend from positive to negative infinity.

Imagine an infinite number of samples of the wave on both sides of value zero. Then the signal looks discrete and aperiodic and option (c) above applies. Alternatively, the fictitious samples can be considered a repetition of the finite signal segment, in which case the signal looks discrete and periodic and option (d) applies. Each of the four Fourier representations can be of two versions: real and complex. The real version uses real numbers for decomposition and synthesis while the complex version uses complex numbers.

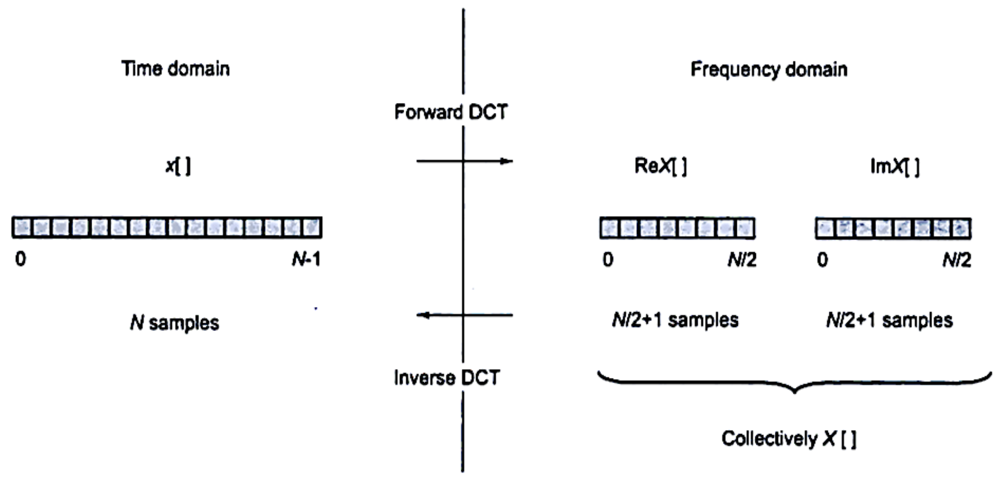
The term transform means an algorithm which can convert a set of values, either continuous or discrete, to another set of values, not necessarily of the same number of even of the same type. When you are computing Fourier representation using a digital computer, you must use option (d). The first two options (a) and (b) dealing with continuous signals cannot be represented in a digital computer, while the third option (c) cannot be used because an infinite number of sinusoids are required to synthesize aperiodic signal. The option (d) is called DFT, in short.

5.5.1 Time Domain and Frequency Domain

Let us consider the input (discrete) signal to be consisting of N samples (e.g. an audio file). This signal is referred as a time domain signal as the samples have been taken over a period of time. The DFT would convert this N sample time domain signal into two signals each of N/2+1 samples long in the frequency domain.

The first component is a cosine wave, also known as the real part and written as Re, and the second component is the sine wave, also known as the imaginary part and written as Im. The process of converting the time domain signal into the frequency domain signal is called the forward DFT, while the reverse process of calculating the time domain signal from the frequency domain signal is referred to as the inverse DFT.

The real and imaginary parts are referred to as ReX[] and ImX[] respectively.



Time Domain and Frequency Domain representations

The output of DFT is a set of number used as scaling factors for a set of sine and cosine waves. The sine and cosine waves have unity amplitude and is called the basic function. These are to be multiplied by necessary scaling factors as follows:

ck[i]=cos(2πki/N)
sk[i]=sin(2πki/N)

where ck[] is the cosine wave for the amplitude held in ReX[k] and sk[] is the sine wave for the amplitude held in ImX[k], N is the total number of samples in the time domain signal and k is the frequency of the sinusoids. The frequency k is the number of complete cycles over the N sample points. Combining together everything said so far, the synthesis equation can be written as:

$$x[i] = \sum_{k=0}^{N/2} \mathrm{Re}\,\overline{X}[k]\cos(2\pi ki/N) + \sum_{k=0}^{N/2} \mathrm{Im}\,\overline{X}[k]\sin(2\pi ki/N)$$

Thus, any N point signal x[i] can be created by adding N/ 2+1 cosine waves and N/ 2+1 sine waves. The amplitudes of the cosine and sine waves are held in the arrays ReX[k] and ImX[k] respectively. The analysis equation is the reverse of the synthesis equation and expresses the frequency domain signals in terms of the time domain signal. This is known as the forward DFT and is given by the relations:

$$\operatorname{Re} X[k] = \sum_{i=0}^{N-1} x[i] \cos(2\pi k i / N)$$

$$\operatorname{Im} X[k] = -\sum_{i=0}^{N-1} x[i] \sin(2\pi k i / N)$$

In words, each sample in the frequency domain is found by multiplying the time domain signal by the sine or cosine wave being looked for, and adding the resulting points.

5.6 PULSE MODULATION

The process of encoding digital information in terms of voltage signal called is called modulation and is done using a device called modulator. The reverse process of decoding the signal back to the digital information is called demodulation & is done using demodulator. The voltage signal is called carrier signal.

Pulse modulation schemes aim at transferring a narrowband analog signal over an analog base band channel as a two-level signal by modulating a pulse wave. Some pulse modulation schemes also allow the narrowband analog signal to be transferred as a digital signal (i.e. as a quantized discrete-time signal) with a fixed bit rate, which can be transferred over an underlying digital transmission system, for example some line code. These are not modulation schemes in the conventional sense since they are not channel coding schemes, but should be considered as source coding schemes, and in some cases analog-to-digital conversion techniques.

Analog-over-analog methods:

- Pulse-amplitude modulation (PAM)
- Pulse-width modulation (PWM) AND Pulse-depth modulation (PDM)
- Pulse-position modulation (PPM)

Analog-over-digital methods:

- Pulse-code modulation (PCM)

➢ **Pulse-amplitude modulation (PAM)**, is a form of signal modulation where the message information is encoded in the amplitude of a series of signal pulses.

- **Pulse-width modulation (PWM)**, **or pulse-duration modulation (PDM)**, is a modulation technique used in communications systems to encode the amplitude of a signal into the width of the pulse (duration) of another signal.

- **Pulse-position modulation (PPM)** is a form of signal modulation in which M message bits are encoded by transmitting a single pulse in one of 2^Mpossible time-shifts.

- **Pulse-code modulation (PCM)** is a method used to digitally represent sampled analog signals. It is the standard form of digital audio in computers, Compact Discs, digital telephony and other digital audio applications. In a PCM stream, the amplitude of the analog signal is sampled regularly at uniform intervals, and each sample is quantized to the nearest value within a range of digital steps.

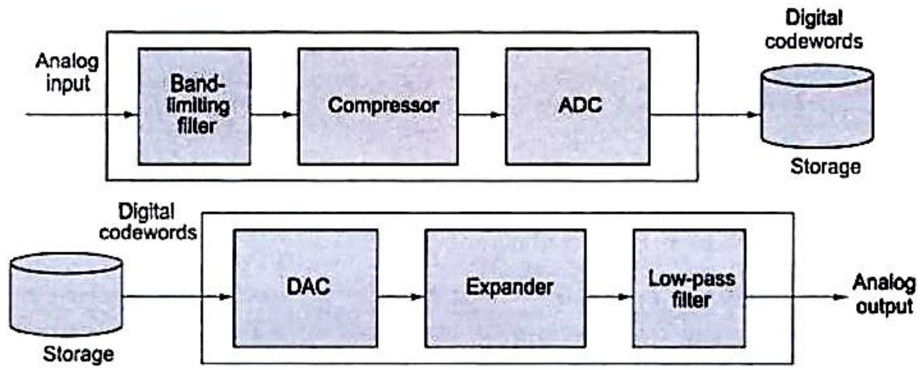
5.6.1 Pulse code modulation [PCM]

Digital numbers are represented by code corresponding to the specific quantization levels.

The signals in PCM are binary; that is, there are only two possible states, represented by logic 1 (high) and logic 0 (low). This is true no matter how complex the analog waveform happens to be. Using PCM, it is possible to digitize all forms of analog data, including full-motion video, voices, music, telemetry, and virtual reality (VR).

To obtain PCM from an analog waveform at the source (transmitter end) of a communications circuit, the analog signal amplitude is sampled (measured) at regular time intervals. The sampling rate, or number of samples per second, is several times the maximum frequency of the analog waveform in cycles per second or hertz. The instantaneous amplitude of the analog signal at each sampling is rounded off to the nearest of several specific, predetermined levels. This process is called quantization. The number of levels is always a power of 2 -- for example, 8, 16, 32, or 64. These numbers can be represented by three, four, five, or six binary digits (bits) respectively. The output of a pulse code modulator is thus a series of binary numbers, each represented by some power of 2 bits.

At the destination (receiver end) of the communications circuit, a pulse code demodulator converts the binary numbers back into pulses having the same quantum levels as those in the modulator. These pulses are further processed to restore the original analog waveform.



Pulse Code Modulation

5.7 IMPORTANCE AND DRAWBACK OF DIGITAL REPRESENTATION

Universal Representation: Any media element for example text, image, and sound can be represented using uniform sequence of binary digits.

Storage: Storage devices like memory chips, hard disk, floppy etc. can be used for all media files.

Transmission: Digital signals are less sensitive to noise as compared to analog signals.

Processing: Software program can be used to analyze, modify, alter and manipulate multimedia data to improve the quality by removal of noises and errors.

The major drawback lies in coding distortion. The process of first sampling and then quantizing and coding the sampled values introduces distortions. As a result the signal generated after digital to analog conversion and presented to the end user has little chance of being completely identical to the original signal.

Another consequence is the requirement of large digital storage capacity required for storing image, sound and video. Fortunately, compression algorithms have been developed to alleviate the problem to a certain extent.

The detection of digital signals requires the communications system to be synchronized, whereas generally speaking this is not the case with analog systems.

Digital communications require greater bandwidth than analogue to transmit the same information.

5.8 SUMMARY

- Two factors affect the potential quality of a digital audio system: its sample rate and bit depth.
- The sample rate defines how often the input voltage is measured, or sampled, with faster rates allowing higher frequencies to be captured.
- The bit depth refers to the size of the numbers used to store the digitized data, with larger numbers giving a lower noise floor.
- The process of quantization is a non-linear process and creates an error, called quantization error (or, sometimes, round-off error).
- Named after French mathematician Jean Baptiste Joseph Fourier, it is mathematical technique which shows that any periodic signal is made up of an infinite series of sinusoidal frequency components.
- There are four types of Fourier representations depending on the nature of the input wave; ***Fourier Transform***: resulting from continuous and a-periodic input signals; ***Fourier Series***: resulting from continuous and periodic input signals; ***Discrete Time Fourier Transform***: resulting from discrete and a-periodic input signals; ***Discrete Fourier Transform***: resulting from discrete periodic input signals
- The forward DFT and is given by the relations:

$$\operatorname{Re} X[k] = \sum_{i=0}^{N-1} x[i]\cos(2\pi ki/N)$$

$$\operatorname{Im} X[k] = -\sum_{i=0}^{N-1} x[i]\sin(2\pi ki/N)$$

- **Pulse-code modulation (PCM)** is a method used to digitally represent sampled analog signals. It is the standard form of digital audio in computers, Compact Discs, digital telephony and other digital audio applications.

5.9 UNIT END EXERCISES

1. What is bit-depth and how is it related, if at all, to sampling rate?
2. What is quantization error?
3. ‘A digital signal is always a degraded version of the original analog signal’. – Explain.

4. Discuss the four types of Fourier representations mentioning the corresponding nature of input waves.
5. Writing the synthesis equation for forward DFT, comment on each part.
6. What is pulse modulation and what are its different variants?
7. Discuss the importance and drawbacks of digital representation.

References:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia

❖❖❖❖

6

UNIT III

TEXT

Unit Structure

6.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define and distinguish different types of text
2. Understand to insert a font
3. Recognize compression techniques
4. Describe different text file formats

6.2 INTRODUCTION

In multimedia presentations, text can be combined with other media in a powerful way to present information and express moods. Text can be of various types: ***Plaintext,*** consisting of fixed sized characters having essentially the same type of appearance. ***Formatted text*****,** where appearance can be changed using font parameters. ***Hypertext,*** which can serve to link different electronic documents and enable the user to jump from one to the other in a non-linear way.

Internally text is represented via binary codes as per the **ASCII table.** The ASCII table is however quite limited in its scope and a new standard has been developed to eventually replace the ASCII standard. This standard is called the **Unicode** standard and is capable of representing international characters from various languages throughout the world. We also generate text automatically from a scanned version of a paper document or image using Optical Character Recognition (**OCR)** software.

6.3 TYPES OF TEXT

There are three types of text that can be used to produce pages of a document:

- Unformatted text
- Formatted text
- Hypertext

6.3.1. Unformatted Text:

Also known as plaintext, this comprise of fixed sized characters from a limited character set. The character set is called **ASCII table** which is short for American Standard Code for Information Interchange and is one of the most widely used character sets. It basically consists of a table where each character is represented by a unique 7-bit binary code. The characters include a to z, A to Z, 0 to 9, and other punctuation characters like parenthesis, ampersand, single and double quotes, mathematical operators, etc. All the characters are of the same height. In addition, the ASCII character set also includes a number of control characters. These include BS (backspace), LF (linefeed), CR (carriage return), SP (space), DEL (delete), ESC (escape), FF (form feed) and others.

Later on as requirements increased an extended version of ASCII table was introduced known as **extended ASCII character set (REFER TABLE 6.2)**, while the original table came to be known as **standard ASCII set (REFER TABLE 6.1)**.

6.3.2. Formatted Text:

Formatted text are those where apart from the actual alphanumeric characters, other control characters are used to change the appearance of the characters, e.g. bold, underline, italics, varying shapes, sizes and colors etc., Most text processing software use such formatting options to change text appearance. It is also extensively used in the publishing sector for the preparation of papers, books, magazines, journals, and so on.

6.3.3. Hypertext:

The term Hypertext is used to mean certain extra capabilities imparted to normal or standard text. Like normal text, a hypertext document can be used to reconstruct knowledge through sequential reading but additionally it can be used to link multiple documents in such a way that the user can navigate non-sequentially from one document to the other for cross-references. These links are called **hyperlinks.** The underlined text string on which the user clicks the mouse is called an **anchor** and the document which opens as a result of clicking is called the **target document.** On the web target documents are specified by a specific nomenclature called Web site address technically known as **Uniform Resource Locators** or URL.

TABLE 6.1 STANDARD ASCII CHARACTER SET

Dec	Hex	Abbr.	Char
0	00	NUL	Null character
1	01	SOH	Start of Header
2	02	STX	Start of Text
3	03	ETX	End of Text
4	04	EOT	End of Transmission
5	05	ENQ	Enquiry
6	06	ACK	Acknowledgment
7	07	BEL	Bell
8	08	BS	Backspace
9	09	HT	Horizontal Tab
10	0A	LF	Line feed
11	0B	VT	Vertical Tab
12	0C	FF	Form feed
13	0D	CR	Carriage return
14	0E	SO	Shift Out
15	0F	SI	Shift In
16	10	DLE	Data Link Escape
17	11	DC1	Device Control 1 (oft. XON)
18	12	DC2	Device Control 2
19	13	DC3	Device Control 3 (oft. XOFF)
20	14	DC4	Device Control 4
21	15	NAK	Negative Acknowledgement
22	16	SYN	Synchronous idle
23	17	ETB	End of Transmission Block
24	18	CAN	Cancel
25	19	EM	End of Medium
26	1A	SUB	Substitute
27	1B	ESC	Escape
28	1C	FS	File Separator
29	1D	GS	Group Separator
30	1E	RS	Record Separator
31	1F	US	Unit Separator

Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
32	20		64	40	@	96	60	`
33	21	!	65	41	A	97	61	a
34	22	"	66	42	B	98	62	b
35	23	#	67	43	C	99	63	c
36	24	$	68	44	D	100	64	d
37	25	%	69	45	E	101	65	e
38	26	&	70	46	F	102	66	f
39	27	'	71	47	G	103	67	g
40	28	(	72	48	H	104	68	h
41	29	)	73	49	I	105	69	i
42	2A	*	74	4A	J	106	6A	j
43	2B	+	75	4B	K	107	6B	k
44	2C	,	76	4C	L	108	6C	l
45	2D	-	77	4D	M	109	6D	m
46	2E	.	78	4E	N	110	6E	n
47	2F	/	79	4F	O	111	6F	o
48	30	0	80	50	P	112	70	p
49	31	1	81	51	Q	113	71	q
50	32	2	82	52	R	114	72	r
51	33	3	83	53	S	115	73	s
52	34	4	84	54	T	116	74	t
53	35	5	85	55	U	117	75	u
54	36	6	86	56	V	118	76	v
55	37	7	87	57	W	119	77	w
56	38	8	88	58	X	120	78	x
57	39	9	89	59	Y	121	79	y
58	3A	:	90	5A	Z	122	7A	z
59	3B	;	91	5B	[	123	7B	{
60	3C	<	92	5C	\	124	7C	|
61	3D	=	93	5D	]	125	7D	}
62	3E	>	94	5E	^	126	7E	~
63	3F	?	95	5F	_	127	7F	DEL

TABLE 6.2 EXTENDED ASCII CHARACTER SET

Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
128	80	Ç	160	A0	á	192	C0	└	224	E0	α
129	81	ü	161	A1	í	193	C1	┴	225	E1	ß
130	82	é	162	A2	ó	194	C2	┬	226	E2	Γ
131	83	â	163	A3	ú	195	C3	├	227	E3	π
132	84	ä	164	A4	ñ	196	C4	─	228	E4	Σ
133	85	à	165	A5	Ñ	197	C5	┼	229	E5	σ
134	86	å	166	A6	ª	198	C6	╞	230	E6	µ
135	87	ç	167	A7	º	199	C7	╟	231	E7	τ
136	88	ê	168	A8	¿	200	C8	╚	232	E8	Φ
137	89	ë	169	A9	⌐	201	C9	╔	233	E9	Θ
138	8A	è	170	AA	¬	202	CA	╩	234	EA	Ω
139	8B	ï	171	AB	½	203	CB	╦	235	EB	δ
140	8C	î	172	AC	¼	204	CC	╠	236	EC	∞
141	8D	ì	173	AD	¡	205	CD	═	237	ED	φ
142	8E	Ä	174	AE	«	206	CE	╬	238	EE	ε
143	8F	Å	175	AF	»	207	CF	╧	239	EF	∩
144	90	É	176	B0	░	208	D0	╨	240	F0	≡
145	91	æ	177	B1	▒	209	D1	╤	241	F1	±
146	92	Æ	178	B2	▓	210	D2	╥	242	F2	≥
147	93	ô	179	B3	│	211	D3	╙	243	F3	≤
148	94	ö	180	B4	┤	212	D4	╘	244	F4	⌠
149	95	ò	181	B5	╡	213	D5	╒	245	F5	⌡
150	96	û	182	B6	╢	214	D6	╓	246	F6	÷
151	97	ù	183	B7	╖	215	D7	╫	247	F7	≈
152	98	ÿ	184	B8	╕	216	D8	╪	248	F8	°
153	99	Ö	185	B9	╣	217	D9	┘	249	F9	∙
154	9A	Ü	186	BA	║	218	DA	┌	250	FA	·
155	9B	¢	187	BB	╗	219	DB	█	251	FB	√
156	9C	£	188	BC	╝	220	DC	▄	252	FC	^{n}
157	9D	¥	189	BD	╜	221	DD	▌	253	FD	^{2}
158	9E	₧	190	BE	╛	222	DE	▐	254	FE	■
159	9F	ƒ	191	BF	┐	223	DF	▀	255	FF	□

6.4 FONT

In traditional typography, a **font** is a particular size, weight and style of a typeface. Each font was a matched set of type, one piece (called a "sort") for each glyph, and a typeface comprised a range of fonts that shared an overall design.

In modern usage, with the advent of digital typography, "font" is frequently synonymous with "typeface", although the two terms do not necessarily mean the same thing.

In particular, the use of "vector" or "outline" fonts means that different sizes of a typeface can be dynamically generated from one design. Each style may still be in a separate "font file"—for instance, the typeface "Bulmer" may include the fonts "Bulmer roman", "Bulmer italic", "Bulmer bold" and "Bulmer extended"—but the term "font" might be applied either to one of these alone or to the whole typeface.

6.5 INSERTION OF TEXT

Text can be inserted in a document using a variety of methods. These are:

6.5.1 Using a keyboard

The most common process of inserting text into a digital document is by typing the text using an input device like the keyboard. Usually a text editing software, like Microsoft Word, is used to control the appearance of text which allows the user to manipulate variables like the font, size, style, color, etc.,

6.5.2 Copying and Pasting

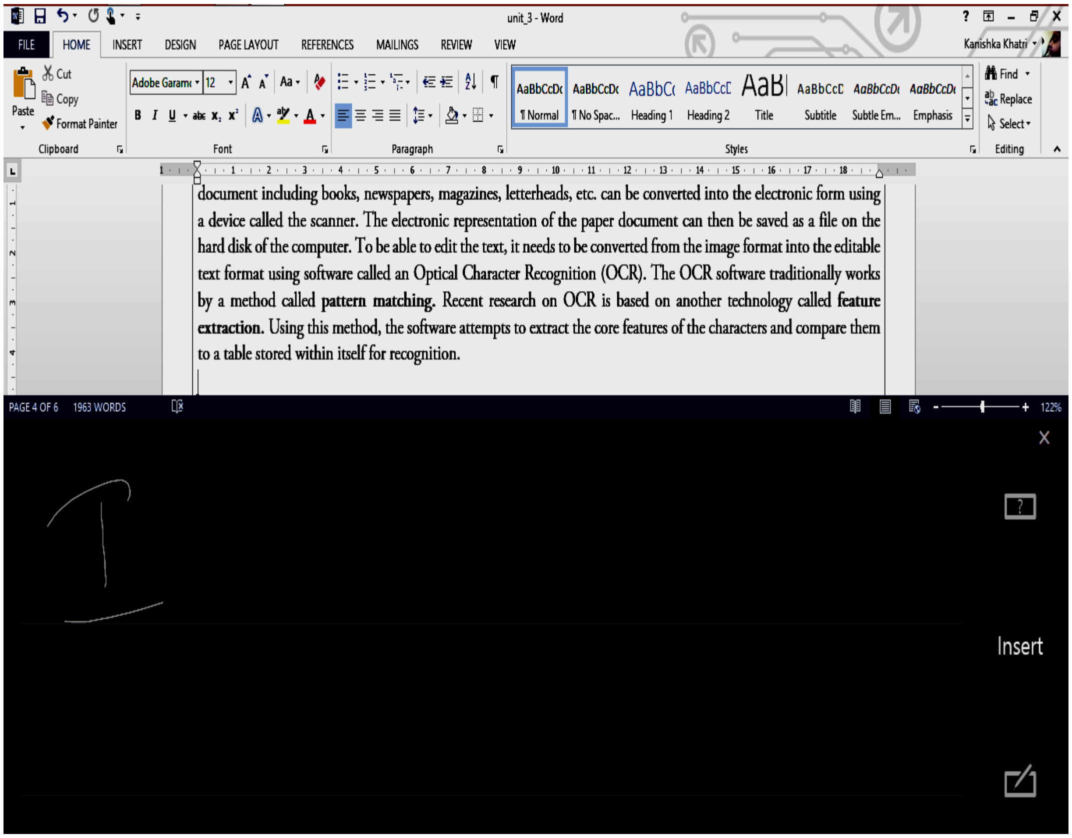
Another way of inserting text into a document is by copying text from a pre-existing digital document. The existing document is opened using the corresponding text processing program and portions of the text may be selected by using the keyboard or mouse. Using the **Copy** command the selected text is copied to the clipboard. By choosing the **Paste** command, whereupon the text is copied from the clipboard into the target document.

