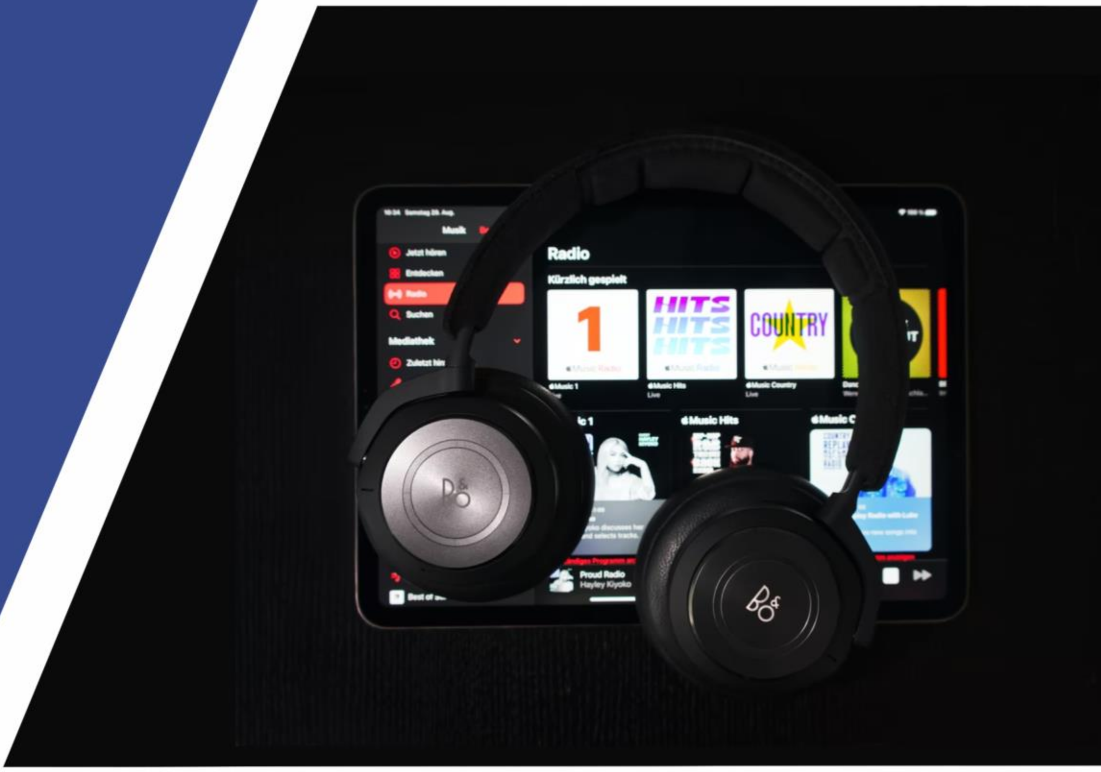
NATIONAL OPEN UNIVERSITY OF NIGERIA

CIT 891



Advanced Multimedia Technology

Module 1

CIT 891
Advanced Multimedia Technology
Module 1

Course Developer/Writer
Dr. Ikhu-Omoregbe, Nicholas, National Open University of Nigeria

Course Coordinator
A. A. Afolorunso, National Open University of Nigeria

Programme Leader
Prof. MoniOluwa Olaniyi, National Open University of Nigeria

Credits of cover-photo: Henry Ude, National Open University of Nigeria

National Open University of Nigeria - University Village, Plot 91, Cadastral Zone, Nnamdi Azikiwe Expressway, Jabi, Abuja-Nigeria.

www.nou.edu.ng centralinfo@nou.edu.ng
oer.nou.edu.ng oerunit@noun.edu.ng OER repository
Published in 2013, 2021 by the National Open University of Nigeria

How to re-use and attribute this content

Unit 1 Multimedia Systems and Requirements

1.0 Introduction

In this unit, we shall provide a foundation for this course on multimedia technologies. In achieving this, we attempt to refresh your skills / knowledge of some basic concepts in computer systems and related courses / topics such as signal processing, computer hardware and software, data communications, computer network, multimedia applications etc. Specifically, we shall describe the components of a multimedia system, explain some desirable features for a multimedia system, and provide other details that will help you understand the remaining parts of this course.

2.0 Objectives

At the end of this unit, you should be able to:
- explain the meaning of multimedia technology
- describe the component of multimedia systems
- explain some desirable features for multimedia systems.

3.0 Main Content

3.1 Overview of Multimedia Technology

Multimedia technology has emerged in the last few years as a major area of research. Multimedia computer systems have opened a wide range of potential applications by combining a variety of information sources such as voice, graphics, animation, images audio, and full motion video. Multimedia technology refers to both the hardware and software, and techniques used to create and run multimedia systems. Multimedia technology has its background in the merging of three industries; computer, communication, and broadcasting industries.

Many applications exist today as a result of the advances in multimedia technology. The mode of delivery for the applications depends on the amount of information that must be stored, the privacy desired, and the potential expertise of the users. Applications that require large amounts of data are usually distributed on Digital Versatile Disk - Read Only Memory (DVD-ROM), while personal presentations might be made directly from a computer using an attached projector. Also some of them can be viewed on the internet from a server or bundled into portable consumer electronics such as Personal digital assistants (PDAs), iPods, iPhones, mp3 payers, etc.

3.2 Definition of Multimedia

Multimedia simply means multiple forms of media integrated together. Media can be text, graphics, audio, video, animation, data etc. An example of multimedia is a blog that has text regarding an owner along with an audio file of some of his music and can even include selected videos of its owner. Besides multiple types of media being integrated with one another, multimedia can also stand for interactive types of media such as games, Digital Versatile Disk (DVD) or Compact Disk - Read Only Memory (CD-ROM) containing computer-aided learning instructions, animations or movies. Other terms that are sometimes used for multimedia include hypermedia, media, etc.

3.3 Multimedia Applications

We have seen a revolution in computer and communication technologies in the twentieth century. The telecommunications industry has experiences some dramatic innovations that allowed analog to digital networking that enabled today's very powerful internet technology. Transition from the analog to the digital world has offered many opportunities in the way we do things. Telecommunications, the Internet, digital entertainment, and computing in general are becoming part of our daily lives. Today, we are talking about digital networks, digital representation of images, movies and video, TV, voice, digital library – all because digital representation of signal is more robust than analog counterpart for processing, manipulations, storage, recovery and transmission over long distances, even across the globe through communication networks.

In recent years, there has been significant advancement in processing of still images, video, graphics, speech, and audio signals through digital computers in order to accomplish different applications challenges. As a result, multimedia information comprising image, video, audio, speech, text, and other data types has the potential to become just another data type. Telecommunications is no longer a platform for peer-to-peer voice communications between two people. Demand for communication of multimedia data through the telecommunications network and accessing the multimedia data through Internet is growing explosively.

In order to handle this pervasive multimedia data it is essential that the data representation and encoding of multimedia data be standard across different platforms and applications. As more portable consumer electronic devices continue to emerge, still images and video data comprise a significant portion of the multimedia data and they occupy the lion share of the communication bandwidth for multimedia communications. As a result, development of efficient image compression technique continues to be an important challenge in multimedia technology research.

With the increasing usage of multimedia systems, it is not uncommon to find them exist as standalone / workstations with associated software systems and tools, such as music composition, computer-aided learning, and interactive video or as distributed systems. The combination of multimedia computing with distributed systems in recent times, have offered greater potentials; new applications based on distributed multimedia systems including multimedia information systems, collaboration and conferencing systems, on-demand multimedia services and distance learning are all made possible today.

Generally, multimedia applications use a collection of multiple media sources e.g. text, graphics, images, sound/audio, animation and/or video. Examples of multimedia applications include:

- World Wide Web (WWW)
- Hypermedia courseware
- Video-on-demand
- Interactive TV
- Computer Games

- Virtual reality
- Digital video editing and production systems
- Multimedia Database systems
- Video conferencing and Teleconferencing
- Groupware
- Home shopping
- Games

3.4 Multimedia Systems

The word multimedia in a computer environment implies that many media are under computer control. Thus a multimedia computer is a system capable of processing multimedia data and applications. In its loosest possible sense, a multimedia computer should support more than one of the following media types: text, images, video, audio and animation. However, that means that a computer which manipulates only text and images would qualify as a multimedia computer. A Multimedia System is characterised by its capability to process, store, generate, manipulate and render multimedia information.