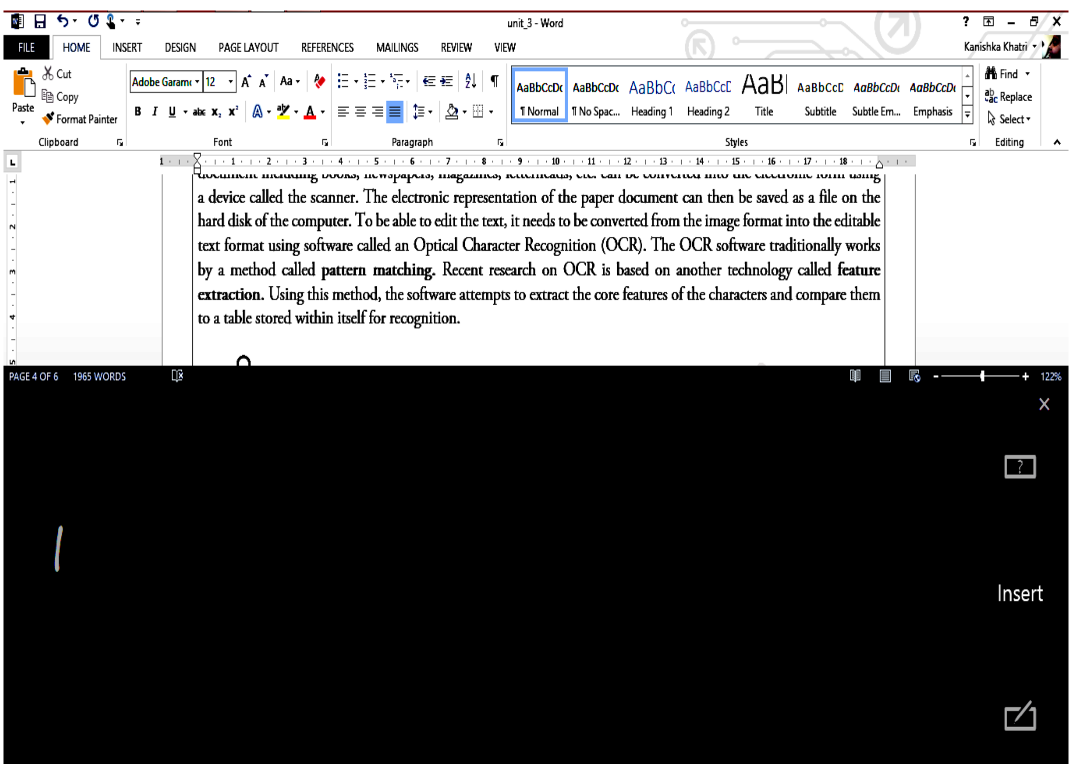
6.5.3 Using an OCR Software

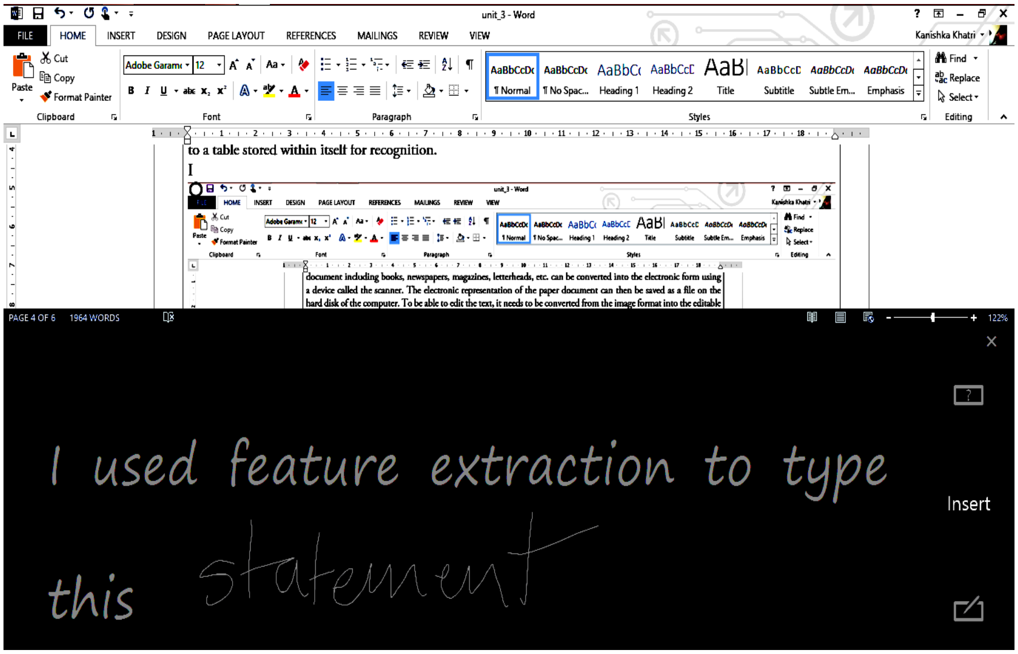
A third way of inserting text into a digital document is by scanning it from a paper document. Text in a paper document including books, newspapers, magazines, letterheads, etc. can be converted into the electronic form using a device called the scanner. The electronic representation of the paper document can then be saved as a file on the hard disk of the computer.

To be able to edit the text, it needs to be converted from the image format into the editable text format using software called an Optical Character Recognition (OCR). The OCR software traditionally works by a method called **pattern matching.** Recent research on OCR is based on another technology called **feature extraction.** Using this method, the software attempts to extract the core features of the characters and compare them to a table stored within itself for recognition.

You may see the feature extraction to type this Statement


unit_3 - Word
FILE
HOME
INSERT
DESIGN
PAGE LAYOUT
REFERENCES
MAILINGS
REVIEW
VIEW
Kanishka Khatri
Cut
Copy
Paste
Format Painter
Clipboard
Font
Paragraph
Styles
Editing
Find
Replace
Select
Normal
No Spac...
Heading 1
Heading 2
Title
Subtitle
Subtle Em...
Emphasis
document including books, newspapers, magazines, letterheads, etc. can be converted into the electronic form using a device called the scanner. The electronic representation of the paper document can then be saved as a file on the hard disk of the computer. To be able to edit the text, it needs to be converted from the image format into the editable text format using software called an Optical Character Recognition (OCR). The OCR software traditionally works by a method called pattern matching. Recent research on OCR is based on another technology called feature extraction. Using this method, the software attempts to extract the core features of the characters and compare them to a table stored within itself for recognition.
PAGE 4 OF 6
1963 WORDS
122%
Insert


unit_3 - Word
FILE
HOME
INSERT
DESIGN
PAGE LAYOUT
REFERENCES
MAILINGS
REVIEW
VIEW
Kanishka Khatri
Cut
Copy
Paste
Format Painter
Clipboard
Font
Paragraph
Styles
Editing
Find
Replace
Select
Normal
No Spac...
Heading 1
Heading 2
Title
Subtitle
Subtle Em...
Emphasis
a device called the scanner. The electronic representation of the paper document can then be saved as a file on the hard disk of the computer. To be able to edit the text, it needs to be converted from the image format into the editable text format using software called an Optical Character Recognition (OCR). The OCR software traditionally works by a method called pattern matching. Recent research on OCR is based on another technology called feature extraction. Using this method, the software attempts to extract the core features of the characters and compare them to a table stored within itself for recognition.
PAGE 4 OF 6
1965 WORDS
122%
Insert



6.6 TEXT COMPRESSION

Large text documents covering a number of pages may take a lot of disk space. We can apply compression algorithms to reduce the size of the text file during storage.

A reverse algorithm must be applied to decompress the file before its contents can be displayed on screen. There are three types of compression methods that are applied to text as explained:

a. Huffman Coding:

This type of coding is intended for applications in which the text to be compressed has known characteristics in terms of the characters used and their relative frequencies of occurrences. An optimum set of variable-length code words is derived such that the shortest code word is used to represent the most frequently occurring characters. This approach is called the **Huffman** coding method.

b. Lempel-Ziv (LZ) Coding

In the second approach followed by the **Lempel-Ziv** (LZ) method, instead of using a single character as a basis of the coding operation, a string of characters is used. For example, a table containing all the possible words that occur in a text document, is held by both the encoder and decoder.

c. Lempel-Ziv-Welsh (LZW) Coding

Most word processing packages have a dictionary associated with them which is used for both spell checking and compression of text. The variation of the above algorithm called

Lempel-Ziv-Welsh (LZW) method allows the dictionary to be built up dynamically by the encoder and decoder for the document under processing.

6.7 FILE FORMATS

The following text formats are commonly used for textual documents.

3.3.1 TXT (Text)

Unformatted text document created by an editor like Notepad on Windows platform. This documents can be used to transfer textual information between different platforms like Windows, DOS, and UNIX.

3.3.2 DOC (Document)

Developed by Microsoft as a native format for storing documents created by the MS Word package. Contains a rich set of formatting capabilities.

3.3.3 RTF (Rich Text Format)

Developed by Microsoft in 1987 for cross platform document exchanges. It is the default format for Mac OS X's default editor Text Edit. RTF control codes are human readable, similar to HTML code.

3.3.4 PDF (Portable Document Format)

Developed by Adobe Systems for cross platform exchange of documents. In addition to text the format also supports images and graphics. PDF is an open standard and anyone may write programs that can read and write PDFs without any associated royalty charges.

3.3.5 PostScript (PS)

Postscript is a **page description language** used mainly for desktop publishing. A page description language is a high-level language that can describe the contents of a page such that it can be accurately displayed on output devices usually a printer.

A PostScript interpreter inside the printer converts the vectors back into the raster dots to be printed. This allows arbitrary scaling, rotating and other transformations.

6.8 SUMMARY

- ***Plaintext*** consists fixed sized characters having essentially the same type of appearance.
- ***Formatted text*'s** appearance can be changed using font parameters*.*
- ***Hypertext*** can serve to link different electronic documents and enable the user to jump from one to the other in a non-linear way.
- Text can be inserted in a document using a variety of methods. These are: **Using a keyboard, Copying and Pasting and Using an OCR Software.**
- There are three types of compression methods that are applied to text as explained: a. **Huffman Coding,** b. **Lempel-Ziv (LZ) Coding and c**. **Lempel-Ziv-Welsh (LZW) Coding.**
- The following text formats are commonly used for textual documents: **TXT (Text), DOC (Document), RTF (Rich Text Format), PDF (Portable Document Format) and Post Script (PS)**
- A reverse algorithm must be applied to decompress the file before its contents can be displayed on screen.
- A PostScript interpreter inside the printer converts the vectors back into the raster dots to be printed.

6.9 UNIT END EXERCISES

1. Differentiate between formatted text and unformatted text.
2. How can a text be inserted within an application?
3. In relation to OCR software, distinguish between pattern-matching and feature extraction.
4. Distinguish between Huffman coding and LZW coding methods of text compression.
5. Discuss in brief about the major file formats.

Reference:

1. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.
2. Wikipedia.

❖❖❖❖

7

IMAGE

Unit Structure

7.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Distinguish types of images
2. Understand different colour models
3. Describe principle and working of several devices
4. Define gamma and gamma correction

7.2 INTRODUCTION

The strength of the radiation from a light source is measured using the unit called the candela, or candle power. The total energy from the light source, including heat and all electromagnetic radiation, is called radiance and is usually expressed in watts.

Luminance is a measure of the light strength that is actually perceived by the human eye.

Radiance is a measure of the total output of the source; luminance measures just the portion that is perceived. Brightness is a subjective, psychological measure of perceived intensity. Brightness is practically impossible to measure objectively. It is relative. For example, a burning candle in a darkened room will appear bright to the viewer; it will not appear bright in full sunshine.

The strength of light diminishes in inverse square proportion to its distance from its source. This effect accounts for the need for high Intensity projectors for showing multimedia productions on a screen to an audience.

Color is the sensation registered when light of different wavelengths is perceived by the brain. It is observed in objects that reflect or emit certain wavelengths of light. Hue distinguishes among colors such as red, green, and yellow. Saturation refers to how far color is from a gray of equal intensity.

Lightness embodies the achromatic notion of perceived intensity of a reflecting object. Brightness is used instead of lightness for a self-luminous object such as CRT.

7.3 TYPES OF IMAGES

Hard Copy vs. Soft Copy

The typical images that we usually come across are the pictures that have been printed on paper or some other kinds of surfaces like plastic, cloth, wood, etc. These are also called **hard copy** images because they have been printed on solid surfaces. Such images have been transformed from hard copy images or real objects into the electronic form using specialized procedures and are referred to as **soft copy** images.

Continuous Tone, Half-tone and Bitone

Photographs are also known as **continuous tone** images because they are usually composed of a large number of varying tones or shades of colors. Sometimes due to limitations of the display or printed devices, all the colors of the photograph cannot be represented adequately. In those cases a subset of the total number of colors is displayed. Such images are called **partial tone** or **half-tone** images. A third category of images is called **bitonal** images which uses only two colors, typically black and white, and do not use any shades of grey.

7.4 COLOR MODELS

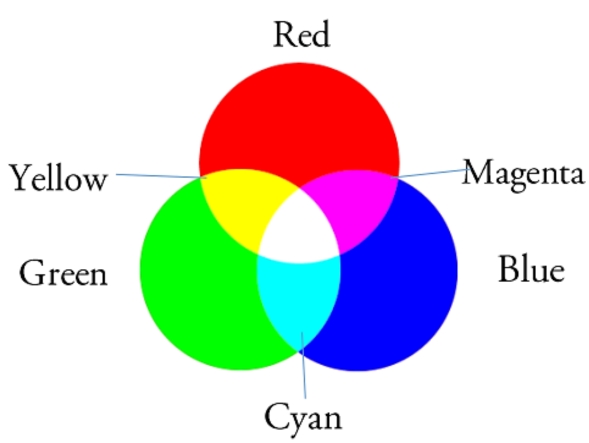
Researchers have found out that most of the colors that we see around us can be derived from mixing a few elementary colors. These elementary colors are known as **primary colors.** Primary colors mixed in varying proportions produce other colors called **composite colors.** Two primary colors mixed in equal proportions produce a **secondary color.** The primary colors along with the total range of composite colors they can produce constitute a **color model.**

7.4.1 RGB Color Model

The RGB color model is used to describe behavior of colored lights like those emitted from a TV screen or a computer monitor. This model has three primary colors: red, green, blue, in short RGB.

Proportions of colors are determined by the beam strength. An electron beam having the maximum intensity falling on a phosphor dot creates 100% of the corresponding color.

All three primary colors at full intensities combine together to produce white, i.e. their brightness values are added up. Because of this the RGB model is called an **additive model.** Lower intensity values produce shades of grey. A color present at 100% of its intensity is called **saturated,** otherwise the color is said to be **unsaturated.**

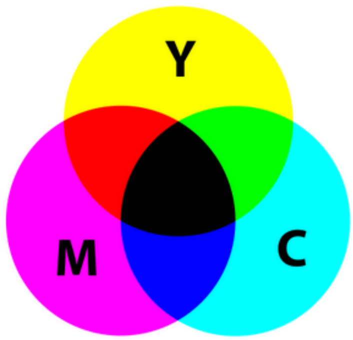


RGB Color Model

7.4.2 CMY Color Model

The RGB model is only valid for describing behavior of colored lights. This new model is named CMY / CMYK model and is used to specify printed colors. The primary colors of this model are cyan, magenta and yellow. These colors when mixed together in equal proportions produce black, due to which the model is known as a **subtractive model**.

Mixing cyan and magenta in equal proportions produce blue, magenta and yellow produce red, and yellow and cyan produce green. Thus, the secondary colors of the CMYK model are the same as the primary colors of the RGB model and vice versa. These two models are thus, known as **complimentary models.**



CMYK Color Model

7.4.3 Device Dependency and Gamut (Lab model)

It is to be noted that both the RGB and the CMYK models do not have universal or absolute color values. But different devices will give rise to slightly different sets of colors. For this reason both the RGB and the CMYK models are known as **device dependent** color models.

Another issue of concern here is the total range of colors supported by each color model. This is known as the **gamut** of the model.

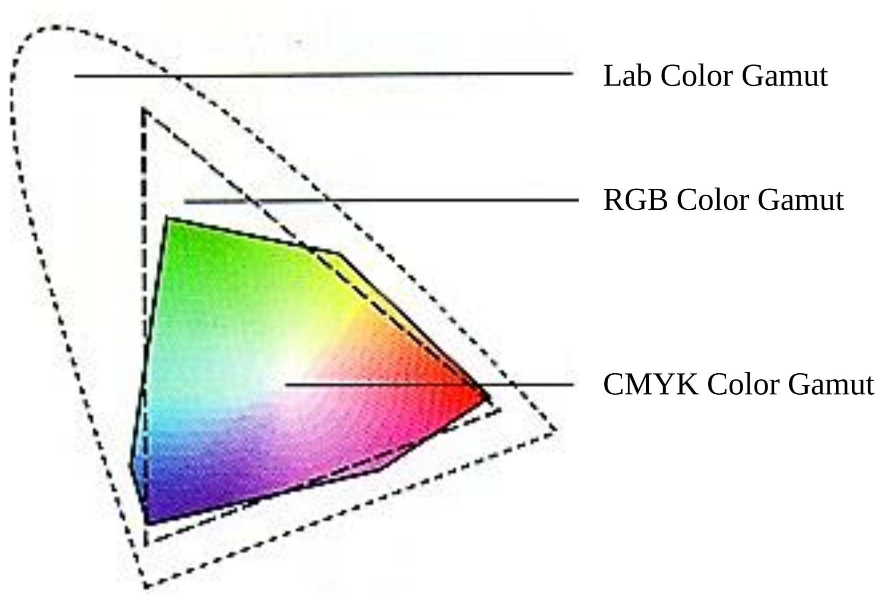


Fig 3.3 Different Color Models with their gamut

<u>Note</u>: *The same color name is represented as different color shades in RGB and CMYK models. For example, RGB Blue is (0,0,255) but is out of the CMYK gamut, the nearest representation being CMYK(88%,77%,0%,0%).*

On the other hand, CMYK Blue is CMYK(100%,100%,0%,0%) and has RGB value RGB(46,49,146). This accounts for the difference in appearances of RGB and CMYK color shades.

7.5 BASIC STEPS FOR IMAGE PROCESSING

Image processing is the name given to the entire process involved with the input, editing and output of images from a system. There are three basic steps:

a. **Input**

Image **input** is concerned with getting natural images into a computer system for subsequent work. Essentially it deals with the conversion of analog images into digital forms using two devices. The first is the scanner which can convert a printed image or document into the digital form. The second is the digital camera which digitizes real-world images, similar to how a conventional camera works.

b. **Editing**

After the images have been digitized and stored as files on the hard disk of a computer, they are changed or manipulated to make them more suitable for specific requirements. This step is called editing.

Before the actual editing process can begin, and important step called **color calibration** needs to be performed to ensure that the image looks consistent when viewed on multiple monitors.

c. **Output**

Image output is the last stage in image processing concerned with displaying the edited image to the user. The image can either be displayed in a stand-alone manner or as part of some application like a presentation or web-page.

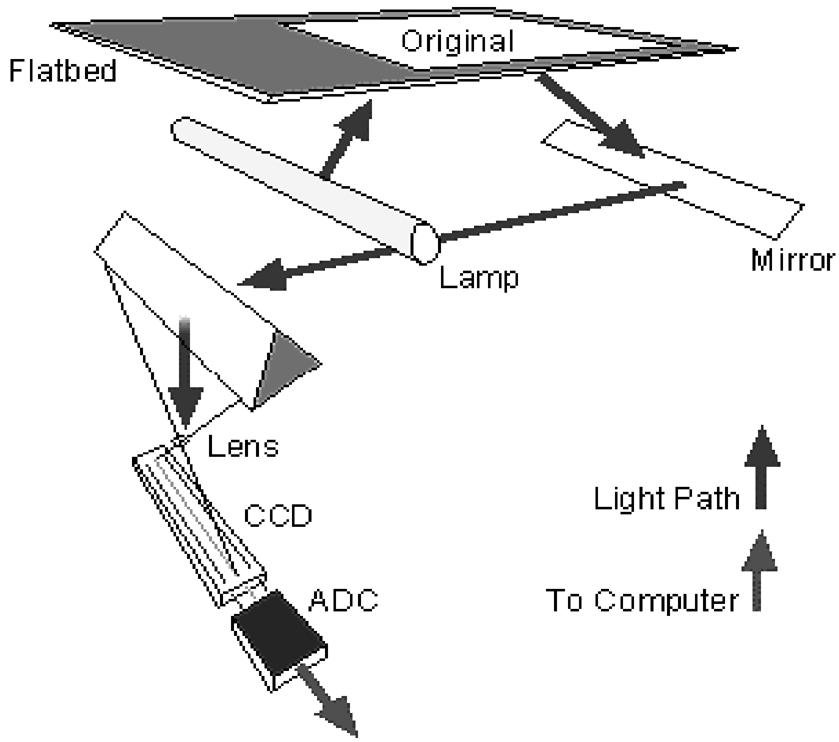
7.6 PRINCIPLE AND WORKING OF SCANNER AND DIGITAL CAMERA

7.6.1 SCANNER

Principle:

For images, digitization involves physical devices like the **scanner** or **digital camera.** The scanner is a device used to convert analog images into the digital form. The most common type of scanner for the office environment is called the **flatbed scanner.**

The traditional way of attaching a scanner to the computer is through an interface cable connected to the **parallel port** of the PC.



Internal Architecture of Flatbed Scanner

Construction and Working:

To start a scanning operation, the paper document to be scanned is placed face down on the glass panel of the scanner, and the scanner is activated using a software from the host computer. The light on getting reflected by the paper image is made to fall on a grid of electronic sensors, by an arrangement of mirrors and lenses. The electronic sensors are called **Charge Coupled Devices (**CCD) are converters of the light energy into voltage pulses. After a complete scan, the image is converted from a continuous entity into a discrete form represented by a series of voltage pulses. This process is called **sampling.** The voltage signals are temporarily stored in a buffer inside the scanner. The next step called **quantization** involves representing the voltage pulses as binary numbers and carried out by an ADC inside the scanner in conjunction with a software bundled with the scanner called the **scanning software.** Since each number has been derived from the intensity of the incident light, these essentially represent brightness values at different points of the image and are known as **pixels.**

7.6.2 DIGITAL CAMERA

Construction and working:

Apart from the scanner used to digitize paper documents and film, another device used to digitize real world images is the digital camera. Unlike a scanner a digital camera is usually not attached to the computer via a cable.

The camera has its own storage facility inside it usually in the form of a memory card. So instead they are **compressed** to reduce their file sizes and stored usually in the JPEG format. This is a lossy compression technique and results in slight loss in image quality. Most of the digital cameras have a **TFT screen** at the back, which serve now important purposes: first it can be used as a viewfinder for composition and adjustment; secondly it can be used for viewing the images stored inside the camera.

The recent innovation of built-in microphones provides for sound annotation, in standard MP3 format. After recording, this sound can be sent to an external device for playback on headphones using an ear socket.

Storage and Software utility

Digital cameras also have a **software utility** resident in a ROM chip inside it which allow the user to toggle between the CAMERA mode and PLAY mode. In the PLAY mode the user is presented with a menu structure having functionalities like: displaying all the images on the memory card, selecting a particular image, deleting selected images, write-protecting the important image for deletion, setting the date and time, displaying how much of the floppy disk space is free and even allowing a floppy to be formatted in the drive.

7.7 GAMMA AND GAMMA CORRECTION

7.7.1 GAMMA

When a picture is displayed on a monitor, its brightness levels can be broadly divided into 3 groups: the darkest regions, the lightest regions and the midtone regions. **Gamma** refers to the **midtone** regions.

When a perfectly linear gradient ranging from black to white is shown on the monitor, the monitor distorts the gradient somewhat. Thus, a brightness of 50% will result in a screen brightness of 20%. This can be represented mathematically by a power function whose value turns out to be approximately 2.7. Thus, 0.5 raised to the power 2.5 results in 0.2. Hence, we may write

$$(\text{output pixel}) = (\text{input pixel})^{\text{gamma}}$$

The image(source: Wikipedia) shows what happens when you increase or decrease gamma.

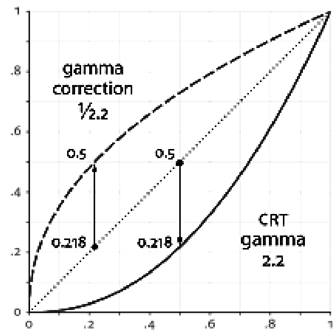
Notice how the absolute white and black points don't change if you don't go overboard. When you increase gamma (pull the curve down) you are mapping more values into the dark regions, and the image goes darker.

When you reduce gamma (pushing the curve up) you are mapping more values into the lighter regions, and the image looks lighter. Whether or not the image will have anything to display depends on its dynamic range and 'strength of data'.

7.7.2GAMMA CORRECTION

To prevent all our pictures on the screen from looking excessively dark a correction is applied to the monitor, which is known as gamma correction. This actually means the inverse of gamma value is applied, according to the relation:

$$(\text{output pixel}) = \{(\text{input pixel})^{1/\text{gamma-correction}}\}^{\text{gamma}}$$



In other words, the transmitted signal is deliberately distorted so that, after it has been distorted again by the display device, the viewer sees the correct brightness.

7.8 IMAGE FILE FORMATS

Image file formats are standardized means of organizing and storing digital images. Image files are composed of digital data in one of these formats that can be rasterized for use on a computer display or printer.

An image file format may store data in uncompressed, compressed, or vector formats. Once rasterized, an image becomes a grid of pixels, each of which has a number of bits to designate its color equal to the color depth of the device displaying it.Including proprietary types, there are hundreds of image file types. The PNG, JPEG, and GIF formats are most often used to display images on the Internet.

GIF: Graphics Interchange Format by the UNISYS and CompuServe

• Initially designed for transmitting images over phone lines

• Limited to 8-bit color images

JPEG: Joint Photographic Experts Group, a standard for photographic (still) image compression.

• Takes advantage of limitations of human vision system to achieve high rates of compression.

• Lossy compression which allows user to set the desired level of quality.

TIFF: Tagged Image File Format by the Aldus Corp.

• Lossless format to store many different types of images.

• No major advantages over JPEG and not user-controllable.

PS/EPS: **Post Script / Encapsulated Post Script**, a typesetting language
• Includes vector/structured graphics and bit-mapped images.
• Used in several popular graphics programs (Adobe).
• No compression, files are large.

BMP: supports 24-bit bitmap images for Microsoft Windows

XBM: supports 24-bit bitmap images for X Windows systems

PNG: **Portable Network Graphics**, created as a free, open-source alternative to GIF.
• Supports 8 bit paletted images and 24 bit true color (16 million colors) or 48 bit true color.

SVG: Scalable Vector Graphics, developed by the World Wide Web Consortium
• Versatile, scriptable and all-purpose vector format for the web and otherwise.