3.5 Components of Multimedia Systems

A multimedia systems is not too different from any other type of computer system except for it ability to process multimedia data. Thus, it should have features that can process audio, video, graphics and animation. Where these data would need to be transmitted, it should have enough memory and support adequate bandwidth / data compression features to minimise delays. Now let us consider the components (Hardware and Software) required for a multimedia system:

3.5.1 Input Devices/Output Devices

An input device is any piece of hardware use to accept any type of multimedia data for processing by the computer while an output device is any piece of computer hardware equipment used to communicate the results of multimedia processing carried out by an multimedia processing system (such as a computer) to the outside world. In computing, input/output, or I/O, refers to the communication between a multimedia processing system (such as a computer), and the outside world. Inputs are the signals or data sent to the system, and outputs are the signals or data sent by the system to the outside world.

The most common input devices used by the computer are the keyboard and mouse. The keyboard allows the entry of textual information while the mouse allows the selection of a point on the screen by moving a screen cursor to the point and pressing a mouse button. The most common outputs are monitors and speakers. Microphone is another input device that can interpret dictation and also enable us to input sound like the keyboard is used for text. A digital camera records and stores photographic images in digital form that can be fed to a computer as the impressions are recorded or stored in the camera for later loading into a computer. The digital cameras are available for still as well as motion pictures. Other capture devices include video recorder, graphics tablets, 3D input devices, tactile sensors, VR devices, etc.

Output devices exist in different forms. A printer for example, receives the signal from computer and transfers the information to paper. Printers operate in different ways, for example, the Dot-matrix printer strikes the paper a line at a time while inkjet sprays ink. The laser printer uses a laser beam to attract ink another example of an output device is the monitor. It is a device for display. It is just like a television set and is measured diagonally from two opposing corners of the picture tube. The standard monitor size is 14 inches. Very large monitors can measure 21 inches diagonal or greater. Another, common device that may be seen with a multimedia system is the amplifier. An amplifier is an electronic device that increases the power of a signal. Amplifiers are used in audio equipment. They are also called power equipment. Speakers with built-in amplifiers have become an integral part of the computers today and are important for any multimedia project.

3.5.2 Systems Devices

These are the devices that are the essential components for a computer. These include microprocessor, motherboard and memory. Microprocessor is basically the heart of the computer. It is a computer processor on a small microchip. When you turn your computer on, it is the microprocessor, which performs some operations. The microprocessor gets the first instruction from the Basic Input/Output System (BIOS), which is a part of its memory. BIOS actually load the operating system into random access memory (RAM). The motherboard contains computer components like microprocessor, memory, basic input / output system (BIOS), expansion slots and interconnecting circuitry. You can enhance the performance of your computer system by additional components to a motherboard through its expansion slot. RAM also called primary memory, locates the operating system, application programs, and data in current use so that the computer's processor reaches them quickly. RAM is called "random access" because any storage location can be accessed randomly or directly. RAM is much faster than the hard disk; the floppy disk, the CD-ROM and any other secondary storage device. But might get slow when used to its limit. That is the reason you need more memory to support multimedia applications.

3.5.3 Storage Devices

Hard disks, CD-ROMs, DVD-ROM, etc Storage device provides access to large amounts of data on an electro magnetically charged surface. Most personal computers typically come with a hard disk that contains several billion bytes (gigabytes) of storage. The popular ones are 40 GB and above. Hard disk contains a part called which is responsible for improving the time it takes to read from or write to a hard disk. The disk cache holds data that has recently been read. The other type of hardware cache inside your computer is cache memory. Cache is the term implies stores something temporarily. For example, Temporary Internet files are saved in Cache. On the other hand Compact Disc, read-only memory can store computer data in the form of text, graphics and sound. To record data into a CD, you need a writer. Normally this type of CD is either (CD-R) or (CD-RW). For the latter you can use the CD as a floppy disk write, erase and again write data into the same disk. In the CD-R, once the data recording is completed, it becomes a CD-ROM and nothing can be deleted. Photo CD: is a standard for storing high-resolution photographic images which can either be a as a pre mastered disc or a CD-WO disc. In the latter the images can be added to it.

A CD-ROM is a read-only, digital medium, whose mastering is expensive, but whose replication is relatively cheap. It current capacity is over 700 MB, it access time are less than 400 m/sec, and its transfer rate is 300 Kbs. A newer technology, the digital versatile disc

(DVD), stores much more in the same space and is used for playing back movies. DVD was originally said to stand for digital video disc, and later for digital versatile disc. DVD is an optical disc technology with a 4.7 gigabyte storage capacity on a single-sided, one-layered disk, which is enough for a 133-minute movie. DVDs can be single- or double-sided, and can have two layers on each side; a double-sided, two-layered DVD will hold up to 17 gigabytes of video, audio, or other information. This compares to 650 megabytes (.65 gigabyte) of storage for a CD-ROM disk. DVD uses the MPEG-2 file and compression standard. MPEG-2 images have four times the resolution of MPEG-1 images and can be delivered at 60 interlaced fields per second where two fields constitute one image frame. (MPEG-1 can deliver 30 noninterlaced frames per second.) Audio quality on DVD is comparable to that of current audio compact discs.

3.5.4 Communication Devices

A modem (modulator-demodulator) modulates digital signals going out from a computer or other digital devices to analog signals for a telephone line and demodulates the analog signal to convert it to a digital signal to be inputted in a computer. Some personal computers come with 56 Kilobits per seconds modems. Modems help your computer to connect to a network, communication networks such as local network, Intranets, Internet, Multimedia server, or other special high speed networks.

3.5.5 Additional Hardware

Having discussed the basic components that you will find on a standard computer system we shall now proceed to mention some additional devices you should expect to see on a system dedicated for multimedia processing. One of such devices is the video capture device. Video capture is the process of converting an analog video signal such as that produced by a video camera or DVD player to digital form. The resulting digital data are referred to as a digital video stream, or more often, simply video stream. Video capture from analog devices like video camera requires a special video capture card that converts the analog signals into digital form and compresses the data.

Video capture card use various components of the computer to pass frames to the processor and hard disk. Video-capture results will depend on the performance and capacity of all of the components of your system working together. For good quality video, a video-capture card must be able to capture full-screen video at a good rate. For example for a full-motion video, the card must be capable of capturing about 35 frames per second at 720 by 480 pixels for digital video and 640 by 480 for analog video. To determine what settings will produce the best results for your projects, you must be careful in defining these parameters. A video adapter provides extended capability to a computer in terms of video. The better the video adapter, the better is the quality of the picture you see. A high quality video adapter is a must for you while designing your developing multimedia applications.

Another device to mention here is the **sound card** which is a device that attaches to the motherboard to enable the computer to input, process, and deliver sound. The sound card generates sounds; records sound from analog devices by converting them to digital mode and reproduce sound for a speaker by reconverting them to analog mode. A standard example of this is Creative Lab's Sound Blaster.

3.6 Multimedia Workstations

A multimedia workstation can be defined as a computer system capable of handling a variety of information format; text, voice, graphics, image, audio, and full motion video. Advances in several technologies are making multimedia systems technically and economically feasible. These technologies include powerful workstations, high capacity storage devices, high-speed networks, advances in image and video processing (such as animation, graphics, still and full-motion video compression algorithm), advances in audio processing such as music synthesis, compression and sound effects and speech processing (speaker recognition, text-to-speech conversion and compression algorithms), asynchronous and ATM networks.

The main subsystem that could differentiate a multimedia workstation from tradition (non-multimedia) workstation include CD-ROM device, video and audio subsystem and multimedia related hardware (including image, audio and video capture storage and output equipment).