Other raster formats

- JPEG XR (New JPEG standard based on Microsoft HD Photo)
- DEEP (IFF-style format used by TV Paint)
- IMG (Graphical Environment Manager image file; planar, run-length encoded)
- ILBM - Interleaved Bitmap, used by Electronic arts
- PLBM - Planar Bitmap, proprietary Amiga format
- ECW (Enhanced Compression Wavelet)
- IMG (ERDAS IMAGINE Image)
- SID (multiresolution seamless image database, MrSID)
- CD5 (Chasys Draw Image)
- FITS (Flexible Image Transport System)
- PGF (Progressive Graphics File)
- PAM (Portable Arbitrary Map)
- Nrrd (Nearly raw raster data)
- XCF (eXperimental Computing Facility format)
- PSD (Adobe PhotoShop Document)
- PSP (Corel Paint Shop Pro)
- VICAR file format (NASA/JPL image transport format)

Compound formats: These are formats containing both pixel and vector data, possible other data, e.g. the interactive features of PDF.

- PDF (Portable Document Format)
- PICT (Classic Macintosh QuickDraw file)
- SWF (Shockwave Flash)
- XAML User interface language using vector graphics for images.

7.9 SUMMARY

- The typical images that we usually come across are the pictures that have been printed on paper or some other kinds of surfaces like plastic, cloth, wood, etc. These are called **hard copy** images.

- Images transformed from hard copy images or real objects into the electronic form using specialized procedures and are referred to as **soft copy** images.

- Photographs are also known as **continuous tone** images because they are usually composed of a large number of varying tones or shades of colors.

- A subset of the total number of colors is displayed. Such images are called **partial tone** or **half-tone** images.

- A third category of images is called **bitonal** images which uses only two colors, typically black and white, and do not use any shades of grey.

- The elementary colors are known as **primary colors.** Primary colors mixed in varying proportions produce other colors called **composite colors.**

- Two primary colors mixed in equal proportions produce a **secondary color.** The primary colors along with the total range of composite colors they can produce constitute a **color model.**

- The RGB model is called an **additive model.**

- A color present at 100% of its intensity is called **saturated,** otherwise the color is said to be **unsaturated.**

- The CMY model is known as a **subtractive model**.

- Both the RGB and the CMYK models are known as **device dependent** color models

- There are three basic steps for processing an image: **Input, Editing and Output.**

- **Charge Coupled Devices (**CCD) are converters of the light energy into voltage pulses.

- **Gamma** refers to the **midtone** regions.

7.10 UNIT END EXERCISES

1. What are different types of images?
2. What is meant by continuous tone, half-tone and bitone images?
3. How does the RGB color model represent color information? Why is it called an additive model?
4. How does the CMYK color model represent color information? Why is it called a subtractive model?
5. What is meant by saturated and unsaturated colors?
6. What is meant by device dependency and gamut of color models?
7. What are the basic steps in image processing?
8. What is a scanner? How does it function? Explain with the help of diagram.
9. What is a digital camera? Compare between the scanner and digital camera in relation to working principle and storage.
10. What is gamma and gamma correction? Why is it necessary?
11. Discuss few Image File Formats.

References:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia.

❖❖❖❖

8

UNIT IV

AUDIO AND VIDEO

Unit Structure

8.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define characteristics of sound waves.
2. Understand Psycho-acoustics
3. Describe Color fundamentals
4. Understand difference between additive and subtractive colors

8.2 INTRODUCTION

Sound is a form of energy capable of flowing from one place to another through a material medium. It is generated from vibrating objects when part of the energy of vibration is converted to sound energy.

The amplitude of the wave corresponds to its loudness, the frequency corresponds to its pitch and the waveform to the quality or timbre.

Sounds of frequencies between 20 Hz to 20000 Hz can be heard by the human ear and constitute the audible frequency range. To process sound the main system components required are microphones, for sound input, amplifiers for boosting the loudness levels and loudspeakers for output or playback of sound.

For audio-CD quality, a sampling rate of 44.1 kHz, a bit-depth of 16-bits and two stereo channels are used. To compress sound files both lossy and lossless compression algorithms can be used.

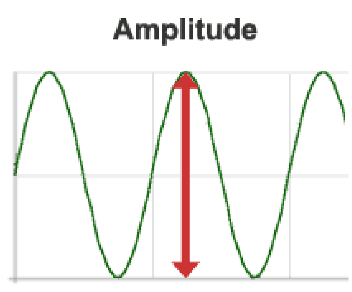
Music generated from synthesizers are pure digital sounds which do not need to be digitized. MIDI compatible instruments and software are capable of playing those sounds. In multimedia presentations, sound is used either as voice-overs or as background music.

8.3 FUNDAMENTAL CHARACTERISTICS OF SOUND

A sound wave is associated with the following physical characteristics: amplitude, frequency, waveform and speed of propagation.

8.3.1 AMPLITUDE

Amplitude of a wave is the maximum displacement of a particle in the path of a wave and is a measure of the peak-to-peak height of the wave. The physical manifestation of amplitude is the intensity of energy of the wave. For sound waves this corresponds to the loudness of sound. Loudness is measured in a unit called decibel (dB).

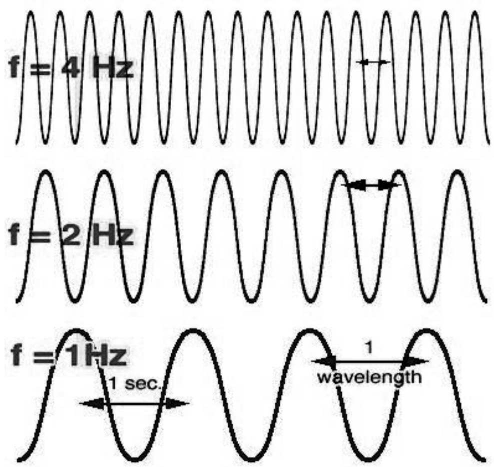


8.3.2 FREQUENCY

This measures the number of vibrations of a particle in the path of a wave, in one second. The physical manifestation of frequency of a sound wave is the pitch of the sound.

A high pitched sound, like that of a whistle, has higher frequency than a dull sound, like the sound of a drum. Frequency is measured in a unit called Hertz (Hz).

The audible frequencies are referred to as ultra-sonic or super-sonic. Frequencies higher than sonic are referred to as ultrasonic while frequencies lower than sonic are called infra-sonic or sub-sonic.



Frequency of Sound Wave

8.3.3 WAVEFORM

This indicates the actual shape of the wave when represented pictorially. Shapes of the waves can be sinusoidal, square, triangular, saw tooth, etc.

The physical manifestation of waveform is the quality or timbre of sound. This helps us to distinguish between sounds coming from different instruments like a guitar and a violin.

8.3.4 SPEED

The speed of sound depends on the medium through which the sound travels, and the temperature of the medium but not on the pressure. The speed is about 240 m/s in air and 1500 m/s in water.

8.4 PSYCHO-ACOUSTICS

Psycho-acoustics is the branch of acoustics which deals with the human auditory perception ranging from the biological design of the ear to the brain's interpretation of aural information.

8.4.1 Decibel (dB)

A decibel is one tenth of a bel, a seldom-used unit named in honor of Alexander Graham Bell. A unit for measuring loudness of sound as perceived by the human ear is decibel.

It involves comparing the intensity of a sound with the faintest sound audible by the human ear and expressing the ratio as a logarithmic value.

$$\text{Power in dB} = 10\log10\,\frac{power\ A}{power\ B}$$

8.4.2 Sound Measurements

There are two popular ways to make acoustical measurements: the direct method and the comparison method. The direct method proposes measuring a set of environmental factors like temperature, humidity, viscosity, echo timings, etc. and using them in appropriate relations to compute sound energy levels.

The comparison method is conducted by measuring sound pressure levels of a reference sound of known energy levels and comparing those pressure levels with the sound being measured.

Sound Pressure is a measure of the pressure deviation from atmospheric pressure caused by a sound wave passing through a fixed point, and expressed in a unit called Pascal (1Pa=1 Newton/m^2).

8.4.3 Hearing Threshold and Masking

The minimum threshold is the least audible sound that the normal human ear can detect and hear. The sensitivity of the human ear is however frequency dependent. The minimum threshold values when plotted against frequency values gives rise to the minimum threshold curve.

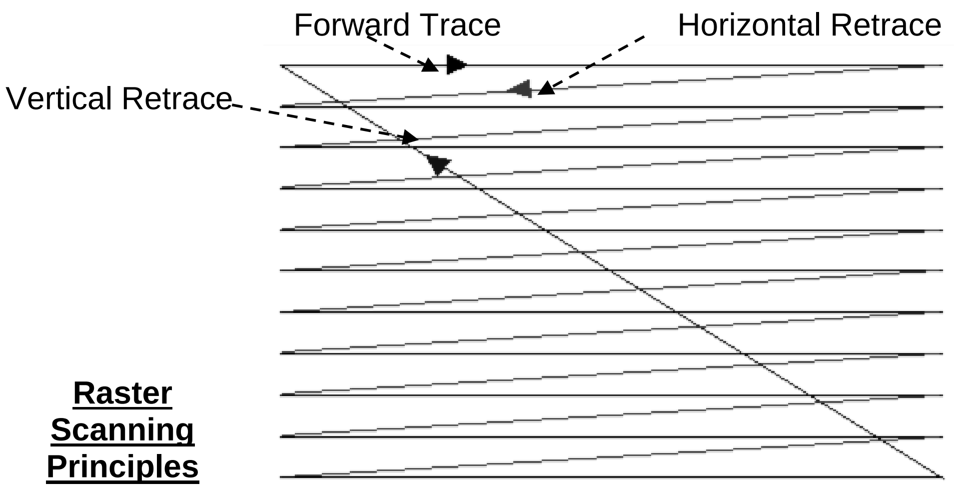
Amplitude masking occurs because an audible sound has a tendency to distort the threshold curve and shift it upwards. Temporal masking occurs when tones are sounded close in time but not simultaneously. A louder tone occurring just before a softer tone makes the latter inaudible.

8.5 RASTER SCANNING PRINCIPLES

To draw an image on the screen the electron beam starts from the upper left corner of the screen and sequentially moves over each pixel row from left to right. This is referred to as **forward trace**. During this phase the electron beam is sometimes switched ON and OFF. When the beam is ON, the pixel over which it is moving gets hit by the electrons and starts glowing. The pixels over which the beams remain OFF remain dark.

At the end of each horizontal line the beam gets switched OFF and retraces diagonally to the beginning of the next row. This phase is called **horizontal retrace**. The switching OFF conserves power and avoids activation of unwanted pixels. At the beginning of the next line, it is again switched ON and begins the next trace.

The process continues until the beam reaches the lower right corner of the screen, after which it is again switched OFF and moves diagonally back to the starting point. This is referred to as **vertical retrace**. The entire process from beginning to end is called raster scanning.



Raster Scanning Principles

8.6 SENSORS FOR TV CAMERA

The television camera is a device that employs light-sensitive image sensors to convert an optical image into a sequence of electrical signals—in other words, to generate the primary components of the picture signal. The first sensors were mechanical spinning disks, based on a prototype patented by the German Paul Nipkow in 1884.

As the disk rotated, light reflected from the scene passed through a series of apertures in the disk and entered a photoelectric cell, which translated the sequence of light values into a corresponding sequence of electric values. In this way the entire scene was scanned, one line at a time, and converted into an electric signal.

Large spinning disks were not the best way to scan a scene, and by the mid-20th century they were replaced by vacuum tubes, which utilized an electron beam to scan an image of a scene that was focused on a light-sensitive surface within the tube.

Electronic camera tubes were one of the major inventions that led to the ultimate technological success of television. Today they have been replaced in most cameras by smaller, cheaper solid-state imagers such as charge-coupled devices. Nevertheless, they firmly established the principle of line scanning (introduced by the Nipkow disks) and thus had a great influence on the design of standards for transmitting television picture signals.

8.7 COLOR FUNDAMENTALS

In general, colors can be reproduced either in self-emitting media (such as displays) or by using reflected or transmitted light (such as in prints and slides).

There are two types of color reproduction: additive and subtractive color reproduction. In an LCD display, red, green and blue filtered sub-pixels are used to reproduce all possible colors for each displayed pixel. Since sub-pixels are very small they are visually unresolved at a typical viewing distance and the light from the LCD light source which is transmitted through them is additively mixed in our retina.

When reflected-light is used, color reproduction is predominantly achieved by subtracting light energy from the source spectrum illuminating the medium.

One example of a subtractive color system is the photographic print that uses a cyan dye that absorbs (i.e. subtracts) long wavelengths corresponding to reddish hues, a magenta dye that absorbs middle wavelengths corresponding to greenish hues, and a yellow dye that absorbs short wavelengths corresponding to bluish hues.

In modern printers, more than three colors may be used (six- or even eight-ink printers are now available and the very latest models by EPSON and Hewlett Packard use 10 or 12 inks to increase the color gamut), although they are still based on a CMY(K) system.

An individual pixel will therefore usually be defined by three (or four, in the case of CMYK) numbers specifying the amount of each primary. These numbers are coordinates in a color space. A color space provides a three-dimensional (usually) model into which all possible colors may be mapped. Color spaces allow us to visualize colors and their relationship to each other.

RGB and CMYK are two broad classes of color space, but there are a range of others, as already encountered, some of which are much more specific and defined than others.

The colorspace defines the axes of the coordinate system and, within this, color gamut of devices and materials may then be mapped; these are the limits to the range of colors capable of being reproduced.

8.7.1 ADDITIVE AND SUBTRACTIVE COLOR MIXING

Additive Colors Combined in Equal Parts		
Blue + Green	=	Cyan
Red + Blue	=	Magenta
Green + Red	=	Yellow
Red + Green + Blue	=	White

Additive Colors Combined in Unequal Parts		
1 Green + 2 Red	=	Orange
1 Red + 2 Green	=	Lime
1 Green + 1 Blue + 4 Red	=	Brown
Subtractive Colors Cyan, Magenta and Yellow		
Combine	**Absorbs**	**Creates**
Blue + Green	Red	Cyan
Red + Blue	Green	Magenta
Green + Red	Blue	Yellow
Combine	**Absorbs**	**Leaves**
Cyan + Magenta	Red + Green	Blue
Cyan + Yellow	Red + Blue	Green
Magenta + Yellow	Green + Blue	Red
Cyan + Magenta + Yellow	Red + Green + Blue	Black

8.8 SUMMARY

- Sound is a form of energy capable of flowing from one place to another through a material medium. It is generated from vibrating objects when part of the energy of vibration is converted to sound energy.

- A sound wave is associated with the following physical characteristics: amplitude, frequency, waveform and speed of propagation.

- Psycho-acoustics is the branch of acoustics which deals with the human auditory perception ranging from the biological design of the ear to the brain's interpretation of aural information.
- A decibel is one tenth of a bel, a seldom-used unit named in honor of Alexander Graham Bell.Power in dB = 10log10 $\frac{power\ A}{power\ B}$
- There are two popular ways to make acoustical measurements: the direct method and the comparison method.
- Sound Pressure is a measure of the pressure deviation from atmospheric pressure caused by a sound wave passing through a fixed point, and expressed in a unit called Pascal (1 Pa = 1 Newton/m2).
- The minimum threshold is the least audible sound that the normal human ear can detect and hear.
- The first sensors were mechanical spinning disks, based on a prototype patented by the German Paul Nipkow in 1884.
- There are two types of color reproduction: additive and subtractive color reproduction.
- In modern printers, more than three colors may be used (six- or even eight-ink printers are now available.
- The colorspace defines the axes of the coordinate system and, within this, color gamut of devices and materials may then be mapped; these are the limits to the range of colors capable of being reproduced.

8.9 UNIT END EXERCISES

1. Discuss the fundamental characteristics of sound.
2. Define decibels. What is meant by saying that a sound has a loudness of 60 dB?
3. Explain the phenomenon of masking using the minimum threshold curve.
4. Explain the raster scanning process. With reference to the same context, explain "The refresh rate of a monitor is 60 Hz".
5. Distinguish between additive and subtractive colors.
6. Write a note on Color Mixing.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia.

❖❖❖❖

9

VIDEO DEVICES

Unit Structure

9.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand working of LCD.
2. Describe construction and working of PDP.
3. Differentiate between different audio / video file formats.

9.2 INTRODUCTION

With digital television becoming the standard of the video industry, new technology is creating a variety of ways to view what is being created. In addition to computers, cell phones, and handheld devices there are new digital display technologies (DDT) being developed.

The original method of displaying video was through the use of the cathode ray tubes (CRT). This type of display is called direct view, with the receiver or monitor being rack-mounted or placed on a stand. The screen size ranges from a few inches up to a fifty-inch picture tube. The newer CRT displays are generally flat screens rather than the curved type that were used in older models.

The newer display types also use projection systems to display an image. Projection systems are available as front-projection and rear-projection types. In a front-projection system, a video projector is used to illuminate a screen in much the same manner as film is displayed.

The screen size can be as large as one hundred inches measured diagonally. Rear-projection systems generally have an internal projector and a mirror to reflect the image on to the viewing screen and are viewed in much the same manner as a direct-view CRT.

The viewing screen is generally forty to eighty inches in size. In the older type of projection systems, both front and rear systems use three CRTs—one red, one blue and one green—that are aligned to create a single image on the viewing screen.

The newer types of displays include Digital Light processing (DLP), liquid crystal display (LCD), liquid crystal on silicon (LCoS), surface conduction electron-emitter display (SED) and plasma.

There is a limit to how large you can make a CRT. As the CRT gets larger and larger, the glass enclosure must get thicker and thicker to maintain strength; this adds a lot of weight to an already heavy set. A new kind of screen is available for showing video that gives a larger viewing area without the size and weight problems of the very large CRTs. The device is only about 4 to 5 inches thick and is sometimes called a flat-screen display, although the proper name is plasma display screen or device.

9.3 LIQUID CRYSTAL DISPLAY (LCD)

LCD stands for liquid crystal display. LCD screens are very complex; this description is therefore going to be very simplified and leave a lot out. The important thing is that you understand the concept of how these screens work.

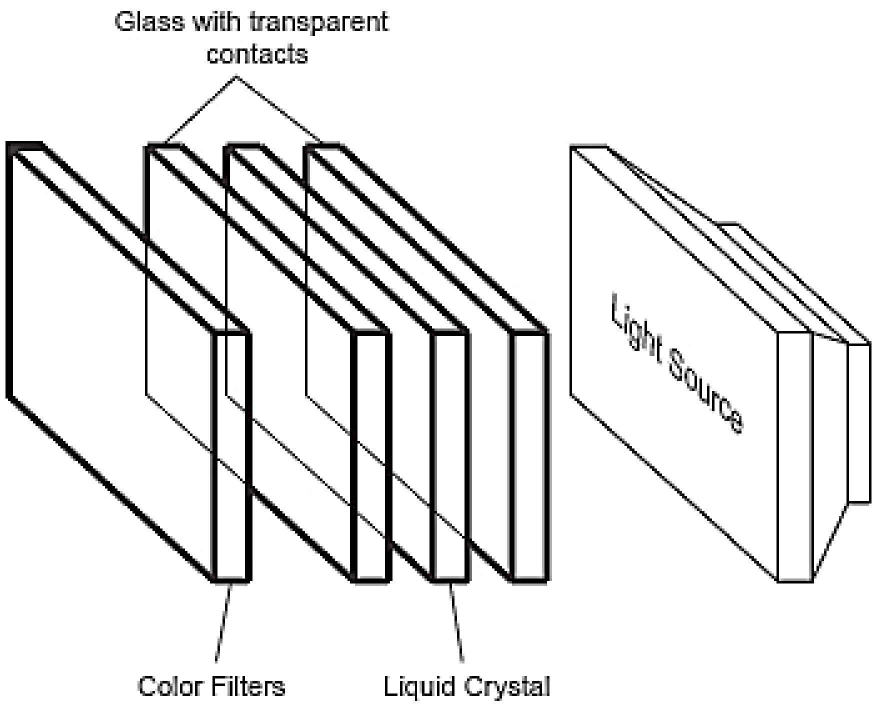
LCD is an unusual substance that has traits of both liquids and solids. The key aspect for us is that when a voltage is applied to the liquid crystal, the crystal part turns within the liquid.

In their normal state, the liquid crystal elements will be flat so that no light can get past them to the colored filters. When an electrical charge is run through them from the glass plate in front to the glass plate behind them, it causes the crystal elements to rotate onto their edges, allowing the light behind them to pass through to the colored filter in front of them. Controlling the electrical charge controls how much the crystals rotate and how much light gets past them. This is how the mix of red, green, and blue is controlled to give us the whole range of colors that we see.

At the time of this writing, LCDs are used for small to medium-sized displays. There are many technical difficulties in producing quality LCD screens as large as some of the current plasma screens. LCD screens are also more expensive than plasma screens of the same size. Engineers are working hard to solve these problems so that one day soon large, high-quality LCD screens will provide the best quality for large-screen viewing. At that time we will be saying goodbye to CRTs and plasma. Both plasma and LCDs provide better pictures that will take up less space in our rooms. CRTs have served us well for the last 60 years, but their days are numbered.

CONSTRUCTION:

An LCD screen is made up of many layers sandwiched together. Among the layers will be a filter layer that has red, green, and blue filters for each pixel element. Then will come a layer of glass that has a transparent contact for each pixel element's color. Then will come the liquid crystal substance. Next will be another glass layer with a transparent contact for each pixel element's color. Finally, there will be a light source — often a fluorescent light.



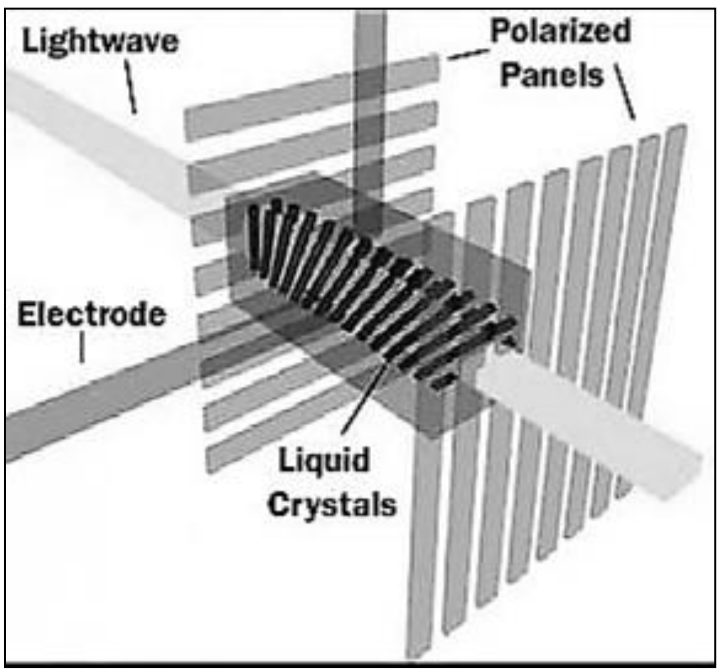
LCD screens are made up of many layers.

WORKING:

By flowing the liquid crystal over a finely grooved surface it is possible to control the alignment of the modules as they follow the alignment of the grooves. A layer of liquid crystal material is placed in a container with two finely grooved surfaces whose grooves are perpendicular to each other and those at the intermediate layers are twisted by intermediate angles. Natural light waves are oriented at random angles and flow along various plane from the light source.

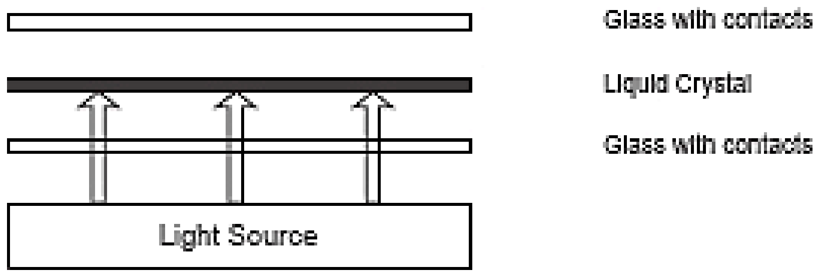
An optical polarizing filter or polarizer can involve a single plane of light from the collection. The filter acts as a net of finely parallel lines blocking all light except those flowing in a parallel to the lines. The light in this condition is said to be polarized. A second polarizer whose lines are perpendicular to the first would block all the polarized light. The container with grooved surfaces is placed in between two perpendicular polarized filters.

Normally light from the first filter would be blocked by the second filter. However in this case, the liquid crystal material placed in between twists the plane of light by 90 degrees as it passes through the material due to its molecular alignments.

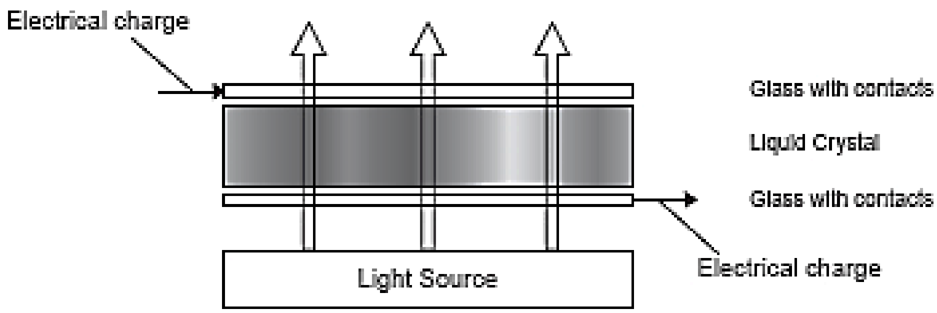


Working of LCD

The light now is parallel to the second filter and comes out wholly through it to the eye of the observer. This constitutes a lighted pixel on the screen. A battery connected across the liquid crystal container generates a current through the liquid crystal, and re-orients its molecules according to the direction of the current flow. This distributes the orderly pattern of the liquid crystal molecules; so that now the molecules at the grooved surfaces are no longer emerge. An observer on the other side of the filter does not see any light coming out. This arrangement creates a dark pixel.