Video Subsystem
A video subsystem for multimedia system is usually composed of a video codec which provides compression and decompression of images and video data. This also performs video capture of TV- type signal (NTSC, PAL, SECAM-these are defined latter in this unit) from camera, VCR and laser disc, as well as playback of full-motion video. The playback part of the system should include logic to decode the compressed video stream and place the result in the display buffer depending on the functionalities supported. Where a video sub system exists for a PC-based multimedia system, it can be connected to a central system to receive data in real-time and allows for the scheduling.

An advanced video subsystem may include additional components for image and video processing. For example, the system can contain output connection for attachment to a monitor, which will allow the user to view live images during the capturing process.

An additional function of the video subsystem may be the mixing of real-time video images with video graphics array (VGA) computers graphics origination from the computer system.

Audio Subsystem
An audio subsystem provides recording, music synthesis and playback of audio data. Audio data is typically presented in one of three forms:

- Analog waveform
- Digital waveforms
- Musical Instrument Digital Interface (MIDI)

Analog waveform: audio is represented by an analog electrical signal whose amplitude specifies the loudness of the sound. This form is used in microphone, cassette tapes, records, audio amplifiers and speakers.

Digital waveform is represented using digital data. The digital audio has considerable advantages over analog audio, such as less indifference to noise and distortion. However, it involves larger processing and storage capacities. Digital devices which use digital waveforms audio format are compact disc, the digital audio tape (DAT) and the digital compact disc (DCD).

MIDI (Musical Instrument Digital Interface) refers to digital encoding of musical information where the sound data is not stored, and only the commands that describe how the music should be played are generated. MIDI gives the highest data compression, is easy for editing, and is widely accepted as a musical data standard. However, it requires additional hardware (music synthesiser) to generate music

An audio system could capture and perform digitisation of external audio signal through an Analog/Digital converter and generation of audio signal through a Digital/Analog converter. The digital signal processor (DSP) performs data compression and some additional processing functions, such as mixing and volume controls. Some advanced multimedia systems combine both video and audio subsystems into a unified / video subsystem.

Self-Assessment Exercise

1. List Five (5) Multimedia applications
2. Describe the three main forms of representing audio data

Multimedia Related Hardware
Multimedia related hardware includes video and audio equipment required at multimedia production and or presentation stages; these equipment (some already mentioned in the previous section) can be divided into:

- Image and Video capture equipment; still and video camera, scanner and video recorders
- Image and video storage equipment: laserdisc, videotapes and optical disks
- Image and video output equipment: displays, interactive display, TV projectors and printers
- Audio equipment: microphones, audio tape, recorders, video tapes recorder, audio mixers, head phones and speakers

There are several home TV distribution standards. These include PAL, SECAM, NTSC

There are three main analog color coding systems: Phase Alternation Line (PAL), Sequential Couleur Avec Memoire (SECAM) and National Television System Committee (NTSC). They differ mainly in the way they calculate the luminance/chrominance components from the Red Green Blue (RGB) components.

PAL - It is a European standard which uses a TV scan rate of 25 frames (50 half-frames) per seconds and a frequency of 625 lines / frames. Other countries where it is used are Australia and South America.

SECAM -It is French standard similar to PAL, but it uses different internal / video and audio frequencies. Besides France, SECAM is also used in Eastern Europe.

NTSC -It is the USA standard which is very different from PAL and SECAM standards. The frame rate in NTSC is 30 frames (60 half-frame per seconds and the frequency is 525 lines per frame. Other countries where this standard is used are Canada, Japan and Korea.

3.7 Desirable Features for a Multimedia Computer

By definition, a multimedia computer processes at least one media type that is either discrete or continuous in nature. Text and images are example of discrete media (i.e., they are time-independent), whereas video and audio are time-dependent, and consequently, continuous. The processing of time-independent media is meant to happen as fast as possible, but this processing is not time critical because the validity of the data does not depend on any time condition. However, in the case of time-dependent media, their values change over time - and, in fact, processing values in the wrong sequence can invalidate (part of) the data. In addition, multimedia systems are presently being faced with some of the following challenges:

- How to render different data at the same time — continuously;
- Sequencing within the media;
- Data retrieval from distributed networks;
- How to strictly maintain the temporal relationships on playback ;
- Retrieving and playing frames in correct order / time frame in video;
- Synchronisation ; inter-media scheduling-E.g. Video and Audio
- How to represent and store temporal information.;
- Multimedia data compression, etc

Given the above challenges the following feature are desirable (if not a prerequisite) for a Multimedia System:

- **Very high processing power.** This is needed to manage large amount of data and real-time delivery of media

- **Efficient and High Input / Output Devices interfaced with the file systems**. This is needed to allow for real-time recording as well as play back of data to users.

- **Special Operating System** —to allow access to file system, quick and efficient processing of data. It is needed to support direct transfers to disk, real-time scheduling, fast interrupt, processing, input/output streaming, etc.

- **Storage and Memory** — large storage units (of the order of hundreds of Gigabytes if not more) and large memory (several gigabytes or more). Large Caches should be provided and high speed buses for efficient management.

- **Network Support** — Required for inter-process controls and Client-server communications in distributed systems

- **Software Tools** — user-friendly tools needed to handle media, design and develop multimedia applications.

- **Bandwidth:** This is probably the most critical area for multimedia system, without sufficient bandwidth, multimedia applications are simply not practical. The challenge is not simply in providing X megabits per second of bandwidth. The bandwidth has to have the right characteristics to ensure the desired quality of service at all times.

4.0 Conclusion

As technology advances, so we expect new multimedia applications / products to be in the market. No doubt, a good number of new media technologies are being used to create complete multimedia experience. For instance, virtual reality integrates the sense of touch with video and audio media to immerse an individual into a virtual world. Other media technologies being developed include the sense of smell that can be transmitted via the Internet from one individual to another. Lots of multimedia entertainment software is available on the internet while others are bundled into portable consumer electronics. No doubt, as computers increase their power, new ways of integrating media will make the multimedia experience extremely exciting

5.0 Summary

In this unit we have covered multimedia systems and technology by learning basic definitions. In addition, the features of the hardware, software and network devices required for a multimedia system to function were discussed. Furthermore the desirable features for a multimedia computer system and some multimedia applications were covered. In the next unit, you shall learn about the elements of multimedia.

6.0 Self-Assessment Exercise

1a. What is the meaning of the term "multimedia"?

b. What do you consider the main requirements for multimedia systems?

2a. List three International television standard and state there features

b. List some additional devices that you expect to find in a multimedia computer system

7.0 References/Further Reading

Acharya,T. , -Sing Tsai,P. (2005). JPEG2000 *Standard for Image Compression* Concept. Algorithm and VLSI

Forouzan,B.A., Fegan,S.C. (2003). *Data Communication and Networking*. Singapore: McGraw Hill Higher Education.

Furht, B. *et al*. (1995). *Video and Image Processing in Multimedia Systems*. Kluwer Academic Publisher.

Heath,S. (1996). *Multimedia and Communications Technology.Oxford:* An imprint of Butterworth-Heinemann, Jordan Hill.

Unit 2 Elements of Multimedia

1.0 Introduction

In the previous unit, you have been introduced to multimedia systems and the basic hardware and software requirements of multimedia systems. You should not forget also that the fundamental characteristics of multimedia systems are that they incorporate and process multimedia such as text, colour, audio/voice, video, and animated graphics. In this unit you will study, the main features of this elements.

2.0 Objectives

At the end of this unit, you should be able to:
- describe the elements of multimedia
- describe some digital video formats.

3.0 Main Content

3.1 Text

There are three types of text that are processed by a multimedia computer and these are: **Unformatted text, formatted text and hypertext**. Texts are captured from the keyboard and hypertext can be followed by a mouse, keyboard, stylus or other devices when it is being run on a computer.

Unformatted text
Unformatted text, also referred to as plain text consists of strings of fixed sized characters from limited character set. An example of a character set that is widely use in computing is ASCII which stands for American Standard Code for Information Interchange. Normal alphabetic, numeric, punctuation and control characters are represented by ASCII character set and constitute the unformatted text.

Formatted text
This is also known as rich text. It enables documents comprising of strings of characters of different size, styles, and shape with tables, images and graphics to be inserted at some points within a document. An example of a formatted text is this course material which was produced by a word processing package. Most word processing packages such as Microsoft Word (MS-WORD) has features that allow a document to be created which consist of characters of different styles and of variable size and shape, each of which can be plain, bold, or italicised. To prepare this course material some features of MS-WORD were used in formatting the texts, preparing tables, inserting graphics, etc, at appropriate positions to make the material more interesting for the reader.

Hypertext
This refers to documents that contain unformatted or formatted text as well as links to other parts of the document, or other documents. The user can move to any section of the document accessible by selecting the link. The linked document may be on a single system or physically distributed across different systems. Hypertext that includes multimedia information such as sound, graphics, and video, is sometimes referred to as hypermedia.

3.2 Graphics/Images

Graphics are visual presentations on some surface, such as a canvas, wall, computer screen, paper, or stone to brand, inform, illustrate, or entertain. Examples are photographs, Line Art, graphs, diagrams, drawings, typography, numbers, symbols, geometric designs, maps, engineering drawings, or other images. Graphics often combine text, illustration, and colour. Graphics are usually generated by a graphics editor program (e.g. Illustrator) or automatically by a program (e.g. Postscript). Graphics files usually store the primitive assembly and do not take up a very high storage overhead. Graphics are usually editable or revisable (unlike Images). Input devices for capturing graphics include keyboard (for text and cursor control), mouse, and trackball or graphics tablet.

Images

Images may be two-dimensional, such as a photograph, screen display, and as well as a three-dimensional, such as a statue. They can be captured by scanner, digital camera for processing by a multimedia computer. In a broader sense, an image can be seen as any two-dimensional figure such as a map, a graph, a pie chart, or an abstract painting. In this wider sense, images can also be rendered manually, such as by drawing, painting, carving, rendered automatically by printing or computer graphics technology, or developed by a combination of methods, especially in a pseudo-photograph. To be more specific a still image is a single static image, as distinguished from a moving image. This phrase is used in photography, visual media and the computer industry to emphasise that one is not talking about movies.

Images are displayed (and printed) in the form of a two dimensional matrix of individual picture elements – known as pixels or sometimes pels. A computer display screen can be considered as being made up of a two-dimensional matrix of individual picture elements (pixel) each of which can have a range of colours associated with it. For example, VGA (Video graphic array) is a common type of display and consist of 640 horizontal pixels by 480 vertical pixels.

3.3 Audio/Sound

Voice and music, for example are by nature analog, so when we record voice or video, we have created an analog electric signal. They can be captured into the computer for processing via microphones and then digitised and stored. If we want to store the recording in the computer or send it digitally, we need to change it through a process called sampling. The term sampling means measuring the amplitude of the signal at equal intervals. After the analog signals is sampled, we can store the binary data in the computer or use line coding (or a combination of block coding and line coding) to further change the signal to a digital one so it can be transmitted digitally. Digital signals are less prone to noise and distortion. A small change in an analog signal can change the received voice substantially, but it takes a considerably change to convert a 0 to 1 or a 1 to 0. Two popular methods of Analog–to-digital conversion are the Pulse amplitude modulation (PAM) and the Pulse code modulation (PCM). A CD Quality Audio requires 16-bit sampling at 44.1 KHz.

3.4 Video

A still image is a spatial distribution of intensity that is constant with respect to time. Video, on the other hand, is a spatial intensity pattern that changes with time. Another common term for video is image sequence, since video can be represented by a time sequence of still images. Video has traditionally been captured, stored and transmitted in analog form. The term analog video signal refers to a one-dimensional (1-D) electrical signal of time that is obtained by sampling the video intensity pattern in the vertical and temporal coordinates and converting intensity to electrical representation. This sampling process is known as scanning.

A typical example of scanning is the Raster scanning. This begins at the top-left corner and progresses horizontally, with a slight slope vertically, across the image. When it reaches the right-hand edge it snaps back to the left edge (horizontal retrace) to start a new scan line. On reaching the bottom-right corner, a complete frame has been scanned and scanning snaps back to the top-left corner (vertical retrace) to begin a new frame. During retrace, blanking (black) and synchronisation pulses are inserted.

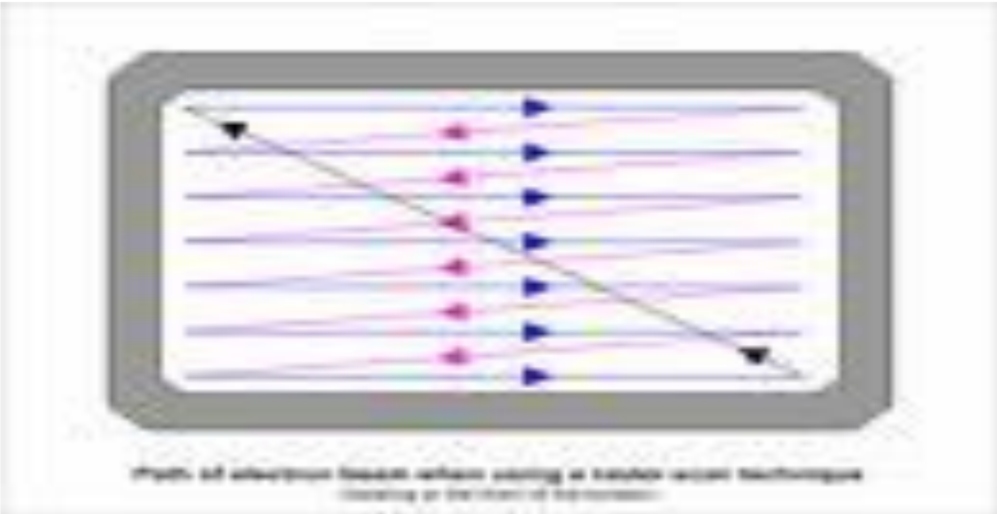


Fig. 2.1: Raster Scan

The aspect ratio, vertical resolution, frame rate, and refresh rate are important parameters of the video signal. The aspect ratio is the ratio of the width to the height of the frame. The vertical resolution is related to the number of scan lines per frame (including the blanking intervals). The frame rate is the number of frames scanned by second. The effect of smooth motion can be achieved using a frame rate of about 25-30 frames per second. However, at these frame rates the human eye picks up the flicker produced by refreshing the display between frames. To avoid this, the display refresh rate must be above 50 Hz.

3.4 Animation

Video may be generated by computer program rather than a video camera. This type of video content is normally referred to as computer animation or sometimes, because of the way it is generated, animated graphics. Animation is the rapid display of a sequence of images of 2-D or 3-D artwork or model positions in order to create an illusion of movement. It is an optical illusion of motion due to the phenomenon of persistence of vision, and can be created and demonstrated in a number of ways. The most common method of presenting animation is as a motion picture or video program, although several other forms of presenting animation also exist. . The typical frame rate required for animation is 15-19 frames per second.

4.0 Conclusion

In this unit, you have been exposed to the fundamental media elements i.e. text, images/graphics, audio, and video/animation.

5.0 Summary

Computer systems have capacities to manipulate multimedia elements. In the next unit, you shall be exposed to how these elements are represented within the computer system

6.0 Self-Assessment Exercise

Explain the following terms:
1. Frame rate
2. Pixel
3. Animation
4. Aspect ratio

7.0 References/Further Reading

Al-Mualla,M.E., Canagarajah,C.N., & Bull,D.R. (2002). *Video Coding for Mobile Communications*. California: Elsevier Science. Academic Press

Bohren,C. (2006). *Fundamentals of Atmospheric Radiation*. An Introduction with 400 Problems. Wiley-VCH.