Liquid crystal at rest blocks light from passing through.



Liquid crystal with an electrical charge lets light pass through.

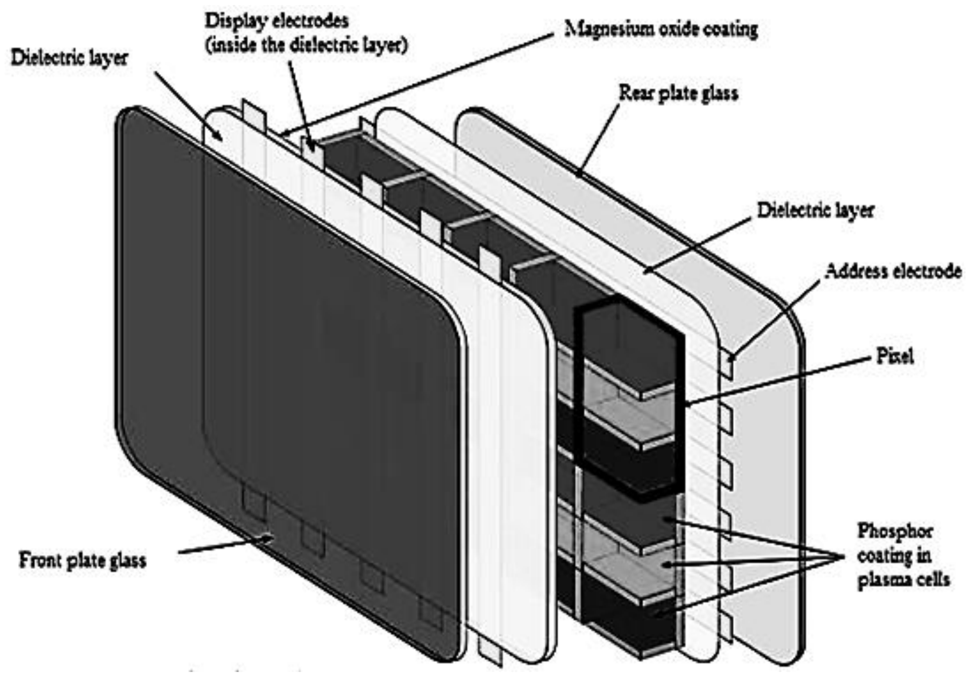
9.4 PLASMA DISPLAY PANEL (PDP)

Plasma is also called ionized gas. Plasma is an energetic gas-phase state of matter, often referred to as "the fourth state of matter", in which some or all of the electrons in the outer atomic orbital have become separated from the atom. The result is a collection of ions and electrons which are no longer bound to each other. Because these particles are ionized (charged), the gas behaves in a different fashion that neutral gas e.g., it can be affected by electromagnetic fields.

Other than CRT and LCD, a third type of display screens often used is known as the plasma Display Panel (PDP). A PDP is an emissive (i.e. discharge of electromagnetic radiation or particles) flat panel display where light is created by phosphor excited by a plasma discharge between two flat panels of gases. The gases used are usually mixture of inert gases like neon and xenon. More over PDP can have large sizes, about 80" diagonally. Plasma displays, like CRT television, are phosphor-based, meaning they rely on glowing phosphor to create their images.

The plasma display screen is made up of two pieces of glass sandwiched together. The back part has ridges going vertically down the glass. In the space between the ridges, alternating columns of red, green, and blue phosphors are laid down. For each color of each pixel there is an electrode on the back that is called the data electrode. The front piece of glass has horizontal ridges across it that will seal off each pixel from the others. The front panel also has two transparent electrodes for each pixel color: a scan electrode and a common electrode.

When the front and back panels are sandwiched together, rare gases (helium, neon, and xenon) are trapped within the pixel compartments. So if we were to look at an individual pixel of the plasma display screen, there would be a red compartment with a trapped rare gas that has a data electrode on the back and transparent scan and common electrodes on the front. There would be a green and a blue compartment set up the same way. The three compartments together would make one pixel.



PDP: Internal Architecture

Construction and working principle:

A plasma panel starts with two sheets of 1/16" glass. Between the sheets of glass are spacers that create their images. Transparent electrodes are then coated inside the front and back pieces of glass (the rear electrode is called "the address electrode"), so that there are two electrodes per cell on the front piece of glass and one electrode per cell on the rear piece of glass. Next each cell is coated with phosphor and filled with a mixture of neon and xenon gases. Note that the entire structure is sealed so that the gas will not escape.

When an electrical charge is added to the front and back electrodes, a voltage differential is created, adding energy to the gas mixture and changing the gas to a plasma state. Once the gas changes into a plasma state, it releases ultraviolet energy which excites the phosphor coated inside each cell, causing it to emit visible light. This is how each cell emits light.

The plasma display is structured so that 3 cells are arrayed side-by-side in close proximity to each other. These 3 cells represent 1 pixel. Each cell within a pixel is phosphor-coated with one of the 3 primary colors of red, green and blue. All other colors that we see are made up of a combination of the 3 primary colors. After activation, a relatively low voltage is required to maintain the plasma state.

9.5 FILE FORMATS

There are several audio file formats in common use. There are fewer video file formats, but audio video file extensions can be confusing. This is just a brief description of what the more commonly used audio and video file formats and systems are.

A video file format is a file format for storing digital video data on a computer system. Video is almost always stored in compressed form to reduce the file size.

A video file normally consists of a container format (e.g. Matroska) containing video data in a video coding format (e.g. VP9) alongside audio data in an audio coding format (e.g. Opus). The container format can also contain synchronization information, subtitles, and metadata such as title etc..

A standardized (or in some cases de facto standard) video file type such as .webm is a profile specified by a restriction on which container format and which video and audio compression formats are allowed.

The coded video and audio inside a video file container (i.e. not headers, footers and metadata) is called the essence. A program (or hardware) which can decode video or audio is called a codec; playing or encoding a video file will sometimes require the user to install a codec library corresponding to the type of video and audio coding used in the file.

Good design normally dictates that a file extension enables the user to derive which program will open the file from the file extension. That is the case with some video file formats, such as WebM (.webm), Windows Media Video (.wmv), and Ogg Video (.ogv), each of which can only contain a few well-defined subtypes of video and audio coding formats, making it relatively easy to know which codec will play the file.

In contrast to that, some very general-purpose container types like AVI (.avi) and Quicktime (.mov) can contain video and audio in almost any format, and have file extensions named after the container type, making it very hard for the end user to use the file extension to derive which codec or program to use to play the files.

It is important to distinguish between the audio coding format, the container containing the raw audio data, and an audio codec. A codec performs the encoding and decoding of the raw audio data while this encoded data is (usually) stored in a container file. Although most audio file formats support only one type of audio coding data (created with an audio coder), a multimedia container format (asMatroska or AVI) may support multiple types of audio and video data.

There are three major groups of audio file formats:

- Uncompressed audio formats, such as WAV, AIFF, AU or raw header-less PCM;

- Formats with lossless compression, such as FLAC, Monkey's Audio (filename extension .ape), Wav Pack (filename extension .wv), TTA, ATRAC Advanced Lossless, Apple Lossless (filename extension .m4a), MPEG-4 SLS, MPEG-4 ALS, MPEG-4 DST, Windows Media Audio Lossless (WMA Lossless), and Shorten (SHN).

- Formats with lossy compression, such asMP3, Vorbis, Musepack, AAC, ATRAC and Windows Media Audio Lossy (WMA lossy).

Here are some common audio/ video file formats:

AAC: [Advanced Audio Coding]:

This is the audio file format used by Apple for the iTunes Music Store, and it may appear with the M4A filename extension. It is better than MP3 for sound quality. It was developed as part of the MPEG4 group owned by Dolby.

AU:

This audio file format is the standard used by Java, Sun and UNIX.

MPEG: [Moving Pictures Experts Group].

There are a number of MPEG, discussed below.

MPEG-1:

This is used in digital cameras and camcorders for small video clips. VHS quality playback can be expected from MPEG-1.

MPEG-2:

Used for digital satellite TV, professional movie recording and recording of home DVD recordings. Provides provision for multi-channel surround sound recordings.

MPEG-3:

MPEG-3 was propose as an entity, but eventually merged into MPEG-2.

MPEG-4:

This is the latest MPEG system and is used for streaming internet content. It is also used in portable video recorders and for internet downloads. Required for DivX. It improves digital broadcasting and interactive graphics and multimedia.

MP3:

Digital audio files, most commonly used to store and playback music. It compresses the files to about 10% of a normal audio file, and a normal music track will be about 5 -6 MB in size. MP3 stands for MPEG-1 Audio Layer 3, not MPEG-3 as many people think. A typical MP3 audio file is near CD quality.

OGG:

"Ogg" is derived from "ogging", jargon from the computer game Netrek, which came to mean doing something forcefully, possibly without consideration of the drain on future resources. An audio file format supporting a variety of codecs, the most popular of which is the audio codec Vorbis. However, MP3 files are much more broadly supported than Vorbis.

RA: [Real Audio].

This format is designed for streaming audio over the Internet. It is a self-contained file format with all the audio information stored within the file itself.

WAV:

The simplest of the audio file formats, developed by Microsoft and IBM, and built into Windows 95. It is an uncompressed audio file format with large file sizes (10 x MP3), and does not need further processing to play. The WAV file consists of three blocks of information: The RIFF block which identifies the file as a WAV file, The FORMAT block which identifies parameters such as sample rate and the DATA block which contains the actual data, or music sample.

WMA: [Windows Media Audio].

A digital system invented by Microsoft, and is used in portable digital audio players. Using WMA, a file can be programmed so that it cannot be copied, and can be used to protect copyright.

WMF: [Windows Media Format].

These are audio-video files comprising WMA and video codecs. They provide high quality and media security for streaming and download and play applications on computers.

WMV: [Windows Media Video]

Used in the Windows media Player, this is used to stream and download and play audio and video content.

When dealing with audio and video file formats, you will sometimes notice the term 'codec'. A codec is simply short for encoder-decoder (or compressor - decompressor).

The main function of a codec is to compress audio or video data streams so that transmission of digital audio samples and video frames can be speeded up and storage space reduced. The objective of all codecs is to reduce the file size to a minimum while maintaining audio and video quality.

A quick indication of the codec's place in the path of transmission and reception is:

Video device (e.g. camcorder) ➔ video capture card ➔ video digitized ➔ codec (compresses digital info) ➔result (MPEG2, AVI, WMV etc.) ➔codec (decompress) ➔video frames ➔display device.

Between the two codecs the compressed result is transferred to the display device transmitted, stored on file, etc.). So to condense the flow even further, we could basically describe it as:

Raw data ➔codec ➔transmit ➔codec ➔play

This is simplistic, but it shows where the codecs are used. Therefore, in order to play a movie, video or piece of music of a certain format, you need a codec in your computer to allow you to decompress the file and play it.

9.6 SUMMARY

- LCD stands for liquid crystal display.
- By flowing the liquid crystal over a finely grooved surface it is possible to control the alignment of the modules as they follow the alignment of the grooves.
- PDP stands for Plasma Display Panel.
- Plasma is also called ionized gas. Plasma is an energetic gas-phase state of matter, often referred to as “the fourth state of matter”, in which some or all of the electrons in the outer atomic orbital have become separated from the atom.
- The plasma display is structured so that 3 cells are arrayed side-by-side in close proximity to each other. These 3 cells represent 1 pixel.
- A video file format is a file format for storing digital video data on a computer system.

9.7 UNIT END EXERCISES

1. What is polarized light and a polarizing filter? How are these utilized within an LCD monitor?
2. What is plasma? Explain the construction and working of a Plasma Display Panel [PDP].
3. Distinguish between the audio coding format, the container containing the raw audio data, and an audio codec.
4. Discuss several audio-video file formats.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia

❖❖❖❖

10

UNIT V

INTRODUCTION TO COMPRESSION

Unit Structure

10.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Define and describe compression.
2. Understand importance of compression techniques.
3. Distinguish between types of compression

10.2 INTRODUCTION

Compression may refer to:
Data compression, the process of encoding digital information using fewer bits

Audio compression (data), the compression of digital audio streams and files

Bandwidth compression, a reduction in either the time to transmit or in the amount of bandwidth required to transmit.

Compression artifact, noticeable defects in audio or video that has been compressed.

Image compression, the application of data compression on digital images.

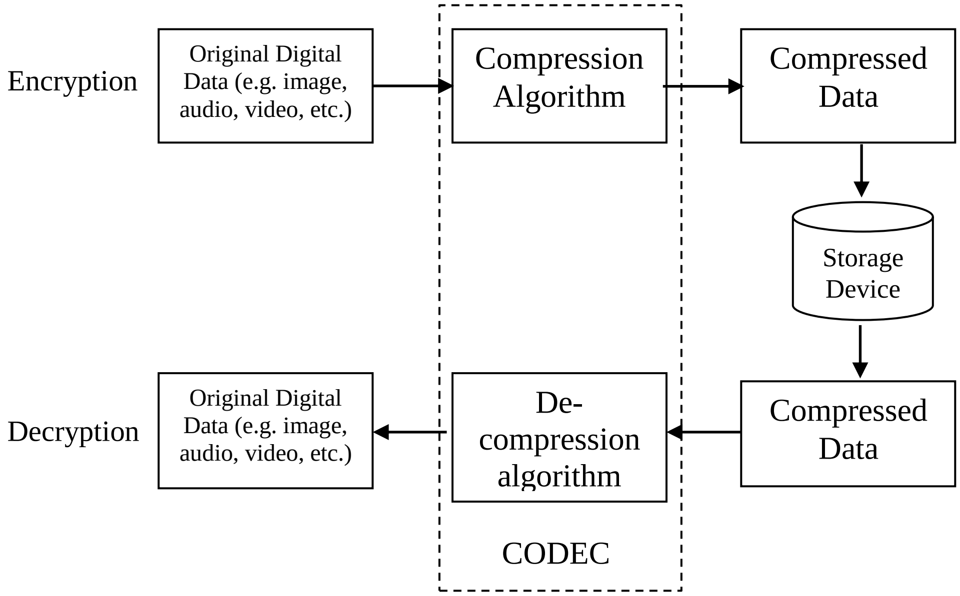
Video compression, the compression of digital video streams and files.

One-way compression function, a cryptographic primitive.

Dynamic range compression, a compression process that reduces the dynamic range of an audio signal.

10.3 WHAT IS COMPRESSION ?

Compression reduces file size of media like text, image, audio, video etc. to a process called CODEC. The amount of compression which is to be achieved depends on original media data and as well as the compression technique applied.



Working of a CODEC

10.4 NEED FOR COMPRESSION

A typical five page document file occupies 100kb (approximate) space, whereas a single image can take about 2MB. A 1 minute audio file can take up to 50MB (approximate) and 1 minute video clip may take 1GB space.

Compression is needed to overcome the following limitations:

1. Due to huge sizes, storage requirement increases rapidly thereby increasing the cost.

2. Even if there is adequate storage space, such file requires large data transfer rate, increasing the load on HDD and Processor.

Compression is useful because it helps reduce resource usage, such as data storage space or transmission capacity. Because compressed data must be decompressed to use, this extra processing imposes computational or other costs through decompression; this situation is far from being a free lunch.

Data compression is subject to a space–time complexity trade-off. For instance, a compression scheme for video may require expensive hardware for the video to be decompressed fast enough to be viewed as it is being decompressed, and the option to decompress the video in full before watching it may be inconvenient or require additional storage.

The design of data compression schemes involves trade-offs among various factors, including the degree of compression, the amount of distortion introduced (e.g., when using lossy data compression), and the computational resources required to compress and uncompress the data.

10.5 TYPES OF COMPRESSION

LOSSLESS V/S LOSSY COMPRESSION

Compression can be broadly divided into two types: lossy and lossless. Lossy compression implies the original data changed permanently during compression unlike lossless compression. CODECS are used for lossless compression to represent the existing information in a more compact form without actually discarding any data.

The advantage of lossless compression is the original data stays intact without degradation of quality. E.g.:-medical images like X-ray plates and ultra sonograph. In lossy compression, parts of original data are completely discarded.

Lossy compression is generally used where media quality may be sacrificed to a certain extent for reducing space requirements like in multimedia presentation and web page content.

INTRAFRAME V/S INTERFRAME COMPRESSION

Intraframe compression is applicable within a still image or a single video frame. In this technique spatial redundancies are detected and exploited to reduce the file size. Such redundancies occur when different portions of an image are identical.

Interframe compression exploits the redundancy between adjacent frames in a video sequence, referred to as temporal redundancy.

10.6 BASIC COMPRESSION TECHNIQUES

RUN LENGTH ENCODING:

Run-length encoding (RLE) is used to reduce the size of a repeating string of characters. The repeating string is referred to as a run. It can compress any type of data regardless of its information content. However, content of data affects the compression ratio. Compression ratio, in this case, is not so high. But it is easy to implement and quick to execute. Typically RLE encodes a run of symbols into two bytes, a count and a symbol.

RLE was developed in the 1950s and became, along with its 2-D extensions, the standard compression technique in facsimile (FAX) coding. FAX is a two-color (black and white) image which is predominantly white. If these images are sampled for conversion into digital data, many horizontal lines are found to be entirely white (long runs of 0's). Besides, if a given pixel is either black or white, the possibility that the next pixel will match is also very high.

The code for a fax machine is actually a combination of a Huffman code and a run-length code. The coding of run-lengths is also used in CCITT, JBIG2, JPEG, M-PEG, MPEG-1/2/4, BMP, etc.

SCENARIO 1:

Example 3.11: Consider the following bit stream:

11111111111111110000000000000000000011

Find the run-length code and its compression ratio.

Solution: The stream can be represented as: sixteen 1's, twenty 0's and two 1's, i.e., (16, 1), (20, 0), (2, 1). Since the maximum number of repetitions is 20, which can be represented with 5 bits, we can encode the bit stream as (10000,1), (10100,0), (00010,1).

The compression ratio is 18:38 = 1:2.11.

SCENARIO 2:

In this method any sequence of repetitive characters may be replaced by a more compact form. A series of *n* successive characters may be replaced be a single instance of the character and the number of occurrences. To depict that the number has special meaning and not part of the normal text a special character is used as a flag.

E.g.: uncompressed data: UNNNNNNIMANNHEIM compressed data: U!6NIMANNEHEIM

This method should only be used if the number of occurrences is equal to or more than a specific number, in this case 4.

Run-length encoding performs lossless data compression and is well suited to palette-based bitmapped images such as computer icons. It does not work well at all on continuous-tone images such as photographs, although JPEG uses it quite effectively on the coefficients that remain after transforming and quantizing image blocks. Common formats for run-length encoded data include Truevision TGA, PackBits, PCX and ILBM.

HUFFMAN CODING:

In computer science and information theory, a Huffman code is an optimal prefix code found using the algorithm developed by David A. Huffman while he was a Ph.D. student at MIT, and published in the 1952 paper "*A Method for the Construction of Minimum-Redundancy Codes*".

The process of finding and/or using such a code is called Huffman coding and is a common technique in entropy encoding, including in lossless data compression. The algorithm's output can be viewed as a variable-length code table for encoding a source symbol (such as a character in a file). Huffman's algorithm derives this table based on the estimated probability or frequency of occurrence (weight) for each possible value of the source symbol.

As in other entropy encoding methods, more common symbols are generally represented using fewer bits than less common symbols. Huffman's method can be efficiently implemented, finding a code in linear time to the number of input weights if these weights are sorted. Although optimal among methods encoding symbols separately, Huffman coding is not always optimal among all compression methods.

10.7 SUMMARY

- Compression reduces file size of media like text, image, audio, video etc. to a process called CODEC.
- Compression is needed to overcome the following limitations: Due to huge sizes, storage requirement increases rapidly thereby increasing the cost. Even if there is adequate storage space, such file requires large data transfer rate, increasing the load on HDD and Processor.
- Compression can be broadly divided into two types: lossy and lossless.

- Lossy compression implies the original data changed permanently during compression unlike lossless compression.

- The advantage of lossless compression is the original data stays intact without degradation of quality.

- Intraframe compression is applicable within a still image or a single video frame.

- Interframe compression exploits the redundancy between adjacent frames in a video sequence, referred to as temporal redundancy.

- Run-length encoding (RLE) is used to reduce the size of a repeating string of characters.

- Run-length encoding performs lossless data compression and is well suited to palette-based bitmapped images such as computer icons.

- Huffman's method can be efficiently implemented, finding a code in linear time to the number of input weights if these weights are sorted.

10.8 UNIT END EXERCISES

1. What is meant by compression and CODEC? Why is compression needed?
2. Distinguish between lossy and lossless compression techniques.
3. Distinguish between intraframe and interframe compression techniques.
4. What is meant by Run Length Encoding?

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia

❖❖❖❖

11

ADVANCED COMPRESSION TECHNIQUES

Unit Structure

11.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand advanced compression techniques.
2. Describe JPEG coding
3. Describe ZIP coding
4. Understand other coding techniques

11.2 INTRODUCTION

A high-definition uncompressed video data stream requires about 2 billion bits per second of data bandwidth. Owing to the large amount of data necessary to represent digital video, it is desirable that such video signals are easy to compress and decompress, to allow practical storage or transmission.

The term data compression refers to the reduction in the number of bits required to store or convey data—including numeric, text, audio, speech, image, and video—by exploiting statistical properties of the data. Fortunately, video data is highly compressible owing to its strong vertical, horizontal, and temporal correlation and its redundancy.

Transform and prediction techniques can effectively exploit the available correlation, and information coding techniques can take advantage of the statistical structures present in video data.

These techniques can be lossless, so that the reverse operation (decompression) reproduces an exact replica of the input.

In addition, however, lossy techniques are commonly used in video data compression, exploiting the characteristics of the HVS, which is less sensitive to some color losses and some special types of noises.

Video compression and decompression are also known as video encoding and decoding, respectively, as information coding principles are used in the compression and decompression processes, and the compressed data is presented in a coded bit stream format.

11.3 JPEG CODING STANDARD

The JPEG (Joint Photographic Experts Group) became an international standard in 1992.

In the sequential mode JPEG compression process is composed of following steps:

- Preparation of data
- Source encoding steps involving forward DCT and quantization
- Entropy encoding steps involving RLE and Huffman encoding

DECOMPRESSION PROCESS

- Entropy decoding steps involving RLD and Huffman decoding
- Source decoding steps involving inverse DCT and de-quantization

11.3.1 BLOCK PREPARATION

An image is represented by 1 or more 2D array of pixel values these blocks are in preparation of next steps where DCT is applied to each block instead of the entire image.

11.3.2 DISCRETE COSINE TRANSFORM (DCT)

The objective of this is to transform each block from the spatial domain to the frequency domain. We know that the synthesis equation of the DFT is given by the relation.

$$x[i] = \sum_{k=0}^{N/2} \mathrm{Re}\,\overline{X}[k]\cos(2\pi ki/N) + \sum_{k=0}^{N/2} \mathrm{Re}\,\overline{X}[k]\sin(2\pi ki/N)$$

where *N* is the total number of samples, *k* is the variable, *i* is the variable indicating the number of input sample considered in the time or(space) domain, $x[i]$ is the actual time(or space) domain signal, $\operatorname{Re}\overline{X}[k]$ is proportional to the k-th entry and the frequency domain. If we ignore the imaginary components and consider the entire time (or space) domain signal to be compiled of real numbers we get

$$x[i] = \sum_{k=0}^{N-1} \operatorname{Re}\overline{X}[k] \cos(2\pi ki / N)$$

The above expression is called the synthesis equation of a one dimensional DCT, also known as inverse DCT. The forward DCT is just the reverse of the above relation, where we express the frequency domain in terms of time (or space) domain signal. Mathematically it can be expressed as

$$\operatorname{Re} X[k] = \sum_{i=0}^{N-1} x[i] \cos[\pi i(k + 1/2) / N)$$

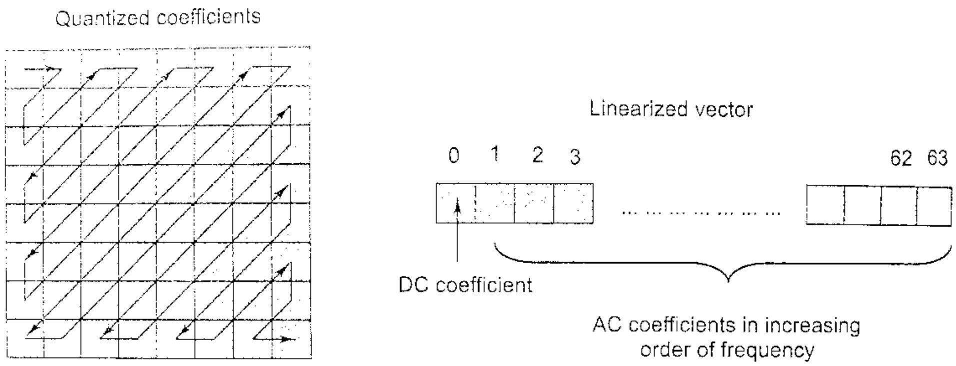
$$or,$$

$$\operatorname{Re} X[k] = \sum_{i=0}^{N-1} x[i] \cos[\pi i(2k + 1) / 2N)$$

The expression for 2D forward DCT is written as

$$F[i, j] = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P[x, y] \cdot \cos[\pi i(2x + 1) / 2N] \cdot \cos[\pi j(2y + 1) / 2N]$$

11.3.3 QUANTIZATION



The next step in the process is quantizing the coefficient numbers that were derived from the luminance and chrominance values by the DCT. Quantizing is basically the process of rounding off the numbers. This is where the file compression comes in. How much the file is compressed depends on the quantization matrix.