Halsall,F. (2001). *Multimedia Communications, Applications, Networks, Protocols and Standard.* Pearson Education.

Unit 3 Multimedia Signal Representation and Processing

1.0 Introduction

In the preceding unit we covered the elements of multimedia. Multimedia technology has in the last few years led to the development of wide range of applications by combining a variety of information sources such as text, images, graphics, colours, animation, images audio, and full motion video. In order to effectively manipulate the varied data element, digitisation process is employed. This entails the conversion from other forms of storage, which are analog, to represent media as a set of binary data which can be manipulated by the machine. The time to access text or imaged-based data is usually short compared to that of audio and video data. Audio and video signals vary continuously with time as the amplitude of the speech, audio and video varies. This type of signal is called analog signal. The duration of applications involving audio / video can be relatively long.

If an application requires a single type of media say images, the basic form of representing the media elements (i.e. pixel) is used. However, in application that involves all media elements (text, picture/images, audio, video/animation) integrated together in some way; the four media type would be represented in digital form for ease of processing by the computer system.

2.0 Objectives

At the end of this unit, you should be able to:

- explain the basic concepts of multimedia element representation
- explain the principles of digitisation
- describe the techniques of analog - digital conversion
- explain the concept of colour representations

3.0 Main Content

3.1 Principles of Digitisation

Both sounds and images can be considered as signals, in one or two dimensions, respectively. Sound can be described as a fluctuation of the acoustic pressure in time, while images are spatial distributions of values of luminance or colour. Images can be described in their RGB (Red Green Blue) or HSB (Hue Saturation Value) components as will be discussed later. Any signal, in order to be processed by numerical computing devices, have to be reduced to a sequence of discrete samples, and each sample must be represented using a finite number of bits. The first operation is called sampling, and the second operation is called quantisation of the domain of real numbers.

The general properties relating to any time-varying analog signal are represented in Figure 3.1. As shown in part (a) of the figure, the amplitude of such signal varies continuously with time. The highest and lowest frequency components of the signal shown in Figure 3.1 (a) may be those shown in Figure 3.1 (b). Our assumption is based on a mathematical technique known as Fourier analysis. As we can use this theory to show that any time-varying analog signal is made up of a possibly infinite number of single-frequency sinusoidal signals whose amplitude and phase vary continuously with time relative to each other. You shall learn more about this theory in Module Two of this course.

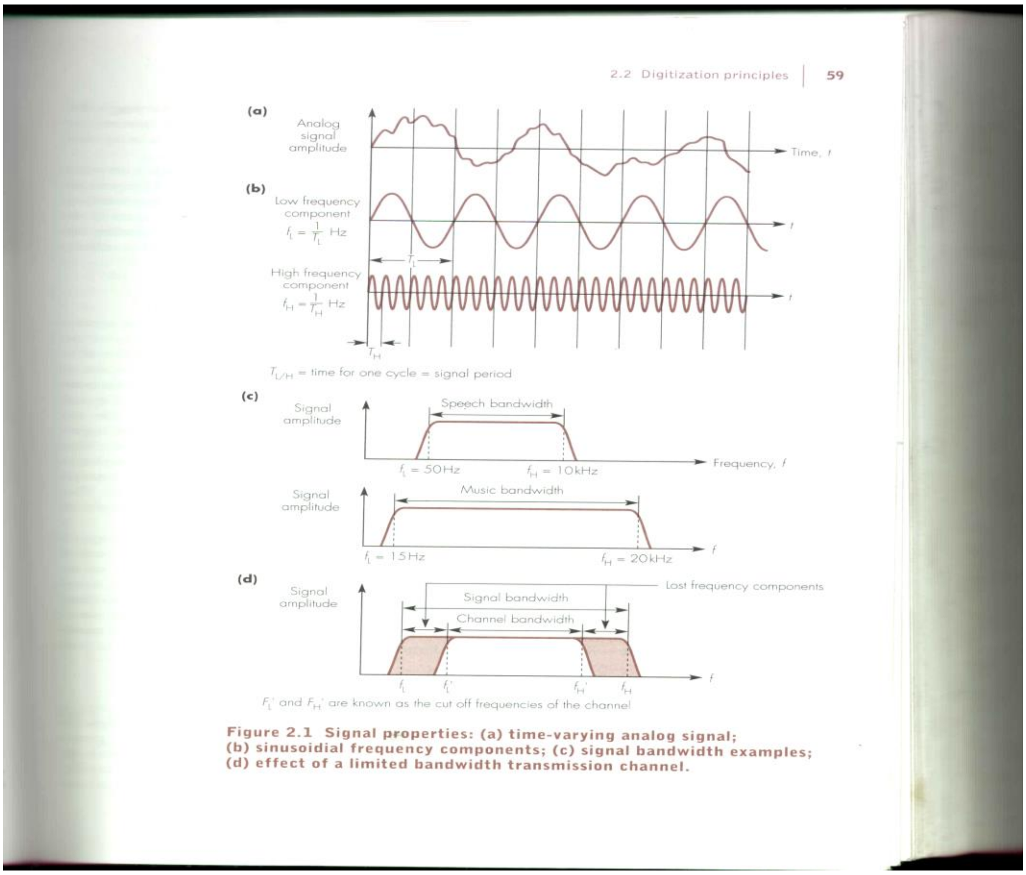


Fig. 3.1 (a – d)

Source: [Halsall,F.(2001).]

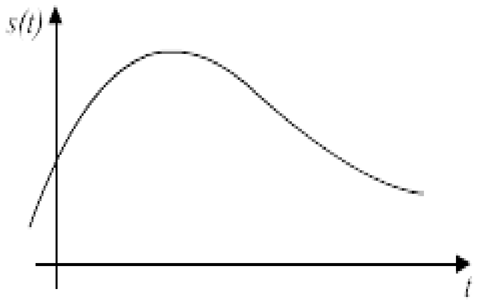
The range of frequencies of the sinusoidal components that make up a signal is called the **signal bandwidth.** We consider two examples as shown in Figure 3.1 (c). Here we are considering audio signals. Our first example is a speech signal and the second a music signal produced by say an orchestral.

Recall that in unit one, we stated that the microphone can be used as an input device in a multimedia system. When human speech is captured by a microphone, what it does is to convert it into electrical signal – that are made up of sinusoidal signal varying in frequency say between 50 Hz and 10kHz. In the case of music signal, however, the range of signal is wider and varies between 15 Hz and 20 kHz, this being comparable with the limits of the sensitivity of the ear.

From Data communication background, we should recall that when an analog signal is being transmitted through a network the bandwidth of the transmission channel – that is the range of frequencies the channel will pass – should be equal to or greater than the bandwidth of the signal. If the bandwidth of the channel is less than this, then some of the low and / or high frequency components will be lost thereby degrading the quality of the received signal. This type of transmission channel is called bandlimiting channel.

3.2 Sampling Rate

The sampling rate or sample rate is a term that defines the number of samples per second (or per other unit) taken from a continuous signal to make a discrete signal. For time-domain signals, it can be measured in samples per second (S/s), or hertz (Hz).



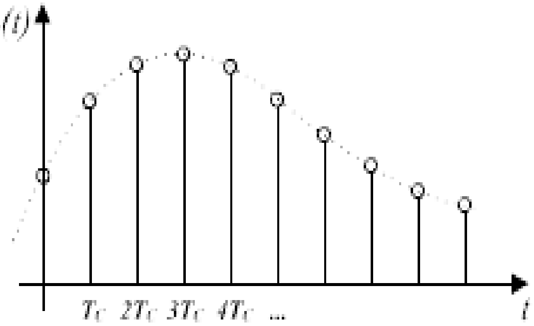


Fig. 3.2 a: Analog Signal **Fig. 3.2b: Resulting Sampled Signal**

In relation to the sampling rate, the Nyquist sampling theorem states that: in order to obtain an accurate representation of a time-varying analog signal, its amplitude must be sampled at a minimum rate that is equal to or greater than twice the highest sinusoidal frequency component that is present in the signal.