The quantization matrix defines how much the information is compressed by dividing the coefficients by a quantizing factor. The larger the number of the quantizing factor, the higher the quality (therefore, the less compression). This is basically what is going on in Photoshop when you save as JPEG and the program asks you to set the quality; you are simply defining the quantizing factor.

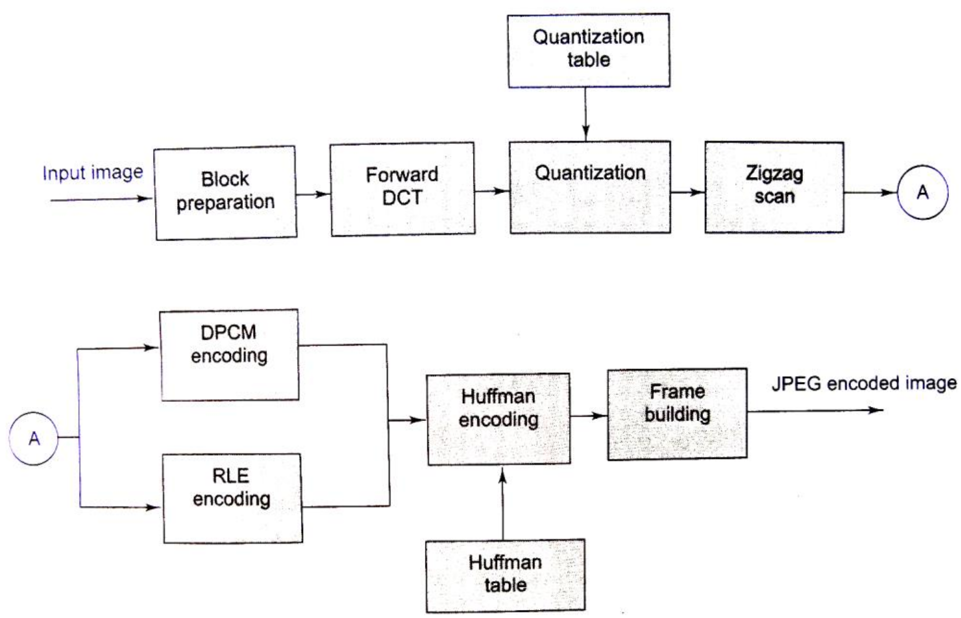
Once the numbers are quantized, they are run through a binary encoder that converts the numbers to the ones and zeros computers love so well. You now have a compressed file that is on average about one-fourth of the size of an uncompressed file.

11.3.4 ZIGZAG SCAN

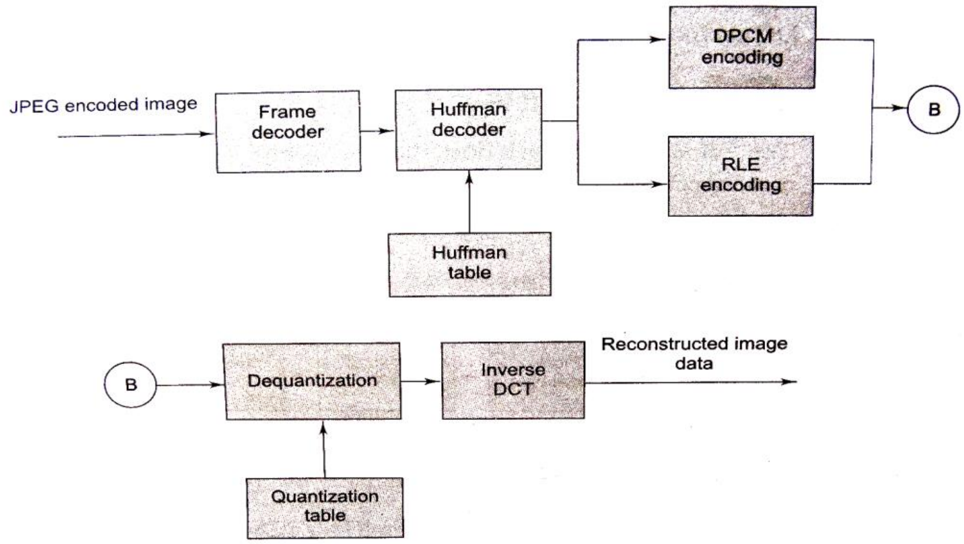
After the DCT stage, the remaining stages involve entropy encoding. The entropy coding algorithms operate on one dimensional string of values i.e. a vector. The output of the quantization stage is a 2D array, hence to apply an entropy scheme, the array is to be converted to a 1D vector. This operation is known as vectoring.

11.3.5 DPCM encoding

There is one DC co-efficient per block because of the small physical area covered by each block, the DC co-efficient varies from one block to the next. To exploit this similarity, the sequence of DC co-efficient is encoded in DPCM mode. This means the difference between the DC co-efficient of each block and the adjacent block is computed and stored.



JPEG Encoder



JPEG Decoder

11.4 ZIP CODING

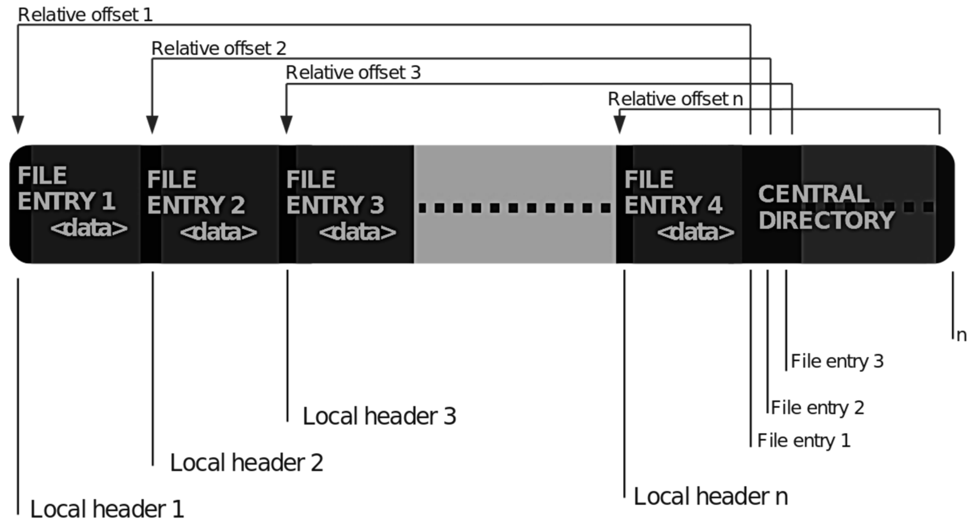
.ZIP files are archives that store multiple files. .ZIP allows contained files to be compressed using many different methods, as well as simply storing a file without compressing it. Each file is stored separately, allowing different files in the same archive to be compressed using different methods. Because the files in a .ZIP archive are compressed individually it is possible to extract them, or add new ones, without applying compression or decompression to the entire archive.

This contrasts with the format of compressed tar files, for which such random-access processing is not easily possible.

A directory is placed at the end of a .ZIP file. This identifies what files are in the .ZIP and identifies where in the .ZIP that file is located. This allows .ZIP readers to load the list of files without reading the entire .ZIP archive. .ZIP archives can also include extra data that is not related to the .ZIP archive.

This allows for a .ZIP archive to be made into a self-extracting archive (application that decompresses its contained data), by prepending the program code to a .ZIP archive and marking the file as executable. Storing the catalog at the end also makes possible hiding a zipped file by appending it to an innocuous file, such as a GIF image file.

The .ZIP format uses a 32-bit CRC algorithm and includes two copies of the directory structure of the archive to provide greater protection against data loss.



Technical representation: ZIP Compression

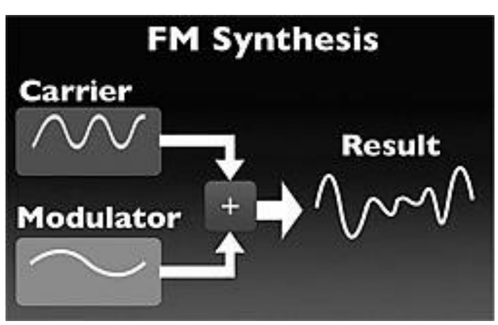
11.5 OVERVIEW OF AUDIO VIDEO COMPRESSIONS

FM Synthesis

In frequency modulation synthesis (or FM synthesis), the timbre of simple waveform (usually a sine wave) called a carrier, is changed by modulating it with another waveform in the audio range (higher than 20Hz). The result is a complex waveform with a different tone.

Video-game developers used FM synthesis in the past because the sounds created by FM synthesis took up less space than those created through PCM.

Nowadays FM synthesis is commonly found in modern DAWs as a plug-in. The sound is also considered ‘retro’ and is still used in certain specialized ‘tracker’ applications to reproduce the sound of a well-loved era of gaming.



A simplification of FM Synthesis, originally developed by John Chowning at Stanford, and patented by Yamaha.

MIDI

The Musical Instrument Digital Interface, or MIDI, is one the most innovative and far reaching standards developed in the last 30 years and is still heavily in use today.

It is an industry protocol enabling different manufacturers to control and trigger events (musical or otherwise) by sending messages back and forth on a serial connection. Keep in mind MIDI isn't audio but it can certainly control audio.

A MIDI interface and MIDI cables. MIDI data can also be used internally and transmitted wirelessly.

Games use MIDI like a player-piano roll to trigger sounds and music passages. When used in conjunction with FM synthesis, for example, MIDI can trigger a significant amount of music and sound for very little file size.
Recently, MIDI achieved industry recognition by winning a technical Grammy.

Dave Smith, the polyphonic synthesizer pioneer, and Ikutaro Kakehashi, a Yamaha engineer, were both honored for the creation of this amazing language spoken by synthesizers, samplers, drum machines, computers and sequencers and much more, that enabled digital music to advance to never-before-seen levels of sophistication.

VIDEO CODING STANDARDS

International video coding standards are defined by committees of experts from organizations like the International Standards Organization (ISO) and the International Telecommunications Union (ITU). The goal of this standardization is to have common video formats across the industry, and to achieve interoperability among different vendors and video codec related hardware and software manufacturers.

The standardization of algorithms started with image compression schemes such as JBIG (ITU-T Rec. T82 and ISO/IEC 11544, March 1993) for binary images used in fax and other applications, and the more general JPEG (ITU-T Rec. T81 and ISO/IEC 10918-1), which includes color images as well. The JPEG standardization activities started in 1986, but the standard was ratified in 1992 by ITU-T and in 1994 by ISO.

The main standardization activities for video compression algorithms started in the 1980s, with the ITU-T H.261, ratified in 1988, which was the first milestone standard for visual telecommunication. Following that effort, standardization activities increased with the rapid advancement in the television, film, computer, communication, and signal processing fields, and with the advent of new usages requiring contributions from all these diverse industries. These efforts subsequently produced MPEG-1, H.263, MPEG-2, MPEG-4 Part 2, AVC/H.264, and HEVC/H.265 algorithms. In the following sections, we briefly describe the major international standards related to image and video coding.

MPEG-4

MPEG-4, formally the standard ISO/IEC 14496, was ratified by ISO/IEC in March 1999 as the standard for multimedia data representation and coding. In addition to video and audio coding and multiplexing, MPEG-4 addresses coding of various two- or three-dimensional synthetic media and flexible representation of audio-visual scene and composition.

As the usage of multimedia developed and diversified, the scope of MPEG-4 was extended from its initial focus on very low bit-rate coding of limited audio-visual materials to encompass new multimedia functionalities.

Unlike pixel-based treatment of video in MPEG-1 or MPEG-2, MPEG-4 supports content-based communication, access, and manipulation of digital audio-visual objects, for real-time or non-real-time interactive or non-interactive applications. MPEG-4 offers extended functionalities and improves upon the coding efficiency provided by previous standards. For instance, it supports variable pixel depth, object-based transmission, and a variety of networks including wireless networks and the Internet.

Multimedia authoring and editing capabilities are particularly attractive features of MPEG-4, with the promise of replacing existing word processors. In a sense, H.263 and MPEG-2 are embedded in MPEG-4, ensuring support for applications such as digital TV and videophone, while it is also used for web-based media streaming.

MPEG-4 distinguishes itself from earlier video coding standards in that it introduces object-based representation and coding methodology of real or virtual audio-visual (AV) objects. Each AV object has its local 3D+T coordinate system serving as a handle for the manipulation of time and space. Either the encoder or the end-user can place an AV object in a scene by specifying a co-ordinate transformation from the object's local co-ordinate system into a common, global 3D+T co-ordinate system, known as the scene co-ordinate system.

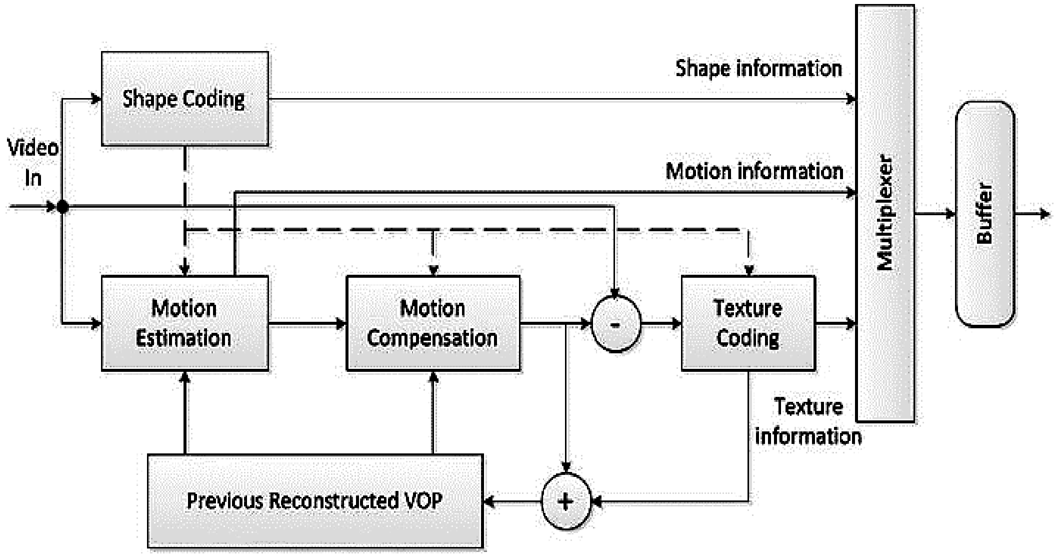
The composition feature of MPEG-4 makes it possible to perform bit stream editing and authoring in compressed domain.

One or more AV objects, including their spatio-temporal relationships, are transmitted from an encoder to a decoder. At the encoder, the AV objects are compressed, error-protected, multiplexed, and transmitted downstream.

At the decoder, these objects are demultiplexed, error corrected, decompressed, composited, and presented to an end user. The end user is given an opportunity to interact with the presentation. Interaction information can be used locally or can be transmitted upstream to the encoder.

The transmitted stream can either be a control stream containing connection setup, the profile (subset of encoding tools), and class definition information, or be a data stream containing all other information.

Control information is critical, and therefore it must be transmitted over reliable channels; but the data streams can be transmitted over various channels with different quality of service.



Video Object encoder structure: MPEG-4

11.6 SUMMARY

- The JPEG (Joint Photographic Experts Group) became an international standard in 1992.
- The larger the number of the quantizing factor, the higher the quality.
- .ZIP allows contained files to be compressed using many different methods, as well as simply storing a file without compressing it.
- In frequency modulation synthesis (or FM synthesis), the timbre of simple waveform (usually a sine wave) called a carrier, is changed by modulating it with another waveform in the audio range.
- MPEG-4, formally the standard ISO/IEC 14496, was ratified by ISO/IEC in March 1999 as the standard for multimedia data representation and coding.
- Unlike pixel-based treatment of video in MPEG-1 or MPEG-2, MPEG-4 supports content-based communication, access, and manipulation of digital audio-visual objects, for real-time or non-real-time interactive or non-interactive applications.

11.7 UNIT END EXERCISES

1. Write a note on ZIP Coding.
2. Write a note on JPEG Compression Technique.
3. What is Discrete Cosine Transformation?
4. Diagrammatically, explain JPEG Encoder – Decoder.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia

❖❖❖❖

12

UNIT VI

MULTIMEDIA AUTHORING

Unit Structure

12.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand types of multimedia authoring metaphor.
2. Describe stages of Multimedia Production.

12.2 INTRODUCTION

In the best case, you must be prepared to choose the tool that best fits the job; in the worst case, you must know which tools will at least “get the job done.” Authoring tools are constantly being improved by their makers, who add new features and increase performance with upgrade development cycles of six months to a year.

It is important that you study the software product reviews in the blogs and computer trade journals, as well as talk with current users of these systems, before deciding on the best ones for your needs. The following sections describe what to look for.

12.3 OVERVIEW

Multimedia authoring tools provide the important framework you need for organizing and editing the elements of your multimedia project, including graphics, sounds, animations, and video clips.

Authoring tools are used for designing interactivity and the user interface, for presenting your project on screen, and for assembling diverse multimedia elements into a single, cohesive product.

Authoring software provides an integrated environment for binding together the content and functions of your project, and typically includes everything you need to create, edit, and import specific types of data; assemble raw data into a playback sequence or cue sheet; and provide a structured method or language for responding to user input.

With multimedia authoring software, you can make

- Video productions
- Animations
- Games
- Interactive web sites
- Demo disks and guided tours
- Presentations
- Kiosk applications
- Interactive training
- Simulations, prototypes, and technical visualizations

Each multimedia project you undertake will have its own underlying structure and purpose and will require different features and functions-learning modules such as those seen on PDAs, MP3 players, and intra-college networks may include web-based teaching materials, multimedia CD-ROMs or web sites, discussion boards, collaborative software, wikis, simulations, games, electronic voting systems, blogs, computer-aided assessment, simulations, animation, blogs, learning management software, and e-mail.

This is also referred to as distance learning or blended learning, where online learning is mixed with face-to-face learning.

The following are several authoring tools for e-learning:

- Adobe Captivate (www.adobe.com/products/captivate)
- Adobe Presenter (www.adobe.com/products/presenter)
- Articulate Storyline and Articulate Studio (www.articulate.com)
- Composica (www.composica.com)
- Claro dominKnow (www.dominknow.com)
- Easygenerator (www.easygenerator.com)
- GoAnimate (http://goanimate.com)

- Harbinger Raptivity (www.raptivity.com)
- iSpring Presenter (www.ispringsolutions.com)
- Lectora Inspire (http://lectora.com)
- Qarbon Viewlet Builder (www.qarbon.com)
- Skilitics Interact and Skilitics Thrive (www.skilitics.com)
- SmartBuilder (www.smartbuilder.com)
- TechSmith Camtasia Studio (www.techsmith.com)
- ZebraZapps (www.zebrazapps.com)

12.4 MULTIMEDIA AUTHORING METAPHOR

Three metaphors are used by authoring tools that make multimedia, based on the method used for sequencing or organizing multimedia elements and events:

- Card- or page-based tools
- Icon- or object-based, event-driven multimedia and game-authoring tools
- Time-based tools

Card- or Page-Based Authoring Tools

Card-based or page-based tools are authoring systems wherein the elements are organized as a stack of cards or pages of a book, respectively. Thousands of pages or cards may be available in the book or stack. These tools are best used when the bulk of your content consists of elements that can be viewed individually, letting the authoring system link these pages or cards into organized sequences.

You can jump, on command, to any page you wish in the structured navigation pattern.

Page- or card-based authoring systems such as Live Code (www.livecode.com) contain media objects: buttons, text fields, graphic objects, backgrounds, pages or cards, and even the project itself. The characteristics of objects are defined by properties (highlighted, bold, red, hidden, active, locked, and so on). Each object may contain a programming script, usually a property of that object, activated when an event (such as a mouse click) related to that object occurs.

Events cause messages to pass along the hierarchy of objects in the project; for example, a mouse-clicked message could be sent from a button to the background, to the page, and then to the project itself. As the message travels, it looks for handlers in the script of each object; if it finds a matching handler, the authoring system then executes the task specified by that handler.

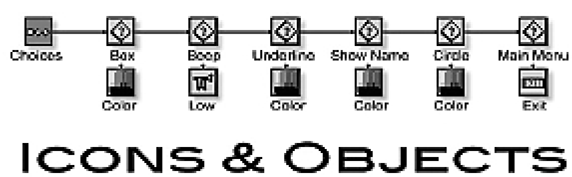
Following are some typical messages that might pass along the object hierarchy of the Live Code authoring system: close Card, close Stack, idle, mouseDown, mouse Still Down, mouseUp, new Background, open Card, open Stack. Now let's look at specific examples. To go to the next card or page when a button is clicked, place a message handler into the script of that button.

Consider the handler, if placed in the script of the card or page, executes its commands when it receives a mouseUp event message that occurs at any location on the card or page—not just while the cursor is within the bounds of a button.

Card- or page-based systems typically provide two separate layers on each card or page: a background layer that can be shared among many cards or pages, and a foreground layer that is specific to a single card or page.

Icon- or Object-Based Authoring Tools

Icon-based or object-based authoring tools are event-driven authoring systems wherein multimedia elements and interaction cues (events) are organized as objects in a structural framework or process. Icon- or object-based, event-driven tools simplify the organization of your project and typically display flow diagrams of activities along branching paths. In complicated navigational structures, this charting is particularly useful during development.



Icon-based, event-driven tools provide a visual programming approach to organizing and presenting multimedia. First you build a structure or flowchart of events, tasks, and decisions, by dragging appropriate icons from a library. These icons can include menu choices, graphic images, sounds, and computations.

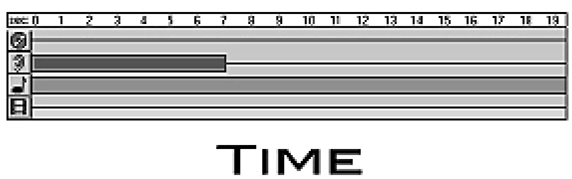
The flowchart graphically depicts the project's logic. When the structure is built, you can add your content: text, graphics, animation, sounds, and video movies. Then, to refine your project, you edit your logical structure by rearranging and fine-tuning the icons and their properties.

With icon-based authoring tools, non-technical multimedia authors can build sophisticated applications without scripting. By placing icons on a flow line, you can quickly sequence events and activities, including decisions and user interactions. These tools are useful for storyboarding, as you can change sequences, add options, and restructure interactions by simply dragging and dropping icons.

You can print out your navigation map or flowchart, an annotated project index with or without associated icons, design and presentation windows, and a cross-reference table of variables.

Time-Based Authoring Tools

Time-based tools are authoring systems wherein elements and events are organized along a timeline, with resolutions as high as or higher than 1/30 second. Time-based tools are best to use when you have a message with a beginning and an end. Sequentially organized graphic frames are played back at a speed that you can set. Other elements (such as audio events) are triggered at a given time or location in the sequence of events. The more powerful time-based tools let you program jumps to any location in a sequence, thereby adding navigation and interactive control.



Each tool uses its own distinctive approach and user interface for managing events over time. Many use a visual timeline for sequencing the events of a multimedia presentation, often displaying layers of various media elements or events alongside the scale in increments as precise as one second. Others arrange long sequences of graphic frames and add the time component by adjusting each frame’s duration of play.

Adobe Flash is a time-based development environment. Flash, however, is also particularly focused on delivery of rich multimedia content to the Web. With the Flash Player plug-in installed in more than 95 percent of the world’s browsers, Flash delivers far more than simple static HTML pages.

Action Script, the proprietary, under-the-hood scripting language of Flash, is based upon the international ECMA Script standard (www.ecmainternational.org) derived from Netscape's original JavaScript.

Adobe Director is a powerful and complex multimedia authoring tool with a broad set of features to create multimedia presentations, animations, and interactive multimedia applications. It requires a significant learning curve, but once mastered, it is among the most powerful of multimedia development tools. In Director, you assemble and sequence the elements of your project, called a "movie," using a Cast and a Score.

The Cast is a multimedia database containing still images, sound files, text, palettes, QuickDraw shapes, programming scripts, QuickTime movies, Flash movies, and even other Director files. You tie these Cast members together using the Score facility, which is a sequencer for displaying, animating, and playing Cast members, and it is made up of frames that contain Cast members, tempo, a palette, timing, and sound information.

Each frame is played back on a stage at a rate specified in the tempo channel. Director utilizes Lingo, a full-featured object-oriented scripting language, to enable interactivity and programmed control.

12.5 MULTIMEDIA PRODUCTION

A multimedia project involves a host of people with specialized skills, such as, an art director, graphic designer, production artist, producer, project manager, writer, user interface designer, sound designer, videographer, and 2D and 3D animators, as well as programmers.

The production process can be organized into three consecutive stages: preproduction, production, and postproduction. Everything from the inception of the project idea to setting up for actual recording is part of the preproduction stage. This includes the writing of a proposal, treatment, and script, and the breakdown of the script in terms of production scheduling and budgeting.

The second major phase of production is the production stage. Everything involved in the setup and recording of visual images and sounds, from performer, camera, and microphone placement and movement to lighting and set design, makes up part of the production stage. Postproduction consists of the editing of the recorded images and sounds, and all of the procedures needed to complete a project in preparation for distribution on various media.

Preproduction

Preproduction consists of the preparation of project proposals, premises, synopses, treatments, scripts, script breakdowns, production schedules, budgets, and storyboards. A proposal is a market summary used to promote or sell a project. A premise is a concise statement or assertion that sums up the story or subject matter. For example, the basic premise of Joan Didion's film The Panic in Needle Park (1971) is "Romeo and Juliet on drugs in New York's Central Park."

A synopsis is a short paragraph that describes the basic story line. Treatments are longer plot or subject-matter summaries in short-story form, which often accompany oral pitches of a premise or concept, and scripts are virtually complete production guides on paper, specifying what will be seen and heard in the finished product.

One can break down a script by listing all equipment and personnel needs for each scene so that a production can be scheduled and budgeted. A budget describes how funds will be spent in each production category. A storyboard provides a graphic visualization of important shots that the camera will eventually record.

Production

Production begins with setup and rehearsal. The film, video, or multimedia director stages and plots the action by rehearsing scenes in preparation for actual recording. Charting the movement of talent on the set is known as performer blocking, while charting the movements of the cameras is called camera blocking.

Every camera placement and movement of the talent must be carefully worked out before recording. If the action cannot be controlled, as in the live transmission of a sporting event or the production of a documentary, the director must be able to anticipate exactly where the action is likely to go and place the camera or cameras accordingly.

During actual production, the entire project is essentially in the hands of the director. In multiple-camera studio or location production, for example, the director often selects the shots by commanding the technical director (TD) to press certain buttons on a device called a switcher, which makes instantaneous changes from one camera to another. In single-camera production, the director remains on the set and communicates directly with the talent and crew.

The script supervisor or continuity person watches the actual recording session with a sharp eye to ensure that every segment in the script has been recorded. Perfect continuity between shots, in such details as a consistent left-to-right or right-to-left direction and identical flow of performer movements (matched action) from one shot to the next, must be maintained so that these shots can be properly combined during editing.

In an audio production or recording session, the producer maintains the same authority and responsibilities as a video or film director: rehearsing the musicians, instructing the engineer, and supervising the actual recording session. In a digital multimedia production or an interactive session, whether for a computer game, CD-ROM, blue-laser disc, or DVD recording, the producer's authority and responsibilities are the same except that the producer may be gathering and working totally with digital material instead of people.

In multimedia and interactive production sessions, the producer may very well perform all aspects of the production—from writing the entire process including preproduction through postproduction to creating the graphics, entering code in order to create the program in a digital form, and performing final editing functions.

Postproduction

Postproduction begins after the visual images and sounds have been recorded (although in live television, production, and postproduction stages occur simultaneously). Possible edit points can be determined during the preview stage, when the recorded images and sounds are initially viewed. Pictures and accompanying sounds are examined and reexamined to find exact edit points before various shots are combined. Separate soundtracks can be added later to the edited images, or the sounds can be edited at the same time as the pictures.

The postproduction stage ties together the audio and visual elements of production and smoothens all the rough edges. The visual and audio elements must be properly balanced and controlled. Sophisticated digital devices help editors and technical specialists mold sounds and images into their final form.

In audio postproduction, the emphasis is placed on choosing the best of the many sound takes and combining the various tracks onto one or, in the case of stereo, two finished tracks, or, as in the case of audio for high-definition television (HDTV) for theaters and home theaters, as many as six or more tracks.

In motion picture production, the sound editor may use as many as 64 or more tracks to complete the production. Games and other interactive and animated productions also require multi-channel audio tracks. Signal processing, including equalization, adding effects, and balancing tracks against each other, is often performed during the sound mix, that is, during the final process of combining various soundtracks. Such processing operations may be performed either in an analog or in a digital format. The tendency is to manipulate audio in a digital format to avoid any degeneration or degradation of the signal.

The three stages of production are separate only in a chronological sense. Proficiency in one stage of the production process necessarily requires some knowledge of all other stages. A director or writer cannot visualize the possibilities for recording a particular scene without having some awareness of how images can be combined during editing.

In short, although the overall organization of this text into three stages (preproduction, production, and postproduction) follows a logical progression, mastery of any one stage demands some familiarity with other stages as well.

12.6 SUMMARY

- Three metaphors are used by authoring tools that make multimedia, based on the method used for sequencing or organizing multimedia elements and events: Card- or page-based tools, Icon- or object-based, event-driven multimedia and game-authoring tools, Time-based tools.

- Card-based or page-based tools are authoring systems wherein the elements are organized as a stack of cards or pages of a book, respectively.

- Card- or page-based systems typically provide two separate layers on each card or page: a background layer that can be shared among many cards or pages, and a foreground layer that is specific to a single card or page.

- Icon-based or object-based authoring tools are event-driven authoring systems wherein multimedia elements and interaction cues (events) are organized as objects in a structural framework or process.

- Time-based tools are authoring systems wherein elements and events are organized along a timeline, with resolutions as high as or higher than 1/30 second.

- Adobe Flash is a time-based development environment.

- Adobe Director is a powerful and complex multimedia authoring tool with a broad set of features to create multimedia presentations, animations, and interactive multimedia applications.
- The Cast is a multimedia database containing still images, sound files, text, palettes, QuickDraw shapes, programming scripts, QuickTime movies, Flash movies, and even other Director files.

- The production process can be organized into three consecutive stages: preproduction, production, and postproduction.

- Preproduction consists of the preparation of project proposals, premises, synopses, treatments, scripts, script breakdowns, production schedules, budgets, and storyboards.

- A synopsis is a short paragraph that describes the basic story line.

- A storyboard provides a graphic visualization of important shots that the camera will eventually record.

- In an audio production or recording session, the producer maintains the same authority and responsibilities as a video or film director: rehearsing the musicians, instructing the engineer, and supervising the actual recording session.

- Postproduction begins after the visual images and sounds have been recorded.

12.7 UNIT END EXERCISES

1. Write a note on types of multimedia authoring metaphor.
2. Describe stages of Multimedia Production.
3. Distinguish between icon metaphor and timeline metaphor.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Multimedia: Making It Work, Ninth Edition, 9th Edition, McGraw-Hill Osborne Media

III. Fundamentals of Multimedia, First Indian Reprint Edition 2004, Ze-Nian Li, Mark S. Drew, Pearson Education, Inc.

❖❖❖❖

13

MULTIMEDIA PRESENTATION AND AUTOMATIC AUTHORING

Unit Structure

13.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand graphic styles and fonts.
2. Learn Externalization versus Linearization.
3. Note Color Principles and Guidelines.
4. Understand Automatic Authoring.

13.2 INTRODUCTION

Multimedia authoring is the creation of multimedia productions, sometimes called "movies" or "presentations". Since we are interested in this subject from a computer science point of view, we are mostly interested in interactive applications. Also, we need to consider still-image editors, such as Adobe Photoshop, and simple video editors, such as Adobe Premiere, because these applications help us create interactive multimedia projects.

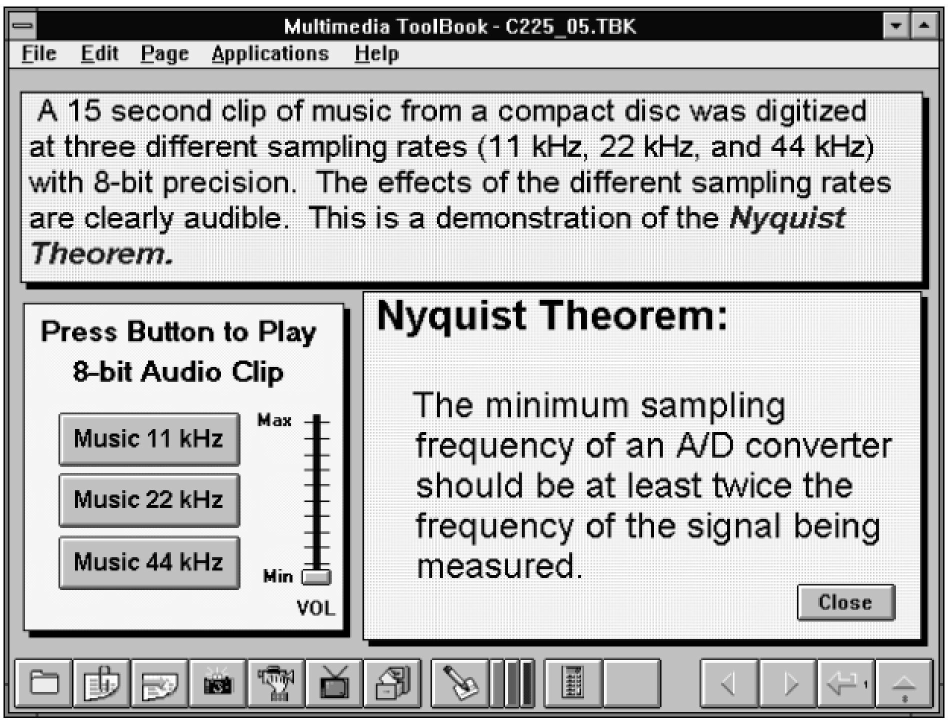
In this unit, we briefly outline some effects to keep in mind for presenting multimedia content as well as some useful guidelines for content design.

13.3 MULTIMEDIA PRESENTATION

Graphics Styles

Careful thought has gone into combinations of color schemes and how lettering is perceived in a presentation. Many presentations are meant for business displays, rather than appearing on a screen.

Human visual dynamics are considered in regard to how such presentations must be constructed. Most of the observations here are drawn from Vetter et al., as in Figure 13.1 (both parts).



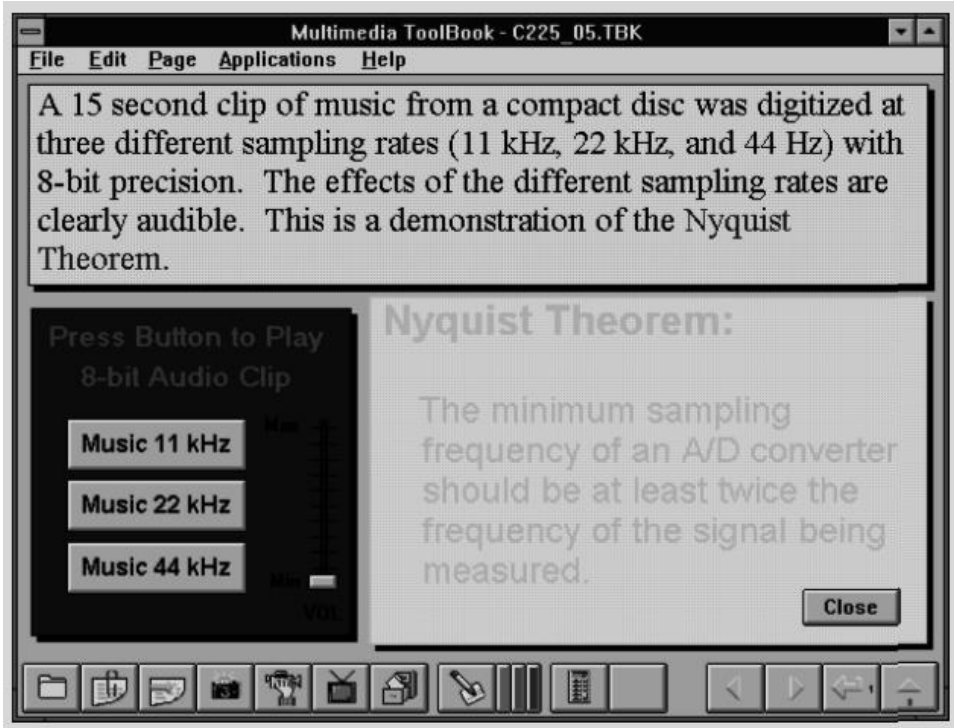


Fig. 13.1: Colors and fonts. (This figure also appears in the color insert section.) Courtesy, Ron Vetter.

Color Principles and Guidelines

Some color schemes and art styles are best combined with a certain theme or style. Color schemes could be, for example, natural and floral for outdoor scenes and solid colors for indoor scenes. Examples of art styles are oil paints, watercolors, colored pencils, and pastels.

A general hint is not to use too many colors, as this can be distracting. It helps to be consistent with the use of color - then color can be used to signal changes in theme.

Fonts

For effective visual communication, large fonts (18 to 36 points) are best, with no more than six to eight lines per screen. As shown in Figure 13.1, sans serif fonts work better than serif fonts (serif fonts are those with short lines stemming from and at an angle to the upper and lower ends of a letter's strokes). Figure 13.1 shows a comparison of two screen projections, (Figure 2 and 3 from Vetter, Ward and Shapiro).

The top figure shows good use of color and fonts. It has a consistent color scheme, uses large and all sans-serif (Arial) fonts. The bottom figure is poor, in that too many colors are used, and they are inconsistent. The red adjacent to the blue is hard to focus on, because the human retina cannot focus on these colors simultaneously.

The serif (Times New Roman)font is said to be hard to read in a darkened, projection setting. Finally, the lower right panel does not have enough contrast - pretty pastel colors are often usable only if their background is sufficiently different.

A Color Contrast Program

Seeing the results of Vetter et al.'s research, we constructed a small Visual Basic program to investigate how readability of text colors depends on foreground color and the color of the background.

The simplest approach to making readable colors on a screen is to use the principal complementary color as the background for text. For color values in the range 0 to 1 (or, effectively, 0 to 255), if the text color is some triple (R, G, B), a legible color for the background is likely given by that color subtracted from the maximum:

$$(R,G,B) \rightarrow (1-R,1-G,1-B)$$

That is, not only is the color "opposite" in some sense (not the same sense as artists use), but if the text is bright, the background is dark, and vice versa.

In the Visual Basic program given below, sliders can be used to change the background color. As the background changes, the text changes to equal the principal complementary color. Clicking on the background brings up a color-picker as an alternative to the sliders.

If you feel you can choose a better color combination, click on the text. This brings up a color picker not tied to the background color, so you can experiment. (The text itself can also be edited.)

A little experimentation shows that some color combinations are more pleasing than others - for example, a pink background and forest green foreground, or a green background and mauve foreground. Figure 13.2 shows this small program in operation.

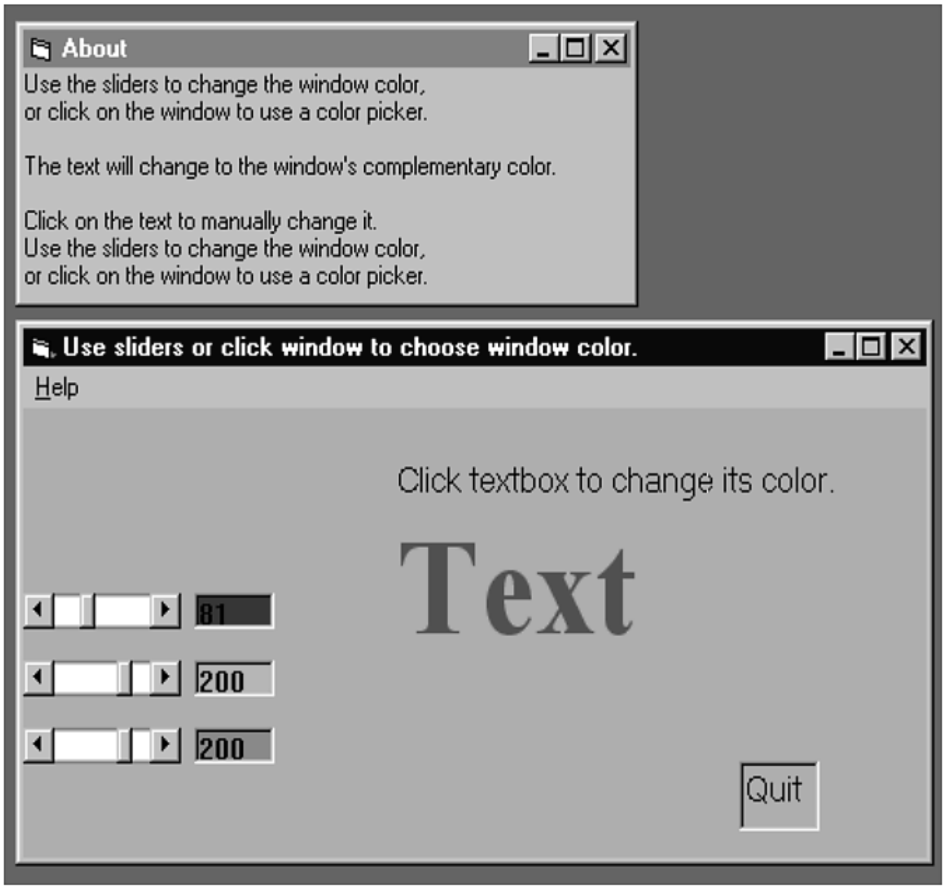


Fig. 13.2:- Program to investigate colors and readability

Figure 13.3 shows a "color wheel", with opposite colors equal to (1-R, 1-G, 1-B). An artist's color wheel will not look the same, as it is based on feel rather than on an algorithm. In the traditional artist's wheel, for example, yellow is opposite magenta, instead of opposite blue as in Figure 13.3, and blue is instead opposite orange.

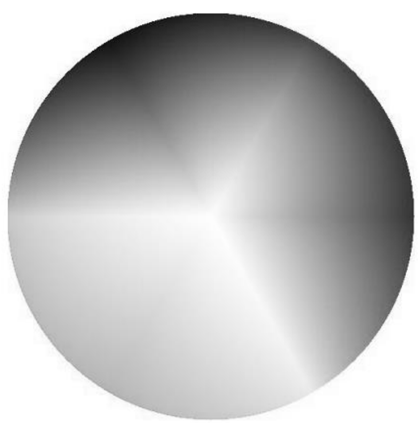


Fig. 13.3:- Color wheel
(This figure also appears in the color insert section.)

Sprite Animation

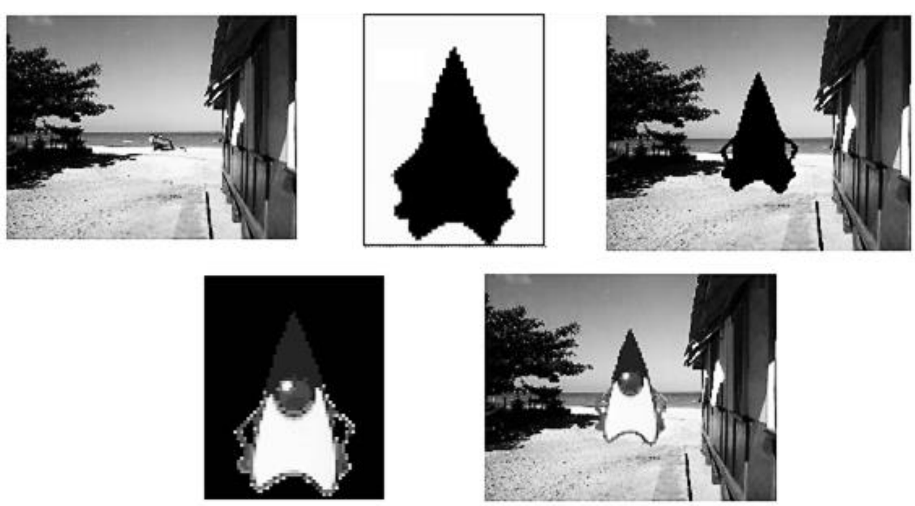
Sprites are often used in animation. For example, in Macromedia Director, the notion of a sprite is expanded to an instantiation of any resource. However, the basic idea of sprite animation is simple. Suppose we have produced an animation figure, as in Figure 13.4(a). Then it is a simple matter to create a 1-bit mask M, as in Figure 13.4(b), black on white, and the accompanying sprite S, as in Figure 13.4(c).

Now we can overlay the sprite on a colored background E, as in Figure 13.5(a), by first ANDing Band M, then ORing the result with S, with the final result as in Figure 13.5(e).Operations are available to carry out these simple compositing manipulations at frame rate and so produce a simple 2D animation that moves the sprite around the frame but does not change the way it looks.

(a) (b) (c)

<u>Fig. 13.4:- Sprite creation: (a) original; (b) mask image M; and (c) sprite S. "Duke"*figure courtesy, Sun Microsystems*.</u>

(a) (b) (c)



(d) (e)

<u>Fig. 13.5:- Sprite animation: (a) Background B; (b) Mask M; (c) B AND M; (d) Sprite S;(e) B AND M OR S.</u>

Video Transitions

Video transitions can be an effective way to indicate a change to the next section. Video transitions are syntactic means to signal "scene changes" and often carry semantic meaning. Many different types of transitions exist; the main types are cuts, wipes, dissolves, fade-ills and fade-outs.

A cut, as the name suggests, carries out an abrupt change of image contents in two consecutive video frames from their respective clips. It is the simplest and most frequently used video transition.

A wipe is a replacement of the pixels in a region of the view port with those from another video. If the boundary line between the two videos moves slowly across the screen, the second video gradually replaces the first. Wipes can be left-to-right, right-to-left, vertical, and horizontal, like an iris opening, swept out like the hands of a clock, and so on.

A dissolve replaces every pixel with a mixture over time of the two videos, gradually changing the first to the second. A fade-out is the replacement of a video by black (or white), and fade-in is its reverse.

Some Technical Design Issues

Technical parameters that affect the design and delivery of multimedia applications include computer platform, video format and resolution, memory and disk space, delivery methods.

Computer Platform

Usually we deal with machines that are either some type of UNIX box (such as a Sun) or else a PC or Macintosh. While a good deal of software is ostensibly "portable", much cross-platform software relies on runtime modules that may not work well across systems.

```
for t in 0..t_max
    for x in 0..x_max
        if ( x/x_max < t/t_max)
            R = R_L ( x + x_max * [1 - t/t_max], t)
                else
                  R = R_R ( x - x_max * t/t_max, t)
```

Fig. 13.8:- Pseudo code for slide video transition

Video Format and Resolution

The most popular video formats are NTSC, PAL, and SECAM. They are not compatible, so conversion is required to playa video in a different format. The graphics card, which displays pixels on the screen, is sometimes referred to as a “video card”. In fact, some cards are able to perform “frame grabbing”, to change analog signals to digital for video. This kind of card is caned a “video capture card”.

The graphics card's capacity depends on its price. An old standard for the capacity of a card is S-VGA, which allows for a resolution of 1280x1024 pixels in a displayed image and as many as 65,536 colors using 16-bit pixels or 16.7 million colors using 24-bit pixels.

Nowadays, graphics cards that support higher resolution, such as1600x1200, and 32-bit pixels or more are common.

Memory and Disk Space Requirement

Rapid progress in hardware alleviates the problem, but multimedia software is generally greedy. Nowadays, at least 1Gigabyte of RAM and 20 gigabytes of hard-disk space should be available for acceptable performance and storage for multimedia programs.

Delivery Methods

Once coding and all other work is finished, how shall we present our clever work? Since we have presumably purchased a large disk, so that performance is good and storage is not an issue, we could simply bring along our machine and show the work that way. However, we likely wish to distribute the work as a product. Presently, rewritable DVD drives are not the norm, and CD-ROMs may lack sufficient storage capacity to hold the presentation. Also, access time for CD-ROM drives is longer than for hard-disk drives.

Electronic delivery is an option, but this depends on network bandwidth at the user side (and at our server). A streaming option may be available, depending on the presentation. No perfect mechanism currently exists to distribute large multimedia projects. Nevertheless, using such tools as PowerPoint or Director, it is possible to create acceptable presentations that fit on a single CD-ROM.

13.4 AUTOMATIC AUTHORING

Thus far, we have considered notions developed for authoring new multimedia. Nevertheless, a tremendous amount of legacy multimedia documents exists, and researchers have been interested in methods to facilitate *automatic authoring*.

By this term is meant either an advanced helper for creating new multimedia presentations or a mechanism to facilitate automatic creation of more useful multimedia documents from existing sources.

Hypermedia Documents

Let us start by considering hypermedia documents. Generally, three steps are involved in producing documents meant to be viewed nonlinearly: information generation or capture, authoring, and publication. A question that can be asked is, how much of this process can be automated?

The first step, capture of media, be it from text or using an audio digitizer or video frame-grabber, is highly developed and well automated. The final step, presentation, is the objective of the multimedia tools we have been considering. But the middle step (authoring) is most under consideration here.

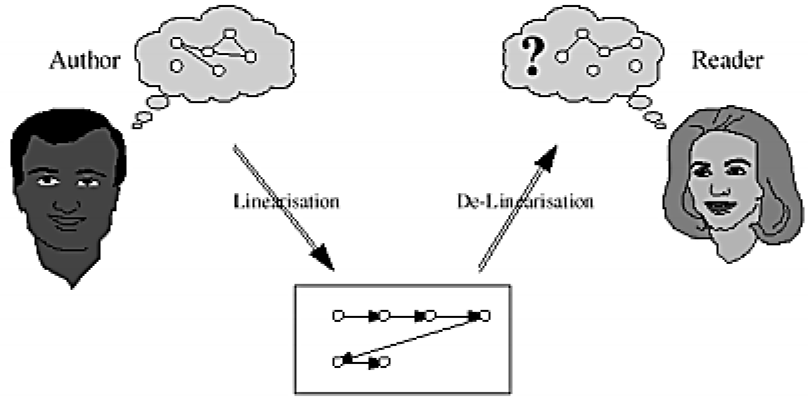
Essentially, we wish to structure information to support access and manipulation of the available media. Clearly, we would be well advised to consider the standard computing science data structures in structuring this information: lists, trees, or networks (graphs).