This is known as the Nyquist rate and is normally represented as either Hz or more correctly, samples per second (sps). Sampling a signal at a rate which is lower than Nyquist rate results in additional frequency components being generated that are not present in the original signal which in turn, cause the original signal to become distorted.

The distortion caused by sampling a signal at a rate lower than the Nyquist rate is best illustrated by considering the effect of undersampling a single-frequency sinusoidal signal as shown in Figure 3.3

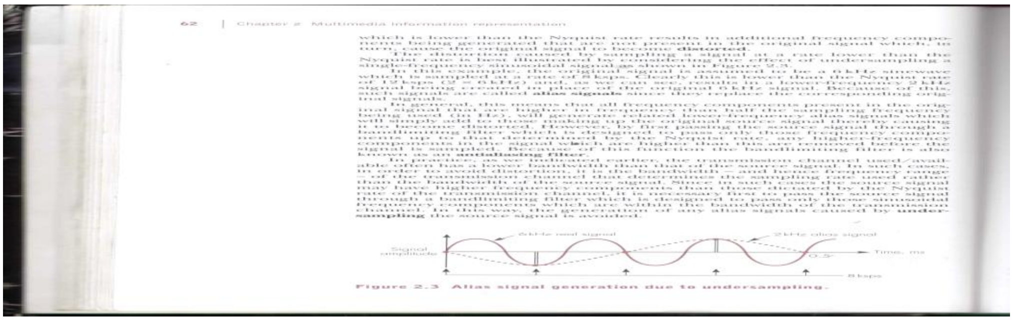


Fig. 3.3: Alias Signal Generation due to Under Sampling

***Source* [Halsall,F.(2001).]**

In the example, the original signal is assumed to be a 6kHz sinewave which is sampled at a rate of 8 kilo samples per second. Clearly, this is lower than the Nyquist rate of 12ksps (2 x 6 kHz) and, as we can see, results in a lower-frequency 2 kHz signal being created in place of the original 6 kHz signal. Because of this, such signal are called alias signals since they replace the corresponding original signals

In general, this means that all frequency components present in the original signal that are higher in frequency than half the sampling frequency being used (in Hz), will generate related low-frequency alias signals which will add to those making up the original source thereby causing it to become distorted. However, by first passing the source signal through a bandlimiting filter which is designed to pass only those frequency components up to that determined by the Nyquist rate, any higher-frequency components in the signal which are higher than this are removed before the signal is sampled. Due to this function the bandlimiting filter is also known as an antialiasing filter.

3.3 Nyquist's Theorem

From the preceding discussions, the sample rate must be chosen carefully when considering the maximum frequency of the analog signal being converted. Nyquist's theorem states that the minimum sampling rate frequency should be twice the maximum frequency of the analog signal. A 4Hz analogue signal would need to be sampled at twice that frequency to convert it digitally. For example, a hi-fi audio signal with a frequency range of 20 Hz to 20 KHz would need a minimum sampling rate of 40 kHz. Higher frequency sampling introduces a frequency component which has to filtered out using an analog filter.

In the process of digitising video, if the frequency content of the input analog signal exceeds half the sampling frequency, aliasing artifacts will occur. a filtering operation is used to bandlimit the input signal and condition it for the following sampling operation. The amplitude of the filtered analog signal is then sampled at specific time instants to generate a discrete-time signal. The minimum sampling rate is known as the Nyquist rate and is equal to twice the signal bandwidth.

The resulting discrete-time samples have continuous amplitudes. Thus, it would require infinite precision to represent them. The quantisation operation is used to map such values onto a finite set of discrete amplitude that can be represented by a finite number of bits.

3.4 Encoders and Decoders for Multimedia Applications

Signal Encoders
The conversion of an analog signal into digital form is carried out using an electrical circuit known as a signal encoder. An encoder is a device used to change a signal (such as a bitstream) or data into a code. Encoders serve any of a number of purposes such as compressing information for transmission or storage, encrypting or adding redundancies to the input code, or translating from one code to another. This is usually done by means of a programmed algorithm, especially if any part is digital, while most analog encoding is done with analog circuitry.

Signal Decoder
Similarly, the conversion of the stored digitised sample relating to a particular media type into their corresponding time-varying analog form is performed by an electrical circuit known as signal decoder. A decoder is a device which does the reverse of an encoder, undoing the encoding so that the original information can be retrieved. The same method used to encode is usually just reversed in order to decode. In digital electronics, a decoder can take the form of a multiple-input, multiple-output logic circuit that converts coded inputs into coded outputs, where the input and output codes are different.

3.4.1 Encoder Design

The conversion of a time-varying analog signal such as an audio signal into digital form is carried out using an electronic circuit known as signal encoders. The principles of an encoder are shown in figure 3.4 and, as we can see in part (a), it consist of two main circuits; a bandlimiting filter and an analog-to-digital converter (ADC), the latter comprising a sample-and-hold and a quantiser. A typical waveform set for a signal encoder is shown in part (b) of the figure. The function of the bandlimiting filter is to remove selected higher-frequency components from the source signal (A). The output of the filter (B) is often then fed to the sample-and-hold circuit which, as its name implies is used to sample the amplitude of the filtered signal at regular time intervals (C) and to hold the sample amplitude constant sample (D). After which, it is fed to the quantiser. This converts each sample amplitude into binary value known as a codeword (E).

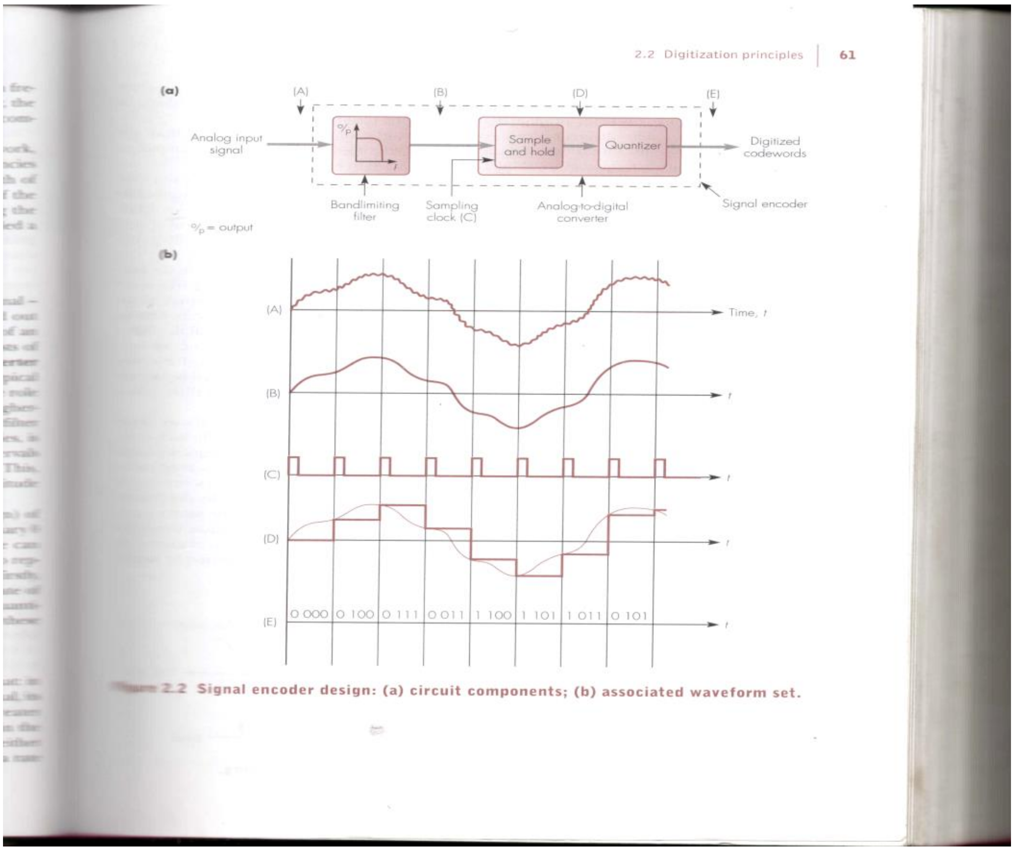
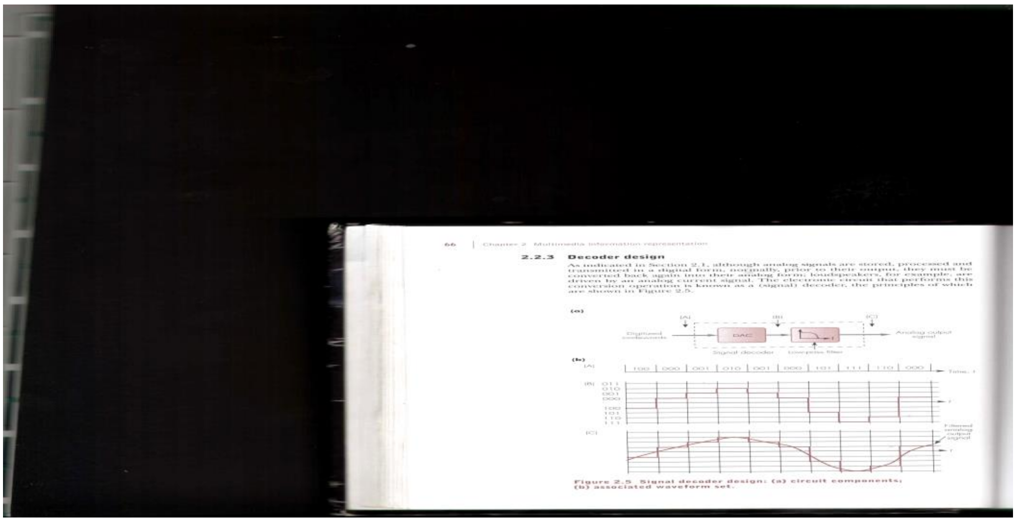