However, here we would like to consider how best to structure the data to support multiple views, rather a single, static view.

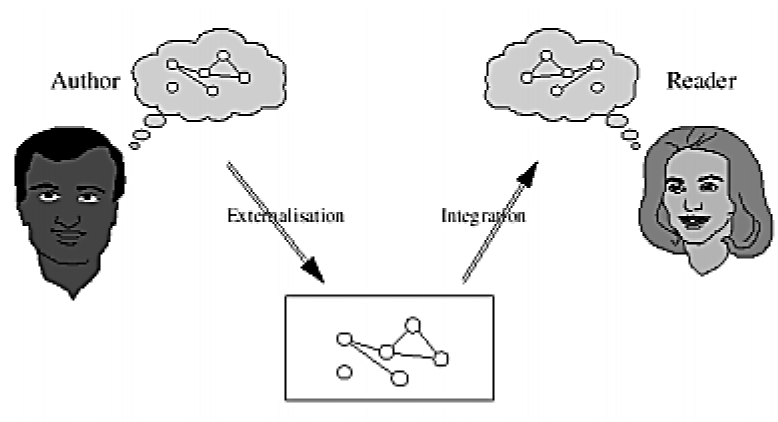
Externalization versus Linearization

Figure 13.9 shows the essential problem involved in communicating ideas without using a hypermedia mechanism: the author's ideas are “linearized” by setting them down in linear order on paper.

In contrast, hyperlinks allow us the freedom to partially mimic the author's thought process (i.e., externalization).



(a)



(b)

Fig. 13.9:- Communication using hyperlinks. Courtesy of David Lowe; (@1995 (IEEE).

Now, using Microsoft Word, say, it is trivial to create a hypertext version of one's document, as Word simply follows the layout already set up in chapters, headings, and so on. But problems arise when we wish to extract semantic content and find links and anchors, even considering just text and not images.

Figure 13.10 displays the problem: while it is feasible to mentally manage a few information nodes, once the problem becomes large, we need automatic assistants.

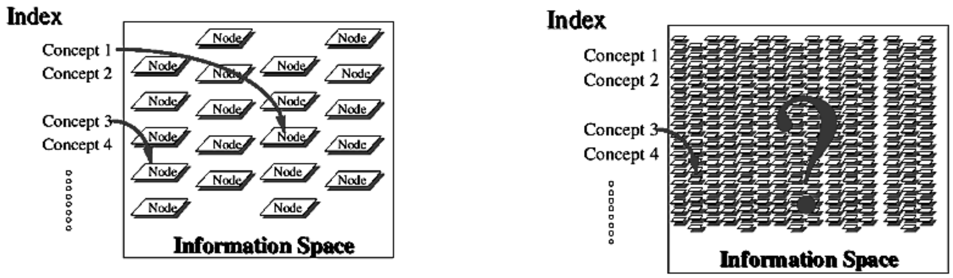


Fig. 13.10:- Complex information space: (a) complexity: manageable;
(b) complexity:overwhelming. *Courtesy of David Lowe; (@1995 IEEE).*

Once a dataset becomes large, we should employ database methods. The issues become focused on scalability (to a large dataset), maintainability, addition of material, and reusability. The database information must be set up in such a way that the "publishing" stage, presentation to the user, can be carried out just-in-time, presenting information in a user-defined view from an intermediate information structure.

Semiautomatic Migration of Hypertext

The structure of hyperlinks for text information is simple: "nodes" represent semantic information and are anchors for links to other pages. Figure 13.11 illustrates these concepts.

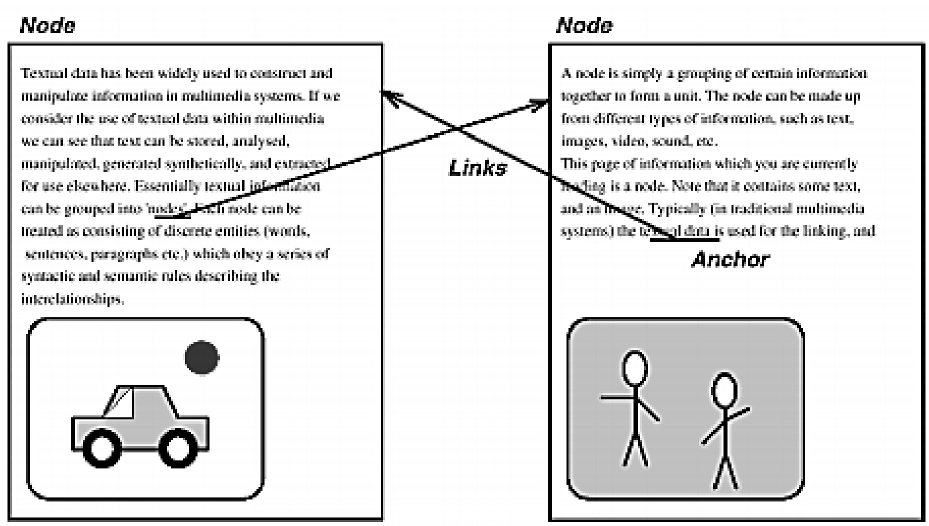


Fig. 13.11:- Nodes and anchors in hypertext. Courtesy ofDavid Lowe.

For text, the first step for migrating paper-based information to hypertext is to automatically convert the format used to HTML. Then, sections and chapters can be placed in a database. Simple versions of data mining techniques, such as word stemming, can easily be used to parse titles and captions for keywords - for example, by frequency counting.

Keywords found can be added to the database being built. Then a helper program can automatically generate additional hyperlinks between related concepts.

A semiautomatic version of such a program is most likely to be successful, making suggestions that can be accepted or rejected and manually added to. A database management system can maintain the integrity of links when new nodes are inserted.

For the publishing stage, since it may be impractical to re-create the underlying information structures, it is best to delay imposing a viewpoint on the data until as late as possible.

Hyper images

Matters are not nearly so straightforward when considering image or other multimedia data. To treat an image in the same way as text, we would wish to consider an image to be a node that contains objects and other anchors, for which we need to determine image entities and rules. What we desire is an automated method to help us produce true hypermedia, as in Figure 13.12.

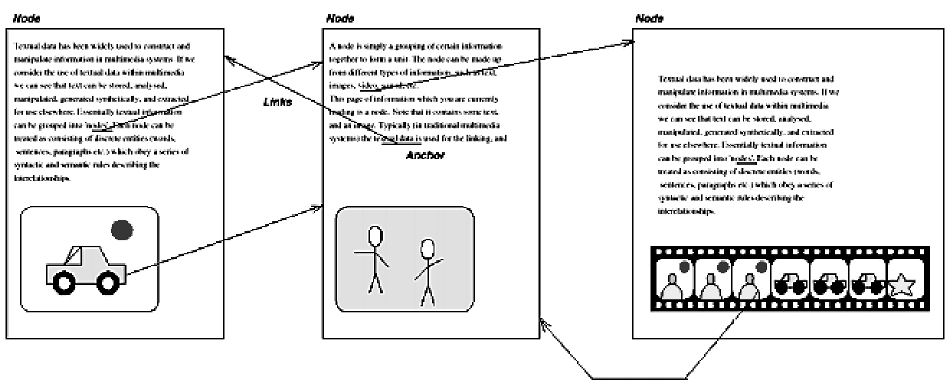


Fig. 13.12:- Structure of hypermedia. *Courtesy of David Lowe*.

It is possible to manually delineate syntactic image elements by masking image areas. These can be tagged with text, so that previous text-based methods can be brought into play.

13.5 SUMMARY

- For effective visual communication, large fonts (18 to 36 points) are best, with no more than six to eight lines per screen.
- The serif (Times New Roman) font is said to be hard to read in a darkened, projection setting.

- The simplest approach to making readable colors on a screen is to use the principal complementary color as the background for text

- *Sprites* are often used in animation.

- Video transitions are syntactic means to signal "scene changes" and often carry semantic meaning

- A cut, as the name suggests, carries out an abrupt change of image contents in two consecutive video frames from their respective clips.

- A wipe is a replacement of the pixels in a region of the view port with those from another video

- A dissolve replaces every pixel with a mixture over time of the two videos, gradually changing the first to the second

13.6 UNIT END EXERCISES

1. Write a note on Multimedia Presentation.
2. What is Sprite Animation?
3. Distinguish between Externalization and Linearization
4. What are Hyper images?

Reference:

I. Fundamentals of Multimedia, 2004 by Pearson Education, Inc.,Ze-Nian Li and Mark S. Drew, Pearson Prentice Hall.

II. Using Color and Text in Multimedia Projections, R. Vetter, C. Ward, and S. Shapiro, IEEE Multimedia, 2(4): 46-54, 1995.

❖❖❖❖

14

DESIGN AND USER INTERFACE

Unit Structure

14.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand design paradigms.
2. Understand working of Adobe Premiere Pro.

14.2 INTRODUCTION

Multimedia user interfaces combine different media such as text, graphics, sound, and video to present information. Due to improvements in technology and decreases in costs, many human factors engineers will soon find themselves designing user interfaces that include multimedia. Since many educators, parents, and students believe that multimedia helps people to learn, one popular application of this technology will be the field of education.

The user interface also embodies the data and functions of computer-based products and provides a basis for the product's usability and commercial success. One of the important challenges to user interface design is how to help the novice user become quickly proficient and eventually become an expert user without the encumbrance of the training aids that were useful for the novice.

Systematic, information-oriented visual communication, or the graphic design of the user interface, is an important part of technologically sophisticated computer-based products as they spread internationally to consumer and business markets.

14.3 DESIGN PARADIGMS AND USER INTERFACE

In traditional programming paradigm, the user begins the program and then the program takes control and prompts the user for inputs as needed. Slowly, as graphical interfaces evolve, “event-driven” programs replaced this paradigm. Events are messages that the user gives to the program, for example, moving the mouse, clicking, double-clicking the mouse at a button on the screen, or typing a letter on the keyboard.

The term user interface in the context of the computer is usually understood to include such things as windows, menus, buttons, the keyboard, the mouse, the sounds that a computer makes, and, in general, all the information channels that allow the user and the computer to communicate. Most successful multimedia systems are used by people who do not have any knowledge of programming. A graphical user interface should be designed with this in mind and should allow ordinary people to use computers daily.

A good graphical user interface uses pictures rather than text to provide users an understanding of how to work on a system. A common understanding of symbols allow us to make systems that can be used with little instruction. Screen objects are also known as Windowing System Components. Some examples of common screen objects are: Windows, menus, controls, dialog boxes, control panels, query boxes, etc.. The design of these screen objects have become standard within the Windows programming environment. Using the same kind of dialog boxes across various programs allows the user to transfer the knowledge gained from using one program to another.

Usability testing, as an emerging and expanding research area of human-computer interface, can provide a means for improving the usability of multimedia software design and development through quality control processes. In the process of usability testing, evaluation experts should consider the nature of users and the tasks they will perform, tradeoffs supported by the iterative design paradigm, and real world constraints in order to effectively evaluate and improve multimedia software.

Usability experts need to be in the field where they can see how real users work with real multimedia software. It is the responsibility of performance technologists, especially multimedia developers’, to make multimedia software simple to use, simple to understand, yet still powerful enough for the task. The issue is no longer whether to conduct usability testing, but how to conduct useful usability testing.

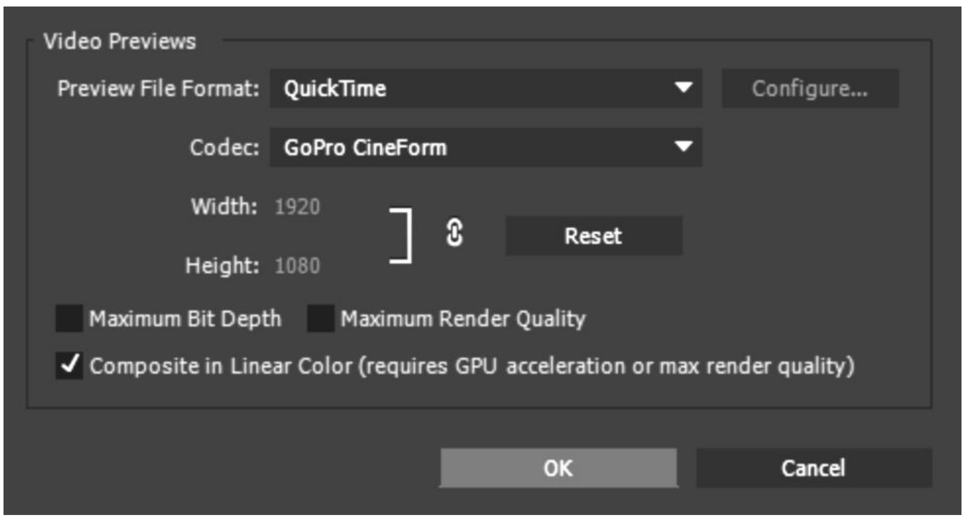
14.4 OVERVIEW OF TOOLS

ADOBE PREMIERE

Adobe Premiere is a very simple video editing program that allows you to quickly create a simple digital video by assembling and merging multimedia components. It effectively uses the score authoring metaphor, in that components are placed in "tracks" horizontally, in a Timeline window.

Premiere Pro provides full smart rendering support for GoPro CineForm files on Windows. You can preview CineForm files in QuickTime format.

1. In the New Sequence or Sequence Settings dialog, select Editing Mode as Custom.
2. Select Preview File Format as QuickTime.
3. Select Codec as GoPro CineForm.
4. (Optional) Save a preset for each combination of height-and-width and frame rate that you commonly use.

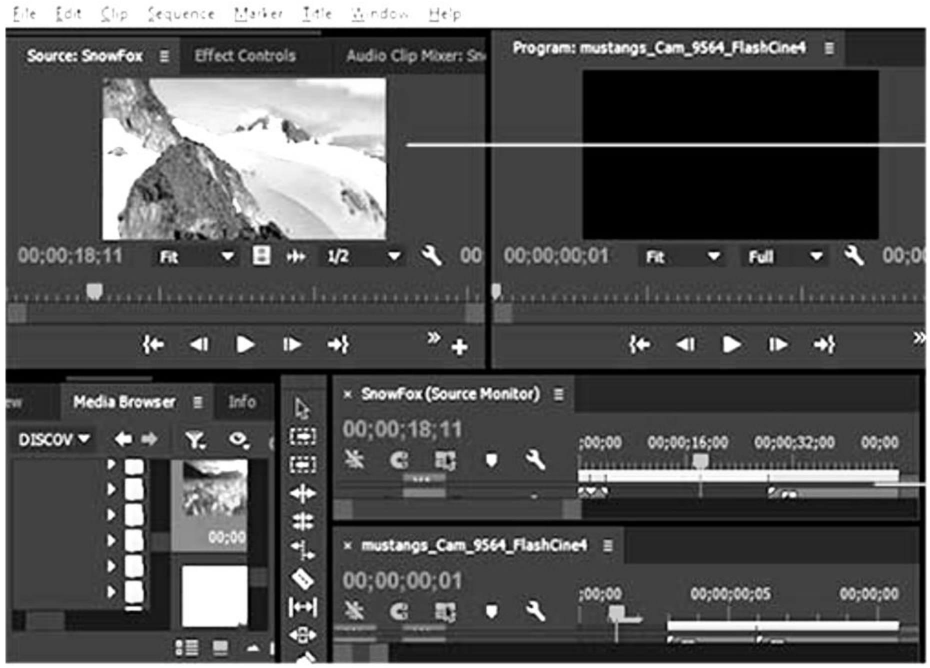


The user interface provides HiDPI support for Apple's Retina displays and Windows 8.1 displays.

The modernized user interface provides cleaner visuals that let you focus more on the content. Apart from the visual enhancements, there are subtle but effective enhancements to the overall user experience. When you select a user interface element, the selected element appears with a blue outline indicating its active state. When deselected, it appears in gray. This high contrast helps you easily distinguish between selected and deselected elements. For a comfortable viewing experience, you can vary the brightness of the user interface from a darker to a lighter tone by using the Appearance preference option.

With HiDPI support, Premiere Pro provides a higher resolution user interface that displays text, icons, and other user interface elements in greater clarity. You can notice an optimal display clarity under various scaling factors. At a 100% scaling, the application displays more real estate for viewing, which means many more panels can be viewed at once. When you change the scaling, the user interface elements scale optimally and continue to appear sharp and clear.

Premiere Pro lets you open and view sequences from an unopened project without importing the sequence into your current project. Using the Media Browser, navigate to the project containing the sequence, and double-click the sequence to open. The new Source Monitor Timeline view opens a second Timeline that displays the contents of the sequence in read-only mode. This second Timeline makes it easy to edit or reuse existing clips, cuts, and transitions from different projects.

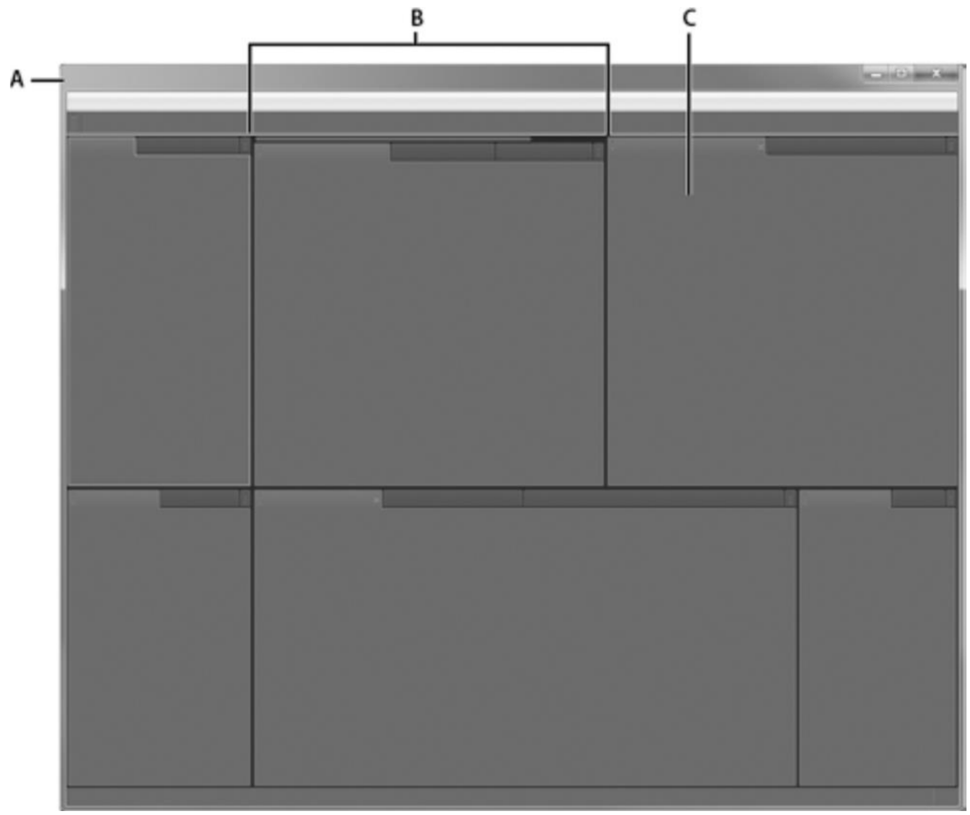


A Source Monitor Timeline View

Premiere Pro now provides an option in the Project Settings dialog that lets you keep all instances of your project items in sync automatically. Select File > Project Settings General to open the Project Settings dialog. In the Project Settings dialog, select *Display The Project Items Name And Label Color For All Instances*. When you select this option, any changes made to a clip in the Project panel ripple to all instances used in sequences. For example, when you change the name of a sequence clip, it ripples up to the master clip and then down to all other sequence clips.

When you send a clip for audio editing in Audition by selecting *Edit > Edit in Adobe Audition > Clip*, the rendered copy of the clip is automatically saved alongside the original media file on disk. Storing the rendered media files along with the original files makes media management easier.

Adobe video and audio applications provide a consistent, customizable workspace. Although each application has its own set of panels (such as Project, Metadata, and Timeline), you move and group panels in the same way across products. The main window of a program is the application window. Panels are organized in this window in an arrangement called a workspace. The default workspace contains groups of panels as well as panels that stand alone. You customize a workspace by arranging panels in the layout that best suits your working style. As you rearrange panels, the other panels resize automatically to fit the window. You can create and save several custom workspaces for different tasks—for example, one for editing and one for previewing.



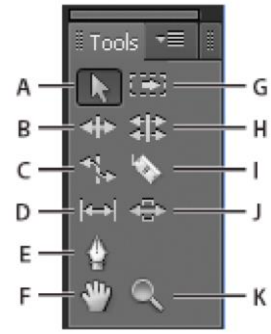
A: Application window B: Grouped panels C: Individual panel

When you open the Options panel, it opens by default in the horizontal docking area running just under the menu bar, forming the Options bar. You can undock, move, and redock the Options panel like any other panel. By default, the Options panel contains a menu of workspaces and a link to CS Services. You can also dock the Tools panel to the Options panel.

TOOLS

When you select a tool, the pointer changes shape according to the selection. For example, when you select the Razor tool and position the pointer over a clip in a Timeline panel, the icon changes to a razor. However, the Selection tool icon can change to reflect the task currently being performed. In some cases, pressing a modifier key (such as Shift) as you use a tool changes its function, and its icon changes accordingly. Select tools from the Tools panel, or use a keyboard shortcut. You can resize the Tools panel and orient it vertically or horizontally.

Note: *The Selection tool is the default tool. It's used for everything other than specialized functions. If the program isn't responding as you expect, make sure that the Selection tool is selected.*



A: SELECTION TOOL B: RIPPLE EDIT TOOL
C: RATE STRETCH TOOL D: SLIP TOOL E: PEN TOOL
F: HAND TOOL G: ROLLING TOOL H: ROLLING EDIT TOOL
I: RAZOR TOOL J: SLIDE TOOL K: ZOOM TOOL

Select any tool to activate it for use in a Timeline panel by clicking it or pressing its keyboard shortcut. Let the cursor hover over a tool to see its name and keyboard shortcut.

Selection Tool

The standard tool for selecting clips, menu items, and other objects in the user interface. It's generally a good practice to select the Selection Tool as soon as you are done using any of the other, more specialized, tools.

Track Selection Tool

Select this tool to select all the clips to the right of the cursor in a sequence. To select a clip and all clips to the right in its own track, click the clip. To select a clip and all clips to its right in all tracks, Shift-click the clip. Pressing Shift changes the Track Selection Tool into the Multi-track Selection Tool.

Ripple Edit Tool

Select this tool to trim the In or Out point of a clip in a Timeline. The Ripple Edit Tool closes gaps caused by the edit and preserves all edits to the left or right of the trimmed clip.

Rolling Edit Tool

Select this tool to roll the edit point between two clips in a Timeline. The Rolling Edit Tool trims the In point of one and the Out point of the other, while leaving the combined duration of the two clips unchanged.

Rate Stretch Tool

Select this tool to shorten a clip in a Timeline by speeding up its playback, or to lengthen it by slowing it down. The Rate Stretch Tool changes speed and duration, but leaves the In and Out points of the clip unchanged.

Razor Tool

Select this tool to make one or more incisions in clips in a Timeline. Click a point in a clip to split it at that precise location. To split clips in all tracks at that location, Shift-click the spot in any of the clips.

Slip Tool

Select this tool to simultaneously change the In and Out points of a clip in a Timeline, while keeping the time span between them constant. For example, if you have trimmed a 10-second clip to 5 seconds in a Timeline, you can use the Slip Tool to determine which 5 seconds of the clip appear in the Timeline.

Slide Tool

Select this tool to move a clip to the left or right in a Timeline while simultaneously trimming the two clips that surround it. The combined duration of the three clips, and the location of the group in the Timeline, remain unchanged.

Pen Tool

Select this tool to set or select keyframes, or to adjust connector lines in a Timeline. Drag a connector line vertically to adjust it. Ctrl-click (Windows) or Command-click (Mac OS) on a connector line to set a keyframe. Shiftclick noncontiguous keyframes to select them. Drag a marquee over contiguous keyframes to select them. For more information about using the Pen Tool, see Select keyframes.

Hand Tool

Select this tool to move the viewing area of a Timeline to the right or left. Drag left or right anywhere in the viewing area. Zoom Tool Select this tool to zoom in or out in a Timeline viewing area. Click in the viewing area to zoom in by one increment. Alt-click (Windows) or Option-click (Mac OS) to zoom out by one increment.

Navigate clips in the Source menu in the Source Monitor

You can set keyboard shortcuts for navigating multiple clips loaded into the Source Monitor. Keyboard shortcuts can speed toggling of clips, skipping to the first or last clip, or closing one or all the clips in the Source Monitor popup menu.

1. Select Edit > Keyboard Shortcuts (Windows) or Premiere Pro > Keyboard Shortcuts (Mac OS). The Keyboard Shortcuts dialog box opens.
2. In the dialog box, click the triangle next to Panels, and then click the triangle next to Source Monitor Panel to reveal the keyboard shortcuts for that panel.