Fig. 3.4: Analog to Digital Signal Conversion

***Source*: [Halsall,F.(2001).]**

3.4.2 Decoder Design

Analog signals stored, processed and transmitted in a digital form, normally, prior to their output must be converted back again into their analog form. The loudspeakers, for example, are driven by an analog current signal. The electronic circuit that performs this conversion is known as a (signal) decoder, the principles of which are shown in figure 3.5 As depicted in the diagram, first, each digital codeword (A) is converted into an equivalent analog sample using a circuit called a digital-to-analog converter (DAC). This produces the signal shown in (B), the amplitude of each level being determined by the corresponding codeword.


66 | Chapter 2 Multimedia information representation

2.2.3 **Decoder design**

As indicated in Section 2.1, although analog signals are stored, processed and transmitted in a digital form, normally, prior to their output, they must be converted back again into their analog form; loudspeakers, for example, are driven by an analog current signal. The electronic circuit that performs this conversion operation is known as a (signal) decoder, the principles of which are shown in Figure 2.5.

Figure 2.5 Signal decoder design: (a) circuit components; (b) associated waveform set.

Fig.3.5

Source Halsall, F.(2001).]

In order to reproduce the original signal, the output of the DAC is transmitted through a low-pass filter which, as its name implies, only passes those frequency components that made up the original filtered signal (C). Usually, the high-frequency cut-off of the low-pass filter is made the same as that used in the bandlimiting filter of the encoder. In this case, the main function of the filter is for reconstruction. Hence, the low-pass filter is known as a recovery or reconstruction filter.

Codec

Many multimedia applications involves audio and video in which the communications channel is usually two-way simultaneous. Thus any terminal equipment used should support both input and output simultaneously. This often informs the need to have the audio / video signal encoder in each of the terminal equipment to be combined into a single unit called audio/ video encoder-decoder or simply an audio/video codec.

3.5 Colour Representations

Colour is the visual perceptual property corresponding in humans to the categories called *red*, *yellow*, *blue* and others. Colour derives from the spectrum of distribution of light energy versus wavelength interacting in the eye with the spectral sensitivities of the light receptors. Electromagnetic radiation is characterised by its wavelength (or frequency) and its intensity. When the wavelength is within the visible spectrum (the range of wavelengths humans can perceive, approximately from 380 nm to 740 nm), it is known as "visible light".

Table 3.5: Colours

Colour	Wavelength interval	Frequency interval
Red	~700-635 nm	~430-480 THz
Orange	~635-590 nm	~480-510 THz
Yellow	~590-560 nm	~510-540 THz

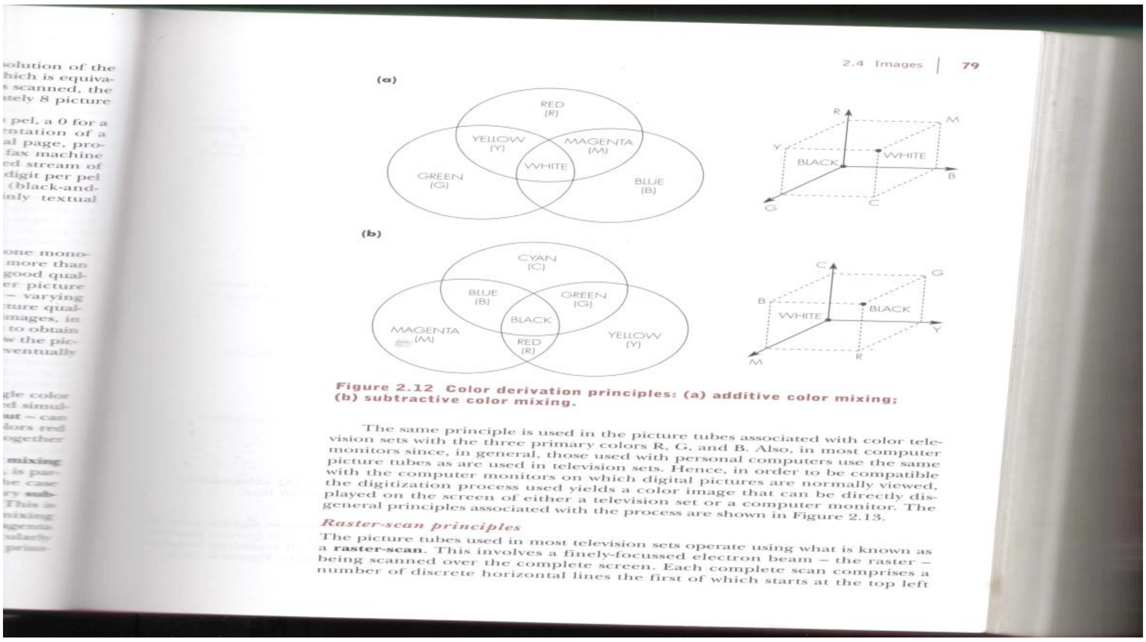
Green	~560-490 nm	~540-610 THz
Blue	~490-450 nm	~610-670 THz
Violet	~450-400 nm	~670-750 THz

Colour categories and physical specifications of colour are also associated with objects, materials, light sources, etc., based on their physical properties such as light absorption, reflection, or emission spectra. Colour space can be used as a model to identify colours numerically; for example, a colour can be specified by their unique RGB and HSV values.

3.6 Colour Principles

It is a known fact, that human eyes see a single colour when a particular set of three primary colours are mixed and displayed simultaneously. The truth is that, a whole spectrum of colours – known as a colour gamut – can be produced by using different proportions of the three primary colours red (R), green (G) and blue (B) . This principle is shown in figure 3.6 together with some example of colours that can be produced.

The missing technique used in part (a) is known as additive colour mixing which, since black is produced when all three primary colours are zero, is particularly useful for producing a colour image on a black surface as is the case in display application. It is also possible to perform complementary subtraction colour mixing operation to produce a similar range of colours. This is shown in part (b) of the figure and as we see, with subtractive mixing white is produced when the three chosen primary colours cyan (C), magenta (M), and yellow (Y) are all zero. Hence this choice of colours is particularly useful for producing a colour image on white surface as the case in printing application.