3. Set keyboard shortcuts for any of the following commands:
   • Source Clip: Close • Source Clip: Close All • Source Clip: First
   • Source Clip: Last • Source Clip: Next • Source Clip: Previous
4. Click OK.

Using the Source Monitor and Program Monitor time controls

The Source Monitor has several controls for moving through time (or frames) in a clip. The Program Monitor contains similar controls for moving through a sequence.

A: Current time display B: Playhead C: Zoom scroll bar
D: Time Ruler E: Duration display

Time rulers

Display the duration of a clip in the Source Monitor and sequence in the Program Monitor. Tick marks measure time using the video display format specified in the Project Settings dialog box. You can toggle the time rulers to display time code in other formats. Each ruler also displays icons for its corresponding monitor's markers and In and Out points. You can adjust the play head, markers, and the In and Out points by dragging their icons in a time ruler. Time ruler numbers are off by default. You can turn the time ruler numbers on by selecting Time Ruler Numbers in the panel menu of the Source or Program Monitors.

Play head

Shows the location of the current frame in each monitor's time ruler. Note: *The play head was formerly called the "current-time indicator" (CTI).*

Current time displays

Show the time code for the current frame. The current time displays are at the lower left of each monitor's video. The Source Monitor shows the current time for the open clip. The Program Monitor shows the sequence's current time. To move to a different time. Alternatively, click in the display and enter a new time, or place the pointer over the time display and drag left or right. To toggle display between full time code and a frame count, Ctrlclick (Windows) or Command-click (Mac OS) the current time in either monitor or a Timeline panel.

Duration display

Shows the duration of the open clip or sequence. The duration is the time difference between the In point and the Out point for the clip or sequence. When no In point is set, the starting time of the clip or of the sequence is substituted. When no Out point is set, the Source Monitor uses the ending time of the clip to calculate duration. The Program Monitor uses the ending time of the last clip in the sequence to calculate duration.

Zoom scroll bars

Zoom scroll bars correspond with the visible area of the time ruler in each monitor. You can drag the handles to change the width of the bar and change the scale of the time ruler below. Expanding the bar to its maximum width reveals the entire duration of the time ruler. Contracting the bar zooms in for a more detailed view of the ruler. Expanding and contracting the bar is centered on the playhead. By positioning the mouse over the bar, you can use the mouse wheel to contract and expand the bar. You can also scroll the mouse wheel in the areas outside of the bars for the same expanding and contracting behavior. By dragging the center of the bar, you can scroll the visible part of a time ruler without changing its scale. When you drag bar, you are not moving the playhead, however, you can move the bar and then click in the time ruler to move the playhead to the same area as the bar. A zoom scroll bar is also available in the Timeline.

Note: *Changing the Program Monitor's time ruler or zoom scroll bar does not affect the time ruler or viewing area in a Timeline panel.*

Before you begin editing

Before you begin editing in Premiere Pro, you will need footage to work with. You can either shoot your own footage, or work with footage that other people have shot. You can also work with graphics, audio files, and more. Many projects you work on do not need a script. However, sometimes you work from or write a script, especially for dramatic projects. You can write your script and organize your production details with *Adobe Story*.

While you shoot, organize your shots and take log notes. You can also adjust and monitor footage as you shoot, capturing directly to a drive. It is important to note that using Adobe Story is not necessary for editing with Adobe Premiere Pro. Writing a script, and making notes on the set are optional steps to help organize a project before you get started.

Get started editing

After you have acquired footage, follow the steps to get started editing with Premiere Pro.

1. Start or open a project

Open an existing project, or start a new one from the Premiere Pro Welcome screen. If you are starting a new project, the New Project dialog launches. From the New Project dialog, you can specify the name and location of the project file, the video capture format, and other settings for your project. After you have chosen settings in the New Project dialog, click OK. After you have exited the New Project dialog, the New Sequence dialog will appear. Choose the sequence preset in the dialog that matches the settings of your footage. First, open the camera type folder, then the frame rate folder (if necessary), and then clicking a preset. Name the sequence at the bottom of the dialog, and then click OK. To open an existing project, click on a link under *Open A Recent Item* in the Premiere Pro Welcome screen. After clicking a link, the project will launch.

2. Capture and import video and audio

For file-based assets, using the Media Browser you can import files from computer sources in any of the leading media formats. Each file you capture or import automatically becomes a clip in the Project panel. Alternatively, using the Capture panel, capture footage directly from a camcorder or VTR. With the proper hardware, you can digitize and capture other formats, from VHS to HDTV. You can also import various digital media, including video, audio, and still images. Premiere Pro also imports Adobe® Illustrator® artwork or Photoshop® layered files, and it translates After Effects® projects for a seamless, integrated workflow. You can create synthetic media, such as standard color bars, color backgrounds, and a countdown.

You can also use *Adobe® Bridge* to organize and find your media files. Then use the Place command in Adobe Bridge to place the files directly into Premiere Pro. In the Project panel, you can label, categorize, and group footage into bins to keep a complex project organized. You can open multiple bins simultaneously, each in its own panel, or you can nest bins, one inside another. Using the Project panel Icon view, you can arrange clips in storyboard fashion to visualize or quickly assemble a sequence.

Note: *Before capturing or importing audio, ensure that Preferences>Audio>Default Track Format* is set to match the desired channel format.

3. Assemble and refine a sequence

Using the Source Monitor, you can view clips, set edit points, and mark other important frames before adding clips to a sequence. For convenience, you can break a master clip into any number of subclips, each with its own In and Out points. You can view audio as a detailed waveform and edit it with sample-based precision.

You add clips to a sequence in a Timeline panel by dragging them there or by using the Insert or Overwrite buttons in the Source Monitor. You can automatically assemble clips into a sequence that reflects their order in the Project panel. You can view the edited sequence in the Program Monitor or watch the full-screen, full-quality video on an attached television monitor. Refine the sequence by manipulating clips in a Timeline panel, with either context-sensitive tools or tools in the Tools panel. Use the specialized Trim Monitor to fine-tune the cut point between clips. By nesting sequences—using a sequence as a clip within another sequence—you can create effects you couldn't achieve otherwise.

4. Add titles

Using the Premiere Pro full-featured Title, create stylish still titles, title rolls, or title crawls that you can easily superimpose over video. If you prefer, you can modify any of a wide range of provided title templates. As with any clip, you can edit, fade, animate, or add effects to the titles in a sequence.

5. Add transitions and effects

The Effects panel includes an extensive list of transitions and effects you can apply to clips in a sequence. You can adjust these effects, as well as a clip's motion, opacity, and Variable Rate Stretch using the Effect Controls panel. The Effect Controls panel also lets you animate a clip's properties using traditional keyframing techniques. As you adjust transitions, the Effect Controls panel displays controls designed especially for that task. Alternatively, you can view and adjust transitions and a clip's effect keyframes in a Timeline panel.

6. Mix audio

For track-based audio adjustments, the Audio Track Mixer faithfully emulates a full-featured audio mixing board, complete with fade and pan sliders, sends, and effects. Premiere Pro saves your adjustments in real time. With a supported sound card, you can record audio through the sound mixer, or mix audio for 5.1 surround sound.

7. Export

Deliver your edited sequence in the medium of your choice: tape, DVD, Blu-ray Disc, or movie file. Using Adobe Media Encoder, you can customize the settings for MPEG-2, MPEG-4, FLV, and other codecs and formats, to the needs of your viewing audience.

Cross-platform workflow

You can work on a project across computer platforms. For example, you can start on Windows and continue on Mac OS. A few functions change, however, as the project moves from one platform to the other.

1. Sequence settings

You can create a project on one platform and then move it to another. Premiere Pro sets the equivalent sequence settings for the second platform, if there is an equivalent. For example, you can create a DV project containing DV capture and device control settings on Windows. When you open the project on Mac OS, Premiere Pro sets the appropriate Mac DV capture and device control settings. Saving the project saves these Mac OS settings. Premiere Pro translates these settings to Windows settings if the project is later opened on Windows.

2. Effects

All video effects available on Mac OS are available in Windows. Windows effects not available on the Mac appear as offline effects if the project is opened on the Mac. These effects are designated “Windows only” in Premiere Pro Help. All audio effects are available on both platforms. Effect presets work on both platforms (unless the preset applies to an effect not available on a given platform).

3. Adobe Media Encoder presets

Presets created on one platform are not available on the other.

4. Preview files

Preview files made on one platform are not available on the other. When a project is opened on a different platform, Premiere Pro re-renders the preview files. When that project is then opened on its original platform, Premiere Pro renders the preview files yet again.

5. High-bit-depth files

Windows AVI files containing either 10-bit 4:2:2 uncompressed video (v210), or 8-bit 4:2:2 uncompressed video (UYVU) are not supported on Mac OS.

6. Preview rendering

The playback quality of un-rendered non-native files is not as high as playback quality of these files on their native platforms. For example, AVI files do not play back as well on Mac OS as they do on Windows. Premiere Pro renders preview files for non-native files on the current platform. Premiere Pro always renders preview files in a native format. A red bar in the timeline indicates which sections contain files needing rendering.

Exporting for the Web and mobile devices

Adobe Premiere Pro lets you create videos that can be exported to the Web or to mobile devices. To export your project, click on the sequence and select *File* > *Export* > *Media*. In the Export Settings dialog box, you can choose the most optimal file format, frame size, bit rate or ready-made presets for faster upload time and better playback quality.

14.5 SUMMARY

- Events are messages that the user gives to the program, for example, moving the mouse, clicking, double-clicking the mouse at a button on the screen, or typing a letter on the keyboard.

- The term user interface in the context of the computer is usually understood to include such things as windows, menus, buttons, the keyboard, the mouse, the sounds that a computer makes, and, in general, all the information channels that allow the user and the computer to communicate.

- A good graphical user interface uses pictures rather than text to provide users an understanding of how to work on a system.

- Usability testing, as an emerging and expanding research area of human-computer interface, can provide a means for improving the usability of multimedia software design and development through quality control processes.

- Premiere Pro provides full smart rendering support for GoPro Cine Form files on Windows.

14.6 UNIT END EXERCISES

1. Describe different tools of Adobe Premiere Pro.
2. What are the pre-requisites of starting a Premiere Project?
3. Explain cross platform workflow.
4. Write a brief note on Adobe Premiere Pro.

Reference:

1. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

2. Introduction to Multimedia Systems, Chapter 16, Gaurav Bhatnager, Shikha Mehta, Sugata Mitra, © 2002 by ACADEMIC PRESS.

3. Designing Interactive Multimedia, paper by Lori L. Scarlatos, http://www.uni-mannheim.de/acm97/papers/Scarlatos/

4. Premiere_pro_reference.pdf available at Adobe Website.

5. Najjar, L. J. (1998). Principles of educational multimedia user interface design. Human Factors, 40(2), 311-323.

❖❖❖❖

15

MORE TOOLS AND BARRIERS

Unit Structure

15.1 OBJECTIVES

After studying this Unit, you will be able to:

1. Understand working of Adobe Flash.
2. Describe Adobe Director.
3. Describe Author ware.
4. Describe Quest.
5. Know the barriers to widespread use.

15.2 INTRODUCTION

In a slide show, interactivity generally consists of being able to control the pace (e.g., click to advance to the next slide). The next level of interactivity is being able to control the sequence and choose where to go next. Next is media control: start/stop video, search text, scroll the view, and zoom. More control is available if we can control variables, such as changing a database search query. The level of control is substantially higher if we can control objects - say, moving objects around a screen, playing interactive games, and so on. Finally, we can control an entire simulation: move our perspective in the scene, control scene objects.

After an introduction to multimedia paradigms, we present some of the practical tools of multimedia content production - software tools that form the arsenal of production. Here we go through the nuts and bolts of a number of standard programs currently in use.

15.3 OVERVIEW OF TOOLS

Adobe Flash Professional

Flash allows users to create interactive movies by using the score metaphor - a timeline arranged in parallel event sequences, much like a musical score consisting of musical notes. Elements in the movie are called symbols in Flash. Symbols are added to a central repository, called library, and can be added to the movie's timeline. Once the symbols are present at a specific time, they appear on the Stage, which represents what the movie looks like at a certain time, and can be manipulated and moved by the tools built into Flash. Finished Flash movies are commonly used to show movies orgames on the web.

The animation guide in Adobe Flash Professional CC helps you enhance the animation you create by defining a path for the objects you animate. This is helpful when you are working on an animation that follows a path that is not a straight line. This process requires two layers to carry out an animation:

- A layer containing the object that you are about to animate.
- A layer defining the path, which the object is supposed to follow during the animation.

Motion tween animation: About tween animation: Before you begin

Before creating tweens, it is helpful to understand the following Flash Pro concepts:

- Drawing on the Stage
- Timeline layers and the stacking order of objects within a single layer as well as across layers
- Moving and transforming objects on the Stage and in the Property inspector
- Using the Timeline, including object lifetime and selecting objects at specific points in time. See Frames and key frames to learn the basics.
- Symbols and symbol properties. Tweenable symbol types include movie clips, buttons, and graphics. Text is also tweenable.
- Nested symbols. Symbol instances can be nested inside other symbols.
- Optional: Bezier curve editing using the Select and Sub select tools. These tools can be used for editing tween motion paths.

Understanding motion tweens

A motion tween is an animation that is created by specifying different values for an object property in different frames. Flash Pro calculates the values for that property in between those two frames. The term tween comes from the words “in between”. For example, you can place a symbol left of the Stage in frame 1, and move it to the right of the Stage in frame 20.

When you create a tween, Flash Pro calculates all the positions of the movie clip in between. The result is an animation of the symbol moving from left to right, from frame 1 to frame 20. In each frame in between, Flash Pro moves the movie clip one 20th of the distance across the Stage. A tween span is a group of frames in the Timeline in which an object has one or more properties changed over time.

A tween span appears in the Timeline as a group of frames in a single layer with a blue background. These tween spans can be selected as a single object and dragged from one location in the Timeline to another, including to another layer. Only one object on the Stage can be animated in each tween span. This object is called the target object of the tween span.

A property key frame is a frame within a tween span where you explicitly define one or more property values for the tween target object. These properties could include position, alpha (transparency), color tint, and so on. Each property you define has its own property key frames. If you set more than one property in a single frame, then the property key frames for each of those properties reside in that frame. You can view each property of a tween span and its property key frames in the Motion Editor. You can also choose which types of property key frames to display in the Timeline from the tween span context menu.

In the preceding example of tweening a movie clip from frame 1 to frame 20, frames 1 and 20 are property keyframes. You can use the Property inspector, the Motion Editor, and many other tools in Flash to define values for properties you want to animate. You specify these property values in the frames of your choosing, and Flash Pro adds the required property keyframes to the tween span. Flash Pro interpolates the values for each of these properties in the frames in between the property keyframes you have created.

Differences between motion tweens and classic tweens

Flash Pro supports two different types of tweens for creating motion. Motion tweens, introduced in Flash CS4 Professional, are powerful and simple to create. Motion tweens allow the greatest control over tweened animation. Classic tweens, which include all tweens created in earlier versions of Flash Pro, are more complex to create. While motion tweens offer much more control of a tween, classic tweens provide certain specific capabilities that some users need. The differences between motion tweens and classic tweens include the following:

- Classic tweens use keyframes. Keyframes are frames in which a new instance of an object appears. Motion tweens can only have one object instance associated with them and use property keyframes instead of keyframes.
- A motion tween consists of one target object over the entire tween span. Classic tween allows tweening between two keyframes, containing instances of the same or different symbols.
- Both motion tweens and classic tweens allow only specific types of objects to be tweened. When you apply a motion tween to non-allowed object types, Flash offers to convert them to a movie clip when the tween is created. Applying a classic tween converts them to graphic symbols.
- Motion tweens consider text a tweenable type and do not convert text objects to movie clips. Classic tweens convert text objects to graphic symbols.
- No frame scripts are allowed on a motion tween span. Classic tweens allow frame scripts.
- Any object scripts on a tween target cannot change over the course of the motion tween span.
- Motion tween spans can be stretched and resized in the Timeline and are treated as a single object. Classic tweens consist of groups of individually selectable frames in the Timeline.
- To select individual frames in a motion tween span, Ctrl-click (Windows) or Command-click (Macintosh) the frames.

Adobe Director

Director uses a movie metaphor to create interactive presentations. This powerful program includes a built-in scripting language, Lingo, which allows creation of complex interactive movies. The "cast" of characters in Director includes bitmapped sprites, scripts, music, sounds, and palettes. Director can read many bitmapped file formats. The program itself allows a good deal of interactivity, and Lingo, with its own debugger, allows more control, including control over external devices, such as VCRs and videodisc players.

Director also has web authoring features available, for creation of fully interactive Shockwave movies playable over the web.

Top new features of Adobe Director 11.5.9

Parallax mapping

Parallax mapping is an enhancement of Normal/Bump Mapping technique. This technique provides more apparent depth and greater realism with less impact on the performance. For more information, see Parallax Mapping.

Cube mapping

Cube mapping is a method of using a six-sided cube as the shape of the map. The model is projected onto the six faces of a cube and stored as six square textures, or unfolded into six regions of a single texture. For more information, see Cube Mapping.

Normal mapping

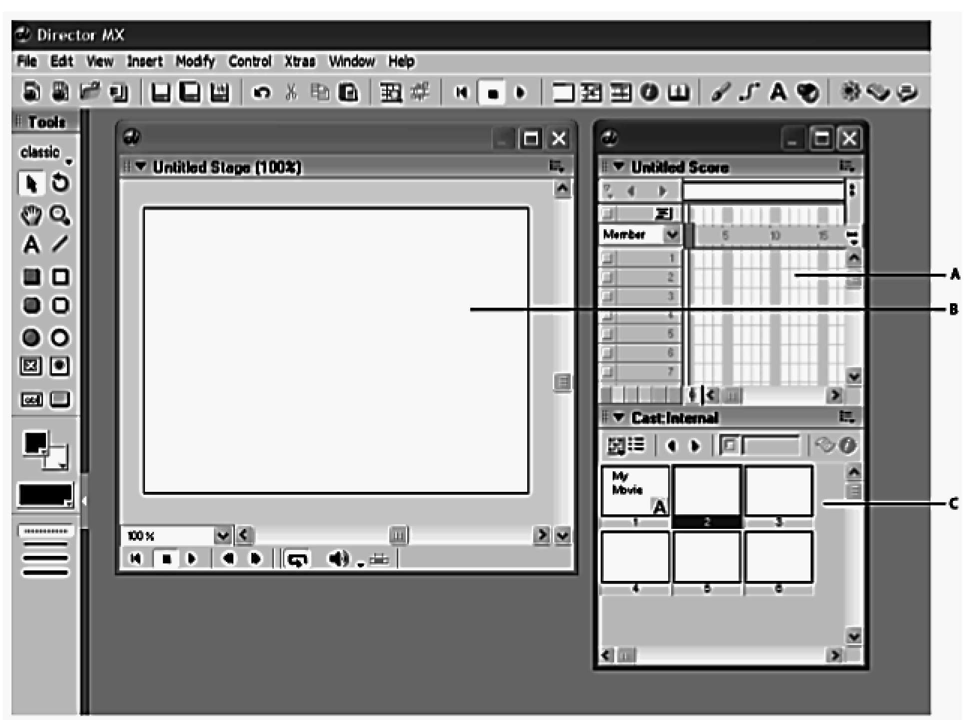
Normal mapping is a technique used for showing finer details like lighting of bumps and dents without using additional polygons.

Render to texture

Render Textures are textures that are created and updated at runtime.

Importing Collada models

Director now provides support for importing Collada-based models.



The Cast Window, the Stage and the Score in Director

A. Score B. Stage C. Cast

The Director User interface is designed around a movie metaphor. Each project you create can be thought of as a movie, with a cast of characters, a Score, a Stage where the action takes place, and a director (you, the author). Each media element that appears in your movie (sound, video, images, text, buttons, and so on) can be thought of as a member of the movie's cast.

In Director, the Cast window is where you view the list of media elements that appear in your movie. As with a real movie, each Director movie has a Score. However, the Score of a Director movie contains more than just music. The Score window in Director contains information about when and where each of the cast members appears on the Stage. The Score describes the action that happens in the movie.

The action in a Director movie takes place in a window called the Stage. To create a Director movie, you add cast members (media elements) to the Cast window by creating them in Director or importing them. Next, you place them on the Stage as sprites. A sprite is simply a copy of a cast member that appears on the Stage. Then you refine the actions of the sprites by editing them on the Stage or in the Score.

Macromedia Authorware

Authorware is a mature, well-supported authoring product that has an easy learning curve for computer science students because it is based on the idea of flowcharts (the so-called iconic/flow-control metaphor). It allows hyperlinks to link text, digital movies, graphics, and sound. It also provides compatibility between files produced in PC and Mac versions. Shockwave Author ware applications can incorporate Shockwave files, including Director Movies, Flash animations, and audio.

Author ware has a modern appearance, to provide an enhanced user experience:

- It's easier to change icon properties with the new Property inspector.
- The Toolbar and Icon palette are movable.
- Knowledge Objects, functions, and variables are easily accessible in panels that dock to the right side of the application window.
- The Preferences dialog box is easier to use, with more Calculation window buttons enabled by default and specific Authorware and JavaScript (JS) tabs.

Quest

Quest, which uses a type of flowcharting metaphor, is similar to Author ware in many ways. However, the flowchart nodes can encapsulate information in a more abstract way (called "frames") than simply subroutine levels. Asa result, connections between icons are more conceptual and do not always represent flow of control in the program.

15.4 BARRIERS TO WIDESPREAD USE

In spite of recent advances in hardware and software technology, there are some substantial barriers to the widespread use and success of authoring and presentation systems. These issues are discussed in the following subsections:

Cost of acquisition, development and delivery of multimedia

Software applications are now available to help a developer plan, track and integrate the multiple information types involved. These tools can boost productivity, and help reduce the number of people and overall number of hours required for development.

As certain products become industry standards, it is likely that developers can be found who are already familiar with the authoring tools selected for the project, reducing the time required for staff training. Similarly, platform costs have dropped while functionality and storage have increased.

Difficulties with production quality

A major issue related to cost of development has been that of production quality for multimedia materials. High-quality video or audio can be expensive because of the need for professional equipment and facilities (such as sound stages and editing suites) and contract employees to create the audio or video. If developers create the materials on their own without professional equipment or contract producers, the quality may not meet the users' expectation.

This situation is changing because of the technology. Equipment for recording high-quality video such as SVHS and Hi-8 mm camcorders is now much more accessible to developers. Digital Audio Tape (DAT) recorders can handle better than CD audio quality. Capture and edit tools for these media are available as off-the-shelf PC products today. Current development limitations include lighting and recording facilities.

Enforcement of Intellectual Property Rights

Due to the significant investment needed to obtain multimedia materials and the large potential value of such information, policies and mechanisms for enforcement of intellectual property rights are needed, otherwise, owners of such materials will be reluctant to make them available for use in multimedia applications.

The growing power of tools for manipulating digital media makes it easier for new content to be derived from existing content.

Cost, availability and ease of use of tools

More software tools have become available for designing and delivering multimedia productions, but problems still exist. Some applications cost several thousand dollars to purchase, and many require substantial learning time for a developer to become a proficient and productive user.

As noted earlier, the more powerful a tool is, the more difficult it is to become an expert, Ease of use will continue to be an important research issue.

Lack of standards for delivery and interchange

The same process that settled the standards issue for video disc developers may solve the notable fragmentation of delivery platforms that exist today. There are many activities underway, from formal standards organizations to industry consortia to vendor-driven. Without such standards, the fragmented market will make it harder to attract publishers of multimedia materials.

Lack of clear vision for multimedia applications

Today there seems to be tremendous interest in all aspects of multimedia, and early adopters are creating prototypes and innovative products. However, there is still a need for better understanding of the potential of new interactive media in areas such as business communications and mass market entertainment, particularly with regard to the forms these new media need to tale.
The new products and services being developed today are causing the most important transition in the industry: the raised awareness of both developers and users of the power and importance of the new media.

15.5 SUMMARY

- Elements in the movie are called symbols in Flash. Symbols are added to a central repository, called a library, and can be added to the movie's timeline.

- A motion tween is an animation that is created by specifying different values for an object property in different frames.

- A property keyframe is a frame within a tween span where you explicitly define one or more property values for the tween target object.

- Director uses a movie metaphor to create interactive presentations.

- Cube mapping is a method of using a six-sided cube as the shape of the map.

- Render Textures are textures that are created and updated at runtime.

- In Director, the Cast window is where you view the list of media elements that appear in your movie.

- The action in a Director movie takes place in a window called the Stage.

15.6 UNIT END EXERCISES

1. Write a note on Adobe Flash.
2. Explain in brief, Adobe Director.
3. What is the use of Authorware?
4. What is a quest application?
5. Differentiate between motion tweens and classic tweens.
6. What are the top new features of Adobe Director 11.5.9?

Reference:

1. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

2. Introduction to Multimedia Systems, Chapter 16, Gaurav Bhatnager, Shikha Mehta, Sugata Mitra, © 2002 by ACADEMIC PRESS.

3. Designing Interactive Multimedia, paper by Lori L. Scarlatos, http://www.uni-mannheim.de/acm97/papers/Scarlatos/

4. Flash_reference.pdf available at Adobe's Website

5. Using_authorware_7.pdf available at Adobe's Website.

6. Director_11.5_help.pdf available at Adobe's Website.

❖❖❖❖