2.4 Images | 79

(a)

(b)

Figure 2.12 Color derivation principles: (a) additive color mixing; (b) subtractive color mixing.

The same principle is used in the picture tubes associated with color television sets with the three primary colors R, G, and B. Also, in most computer monitors since, in general, those used with personal computers use the same picture tubes as are used in television sets. Hence, in order to be compatible with the computer monitors on which digital pictures are normally viewed, the digitization process used yields a color image that can be directly displayed on the screen of either a television set or a computer monitor. The general principles associated with the process are shown in Figure 2.13.

Raster-scan principles

The picture tubes used in most television sets operate using what is known as a **raster-scan**. This involves a finely-focussed electron beam – the raster – being scanned over the complete screen. Each complete scan comprises a number of discrete horizontal lines the first of which starts at the top left

Fig 3.6: Colour Derivative Principles a) Additive Colour Mixing b) Subtractive Colour Missing Source [Halsall,F. (2001).]

This is the principles used in the picture tubes associated with colour television set and most computer monitors in the formation of images. The three main properties of a colour source that the eye makes use of are:

- Brightness: this represents the amount of energy that simulates the eye and varies on a gray scale from black (lowest) through to white (highest). It is thus independent of the colour of the source;
- Hue: this represents the actual colour of the source, each colour has a different frequency / wavelength and the eye determines the colour from this;
- Saturation: this represents the strength or clarity of the colour, a pastel colour has a lower level of saturation than a colour such as red. Also a saturated colour such as red has no white light in it.

Another term we must define is luminance. This term refers to the brightness of a source.

As we saw in section 3.6 a range of colours can be produced by mixing the three primary colours R, G, and B. In a similar way a range of colours can be produced on a television display screen by varying the magnitude of the three electrical signals that energise the red, green, and blue phosphorous. For example, if the magnitude of the three signal are in the proportion 0.299R + 0.587G + 0.114B

Then the colour white is produced on the display screen. Hence, since the luminance of a source is only a function of the amount of white light it contains, for any colour source its luminance can be determined by summing together the three primary components that make up the colour in the proportion. That is;

$Y_S = 0.299R_S + 0.587G_S + 0.114B_S$

Where Y_S is the amplitude of the luminance signal and R_S, G_S, and B_S are the magnitude of the three colour component signals that make up the source. Thus, since the luminance signal is a measure of the amount of white light it contains, it is the same as the signal used by a monochrome television. Two other signals, the blue chrominance (Cb) and the red chrominance (Cr), - are then used to represent the colouration – hue and saturation – of the source. These are obtained from the two colour difference signals:

Cb = Bs- Ys and Cr = Rs – Ys

Which, since the Y signal has been subtracted in both cases, contain no brightness information. Also, since Y is the function of all three colours, then G can be readily computed from these two signals. In this way, the combination of the three signals Y, Cb and Cr contains all the information that is needed to describe a colour signal while at the same time being compatible with monochrome television which use the luminance signal only.

Chrominance Components

In practice, although all colour television systems use this same basic principles to represent the colouration of a source, there are some small differences between the two systems in terms of the magnitude used for the chrominance signals.

3.7 Colour Models

A colour model is a method for specifying colours in some standard way. It generally consists of a three dimensional coordinate system and a subspace of that system in which each colour is represented by a single point. We shall investigate three systems.

RGB

In this model, each colour is represented as three values R, G, and B, indicating the amounts of red, green and blue which make up the colour. This model is used for displays on computer screens; a monitor has three independent electron "guns" for the red, green and blue component of each colour. Some colours require negative values of R, G, or B. These colours are not realisable on a computer monitor or TV set, on which only positive values are possible. The colours corresponding to positive values form the RGB gamut; in general a colour "gamut" consists of all the colours realisable with a particular colour model.

The RGB colour model is an additive colour model in which red, green, and blue light are added together in various ways to reproduce a broad array of colours. The name of the model comes from the initials of the three additive primary colours, red, green, and blue.

The main purpose of the RGB colour model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB colour model already had a solid theory behind it, based on human perception of colours. RGB is a device-dependent colour space: different devices detect or reproduce a given RGB value differently, since the colour elements (such as phosphors or dyes) and their response to the individual R, G, and B levels vary from manufacturer to manufacturer, or even in the same device over time. Thus an RGB value does not define the same colour across devices without some kind of colour management. Typical RGB input devices are colour TV and video cameras, image scanners, and digital cameras. Typical RGB output devices are TV sets of various technologies (CRT, LCD, plasma, etc.), computer and mobile phone displays, video projectors, multicolour LED displays, and large screens as JumboTron, etc. Colour printers, on the other hand, are not RGB devices, but subtractive colour devices.

HSV

HSV stands for Hue, Saturation and Value. These terms have the following meanings:

Hue: The "true colour" attribute (red, green, blue, orange, yellow, and so on).

Saturation: The amount by which the colour is been diluted with white. The more white in the colour, the lower the saturation. So a deep red has high saturation, and a light red (a pinkish colour) has low saturation.

Value: The degree of brightness: a well lit colour has high intensity; a dark colour has low intensity.

This is a more intuitive method of describing colours, and as the intensity is independent of the colour information, this is a very useful model for image processing.

Note that a conversion can be made between RGB and HSV.

YIQ
This colour space is used for TV/video in America and other countries where NTSC is the video standard (Australia uses PAL). In this scheme Y is the "luminance" (this corresponds roughly with intensity), and I and Q carry the colour information. The conversion between RGB is straightforward

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

and

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} = \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

The two conversion matrices are of course inverse of each other. Observe the difference between Y and V

Y= 0.299R + 0.587G + 0.114B

V=max {R, G, B}

This reflects the fact that human visual system assigns more intensity to the green component of an image than the red and blue components.

Self-Assessment Exercise

i. In your own words, explain the following terms;
   a. Luminance
   b. Chrominance
ii. Explain the Nyquist's Theorem

4.0 Conclusion

Multimedia technology has in the last few years lead to the development of wide range or applications by combining a variety of information sources such as text, images, graphics, colours, animation, images audio, and full motion video. In order to effectively manipulate the varied data element, digitisation process is employed. This entails the conversion from other forms of storage, which are analog, to represent media as a set of binary data which can be manipulated by the machine. There is one level at which text has been digitalised since the advent of computers, although text with print-like features did not become available until much later. Audio came next with the advent of digital studio, then still images began to move to digital formats, and the same process is now being used for video and animated graphics. The possibilities offered by digitisation have no doubt allowed a number of new applications and services to be introduced. Examples of such services include; Video on demand, high definition television, videoconferencing, medical imaging, surveillance, flight simulation etc.

5.0 Summary

All types of multimedia information are stored and processed within the computer in a digital form. Textual information captured via the keyboard made up of characters is represented by a unique combination of a fixed number of bits – known as codeword. Images such as a line and arc are represented by pixel indicating the start and end coordinates of the line relative to the complete image. Audio and video data captured by microphones and digital camera respectively produce electrical signals whose amplitude vary continuously with time known as an analogue signals. This is usually converted to digital signal for ease of processing by the computer system and reconverted to analog signals for display on desired devices/display. We also covered colour principles in this unit.

6.0 Self-Assessment Exercise

1. Explain the meaning of the following terms:
   a.Codeword
   b. analog signal
   c. signal encoder
   d. signal decoder
2. Explain the three popular models for color specification

7.0 References/Further Reading

Halsall,F. (2001). *Multimedia Communications, Applications, Networks, Protocols and Standard.,* Pearson Education.

Heath,S. (1996). *Multimedia and Communications Technolog.Oxford:,* An imprint of Butterworth-Heinemann. Jordan Hill.

McAndrew,A. (2004). *An Introduction to Digital Image Processing with Matlab*. School of Computer Science and Mathematics, Victoria University of Technology.

http://upload.wikimedia.org/wikipedia/commons/6/6e/Sampled_signal.png.